AI and IoT for Production Data Analytics in SMEs

Hansen, Emil Blixt

DOI (link to publication from Publisher):
10.54337/aau517403146

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

AI and IoT for Production Data Analytics in SMEs

BY
Emil Blixt Hansen

DISSERTATION SUBMITTED 2022

AALBORG UNIVERSITY
DENMARK

AI and IoT for Production Data Analytics in SMEs

PhD Thesis

Emil Blixt Hansen

Thesis submitted November, 2022

Dissertation submitted:	November, 2022
PhD supervisor:	Assoc. Prof. Simon Bøgh Aalborg University
Assistant PhD supervisor:	Assoc. Prof. Nadeem Iftikhar University College of Northern Denmark
PhD committee:	Associate Professor Peter Nielsen (chair) Aalborg University, Denmark Associate Professor Elias Ribeiro da Silva University of Southern Denmark, Denmark Professor Heri Ramampiaro Norwegian University of Science and Technology, Norway
PhD Series:	Faculty of Engineering and Science, Aalborg University
Department:	Department of Materials and Production

ISSN (online): 2446-1636
ISBN (online): 978-87-7573-793-2

Published by:
Aalborg University Press
Kroghstræde 3
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Stibo Complete, 2022

Curriculum Vitae

Emil Blixt Hansen

Emil Blixt Hansen grew up in the town of Hobro, Denmark. In 2014 he started on the newly created Robotics Bachelor education at Aalborg University. He received his BSc. diploma in 2017 and started on the Master's education in Manufacturing Technology at Aalborg University. In 2019 he received his MSc. diploma and started as a research assistant at Aalborg University. In September of 2019, he started as a PhD student at Aalborg University in the Robotics and Automation group. The PhD was a part of the Innovation Factory North project, which included research with danish manufacturing SMEs. During his PhD, he had an external collaboration and stayed abroad at the Norwegian University of Science and Technology, Norway, in the department of Computer Science. Professor Helge Langseth hosted the collaboration.

Emil's research is mainly on AI and IoT within manufacturing SMEs. A particular focus has been on understanding the challenges of SMEs along with data-driven anomaly detection and IoT implantation.

Abstract

With the rise of Industry 4.0 and its technologies, manufacturing companies have many new opportunities and methods to optimise their business. Larger enterprises commonly have dedicated teams to explore new technologies and take advantage of them. They have the benefits of in-house knowledge and financial resources. Small and medium-sized enterprises (SMEs) do not have these benefits. They often lack knowledge of what new technologies are available and how to use them, and even more so, they lack financial resources. Some of these technologies are the internet of things (IoT) and artificial intelligence (AI), which have shown many benefits for manufacturing companies. Nonetheless, even with the addressed benefits, SMEs general neglect these technologies.

This PhD investigates the topic of AI and IoT within SMEs. Firstly this is done by a literature review of AI and IoT used in SMEs. This literature review discovered different patterns, and thus, more research on why and how SMEs could use these technologies was conducted.

The PhD proposes two different types of AI architectures for SMEs. The first one is an easy-to-use machine learning platform where all the complex model parameters are hidden from the operator. The other is a novel approach to building an arbitrary machine's health indicator, again with the underlying model hidden from the operator. The PhD also contributes with a new publicly available dataset and general guidelines on anonymising such datasets.

Implementing IoT infrastructure can be done in numerous ways. This PhD study presents a field study on how it can be done at an SME. Further experiments are conducted on how this data can be used to control critical process control parameters. Finally, the perspective and implications of the company are also presented.

Resumé

Med Industri 4.0 og alle teknologierne der følger med, har produktionsvirksomheder mange nye muligheder og metoder til at optimere deres forretning. Større virksomheder har ofte dedikerede teams til at udforske nye teknologier og drage fordel af dem. De har f.eks. fordelene med mere intern viden og er økonomiske ressourcemæssigt stærke. De små og mellemstore virksomheder (SMV'er) har ikke disse fordele. De mangler ofte viden om, hvilken type nye teknologier der er tilgængelige, og hvordan de kan bruges, og ydermere er de ressourcemæssigt svage. Nogle af teknologierne er internet of things (IoT) og kunstig intelligens (AI), som har vist sig at have mange fordele for produktionsvirksomheder. Ikke desto mindre forsømmer SMV'er generelt disse teknologier.

Denne Ph.D. undersøger emnet AI og IoT inden for SMV'er. Det startes med en litteraturreview af kunstig intelligens og IoT brugt i SMV'er. Denne litteraturreview opdagede forskellige mønstre, og derfor blev der udført mere forskning i, hvorfor og hvordan SMV'er kunne bruge disse teknologier.

Ph.D.'en kommer med to forskellige bud på AI-arkitekturer til SMV'er. Den første er en nem at bruge maskinlæringsplatform, hvor alle de komplekse modelparametre er skjult for operatøren. Den anden er en ny tilgang til, hvordan man kan bygge en sundhedsindikator til en vilkårlig maskine, igen med den underliggende model skjult for operatøren. Ph.D.'en bidrager også med et nyt offentligt tilgængeligt datasæt sammen med generelle retningslinjer for, hvordan man kan anonymiserer datasæt.

Implementering af IoT-infrastruktur kan gøres på mange måder. Dette Ph.D.-studie præsenterer et feltstudie i, hvordan det kan gøres i en SMV. Derudover, udføres der eksperimenter i, hvordan disse data kan bruges til at kontrollere kritiske processtyringsparametre. Til sidst præsenteres også perspektivet og hvilke implikationerne der har for virksomheden.

Contents

Thesis Details

Thesis Title: AI and IoT for Production Data Analytics in SMEs
PhD Student: Emil Blixt Hansen
Supervisors: Assoc. Prof. Simon Bøgh, Aalborg University
Assoc. Prof. Nadeem Iftikhar, University College of Northern Denmark

The main body of this thesis consists of the following papers.

[Paper A | [1]] Emil Blixt Hansen and Simon Bøgh, "Artificial intelligence and internet of things in small and medium-sized enterprises: A survey", *Journal of Manufacturing Systems*, vol. 58, pp. 362–372, 2021.

[Paper B | [2]] Emil Blixt Hansen, Nadeem Iftikhar, and Simon Bøgh, "Concept of easy-to-use versatile artificial intelligence in industrial small & medium-sized enterprises", *Procedia Manufacturing*, Vol. 51, pp. 1146–1152, 2020.

[Paper C | [3]] Emil Blixt Hansen, Emil Robenhagen van der Bijl, Mette Busk Nielsen, Morten Svangren Bodilsen, Simon Vestergaard Berg, Jan Kristensen, Nadeem Iftikhar, and Simon Bøgh, "A New Authentic Cloud Dataset from a Production Facility for Anomaly Detection", *Towards Sustainable Customization: Bridging Smart Products and Manufacturing Systems*, pp. 415–422, 2022.

[Paper D | [4]] Emil Blixt Hansen, Helge Langseth, Nadeem Iftikhar, and Simon Bøgh, "A data-driven modular architecture with denoising autoencoders for health indicator construction in a manufacturing process", *IEEE Access*, in peer review, submitted August 2022.

[Paper E | [5]] Emil Blixt Hansen and Simon Bøgh, "Artificial Intelligence and Machine Learning", *The Future of Smart Production for SMEs*, pp. 323–326, 2023.

[Paper F | [6]] Emil Blixt Hansen, Nadeem Iftikhar, and Simon Bøgh, "On the topic of anonymising production data for machine learning", *Journal of Intelligent Manufacturing*, in second peer review, submitted April 2022.

[Paper G | [7]] Emil Blixt Hansen, Kent Lindegaard Hansen, Nadeem Iftikhar, and Simon Bøgh, "An in-depth investigation of machine learning and IoT adoption at a manufacturing SME: A field study", *32nd International Conference on Flexible Automation and Intelligent Manufacturing (FAIM)*, submitted October 2022.

In addition to the main papers, the following publications have also been made.

[1] Emil Blixt Hansen, Rasmus Eckholdt Andersen, Steffen Madsen, and Simon Bøgh, "Transferring Human Manipulation Knowledge to Robots with Inverse Reinforcement Learning," *IEEE/SICE International Symposium on System Integration (SII)*, pp. 933-937, 2020.

[2] Alexander S. Staal, Carolina G. Salvatierra, Ditte D. Albertsen, Mathiebhan Mahendran, Rahul Ravichandran, Rasmus F. Thomsen, Emil Blixt Hansen, and Simon Bøgh, "Towards a Collaborative Omnidirectional Mobile Robot in a Smart Cyber-Physical Environment," *Procedia Manufacturing*, vol. 51, pp. 193-200, 2020.

[3] Bence Bejczy, Rohat Bozyil, Evaldas Vaičekauskas, Sune B. Krogh Petersen, Simon Bøgh, Sebastian Schleisner Hjorth, and Emil Blixt Hansen, "Mixed Reality Interface for Improving Mobile Manipulator Teleoperation in Contamination Critical Applications," *Procedia Manufacturing*, vol. 51, pp. 620-226, 2020.

[4] Charles Møller, Andreas Kornmaaler Hansen, Dan Palade, Daniel Grud Hellerup Sørensen, Emil Blixt Hansen, Jonas Nygaard Uhrenholt, and Maria Stoettrup Schioenning Larsen, "An Action Design Research Approach to Study Digital Transformation in SME," *The Future of Smart Production for SMEs*, pp. 51-65, 2023.

[5] Charles Møller, Andreas Kornmaaler Hansen, Dan Palade, Daniel Grud Hellerup Sørensen, Emil Blixt Hansen, Jonas Nygaard Uhrenholt, and Maria Stoettrup Schioenning Larsen , "Innovation Factory North: An Approach to Make Small and Medium Sized Manufacturing Companies Smarter," *The Future of Smart Production for SMEs*, pp. 113-126, 2023.

[6] Simon Bøgh, Daniel S. Hain, Emil Blixt Hansen, Simon Buus Jensen, Torben Tvedebrink, and Roman Jurowetzki, "Predictive analytics appli-

cations for small and medium-sized enterprises (SMEs) – A mini survey and real-world use cases," *The Future of Smart Production for SMEs*, pp. 263-279, 2023.

This thesis has been submitted for assessment in partial fulfilment of the PhD degree. The thesis is based on the submitted or published scientific papers listed above. Parts of the papers are used directly or indirectly in the extended summary of the thesis. As part of the assessment, co-author statements have been made available to the assessment committee and are also available to the faculty. The thesis is not in its present form acceptable for open publication but only in limited and closed circulation as copyright may not be ensured.

Preface

Normally, the preface is the last thing you will write in a book. You will use it to write how you ended up writing the book, what your motivation was and justify your authorship. This part I will try to make short. Before writing the conclusion and abstract, I write this preface, reflecting upon the thesis and the PhD itself. I got into doing a PhD by a collection of different factors I would never be able to sum up – although some factors had greater influence than others. To quote Mike Shinoda's song "Remember the Name": *...This is 10 percent luck, 20 percent skill, 15 percent concentrated power of will, 5 percent pleasure, 50 percent pain...* this is what a PhD could be described as. While the exact percentage is probably off, the level of *luck* is a significant factor for me even to be given this opportunity. In general, people who reach great success seldom acknowledge the level of luck. I have doubted that I was fit for a PhD position, sometimes even being afraid of being called a fraud. If it is some miniature version of the impostor syndrome or just the good old danish "Jantelov" playing mind tricks, I do not know. What I know is that a lot of effort and time was spent reading papers, writing papers, re-writing papers, conceptualising, programming, fixing bugs, learning and teaching others. Reflecting upon it all fills me with joy, and now that eight years (three bachelor years, two master years, three PhD years) of presence at Aalborg University is coming to an end, it is an emotional feeling.

With that brief summary of how I ended up doing a PhD out of the way, I will spent the rest of this "free space" to address something important. With the digitalisation and the fast knowledge sharing possibilities of the 21st century, the research community is still anchored in the past. This is not an attack on all the researchers out there; quite the opposite. So much excellent research is being conducted and published every day around the globe. Researchers built their research on prior research to, e.g. improve technology or understand complicated behaviour. All in all, research is a foundation for improving the quality of life for all people on earth – even though some research is only focused on, e.g. improving SMEs in Denmark. The different government bodies know this, and they pour a considerable amount of money into research. Since the research is for the common good and a large percentage is paid from government bodies, the ordinary Joe should be able to read this research, right? No, as it is now and always has been, research is for the few. It is for the researchers who research at wealthy research institutions and can afford to buy the costly publisher outlets. If you work at a less wealthy research institute, you might not have access to all of the publishers. Even more so, if you are a private person, you depend on whether researchers paid extra to the publisher to make it open access or if the researchers made it available by sharing e.g. a pre-print. While the peer review process is paramount for securing research integrity, it is not mutually exclusive with the open access principles. I will be the first to admit that I have not exclusively published in open access, but I encourage my fellow researcher to do better than me. I encourage them to do what is necessary to make their research public and look further than simple outlets metrics such as impact factor – which again is a bad name for a flawed metric.

Acknowledgements

First and foremost, I would like to thank my main supervisor, Simon Bøgh, for firstly introducing me to this specific PhD position. Thank you for guiding and helping me throughout the study and for all the talks and discussions we had, research-related or not. I could not have wished for anything more from a supervisor. I would also like to thank my co-supervisor, Nadeem Iftikhar, for helping me and guiding me through the process.

My gratitude also goes to Helge Langseth from NTNU for welcoming me to your research group and for the many discussions and conversations related to my research. Thank you for being patient with the many disturbance to my stay-abroad plan caused by Covid-19.

I want to thank all of the researchers in the Innovation Factory North (IFN) project for the collaboration and conversation between our different research areas. Moreover, thanks to all of the companies involved with IFN. It has given me a priceless insight into the manufacturing SMEs of Denmark. I would also like to thank two companies and contact persons in particular. Jan Kristensen from Körber Supply Chain for letting me experiment with their machines and equipment. My father, Kent Lindegaard Hansen, from Almo Sil, for allowing me to implement a complete IoT infrastructure with marginal intervention.

Lastly, I owe my family and friends the most incredible gratitude. My parents, Dorte and Kent, for believing in me and helping me pursue my dreams. My friends, new and old, for always being there when needed, guiding and helping me with the research and this thesis, and not least, distracting me when needed. Finally, the last thank you goes to AAU's Robotics and Automation research group. I could not have dreamed of being in a bet-

ter research group – it is not a mistake that I placed you in the family and friends paragraph.

Reader's Guide

This PhD thesis is written as a collection of papers. The thesis starts with Part I, which will present the general motivation for the thesis, followed by the research question and state of the art. In Part II, dedicated chapters are presented for each of the papers, which gives an extended summary and connects it to the relevant research question. The part ends with the conclusion in Chapter 11. Lastly, in Part III all of the papers are presented. As stated, this PhD is a collection of papers though some are already published, and some are in press or under review. The extended summaries can be read without reading the papers though some general background and details will be left out. Throughout the thesis, when referring to a paper belonging to this thesis, it would look like this: [Paper A | [1]]. The first part (Paper A) is a hyperlink to the included paper in Part III. The second part ([1]) is a hyperlink to the actual reference in the bibliography, along with the rest of the references used in this thesis. The appended papers follow the style of the outlet but have been cropped to fit the thesis. A glossary and acronyms list can be found in the Glossary chapter in the end of Part II.

Emil Blixt Hansen
Aalborg University, November 13, 2022

Part I

Introduction

Chapter 1

Motivation

Since the start of the 2010s, the manufacturing sector has started the transformation of how they can improve their business. This includes changing the manufacturing process with the maturity of technologies such as 3D printing, robotics and information technologies (IT). The collection of these technologies is often referred to as the fourth industrial revolution or Industry 4.0. The whole concept of Industry 4.0 began in Germany at the Hanover fair in 2011 under the German name *Industrie 4.0*. The name Industrie 4.0 is polysemous since, in the beginning, it meant the strategic decision by the German government to enhance the competitiveness of their manufacturing sector [8]. It was also meant as a synonym for the fourth industrial revolution, which is what is currently being used today. As seen in Figure 1.1, the other industrial revolutions centred around a revolution of new technologies (steam power, electricity and computers). Industry 4.0 is more a product of new IT technologies, business methods and a pull from customers [9, 10].

Since Industry 4.0 is more of an evaluation containing different technologies and methods, a clear global definition does not exist. However, some technologies are often referred to as the main contributors. This can be cyber-physical system (CPS) [12], internet of things (IoT) [13], and smart factories [14]. In Figure 1.2, a suggestion of nine essential technologies of Industry 4.0 can be seen from Boston Consultancy Group (BCG). These technologies have been widely used in larger enterprises [15], however, the same cannot be said for small and medium-sized enterprises [16].

1.1 Small and Medium-sized Enterprises

A small and medium-sized enterprise (SME) defines a company or enterprise with certain characteristics. An SME is sometimes called a small and medium-sized business (SMB). Even though there is a slight distinguish be-

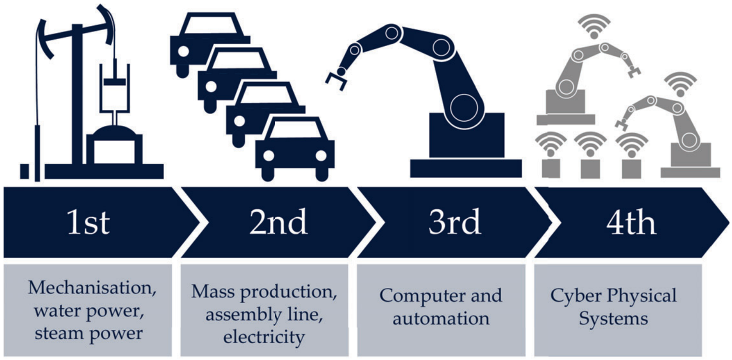


Fig. 1.1: The four industrial revolutions. Adopted from BRICS [11].

tween *enterprise* and *business*, they are often used interchangeably and thus no distinguishing is made between *SME* and *SMB* in this thesis. Moreover, SMEs include all company types in different sectors such as manufacturing, agriculture and services. In this thesis, the focus is on the manufacturing industry; thus, when an SME is described, it is a manufacturing SME unless specified otherwise. As stated, to be qualified as an SME, the company has to have a certain set of characteristics. The characteristics are different depending on the geographical location. For example, in Australia, different requirements depend on the government agency. For some Australian government agencies, the company has to have less than 200 employees while others it is 250 employees. The same goes for turnover, where for some, it is below A$ 20 million and for other government agencies, it is below A$ 150 million [18]. In the USA, there exist no clear definite either [19]. The European Commission has defined a specific set of characteristics which defines an SME [20]. These characteristics can be seen in Table 1.1.

Table 1.1: The European Commission definition of an SME. The **Turnover** and **Total Balance** columns only requires one of them to be fulfilled and **M** denotes millions [20].

Category	**Staff**	**Turnover**	*or*	**Total Balance**
Medium	< 250	$\leq$ € 50 M		$\leq$ € 43 M
Small	< 50	$\leq$ € 10 M		$\leq$ € 10 M
Micro	< 10	$\leq$ € 2 M		$\leq$ € 2 M

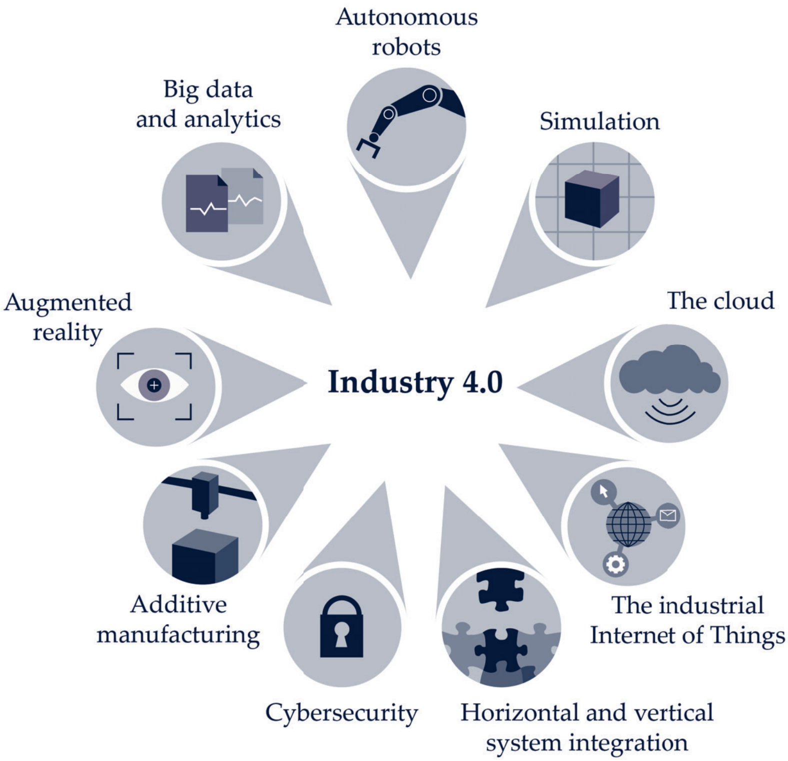


Fig. 1.2: The nine technologies transforming the manufacturing paradigm identified by BCG. Adopted from BCG [17].

1.1.1 Manufacturing Industry

The manufacturing industry is about turning raw material (or partly finished product) into a finished product. When the term manufacturing industry or manufacturing sector is used, it is mainly used to encompass all the companies that manufacture physical goods in some way. Humans have manufactured items for millenniums [21], but only within the last few centuries breakthroughs have happened that changed how manufacturing is thought of and used. As stated at the beginning of this chapter, these breakthroughs are often referred to as the industrial revolutions as shown in Figure 1.1. Even though we are in the fourth industrial revolution, companies still strive to improve their respective production. The motivations depend on each company, but it can commonly be cost and lead time reduction and quality increase. Not surprisingly, how this can be accomplished depends on the company and their motivation. Some of the focus areas could be material selection, value chain management, business models and process optimisation.

When automation is applied in a manufacturing company, it typically follows the *automation pyramid* scheme, as seen in Figure 1.3. The automation pyramid is a visual example of how different parts of the company's automation stack integrate. The bottom layer, *field level*, is where all of the low-level hardware is placed. This includes sensors, actuators, valves and motors. The next layer, *control level*, primarily consists of devices which control the devices from the field level. These are mainly programmable logic controllers (PLCs) and proportional integral derivative (PID) controllers. On top of the control level, a *supervisory level* is commonly used. This level interacts directly with the operators through human-machine interfaces (HMIs). Moreover, supervisory control and data acquisition (SCADA) systems can be used to control and monitor multiple controllers. The next level, *planning level*, is where manufacturing, including its resources, is planned and monitored. Commonly, this is done with a manufacturing execution system (MES). The final level, *management level*, is where all of the company's resources are managed. This is often done with an enterprise resource planning (ERP) system. These five levels is part of the *ANSI/ISA-95 Enterprise-Control System Integration* (also known as IEC/ISO 62264) standard [22]. Even though the standard explains how the different layers are connected and what they contain, it is still up to the individual company to adopt it. Moreover, research within the manufacturing field is starting to question the pyramid structure with the rise of Industry 4.0 and its subsequent technologies [23–25].

Fig. 1.3: The automation pyramid for a classical manufacturing company. Adopted from Rahman et al. [26].

With the advancement of IT technologies and Industry 4.0, manufacturing companies are faced with new challenges. Khan and Turowski [27] found the three main challenges for manufacturing companies to be:

Data integration:
When data is collected throughout a company, they are often collected in different ways. Therefore, it is possible that *data silos* occurs. These data silos make it difficult for a company to utilise the data beneficially. Furthermore, this challenges the real-time inspection of machinery. Analysing the data is done on a per-need basis; thus, a significant effort is put into gathering the data and analysing it each time.

Process flexibility:
With the reduced life cycle of new products, more requirements for the different processes are made. The problem is that much equipment and machinery are not designed for the needed flexibility. It is also a problem from the software perspective regarding sensor data and database management.

Security:
Since companies have an ever-growing inventory of smart devices, be it phones, tablets, computers, or sensors, the topic of cyber security is also a growing concern. Companies must keep their devices updated regarding software and have a structured plan for their IT infrastructure.

Even though Khan and Turowski's identified challenges are not explicitly focused on SMEs, the same issues apply to them, albeit even more so.

1.1.2 Opportunities and Challenges

An SME, compared to larger enterprises, has different sets of characteristics and thus different opportunities and challenges. Various studies have found several characteristics of manufacturing SMEs that bring them opportunities. These characteristics have been described as the role of the SME's manager and the short hierarchical line [28]. And also, the working environment of SMEs is, in general, welcoming entrepreneurship with the inherent culture and informal environment [29, 30]. This means that an SME often has less bureaucracy and, therefore, less time from idea to action.

The challenges for manufacturing SMEs in Industry 4.0 is often condensed down to lack of knowledge, technology awareness limitation, and financial limitation [31]. These challenges coexist with the already manufacturing challenges identified by Khan and Turowski. With these SME-specific challenges, different government authorities, consultancies and research institutions are providing aid differently. For example, the EU have the initiative *I4MS*[1] to

[1] https://i4ms.eu/

expand digital innovation of manufacturing SMEs. Here, SMEs can apply for financial and technical support. In Denmark, the Danish authorities have *Innovation Fund Denmark*[2] where companies can apply for financial support for, e.g. digitalisation projects. Besides financial support, initiatives such as the danish *MADE*[3] are constructed to bring digital innovation to manufacturing companies in Denmark through collaboration between the industry and research institutions. Moreover, specifically tailored towards manufacturing SMEs is the *Innovation Factory North* (IFN) project [32]. This project focuses on bringing awareness of Industry 4.0 to SMEs and demonstrating and utilising said technologies.

This PhD study is a part of the IFN project. Subsequently, this means that the focus is on the manufacturing SMEs primarily located in the northern part of Denmark. As the IFN project revolves around the challenge regarding lack of technology awareness, this project will focus on that. Thus the financial aspect and managerial level are only briefly touched upon.

1.2 Internet of Things

The internet of things (IoT) covers the subject of devices that are interconnected with each other. In this case, devices are physical equipment such as sensors, computers or other hardware with software. A device is an IoT device when it can communicate with other devices over a network protocol. In the industry, IoT is often called the industrial internet of things (IIoT). BCG also ranks IIoT as one of the pillars of Industry 4.0. While IIoT is tailored towards the industrial industry, there is no real difference between it and IoT. Thus for this PhD thesis, the IoT term will be used.

As stated, for IoT to work, it requires both the a physical device with software and a communication protocol. Even though *internet* is in its name, it does not need access to the whole internet. A local network can be sufficient for the task. IoT devices commonly connect to a cloud solution where the data is stored. Such a cloud solution can be placed locally or remotely. With a cloud solution, it is possible for the company to both gain insight into the production and act according to changes within the data. A large amount of data collected in the cloud is often called big data. An IoT and cloud infrastructure can be built differently. An example of a method is the service-oriented architecture for IoT as shown in Figure 1.4.

Implementing IoT within a manufacturing company is generally agreed upon to benefit the company, although the same benefits are hard to calculate [34]. These benefits come in different forms and depend on the individual

[2] https://innovationsfonden.dk/en/p/grand-solutions

[3] https://en.made.dk/

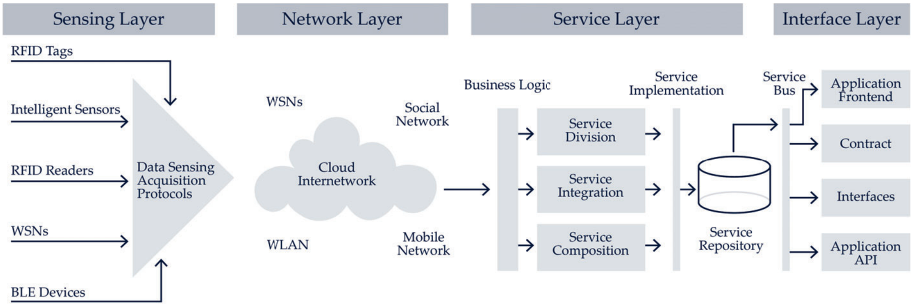


Fig. 1.4: Example of a service-oriented architecture for IoT. Adopted from Li et al. [33].

company and its implementation. Fantana et al. [34] have identified numerous such as:

- Right information providing or collecting
- Visibility identification and location tracking
- Reduced production losses
- New type of maintenance and lifetime approaches

In a white paper, McKinsey estimated that IoT within the manufacturing sector would bring \$ 1.2 to \$ 3.7 trillion per year from 2025 [35], further highlighting the potential.

1.3 Artificial Intelligence

When BCG created the nine pillars of Industry 4.0, they did not use the term artificial intelligence (AI). It was part of the *Big Data and analytics* pillar. Since then, AI has become a much more utilised and sought-out technology [36]. The term itself, AI, is ambiguous and covers many different methods and applications. Overall AI can be split up in to two major categorise, *artificial narrow intelligence* (ANI) and *artificial general intelligence* (AGI). AGI is where the program has a general intelligence level, i.e. it is not specifically trained for one task. This is the AI often depicted in science fiction. AGI is not close to being a reality even though research is ongoing [37]. Currently, all of the breakthroughs in AI have been happening in the ANI category. That is when an AI is specifically trained or applied to one task. In Figure 1.5, an overview of different areas AI can be applied to is shown.

From Figure 1.5, it should be noted that *machine learning* is depicted as an area along with, e.g. vision and speech. Machine learning is the commonly

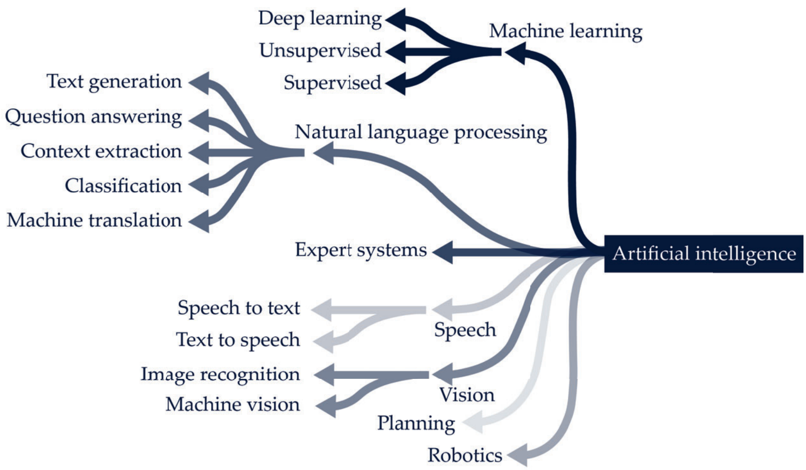


Fig. 1.5: An overview of all the topics covered by artificial intelligence.

used tool to apply AI in different areas. Machine learning is commonly used to describe classical methods such as support vector machines (SVM) and regression methods such as ordinary least squares (OLS). More recently, the subdivision of machine learning called *deep learning* has become the more prominent method. Deep learning uses artificial neural networks (ANN) to mimic the human brain's workings. ANN is also often just called neural networks (NN). With the general advancement in IT technologies, such as the rapidly increased compute power available and the increasing data available, NNs have succeeded in outperforming the classical machine learning methods when enough data is available [38, 39]. In Figure 1.6, an overview of the different machine learning techniques is shown. Machine learning is often characterised into three paradigms: supervised, unsupervised, and reinforcement learning.

Supervised leaning:
: Supervised learning methods are used when the training data also consist of labels. These labels can either be class labels (classification problem) or values to predict (regression). Supervised learning is typically used in image classification, system value forecasting and natural language processing (NLP) problems.

Unsupervised leaning:
: Unsupervised is used data is present without a specific label. This can be a clustering method that tries to split the data into underlying classes based on the data. It can also be dimensionality reduction to reduce the amount of data available. Common use cases are recommender systems, big data visualisation, and anomaly detection.

Reinforcement learning:

Reinforcement learning is used when e.g. a sequence of decisions is required. This is commonly done with an *agent* which is learning by performing actions in an environment and thus gets rewarded/punished for its actions. This method is commonly used in navigation, robot manipulation and real-time decision-making.

The examples of each of the three paradigms are not exclusive since the same problem can be solved with more than one paradigm depending on the setting and problem formulation.

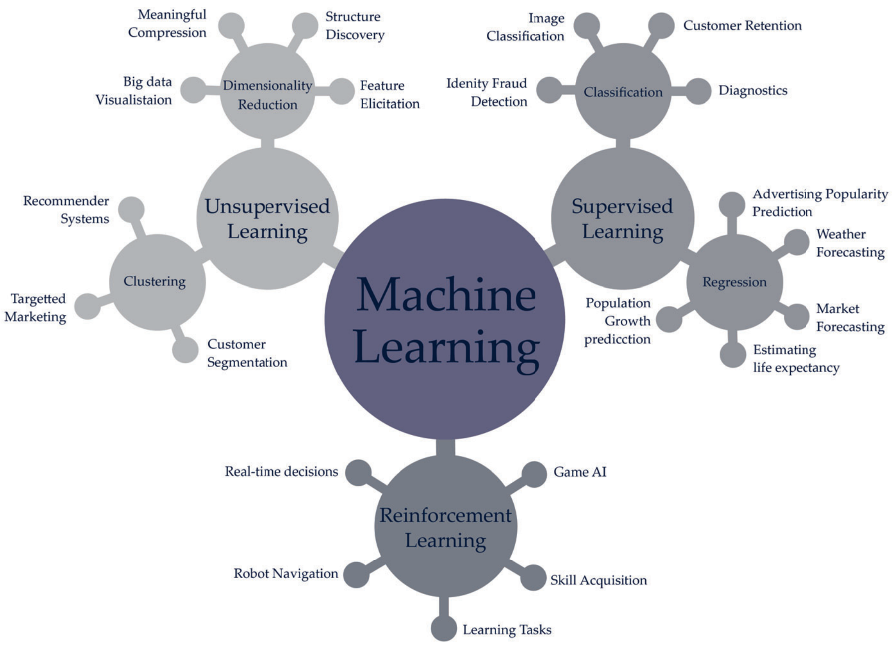


Fig. 1.6: The three paradigms of machine learning along with their methods and applications.

As shown, the topic of AI covers many different areas and methods. For the rest of this thesis, when the term AI is used, it mainly concerns the machine learning topic. Moreover, no directly distinguish is made between deep learning and machine learning. This is because deep learning is a part of machine learning, and thus the latter term is mainly used.

Chapter 2

Research Questions and Approach

Based upon the background described in the prior chapter, this section will outline the research questions for this thesis. This section will also explain the research methodology and how the rest of the PhD thesis is structured. Moreover, this section will describe the different scientific paper publications and submissions conducted as a part of this thesis. Lastly, a description of how the knowledge gained during the PhD has been disseminated.

2.1 Research Questions

As described in Chapter 1, this PhD thesis consists of three main topics: AI, IoT and manufacturing SMEs. Three research questions are composed to conduct further research on these three topics. These research questions also consist of sub-research questions to elaborate the questions further.

RQ1: What is the current state of IoT and AI adoption in SMEs?

RQ1.1: On what level of integration is the current IoT and AI solutions in SMEs?

RQ1.2: What are the challenges the SMEs face when integrating IoT and AI solutions?

RQ1.3: Why should the SMEs adopt the use of IoT and AI and what are the benefits?

RQ2: How can modern digital solution be designed to aid the adoption within SMEs?

RQ2.1: What are the challenges of using and sharing data for SMEs and why should they do it?

RQ2.2: How can an algorithm be designed to overcome the challenges of the SME and thus be used without expert knowledge?

RQ3: How can IoT and AI solutions be integrated in an SME to enhance the production?

RQ3.1: What is the preferred starting methodology for SMEs beginning their utilisation of IoT and AI?

RQ3.2: How can an IoT and AI setup be at an SME, both hardware and software-wise?

These research questions are answered by combining the contributed papers found in Part III and the research summary in Part II. Specifically, the research questions are answered with the following papers:

RQ1: is answered in Paper A, E and G

RQ1.1: is answered in Paper A and E

RQ1.2: is answered in Paper A, E and G

RQ1.3: is answered in Paper A and E

RQ2: is answered in Paper B, C, D and F

RQ2.1: is answered in Paper C and F

RQ2.2: is answered in Paper B and D

RQ3: is answered in Paper A, B and G

RQ3.1: is answered in Paper A and G

RQ3.2: is answered in Paper B and G

2.2 Research Methodology

The main research methodology used in this PhD study was design science research (DSR). While the research in this PhD thesis was conducted iteratively, the next iteration was always created on gaps identified in the prior work. In DSR, the main focus is to enhance knowledge by creating innovative artefacts for real-world problems [40]. Different forms of DSR exist, but they all consist of the same general steps to carry it out. The basic steps include the following phases: background phase, design phase, demonstration phase and evaluation phase. Peffers et al. [41] defined the following 6 steps:

identify, *define*, *design*, *demonstrate*, *evaluation* and *communicate*. While the research conducted in this thesis did not follow the specific steps by Peffers, its core ideas were used. That is, the research directions for this PhD thesis was *identified* and *defined* through the first research question (**RQ1**). The next two research questions were established with the research found in **RQ1**. Research question 2 (**RQ2**) focuses on the IoT and AI design element for SMEs. This research question follows the *design* and *demonstrate* steps. The last research question (**RQ3**) is focused on how such implementation at an SME could look along with evaluation. Therefore, **RQ3** is focused on the last steps such as *demonstration* and *evaluation*. The last step, *communication*, is done primarily through this thesis.

The order of the research questions closely resembles how the research was conducted in this PhD. However, there are some overlaps between the papers and research questions.

2.3 PhD. Study Publications and Submissions

The papers published and submitted for this PhD thesis can be seen in Table 2.1. The paper's dedicated letter is shown in the table, along with each paper's title and objectives.

Table 2.1: The papers published in this thesis along with the objectives.

	Paper Name	**Objectives**
A	Artificial intelligence and internet of things in small and medium-sized enterprises: A survey [Paper A	[1]]	Identify and present an overview of the current state of AI and IoT in SMEs according to the scientific literature. Analyse the drivers behind the implementation and what made the different implementations successful. Finally, it presents directions for future research within the area.
B	Concept of easy-to-use versatile artificial intelligence in industrial small & medium-sized enterprises [Paper B	[2]]	Design and demonstrate how an easy-to-use AI system could be like in an industrial context. The design was focused on not requiring expert knowledge to use and set up, along with how the architecture could handle different types of machine learning problems. It was tested on two other datasets, one from AAU smart laboratory[1] and one from an industrial partner.

Continued on next page

[1] https://www.smartproduction.aau.dk/Laboratory/

Table 2.1 – continued from previous page

	Paper Name	Objectives
C	A new authentic cloud dataset from a production facility for anomaly detection [Paper C	[3]]	Data was collected and shared in a new publicly available dataset. The data was collected at a real production site and contained errors during operation. The dataset was made public to help fill the publicly available production-related datasets gap.
D	A data-driven modular architecture with denoising autoencoders for health indicator construction in a manufacturing process [Paper D	[4]]	A modular approach to construct a health indicator (HI) value for an arbitrary process. The modular approach does not require historical data before initial setup and is not built to a specific process as it is data-driven. It is designed to be modular in the sensor input aspect, regarding numbers and types.
E	Artificial intelligence and machine learning [Paper E	[5]]	Basic explanation of AI and machine learning and showcasing different use cases from different companies.
F	On the topic of anonymising production data for machine learning [Paper F	[6]]	Explaining the need and idea behind anonymising data. Focuses on specific types of data categories related to data gathered in a production environment. A six-step approach is presented to make it more straightforward for others to anonymise a production dataset.
G	An in-depth investigation of machine learning and IoT adoption at a manufacturing SME: A field study [Paper G	[7]]	A demonstration and evaluation of an IoT and AI setup implemented at a Danish manufacturing SME. The IoT setup is described and discussed on both software and hardware level. The AI part is implemented to control critical control parameters and is tested on data collected over six months. A company perspective is also presented and discussed.

As the research is conducted with DSR, most of the papers are built on knowledge and gaps identified in the prior work. To give a better overview of which papers are built on top of which, see Figure 2.1 for an illustration.

2.4 PhD. Knowledge Dissemination

During the PhD study, one important aspect of the research was also conducted, namely knowledge dissemination. Besides publications and conferences, knowledge dissemination was conducted in the following ways:

- Bachelor level teaching
- Bachelor and master student supervision

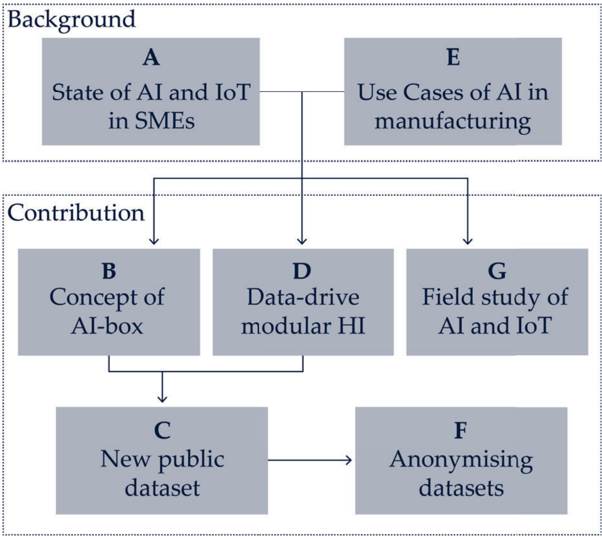


Fig. 2.1: The thought and knowledge flow between the papers.

- External lecture
- Sparing partner for various companies in IFN

During the PhD, different teaching activities were carried out. The teaching activities were conducted at UCN for bachelor students in the software engineering education. As they are purely software students, the teaching was centred around AI, the manufacturing industry and Industry 4.0. The motivation behind it was to give the students a broader knowledge of applied AI and the challenges and opportunities within the manufacturing industry.

Supervision was conducted both at UCN and AAU. The supervision was centred around different topics such as software testing, robotic manipulation, data collection and artificial intelligence. Some of the students managed to publish their work. In one project, the students created a virtual reality interface for remote robot manipulation [42], and in another, the students built an industrial omnidirectional mobile robot platform [43]. Another student group performed the data collection and initial analysis of the data, which was later published as the open dataset AICD [Paper C | [3]].

An external lecture was conducted at the yearly danish *datamatikerlærer foreningen*[2] (computer science teacher's association) gathering. The lecture concerned AI and Industry 4.0, along with ongoing research and an outlook into the future.

[2]https://www.dmlf-cms.dk/

Lastly, as this PhD study was conducted under the IFN project, different types of knowledge dissemination were conducted in that context. The IFN project's main goal is to bring awareness, demonstrate and bring internal changes to the participating companies. For IFN, most dissemination was consulting the companies with their problems. The problems were commonly either digitisation or digitalisation of their production, whereas *paperless production* was a common trend. Moreover, a demonstration of how machine learning can be used for production data and with actual company data was also carried out.

2.5 Outline of Thesis

This thesis is concerning the three main topics of *IoT*, *AI* and *SMEs* with three subsequent research questions. To fully understand the current state of the art within the two technologies, IoT and AI, and certain advancements within manufacturing, Chapter 3 presents the state of the art. After that, the thesis moves into Part II *Research Summary*, which covers all publications made and answers the research questions. Part II is structured such that each publication has its own chapter, as Chapter 4 concerns Paper A, Chapter 5 concerns Paper B and so forth. Each chapter in the *Research Summary* has the same structure. They start with informing what paper it covers and which research question it answers, following general information for that specific paper and research area. After that, an extended abstract is presented, describing the key background information, methods and results. Each chapter ends with an *implications* section, which connects the paper with the research questions and describes what it means in the broader scope. As these chapters describe the same information which is presented in the papers, repetition of context, phrasing, results, figures, and tables are to be expected.

Chapter 3

State of the Art

Before answering the three research questions, it is paramount to understand the state of the art. As the motivation for SMEs has been established in Chapter 1, this chapter will focus on the two technologies, IoT and AI, and trends in the manufacturing sector. Specifically, to answer the research question, this chapter will investigate different aspects of the technologies regarding architecture, usages and different deployment types. Both of these topics encompass many different sub-areas; thus, only relevant areas will be described in this chapter. The context of SMEs in regard to the technologies will not be considered in this chapter.

3.1 Internet of Things

The internet of things (IoT) can have different architectural layouts depending on various factors. Kumar et al. [44] summed it up to a four-stage setup. The first stage was the sensor and actuation layer, where all the interaction with the real world happens regarding sensing and interaction. The second stage is where the data acquisition happens, and data is transformed in a meaningful manner and then distributed further. The third stage is edge computing, where local analysis and interaction with the data, such as AI methods, happens. The fourth and final stage is the cloud storage, where data is archived, and long-term statistics and analysis is performed. An overview of these four stages can be seen in Figure 3.1.

3.1.1 Sensors

One integral part of IoT is sensors and actuators. Sensors are small devices which can measure physical events such as pressure, light, temperature and weight. Wireless sensor networks (WSN) have gained traction with the rise

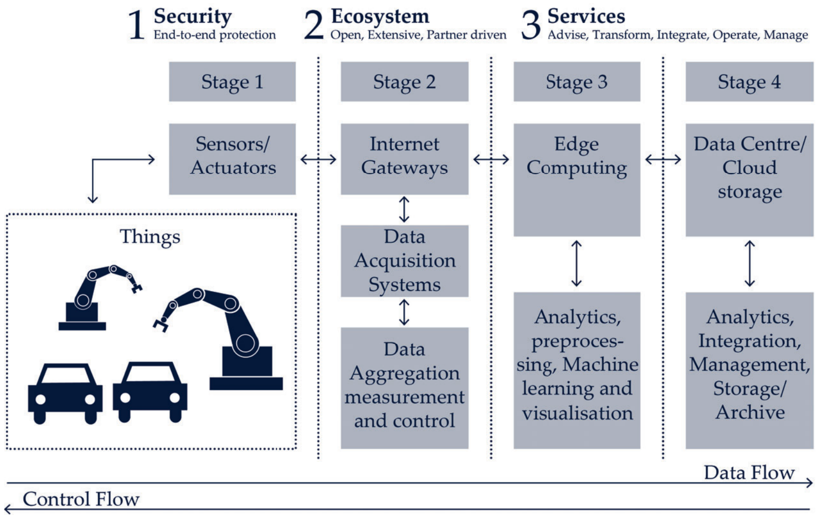


Fig. 3.1: An example of how an IoT architecture can be setup. Adopted from Kumar et al. [44].

of IoT [45]. WSN sensors are small sensors with limited processing power and wireless communication protocols built in. These sensors are generally applied in an application where the infrastructure is limited and is generally decided by the implementation. Li and Kara [46] describe how WSN is an essential part of IoT in an Industry 4.0 context. They demonstrate that WSN should be used as the sensing layer, corresponding to Kumar et al. [44] stage 1. They argue that this would reduce the initial investment cost compared to traditional methods where the analogue sensors are directly attached to a data acquisition computer. Kandris et al. [47] presented a comprehensive review of WSN where they identified three main areas it can be used in an industrial context. These were logistics, robotics and machine health monitoring.

Within the manufacturing sector, machinery is often used as long it is feasible. Consequentially this means that some machinery can be several years old and thus might not be compatible with modern methods and technologies such as IoT. Therefore, companies can *retrofit* new sensors to compensate for it. Jaspert et al. [48] presented a systematic review of smart retrofitting in manufacturing. They found that the main drivers were to ensure competitiveness, increase efficiency, market responsiveness and regulations compliance. An example of sensors used is an accelerometer for equipment state [49]. Accelerometer and temperature sensors have also been used in conjunction for health monitoring [50] and with the correlating PLC data [51]. An example of attaching an energy sensor to a CNC machine also gets insight into the machine's overall equipment effectiveness (OEE) [52].

3.1.2 Communication Protocols

For sensors to connect and be used for other than local measurement a communication medium is required. Physical communication can either be wire or wireless. Historically, wired has been the go-to method for stability, latency and security reasons. However, with the ongoing need for more interconnected sensors and devices, connecting the sensors by wire can become infeasible, and thus wireless communication protocols are used [53].

Besides the commonly used WiFi, there exist many other protocols. These range from low-power and low-throughput standards such as ZigBee and Z-Wave [54] and high-power and high-throughput such as 5G [55]. Protocols such as LoRa also deliver long-range, and low-power consumption [56].

To send the sensor readings, typically, a higher abstraction level is needed on top of the already mentioned network protocols. Here different types of communication protocols can be used. For many years protocols like Modbus, Profibus and Profinet have been used as the communication protocol between PLCs and other devices. Nonetheless, with the expansion of new devices, other communication protocols are gaining traction. OPC UA is an example of a communication protocol which is being used. It is found to integrate well into communication between classical machinery (e.g. CNC machines) and PLCs with low sampling rate [54]. It also has the semantic modelling of the data, making it easier to interpret the data. Within manufacturing, the messaging protocol MQTT is also being used. It has the benefit of being widely used, and thus it can be implemented on a wide variety of devices [54]. This makes it suitable as the communication platform for retrofitted sensors. MQTT lacks semantic data modelling compared to OPC UA, making it harder to scale and understand in complex scenarios [57].

3.1.3 Databases

In manufacturing, databases are commonly used in most of the automation pyramid. The data can be stored in the same database or multiple databases according to the subsequent automation layer. Databases have existed for many decades, and while their main objective of storing data has stayed the same, the way it is done and the usage of the data have changed [58]. Now there are different databases for different types of data and usages, which the users operate through a database management system (DBMS). A relational database where SQL is used to query data is commonly used in the industry to store data. This data can e.g. be inventory, work info and events. Where relational databases use tables and relations to structure the data, NoSQL (often called *Not only SQL*) can work with many different database models. These types can e.g. be relational, graph, and documents [59]. With the rise of Industry 4.0 and IoT, data collection has also risen, and when a lot of data is collected, it is often called big data. The data collected is now also continuously data from various sources such as temperature, vibration, and events. These types of measurement are often associated with time, thus called time series data. As time series data can be collected with a high frequency, it can be beneficial to use dedicated time series databases to store it. These time series databases are constructed in such a way that it requires less space and computing power to store and read the data, along with other built-in methods for reading the data, such as mean aggregation over a certain window period. An example of a time series database is InfluxDB, one of the most popular time series databases [60].

Besides databases with accompanying DBMSs, other storage mediums exist. Examples of these mediums are Excel, HDF5, CSV, Zarr and JSON files. These mediums are often just a single file with no direct interaction method like DBMS. Instead, they are commonly used to share the data or train, e.g. machine learning algorithms.

3.1.4 The Cloud

A cloud solution is where the data is stored on a server, commonly at a cloud provider. Besides acting as an offsite backup, it is possible to inspect, analyse and act upon the data. As seen in Chapter 1, BCG sees the cloud as an entire pillar of Industry 4.0 by itself. In a manufacturing context, the cloud can also be referred to as cloud-based manufacturing (CBM). CBM can be described as a network of manufacturing models which use on-demand access to a shared pool of diverse and distributed manufacturing resources. Here it is used to create temporary, re-configurable CPS that can improve efficiency and lower the cost of product lifecycles in response to customer demand [61]. A cloud solution can be designed as service-oriented architecture (SOA), which, e.g. makes components reusable and thus reduce costs [62]. Often the cloud providers offer different types of *services* in the type of infrastructure as a service (IaaS), platform as a service (PaaS) and Software as a service (SaaS) [63]. The three **aaS* distinguish the level of control/manageable items between the customer (the company paying for the service) and the provider. In IaaS, the provider only manages the infrastructure, i.e. servers, networks and virtualisation. The customer then handles the rest. In PaaS, the provider also handles the OS and middleware for the servers. Here the user only is responsible for deploying and maintaining the application they need. In SaaS, the provider takes care of the entire stack, making it the easiest to use for the customer but also the least flexible of them. An overview of this in Figure 3.2 is shown.

Additionally to the cloud solution, companies are also bringing some of the functionality locally, this is called *fog computing*. Fog computing can per-

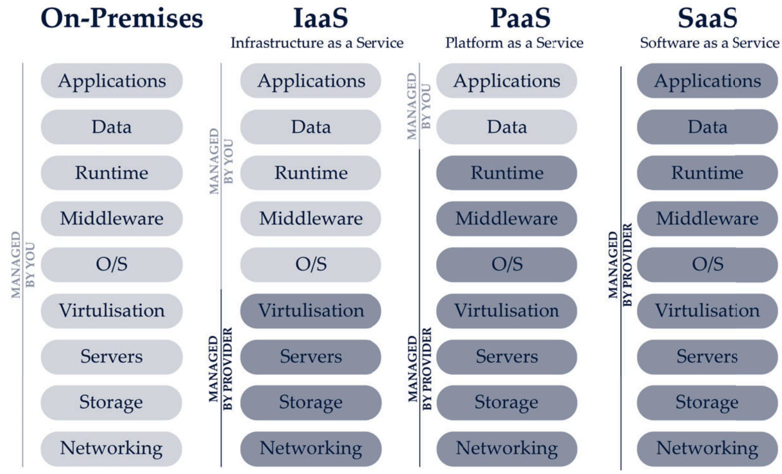


Fig. 3.2: Management of IaaS, PaaS SaaS. Adopted from Saraswat and Tripathi [64].

form calculations and sensor data selection before being sent to the cloud. Thus it is possible to decrease the amount of data sent to the cloud, which can reduce the cost [65]. It is also possible to act on the data quicker than because of the reduced latency compared to the cloud. AI models have also been deployed in the fog, which both impact the increased response time and enhanced security regarding the computation happening on site [66].

3.2 Artificial Intelligence

With increased interest in artificial intelligence (AI) and the possibility it brings, much research and advancement have happened in the last decade. Examples of advancement are IBM Watson beating human opponents in Jeopardy, AlphaGo Zero beating the prior version AlphaGo Lee without training on labelled data, and Tesla's self-driving hardware being implemented in all of their vehicles since 2016 [67]. As stated at the start of this thesis, the AI topic is vast and covers many different sub-areas. For this PhD thesis, it would be irrelevant to describe all of them; thus, only the relevant areas would be described.

3.2.1 Image Classification

Within AI, image classification and object detection have various use cases, such as quality inspection of solar cells [68] and autonomous train stopping [69]. When image classification is the objective, convolutional neural network (CNN) is the go-to method as it mainly outperforms other methods such as

ANN and SVM [70]. A CNN network normally consists of one or more convolutional layers and sub-sampling layers (e.g. max-pooling) to reduce the dimensions. Following the convolutional and sub-sampling layers, normal fully connected layers (also known as dense layers) would be placed. In this type of setup, the convolution layers will learn the features of the training data, e.i. they would indirectly perform the feature extraction, and then the fully connected layers would perform the classification. A common example of this network is AlexNet which was a breakthrough with the use of deep CNN trained on GPUs [71]. In Figure 3.3, a graphical representation of AlexNet can be seen.

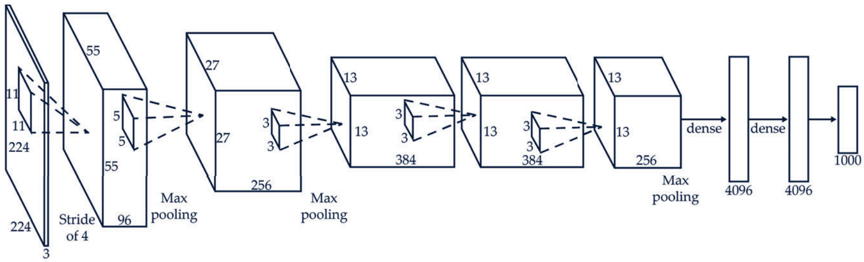


Fig. 3.3: AlexNet architecture. Adopted from Krizhevsky et al. [71].

Since AlexNet, many types of CNN architecture with different techniques have emerged. Iandola et al. [72] presented SqueezeNet, which has the same accuracy as AlexNet but with a significant lower model size. They achieved this by replacing the fully connected layers with global average pooling and their module named *fire module*. He et al. [73] introduced residual blocks which ease the training of deep networks. This idea of the residual block was used by [74] to build ResNeXt. Dosovitskiy et al. [75] used *transformers* to perform image classification, which achieved similar or better results compared to state-of-the-art CNN architectures. Besides image classification, object detection is the objective of locating an object within the image and image segmentation finds all the pixels which belong to the object. Examples of object detection is YOLO (you only look once) [76] and SSD (single shot multibox detector) [77]. Meta AI has created image segmentation named Mask R-CNN, allowing users to estimate other aspects such as human poses [78].

3.2.2 Time Series Classification and Regression

As described in Section 3.1.3, a time series is a measurement with associated time stamps or is collected over time. If more than one measurement is available, this is referred to as multivariate time series. See Figure 3.4 for an illustration of a multivariate time series. Within research, time series classification and regression have been used to classify hand motions of surgeons [79] and forecasting, e.g. energy consumption [80].

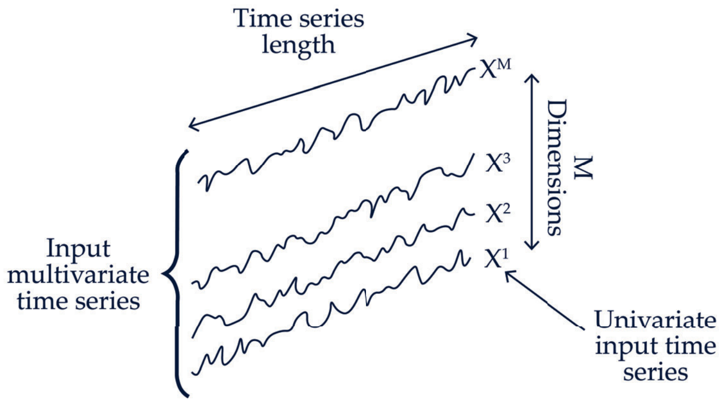


Fig. 3.4: Time series data.

CNN architectures have also been used in time series classification, as it is possible to represent the time series as an image and apply the successful CNN methods [81]. Fawaz et al. [82] showed that transfer learning in CNN time series classification problems could be beneficial and a hurdle if the correct dataset is not used. They identified that the dataset used for pretraining should be similar to the target dataset. A commonly used architecture for time series is a recurrent neural network (RNN). They work in the way that each cell in the RNN layers remembers prior information and, depending on the method, can share information bi-directionally. One of the most used RNNs is long short-term memory (LSTM) [83]. They have e.g. been used to forecast petroleum production through a neural network of LSTMs [84]. Essien and Giannetti [85] proposed stacked convolutional-LSTM layers with bidirectional stack LSTM layers time series forecasting for industrial machinery.

3.2.3 Anomaly Detection

Anomaly detection covers the subject of detecting anomalies in the applied area. These areas can be detecting spam emails and faulty bearings in machinery. One-class support vector machines (SVM) have been used to outperform other classical machine learning methods in network intrusion detection [86]. More recently, with NNs, autoencoders (AE) have become a successful method in anomaly detection. As shown in Figure 3.5, an AE works by compressing the input data through the *encoder* part into a *code* part (also called latent space). Then the data is decompressed through the *decoder* into an output matching the input. If the AE is only trained on good/correct data, then the reconstruction loss would become high if abnormal data is present. The use of AE as anomaly detection is further described in Chapter 7. Sakurada and Yairi [87] showed that an AE could outperform a linear principal component analysis (PCA) and a kernel PCA. Denoising autoencoders (DAE) is where noise is added to the network, typically through methods such as dropout. DAE has, e.g. been used to detect anomalies in wind turbines [88]. Variational autoencoders (VAE) are another type of AE, where the latent states are the statistical features mean and variance, and thus it describes a probability distribution of the input. VAE has also been successfully used in anomaly detection [89].

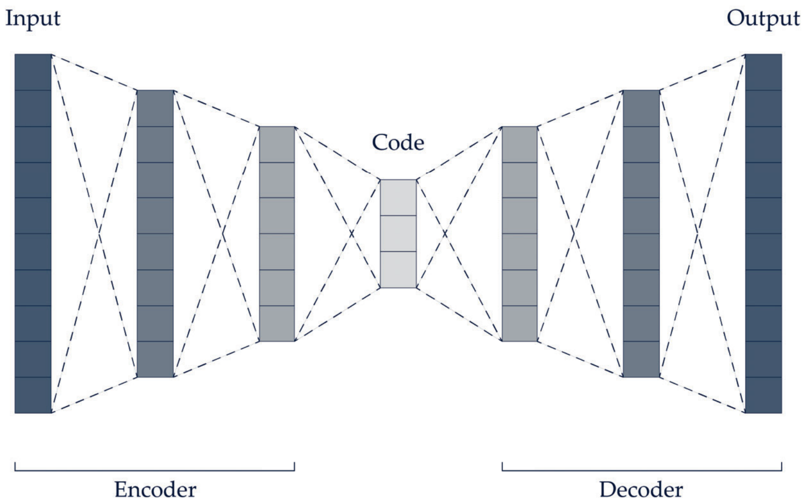


Fig. 3.5: An autoencoder with the three parts encoder, decoder and code.

3.3 Manufacturing

So far, the subjects covered in this chapter have been done in a general sense and in relation to the industry. Aside from the general introduction given in Chapter 1, specific industry-relevant advancement is further investigated here.

3.3.1 Predictive Maintenance

Maintenance is a crucial factor in any production. Failure to perform maintenance on equipment can lead to them failing unexpectedly and thus create unforeseen downtime and potential damage to products and machinery. Doing maintenance too often can also lead to unnecessary maintenance costs in terms of personnel and equipment costs, along with unnecessary downtime. Commonly, four analytics levels are described as descriptive, diagnostic, predictive and prescriptive [90]. How these four levels can be interpreted can be seen in Figure 3.6. These levels can also be applied in a maintenance context, e.g. PLC alarms can be descriptive or diagnostic. Predictive maintenance is when the maintenance process is performed in a predictive manner in some way, e.g. with expertise or machine learning. Prescriptive analytics is where the system can alter its settings to get the desired output. PWC [91] found in 2017 that only 3% of the companies in the survey did not perform any predictive maintenance, while 63% did it with visual inspections and instrument readouts. Only 22% used real-time condition monitoring with alerts at specific values. At last, only 11% used regression and machine learning methods based on big data to predict maintenance. Zonta et al. [92] presented a comprehensive literature review concerning predictive maintenance in Industry 4.0. They found that data-driven approaches is the most promising approach based on the increasing level of data collection. Moreover, if the objective is to make a time-based prediction, such as remaining useful life (RUL), data-driven approaches are needed because of the need for historical data. In another survey of predictive maintenance by Çinar et al. [93], they observed the same as Zonta et al. [92], but also argued that future directions for predictive maintenance are:

- Automating predictive maintenance may be made possible by using intelligent data collecting systems to extract real-time data.
- Combining more than one machine learning model can provide better predictions.
- Machine learning models implemented in the cloud can be further studied.
- Combining classification and anomaly detection algorithms can maintain the classification precision while keeping anomaly detection. This could reduce the need for larger datasets.

3.3.2 Health Indicator

One method to perform predictive maintenance is with health indicators (HI). HI is a number describing the specific machine's general "health". This number could be in percentage, where 100% means the machine is fully functional. Then when the machine is used, and parts are worn down, its health will decrease. HI, scores have been constructed for bearings with RNNs with

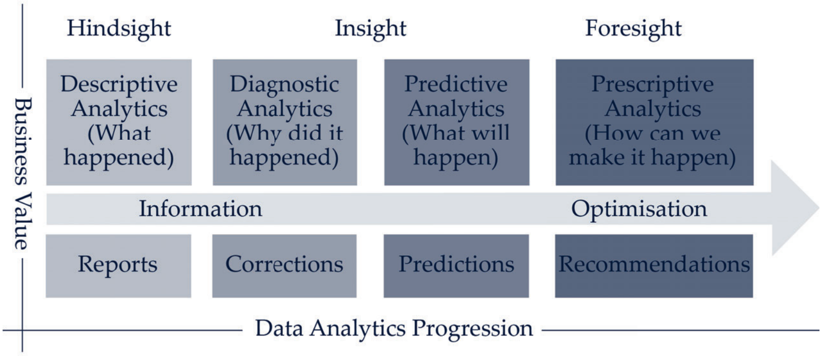


Fig. 3.6: The four analytic capabilities. Adopted from [90].

HI mapped to the range 0 to 1 [94]. The same bearing dataset what also used to construct HI score using CNN [95] and CNN with residual blocks followed by an RNN layer [96]. While these examples were all data-driven, examples of using expert knowledge of the system also exist [97]. Lei et al. [98] presented a comprehensive review of HI and RUL prediction. Among the different ways of constructing the HI score, they also discussed using health stages, which breaks up health into stages. Here it is possible to have, e.g. two stages (healthy or unhealthy) or three stages (healthy, degradation or critical). The HI score is often used to estimate the RUL, as seen in [94, 98]. However, all of them assume either that historical run-to-failure data already exist or simulation thereof. Hu et al. [99] proposed an RUL prediction with now historical data with the use of Kalman filtering and particle filters, and physical models.

3.3.3 Open Innovation

Sharing ideas and knowledge between companies is a practice with its origins in high-tech companies [100]. This practice is called *open innovation* and has grown in popularity along with Industry 4.0. Open innovation has also seen a rise in popularity within SMEs where two areas are specifically in focus: technology exploitation and technology exploration [101]. Technology exploitation refers to enhancing technological capabilities externally, and technology exploration refers to gathering external knowledge and drawing benefits from it, technological-wise. Researchers have proposed local open innovation networks for SMEs, which evolve both SMEs companies, the public sector, universities and research institutions [100, 102]. This methodology of local open innovation neatly follows the core idea behind the IFN project described in Section 1.1.2. Besides all the benefits, SMEs are also holding back on embracing the open innovation idea. Leckel et al. [102] identified the concerns as the protection of intellectual property (IP) and the lack of resources, money and knowledge.

Part II

Research Summary

Chapter 4

The State of AI and IoT in SMEs

The first paper, Paper A, is titled: *Artificial intelligence and internet of things in small and medium-sized enterprises: a survey*. The paper was submitted and published at the Journal of Manufacturing System, Elsevier in 2021. The paper relates to **RQ1** and **RQ3** and answers them through the sub-questions:

> **RQ1.1:** On what level of integration is the current IoT and AI solutions in SMEs?
>
> **RQ1.2:** What are the challenges the SMEs face when integrating IoT and AI solutions?
>
> **RQ1.3:** Why should the SMEs adopt the use of IoT and AI and what are the benefits?
>
> **RQ3.1:** What is the preferred starting methodology for SMEs beginning their utilisation of IoT and AI?

As this paper was the first paper written for this PhD thesis, it was meant as a means to lay the foundation for the subsequent papers. The paper was carried out as a structured literature review to understand how AI and IoT are used in SMEs. As this chapter is an extended abstract of the mentioned paper, [Paper A | [1]], repetition of context, phrasing, results, figures and tables are to be expected and is from that source.

4.1 Extended Abstract

Introduction and Method

When Industry 4.0 is discussed, the nine technology pillars by BCG's are often mentioned. These pillars include technologies such as IoT and AI [17]. While larger enterprises exploit these technologies, SMEs often lack the resources and knowledge to create a dedicated strategy for this transformation [103]. Real-time communication through IoT, big data, and analytics have been identified as some of the main drivers of Industry 4.0 in SMEs. These are said to bring new insight into the productions [17, 104]. A study by Moeuf et al. [28] showed that 90% of an expert group concur that IoT is essential for the industrial performance of SMEs and more than 55% agree that big data is essential to enhancing business performance. None of the experts said that SMEs possess the knowledge and skills necessary to use AI. Lastly, 75% of the experts said that research teams should encourage the adoption of Industry 4.0 in SMEs. As stated, SMEs have certain characteristics compared to larger enterprises. This study sums the characteristics up as follows:

- Culture and leadership
- Process innovation
- Company strategy
- Customer relationship
- Flexible and informal environment

To better understand SMEs and the technologies before the structured literature review, the study described both SMEs, AI and IoT. These descriptions are similar to what was already described in the motivation (Chapter 1) and state of the art (Chapter 3) chapters of this thesis.

The structured literature review was conducted on Scopus and Web of Science. For the search, the following search query was used:

> *"Small and Medium Sized Enterprise" and "Internet of Things"; "Small and Medium Sized Enterprise" and "The Cloud"; "Small and Medium Sized Enterprise" and "Machine Learning"; "Small and Medium Sized Enterprise" and "Deep Learning"; "Small and Medium Sized Enterprise" and "Neural Networks"; "Small and Medium Sized Enterprise" and "Artificial Intelligence"; "Small and Medium Sized Enterprise" and "Digital Twin";*

A certain set of criteria was needed for the search results to be deemed relevant for the study. These inclusion criteria were especially relevant in this study as the search query included broad terms such as *artificial intelligence*. The criteria can be seen in Table 4.1.

Table 4.1: Criteria of inclusion. **DT:** digital twin [Paper A | [1]].

No.	Criteria	Reason for inclusion
1	SME	The publication has to be relevant for SMEs.
2	IoT/AI/DT	Industry 4.0 is a collection of different technologies and thus different terminologies are often used; however, the publications should fall within at least one of these technologies
3	2010 - 2019	Industry 4.0 as a terminology was first used by the German government in 2011 [105] and since these technologies also first started to be relevant at that time, the publications should be post-2010.
4	Manufacturing	The publications should be relevant to the manufacturing industry.
5	English material	Because of the global aspect of Industry 4.0 and a way to avoid national biases, only English material is considered.

Results

Since some of the terms are general, many misfits were captured. Initially, 155 publications were captured; out of those, only 37 were deemed relevant. The review process can be seen in Figure 4.1. During 2010-2019 the study showed an increased frequency of relevant publications, illustrating the increased interest. Along with the increased interest, Europe and Asia were the continents with the most publications. The literature review revealed five areas which had a high focus. These were: IoT, AI, cloud, digital twin and business. The distribution of these five focus areas can be seen in Figure 4.2.

The IoT focus area found that the term *IoT* was often also used as a synonym for Industry 4.0. From the publications which did use IoT, it has been used to decrease energy consumption [106], machine monitoring and utilisation [107, 108], and room heat control [109]. Low-cost IoT solution was also the focus, where some communication means such as OPC UA and Zigbee were used [110, 111]. Other surveys found that the IoT adoption level is still low [112] and that the push for IoT solutions often comes from internal motivation [113, 114]. Within the AI focus area, only a handful of publications used or discussed AI. Thereof, only two publications used AI directly relevant to the production with machinery status detection [107] and climate and lightning control [109]. Digital twins were the focus area with the least amount of publications with only one use case in textile production [106]. The cloud was the focus area that attracted the most publication. The

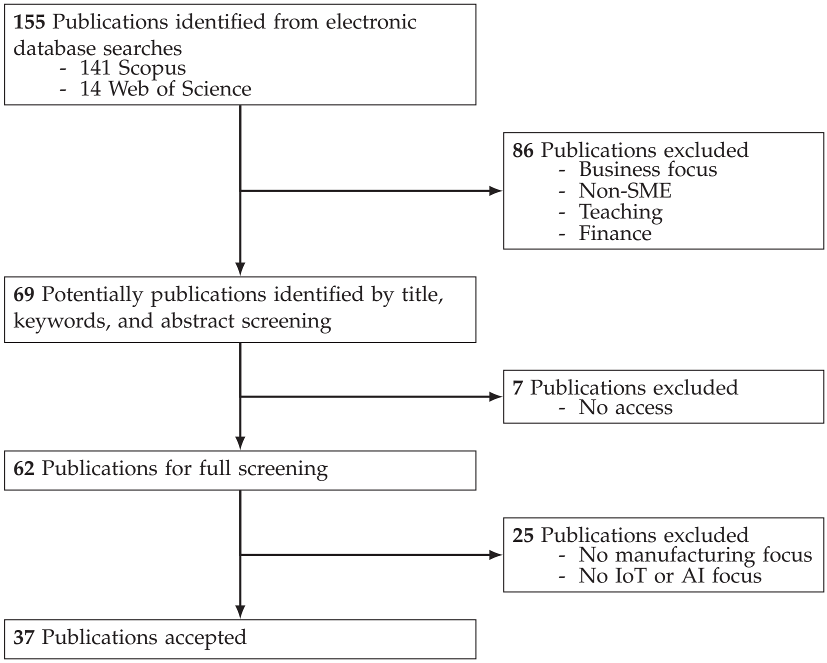


Fig. 4.1: The systematic literature review approach [Paper A | [1]].

study found that the cloud mainly covered the sub-areas of cloud computing, decision support systems (DSS), ERP, MES and security. Where study has stated that manufacturing SMEs should focus on ERP with DSS [115], whereas another study gave an example of how it could be designed [116]. It has been found that SMEs are especially vulnerable to cyber-attacks, with research showing that up to 91% of the attacks are external attacks [117, 118]. Moreover, it has also been found that the current cloud solutions are unsuitable for SMEs, and thus, specifically tailored solutions should be made [119]. Within the business focus area, it was found that the innovation process is often started internally within the SMEs [114]. Moreover, it has also been found that SMEs should be ready to adopt new business models to stay competitive [120] and thus embrace ideas such as open innovation [121].

Besides the found focus areas, the study also tried to map the characteristics of SMEs to the publications. Unfortunately, many of the publications did not state what the motivations behind the adoption were. Some were surveys, and some only focus on the technological aspect. Therefore, only a few were able to map to a characteristic. One of the characteristics that worked as the main driver was process innovation. This motivation was mainly to reduce cost [106, 110], increase production transparency [107] and shorten changeover time [122]. The other characteristic found by the study

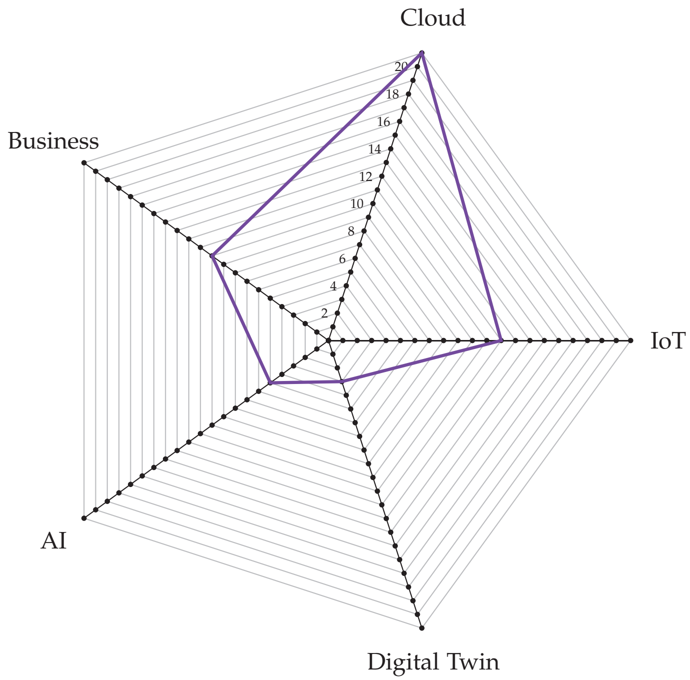


Fig. 4.2: The five identified focus areas [Paper A | [1]].

was company strategy. It was found that SMEs have changed their strategy to meet new markets [121], and most changes in company strategy comes from internal influence [114].

Conclusion

The literature review showed that cloud solution is one of the most used Industry 4.0 technologies in SMEs, even though other studies call for more SME-tailored versions. It also found that IoT is a slow adoption level and is primarily machine-wise implementation, e.i. only a single machine is connected. AI has the lowest level of adoption, even though successful implementations exist. Only a few studies exist on the characteristics aspects; thus, more research is needed in the area to better understand SMEs' driving forces.

One of the reasons why cloud solutions are the most adopted Industry 4.0 technology is that it is comparable easy to both get started with and understand the benefits. This finding agrees with Moeuf et al. [28]. Therefore, future research should focus on making IoT and AI easier for SMEs to utilise by reducing the complexity and need for expert knowledge.

4.2 Implications

Understanding the current state of AI and IoT in SMEs is paramount to lay the road ahead. The paper summarised the use of these technologies in SMEs over the last decade, where a clear pattern of increased interest was shown. Moreover, it serves as a foundation for the following research in this PhD study. While technology-wise, only a few were mentioned, such as OPC-UA and Zigbee as communication protocols, the study showed that most of the changes from the SMEs come from within. Since the changes come from within, the SMEs are potentially more invested in the technology and thus have a higher chance of success. The motivation ties well with the IFN project as companies are signing up by themselves and are invested in gaining knowledge about Industry 4.0. When it comes to using IoT, the majority were machine-wise implementation, which is a great starting point but can be complicated to scale throughout the factory. The area of AI is even less utilised within SMEs. A conclusion can be made that knowledge of AI and IoT and more accessible methods are needed if SMEs are expected to take advantage of them. It can, e.g. be done through external collaboration with other SMEs or research institutions. The contributions of this paper can be summarised as follows:

1. The current adoption and integration of IoT and AI in SMEs are low but rising.
2. Digital twins are not feasibly for SMEs at the current time.
3. The use of AI and IoT shows promising results, both in larger enterprises and SMEs.
4. IoT and AI solutions should be made more easily available for SMEs to increase adoption.

Chapter 5

Concept of AI-Box in SMEs

The second paper, Paper B, is titled: *Concept of easy-to-use versatile artificial intelligence in industrial small & medium-sized enterprises*. The paper was submitted for presentation at the 30th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2021) in 2021. Afterwards, it was published in Procedia Manufacturing, Elsevier. The paper relates to **RQ2** and **RQ3** and answers them through the sub-questions:

> **RQ2.2:** How can an algorithm be designed to overcome the challenges of the SME and thus be used without expert knowledge?
>
> **RQ3.2:** How can an IoT and AI setup be at an SME, both hardware and software-wise?

This paper explores how a software architecture concept can be made to give AI capabilities to SMEs. Specifically, the software is deployed on an NVIDIA Jetson, and all the classical machine learning architecture setups are hidden from the user. The concept is tested on two datasets, one from the AAU smart laboratory and one from an industrial partner. As this chapter is an extended abstract of the mentioned paper, [Paper B | [2]], repetition of context, phrasing, results, figures and tables are to be expected and is from that source.

5.1 Extended Abstract

Introduction and Concept

As larger enterprises are using the technologies from Industry 4.0 [123, 124], SMEs have not had the same boost where e.g. maturity assessment is not suited for SMEs [125]. There exists *smart devices* which individually try to solve different aspects on a higher level of abstraction. These devices can be OEE monitoring[1], live data-streaming[2] and visual inspection[3]. To contribute to the development of these smart devices and bring easy-to-use AI devices to SMEs, this study presents the concept behind the AI-Box. The AI-Box concept is built with three use cases in mind:

1. False alarm with a visual inspection
2. Audible error detection
3. Unknown error detection through vibration signals

The operators would train these three use cases without them knowing it by answering questions through the web browser. The AI-Box is intended to be connected to the PLC, and thus the operator controls both the machine and the training process of the models. The hardware for the AI-Box, an NVIDIA Jetson AGX Xavier, was chosen. The Jetson have a built-in GPU with CUDA cores, making it suitable for NN deployments and possible to train models directly on it. Python 3.6 was chosen for the software for rapid development, along with Django to handle the web interface and Tensorflow 2.0 for the NN framework. Two different types of internal architectures were used. The first one was the model-view-controller (MVC) [126]. The MVC keeps the underlying model (in this case, both NN and program information) filtered for the user so only the correct information is presented and accessible. To keep the AI-Box flexible and easy to implement new NN models and sensor inputs, the second architecture was based on the *layer pattern* architecture. This pattern enables each component to be independent and thus only used to communicate in between. In Figure 5.1, the system architecture for the AI-Box can be seen. Here the top *GUI* layer is where the operator interacts with the AI-Box. The *global state handler* ensures that the different aspects, such as the attached devices, are in a correct state. The *model handler* initialises the correct parameters in both the *loop controller* and the *sensor handler* along with their intercommunication. The *loop controller* handles how the NN model is running along with the training data. The *sensor handler* handles the communication with the correct sensor.

Experiments

To test the AI-Box, two different types of experiments were conducted. Use case experiments and NN model architecture experiments.

The use case experiments tested the first type of use case (*False alarm with visual inspection*). One experiment was conducted at an industrial partner,

[1] Factbird `https://blackbird.online/product-overview/`
[2] M-Box `https://www.monitor-box.com/`
[3] DTI Vision Box `https://www.dti.dk/quality-control-and-vision/38108`

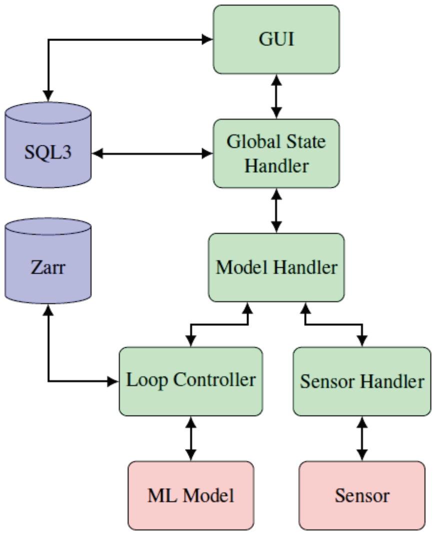


Fig. 5.1: The system architecture of the AI-Box. The green squares indicate that they are static models, the blue is databases, and the red is non-static model depended layers [Paper B | [2]].

and one was conducted at the FESTO CP factory at AAU smart laboratory. The industrial partner has a robot which picks layers of goods from a pallet and places it elsewhere. Between each layer, a slip-sheet is placed to keep the pallet steady. Sometimes, this slip-sheet is stuck under the layer, and no alarm is triggered. The AI-Box was placed beside the robot with a webcam mounted to it. Then the AI-Box started to take images beneath the picked-up layer. The NN was chosen to be the CNN architecture AlexNet [71]. After the first 30 captured images, the *loop controller* was triggered to start training. On each batch in training, a random generator was used to apply augmentation to each image to compensate for the low sample size. There is a probability of 0.6 that one or more augmentation is performed on each image. The available augmentations are brightness, contrast, flipping, hue, saturation, quality change, rotation, blurring, and cropping. A total of 87 samples were captured. The same procedure was followed at the FESTO CP factory, where it instead was to distinguish between black and blue phone cases. Here 41 samples were collected.

The second experiment was conducted to identify the most suitable CNN architecture and to test the augmentation method. The study tested three different types: AlexNet [71] , ResNeXt [74] and SqueezeNet [72]. Some small changes were made to AlexNet and ResNeXt to make them able to fit into the memory of the Jetson. The architecture of all three can be seen in Table 5.1. The models were tested on the same dataset from the FESTO CP factory with only 41 samples. Eight were used for the validation set, and a larger dataset from the industrial partner with a total of 1050 images, where 210 were used for validation. The experiments were tested with two different learning rates (0.001 and 0.0001) and were only allowed to run for 200 epochs. This was done to validate the quick learning requirement for the AI-Box. Thus, two different learning rates were tested.

Table 5.1: The different specifications for the implemented architectures. For AlexNet, a dropout is placed after all convolutional operations. In ResNeXt, no-bias is used in the convolutional operations. All models use Adam optimiser with default hyperparameters [Paper B | [2]].

AlexNet	ResNeXt	SqueezeNet
2 x Conv(32,3,1)	ZeroPad(3)	Conv(96,7,2)
MaxPool(2)	Conv(64,7,2)	MaxPool(3)
Conv(64,3,1)	BatNorm(1.001e-5)	2 x F.Module(16,64,64)
MaxPool(2)	ZeroPad(1)	F.Module(64,128,128)
Conv(128,3,1)	MaxPool(3)	MaxPool(3)
MaxPool(2)	3 x R.Block(64,1,32)	F.Module(32,128,128)
Dense(128)	3 x R.Block(128,2,32)	2 x F.Module(48,192,192)
Dense(68)	3 x R.Block(256,2,32)	F.Module(64,256,256)
Dense(2)	2 x R.Block(512,2,32)	MaxPool(3)
	GlobalAvgPool	F.Module(64,256,256)
	Dense(2)	Conv(2,1,1)
		GlobalAvgPool
Parameters	**Parameters**	**Parameters**
16,888,226	22,576,706	736,450

Results

The results from the experiment at the industrial partner showed that the test accuracy started to converge at 100% accuracy after 2 minutes of training on the Jetson. The FESTO CP Factory test showed it started to converge at 100% after only 17 epochs. In both tests, augmentation was used on the training data, while no augmentation was done on the test data. Because of the shallow data availability, the test data was the same as the training data in these two experiments. This is visible in training and test accuracy where the test data (no random augmentation) converged at 100% and the training data (random augmentation) never converged at 100%.

The results from the three different architectures showed that not all of the models and combinations of data were able to learn. Most noticeable were AlexNet, which outperformed the two others. Only one experiment did not converge, industrial partner data, high learning rate and no augmentation. Other results were that ResNeXt could learn the larger, more complex industrial partner dataset but not the small FESTO CP dataset. SqueezeNet was only able to learn the industrial dataset with a low learning rate. The random augmentation, as expected, prolonged the learning rate but enabled some of the models to learn the data.

Conclusion

The study showed that the concept of the AI-Box can be deployed in different scenarios and handle two different vision datasets. Moreover, the underlying concept works in an intended way with setup and control over the AI-Box and models. The experiments showed that currently, AlexNet is the most suited with random augmentation applied. As the study's main focus was the concept of the design and use case of the AI-Box, more work is needed further test the different aspects. This includes handling the other use cases and data types. Better functionality could also be added to the vision part. One example of such enhancement is the use of Grad-CAM [127]. With Grad-CAM, the operator can see what part of the images the model is learning and thus act upon it if it learns a wrong part of the images such as the background.

5.2 Implications

The AI-Box is an architectural concept of how a small smart device can be built for manufacturing SMEs. This study showed that it is possible to design such a smart device, which hides the underlying model and data handling from the operator. In contrast, the model is still able to perform on a satisfactory level. The experiments show the usability in two environments and validate the suited CNN model. Moreover, it also showed how random augmentation could increase the performance of a model. The users of the system were fond of the idea of hiding the underlying NN model and architecture of the operator. It resolved in that the industrial partner took the next step and incorporated a part of the system into their product. The contributions of this paper can be summarised as follows:

1. It is possible to design and use an NN device tailored towards non-experts.
2. For shallow datasets, simple models like AlexNet outperform more complex models.
3. Image augmentation is a favourable method, especially with a shallow dataset.

Chapter 6

Public Dataset from a Production Facility

Paper C has the title *A new authentic cloud dataset from a production facility for anomaly detection*. The paper was submitted to and presented at CARV 2021, and was afterwards published in: Towards Sustainable Customization: Bridging Smart Products and Manufacturing Systems, Springer. The paper relates to **RQ2** by answering the following sub-question:

> **RQ2.1:** What are the challenges of using and sharing data for SMEs and why should they do it?

This paper presents a new open dataset called authentic industrial cloud data (AICD). The paper starts by presenting the reasons behind the new dataset, along with how it was collected. Furthermore, a baseline experiment with the dataset is presented. The AICD dataset is available for download at: `https://www.kaggle.com/emilblixthansen/aicd-dataset`. As this chapter is an extended abstract of the mentioned paper, [Paper C | [3]], repetition of context, phrasing, results, figures and tables are to be expected and is from that source.

6.1 Extended Abstract

Introduction and Data Collection

It is well known that a larger dataset benefits training machine learning models and even more so in training NNs [128, 129]. With big data, the publicly available datasets greatly benefits new research methods and improves performance in the different sectors. Nonetheless, with the available datasets, only a few are relevant to the manufacturing industry [130–132]. Even though these datasets are relevant to the manufacturing industry, they have some drawbacks. The main drawback is that the data is augmented in different ways. The augmentation are manually introduced faults [130], run-to-failure is simulated [132] or the complete dataset is simulated [131]. Thus there is a gap for publicly available multivariate authentic manufacturing datasets, which this study contributes to in the form of authentic industrial cloud data (AICD).

The presented AICD concerns a pick-and-place operation of large items performed by a robot manipulator. The robot handles various goods, and sometimes, the robot drops the goods. The main objective is to identify when the robot is about to drop an item. This will enable the robot to stop its operation and thus avoid downtime in the form of cleanup. The data was extracted from an existing cloud solution from one robot in an operating production in Europa. The data was extracted with an interval of 10 ms and took place over a time span of 2 days. Before the data could be published, it was a requirement from the company that the data was anonymised. This was done by changing the names of some of the PLC tags and not disclosing which exact robot type, production and whereabouts. Moreover, the dates in the dataset have also been altered. The sensor readings have not been altered; thus, the data integrity is still present.

Dataset Content

The final dataset is distributed in CSV file format and split into five files. The data is also available in a Python pickle file for quicker loading into Python. The data is structured in a classical way, with each column being a feature (PLC tag in this case), and each row is a sample with 10 ms between them. As one of the purposes of this dataset is to be authentic to how data is out in the industry, no preprocessing has been conducted. Moreover, the data is not split up into train and test sets. One feature indicates that the dataset is meant to be used to discover when a drop is happening. The feature called *Alarm.ItemDroppedError* is set to *1* when a drop has been detected. In Figure 6.1 the dropped signal is highlighted over three measurements. The dataset contains in total of 16,990,692 samples of 96 features.

Baseline Experiment and Results

A baseline experiment was conducted to demonstrate how the AICD can be used. The objective of the experiment was to demonstrate the detection of drops. This was done through a DAE with LSTM layers. Only data from the first CSV file was used for the training dataset. The training data was preprocessed, so measurements 2 minutes before and after a drop was removed.

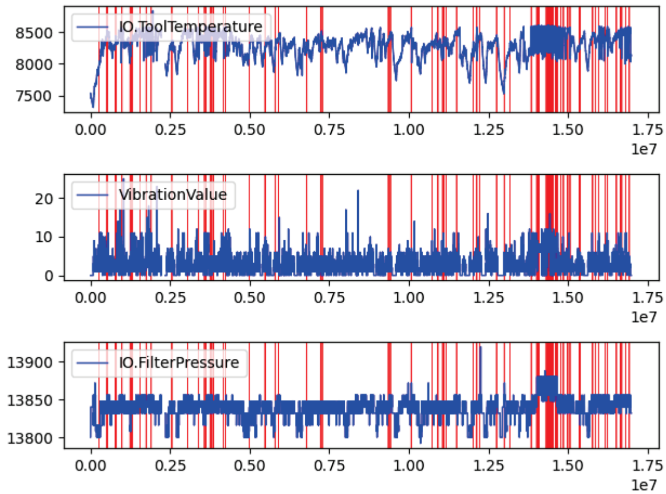


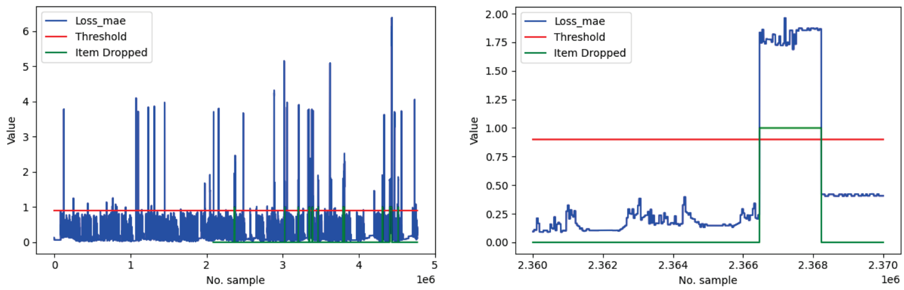
Fig. 6.1: All 16.9 million sensor measurements from three different sensors, measuring suction, vibration, and pressure. The red lines are when the drop detection sensor has been triggered [Paper C | [3]].

Moreover, features with static data and features rising to infinity were also removed, leaving the training data with 48 features. The training data was then standardised with Equation 6.1, where z is the standardised data, x is the original training data, μ and s are the mean and standard deviation of the training data, respectively. The test data was chosen to be the fourth CSV file, where the same features were removed from and was standardised with Equation 6.1, where μ and s are the values from the training data.

$$z = \frac{x - \mu}{s} \tag{6.1}$$

For the training of the DAE, the loss function was chosen to be the mean absolute error (MAE) and was trained with Adam optimiser [133] for 40 epochs. After the training and an inspection of the MAE distribution, a threshold of 0.9 was chosen as the boundary to flag drops.

The test result can be seen in Figure 6.2. In Figure 6.2a, the training and test data are combined, and the test data start where the green *Item Dropped* line starts. It can be seen that every time the item dropped is 1, the loss value also exceeds the threshold line. A closer inspection of the first drop can be seen in Figure 6.2b. Nonetheless, there are also a high amount of false positives, and thus the baseline experiment can be improved.



(a) Combined train and test data. **(b)** The snapshot of the first drop in the test data.

Fig. 6.2: The MAE loss is illustrated in blue, the red line is the threshold of 0.9, and the green line starts with the test data and is the binary item dropped signal. Both figures and captions are from [Paper C | [3]].

Conclusion

The study presented a new publicly available manufacturing dataset named AICD. At the time of publishing, it is one of the only real-world manufacturing datasets focused on anomaly detection and does not contain any augmented or simulated failures. As identified by Wuest et al. [134], there is a need for more available relevant data to overcome challenges in using machine learning in a manufacturing context.

6.2 Implications

As this paper shows, along with others, there is a need for more manufacturing datasets. Moreover, those existing datasets are often augmented or simulated in some manner. As the company requested, the data should be anonymised before it could be published. At the time of writing the paper, no such guidelines exist tailored toward manufacturing data. Lastly, this paper contributed to an authentic publicly available dataset, which was recorded at an actual operational manufacturing site. The contributions of this paper can be summarised as follows:

1. Summarised the lack of authentic manufacturing publicly available dataset.
2. A newly available dataset in the form of AICD.
3. The need for better guidelines on how to anonymise manufacturing data.

Chapter 7

Data-driven Modular HI

Paper D has the title *A data-driven modular architecture with denoising autoencoders for health indicator construction in a manufacturing process*. The paper has been submitted to the open-access journal IEEE Access, IEEE. The paper relates to **RQ2** and answers it through the following sub-question:

> **RQ2.2:** How can an algorithm be designed to overcome the challenges of the SME and thus be used without expert knowledge?

This paper presents a novel method of construction health indicators (HI) for machinery in a manufacturing setting named ModularHI. Firstly, the paper gives an introduction to why HI is relevant. Hereafter the system architecture is explained in depth. ModularHI is tested on two different types of datasets with a total of 13 different experiments. As this chapter is an extended abstract of the mentioned paper, [Paper D | [4]], repetition of context, phrasing, results, figures and tables are to be expected and is from that source.

7.1 Extended Abstract

Introduction and System Architecture

With Industry 4.0 and its technologies, SMEs often have the drawback of old equipment, which negatively impacts them in regards to their competitiveness and innovation [135]. The topic of prognostics and health management (PHM) concerns the area of monitoring of equipment. One such method is with HI, which estimated the *health* of machinery or process and can, in general, be created by three methods: model-based, data-driven, or hybrid [136]. The model-based approach is made by using the system's physical properties to calculate its health, i.e. the underlying systems and properties are needed to calculate it. The data-driven approach is calculated using relevant sensor readings and typically historical data to calculate the HI score. The hybrid approach is a mixture of both. AE is commonly used for constructing an HI score, and DAE has also been shown to outperform other methods in multivariate time series reconstruction problems [137]. As SMEs generally lack knowledge and resources within the area of Industry 4.0 it can not be assumed that they can set up an HI system for each of their machines. Therefore, a general purpose easy-to-use data-driven HI system would benefit them. This paper presents ModularHI, which is a data-driven modular HI scoring system. ModularHI is not built with a single type of machine in mind. Instead, it works with an arbitrary number of sensor types and inputs through the built-in modular architecture. Moreover, it does not require historical data during set-up. The only requirement is that the machine is fully functional and has recently been maintained. As ModularHI is built to aid SMEs and is data-driven, it is not expected to outperform specific engineered HI systems. Instead, it is meant as a tool for SMEs and alike. ModularHI is built with three types of states: *setup*, *burn-in* and *inference* and consists of two main parts: *component models* and the *aggregator*. The overall execution flow can be seen in Figure 7.1.

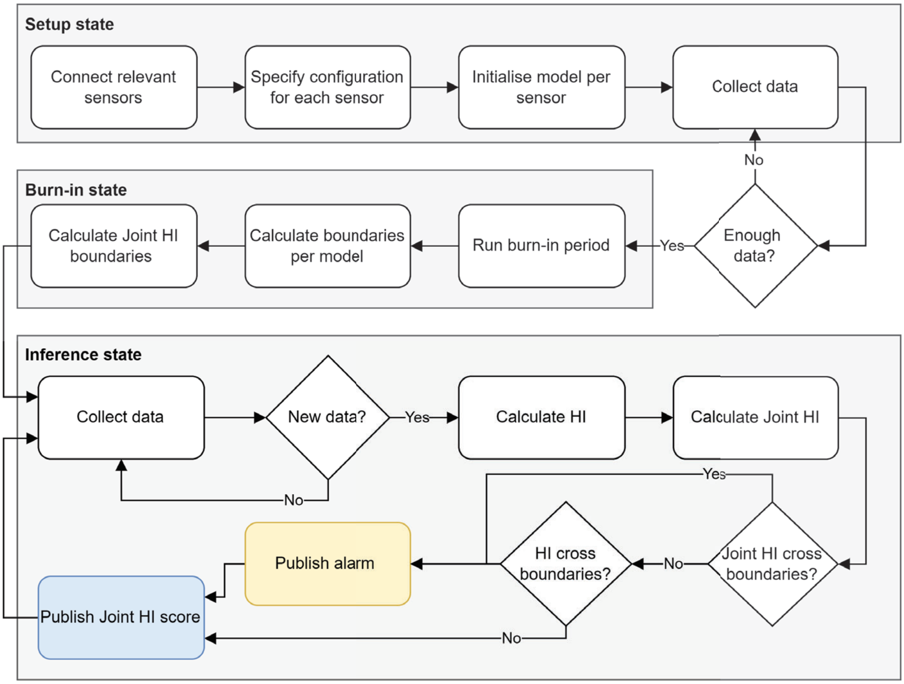


Fig. 7.1: The execution flow of the three states: setup, burn-in, and inference [Paper D | [4]].

The *setup state* is for the first stage of the execution flow. Here the different sensors for the machinery are specified. This specifies which component models to use. In this state, the initial data collection is also started, and when enough data has been collected, the *burn-in state* is started. The first step of the *burn-in state* is to train the component models. When each model has been "burned in", each component model's boundaries are calculated. With the individual boundaries calculated, a joint HI (HI^j) is calculated for the entire system. Hereafter, the *inference state* is executed as a continues operation. When a new data sample is collected, a new HI score is calculated first for each component model and then HI^j. The HI scores are then checked against the boundaries, and if any exceed them, an alarm is published, indicating maintenance could be needed.

As ModularHI was built to be used by SMEs, it is also built to handle different data types. This is why it consists of the so-called component models. Each selected sensor input has its dedicated component models specifically for that type of measurement data. These different component models are already pre-trained on a relevant dataset, which should improve their performance through the method of transfer learning [138]. For ModularHI, the HI scores for each component model were chosen to be calculated with individual DAEs. Since each component model handles univariate data, the size of each DAE is also limited in the number of parameters. Moreover, as the data is time series data, the DAEs were built with LSTM layers. The general DAE architecture can be seen in Table 7.1.

Table 7.1: The LSTM DAE for the component models [Paper D | [4]].

Layer type	Specification
Input	Shape (batch size, window size, 1)
LSTM	8 units, returned sequence
Dropout	Probability 0.5
LSTM	4 units
Repeat vector	8 times
LSTM	4 units, returned sequence
Dropout	Probability 0.5
LSTM	8 units, returned sequence
Output	Time distributed of window size

Even though all the different sensor inputs uses the same DAE architecture, they are as stated individual pre-trained on relevant data. This means that temperature sensors are pre-trained on other temperature data and the same is true with e.g. vibration. The HI score for each component model is based upon the MAE of the reconstruction of the input-signal as seen in

Equation 7.1.

$$\mathrm{HI}_k^m = \frac{1}{n}\sum_{i=1}^{n} |x_i - \tilde{x}_i| \, . \tag{7.1}$$

HI^m is the HI score of the k'th component model m of M component models for a certain setup case. Here $\boldsymbol{x}$ referrers to a vector of sensor measurement with a window-size n. $\tilde{\boldsymbol{x}}$ is the reconstructed measurements with the same window-size n. Each component model has its own upper-boundary. This boundary is calculated in the *burn-in state* by finding the standard deviation σ from the sample mean. The boundary is then set to $9\,\sigma$ and the lower is set to 0 as MAE is a non-negative real number.

The *aggregator's* main task is to combine all of the component models HI score and calculate the joint HI (HI^j). It also evaluates both the HI scores from the component models and joint HI against the boundaries. For this study, it was chosen that HI^j was calculated as mean of all the component models N HI scores:

$$\mathrm{HI}^j = \frac{1}{N}\sum_{i=1}^{N} \mathrm{HI}_i^m \tag{7.2}$$

In order to give the operators more influence over the system it is possible to specify weights for each component models HI score. This can be beneficial if the operator know that e.g. the temperature measurement is crucial for the stability of the machinery. Thus HI^j is calculated as Equation 7.3. By default, all the weights are set to 0.5 and thus HI^j is calculated as Equation 7.2.

$$\mathrm{HI}^j = \frac{\sum_{i=1}^{N} w_i \cdot \mathrm{HI}_i^m}{\sum_{i=1}^{N} w_i} \tag{7.3}$$

The upper-boundary for HI^j is calculated during the burn-in state as $9\,\sigma$ after all the component models finished their HI calculations. Where σ is calculated from the sample mean of all HI^j scores from the burn-in state. In the inference state for each new data point x_t all of the component models will use a sliding window of the last n samples including the new sample $\boldsymbol{x} = \{x_{t-n+1}, \ldots, x_t\}$. Hereafter, all HI^m HI scores will be used to calculated the systems current HI^j. When HI^j is calculated all of the boundaries are checked to check if any is exceeded. If any boundary is exceeded an alarm is published. In Figure 7.2 this inference state is visualised.

Experiments

To validate ModularHI, a relevant dataset was needed. Unfortunately, the amount of suitable manufacturing datasets is limited; thus, the two jet engines dataset from NASA was chosen. Specifically, CMAPSS [131] and N-CMAPSS [132]. These were deemed relevant as they have continuous measurements until failure with various different sensor readings and types. The

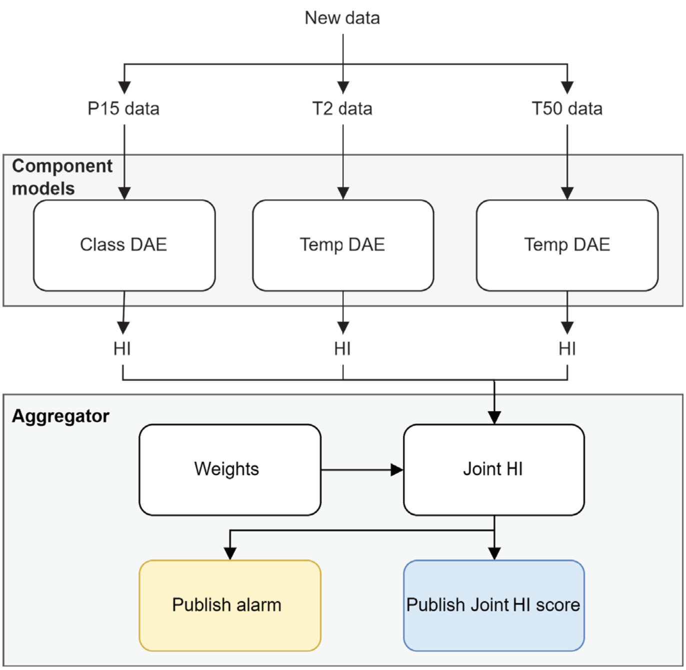


Fig. 7.2: The combination of the different component models and the aggregator from CMAPSS T7 [Paper D | [4]].

CMAPSS dataset contains a single number for each sensor per flight until failure. The N-CMAPSS contains sensor measurements throughout complete flights (including ascending and descending) until failure.

For the CMAPSS dataset, the experiments were only conducted with single jet engine recordings from the dataset. The study used the data from engine number 1 from the dataset FD001, which contains 220 samples. The hyperparameters for ModularHI are mainly the burn-in period and window-size. For the experiments on CMAPSS, these were chosen as 78 and 8, respectively. The dataset contains temperature readings and other time series data types (e.g. flow and pressure). The component models for the temperature sensors were pretrained on the temperature readings from the historic weather dataset[1]. The other component models were pretrained on accelerometer data from smart laboratory at AAU. A total of 8 tests were performed on the CMAPSS dataset, as can be seen in Table 7.2. The first two tests examine the usability of ModularHI, where the sensors have clear degradation in them. Tests 3, 6 and 8 examine ModularHI when there is a mix of sensors that have degradation and some that do not. Tests 4 and 7 test

[1] `https://www.kaggle.com/budincsevity/szeged-weather`

the same but with applied weights to the sensors with degradation. Lastly, test 5 is with a single sensor with no clear degradation.

Table 7.2: The eight test setups for CMAPSS dataset. Each sensor has its own model; if several sensors are mentioned together with only one model, then all sensors have that type of model. The same is true with the assigned weights. The model *T* is pre-trained temperature data and *C* is a generic model pre-trained on accelerometer data [Paper D | [4]].

Tests	Sensors	Models	Weights
CMAPSS - T1	T50	T	0.5
CMAPSS - T2	T30, T50	T	0.5
CMAPSS - T3	T2, T30, T50	T	0.5
CMAPSS - T4	T2, T30, T50	T	0.6, 0.2, 0.2
CMAPSS - T5	P15	C	0.5
CMAPSS - T6	P15, T2, T50	C, T, T	0.5, 0.5, 0.5
CMAPSS - T7	P15, T2, T50	C, T, T	0.2, 0.2, 0.6
CMAPSS - T8	P2, P15, epr, farB, Nf_dmd, PCNfR_dmd, T50	T50: T, Rest: C	0.5

To test ModularHI on a more complex dataset, the second dataset is the N-CMAPSS. Once again, only a single engine was chosen, which was engine number 2 from dataset DS01. As the N-CMAPSS contains both ascend and descend, there is a high variance in the data. ModularHI is built to detect deviation from the stable operation. Thus the study decided only to keep the cruising path between 20,000 and 30,000 feet along with a minimum of 1024 observations. The cleaned dataset contained four sensor measurements, and the first 105,876 was chosen as the burn-in period. A window-size of 1024 was chosen. A total of five tests on the dataset were conducted. As the data changed abruptly between flights, all mini-batches only contained one flight. The tests conducted can be seen in Table 7.3. The first two tests were only on a single measurement, whereas test 2 had clear degradation. Tests 3 and 4 had more sensors, and the last test utilised different weights.

Results

For the experiments for the CMAPSS dataset, the first two tests could detect the degradation and publish alarms. Test 2 result of HI^j can be seen in Figure 7.3. Here it can be seen that the MAE reconstruction loss, e.i. the HI, is starting to grow, indicating a degradation. It can also be seen that an alarm is published at time-step 212. Tests 1, 3, 4, 5, 6 and 7 could detect the degradation in time. Test 8 contained seven measurements, but only one (T50) had a visual degradation. Test 8 also detected degradation as T50 exceeded its

Table 7.3: The five test setups for N-CMAPSS dataset. Each sensor has its own model; if several sensors are mentioned together with only one model, then all sensors have that type of model. The same is true with the assigned weights. The model *T* is pre-trained temperature data and *C* is a generic model pre-trained on accelerometer data [Paper D | [4]].

Test	Sensors	Models	Weights
N-CMAPSS - T1	T40	T	0.5
N-CMAPSS - T2	SmLPC	C	0.5
N-CMAPSS - T3	T40, SmLPC, SmHPC	T, C, C	0.5
N-CMAPSS - T4	T2, SmLPC	T, C	0.5
N-CMAPSS - T5	T2, SmLPC, SmHPC	T, C, C	0.6, 0.2, 0.2

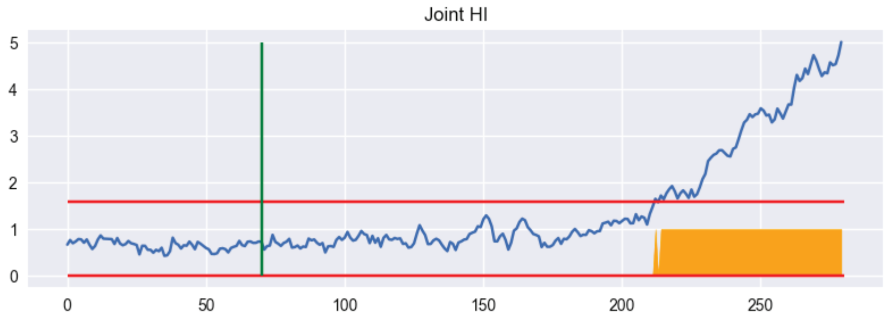


Fig. 7.3: The HI^j result from Test 2. The green vertical line indicates the end of the burn-in state. The horizontal red lines indicate the calculated HI^j boundaries. The orange area is when an alarm is published [Paper D | [4]].

boundary, which triggered the aggregator to publish an alarm. Moreover, test 8 HI^j also exceeded its boundaries at time-step 243.

For the N-CMAPSS dataset, ModularHI detected a few instances of boundary crossing and degradation in the HI^j for tests 3 and 5. However, it was not nearly as stable in publishing the alarm as was the case with the CMAPSS dataset.

Conclusion and Future Work

The study proposed the novel method ModularHI to construct HI values for an arbitrary machine. The experiments on the CMAPSS showed that the system is applicable to a machine which has a general stable baseline with low variance. The experiments and the more complex dataset, N-CMAPSS, showed that the system is currently not applicable to a system with a high variance in the sensor readings. As of this study, two main aspects could use more research, which are the two hyperparameters: burn-in period and window-size. For these tests, they were set based on knowledge about the data. However, that can not be assumed to be possible for SMEs. Therefore, methods should be implemented to do this automatically. Future research should also focus on the design of the boundary, which currently is fixed to 9σ. Currently, the model has only been tested on classical time series data. For future work, ModularHI should also be tested on other data types, such as image data in conjunction with sensor data. ModularHI should also be tested in a real production environment and data to validate it further. Lastly, other methods for the *aggregator* could be investigated to handle the data. In principle, the aggregator could be any form of function approximation, such as an NN. With the results of CMAPSS, ModularHI showed it could be used to construct suitable HI scores for devices without expert knowledge and is thus a start in aiding SMEs for better PHM methods.

7.2 Implications

The paper proposed a novel method of constructing HI scores for arbitrary machinery. ModularHI is a further extension of the underlying architecture presented in [Paper B | [2]]. It is meant as a method for SMEs to get PHM functionality within their production without having experts within that field. This will enable SMEs in the future to deploy such systems with low effort and monitor their process more closely, which would reduce their downtime and cost. ModularHI is not designed to beat other methods engineered for specific devices but to be more versatile. The experiments showed it works as expected on a "simple" CMAPSS dataset. It can also detect degradation and publish alarms even when multiple sensors do not register any changes. Nonetheless, ModularHI struggled in detecting degradation in more "complex" datasets with high variance in the data, such as N-CMAPSS, and thus is not suited for all types of machinery. The contributions of this paper can be summarised as follows:

1. A novel method of constructing HI scores for an arbitrary machinery.
2. Completely data-driven method and does not require expert knowledge nor historical data to use and thus is suited for SMEs.
3. Built-in modules were both component models, and the aggregator can be changed.

Chapter 8

Use Cases of AI in Manufacturing

Paper E has the title *Artificial intelligence and machine learning*. The paper has been submitted and published as a book chapter to the book The Future of Smart Production for SMEs, Springer in 2022. The book chapter relates to **RQ1** and answers it through the following sub-questions:

> **RQ1.2:** What are the challenges the SMEs face when integrating IoT and AI solutions?
>
> **RQ1.3:** Why should the SMEs adopt the use of IoT and AI and what are the benefits?

This book chapter is for the book *The Future of Smart Production for SMEs*, where its main contribution is to give a short overview of what AI has been used for within the industry. It briefly introduces AI and machine learning, followed by examples of AI in SMEs and larger enterprises. As this chapter is an extended abstract of the mentioned paper, [Paper E | [5]], repetition of context, phrasing, results, figures and tables are to be expected and is from that source.

8.1 Extended Abstract

Introduction

It is known that the use of AI is expanding across different sectors, including manufacturing. AI is used in different fields such as planning, computer vision and robotics. To use AI within these fields often machine learning is used as the main tool. In order to solve a problem with machine learning, the following three components are needed:

- A decision process
- An error function
- An optimiser

The first component, *the decision process*, is the algorithm's output, e.g. from an NN. Then the *error function* is used to calculate the difference between the output and the true value. Lastly, the *optimiser* is used to minimise this difference by making changes to the decision process. It is, in general, required to have a large amount of data to get satisfactory results with machine learning. Though there are newer methods to reduce the required hardware and data [139].

Use cases of AI

The results of the survey [Paper A | [1]] found that only in five publications of AI was utilised or discussed in a manufacturing SME context. One of the reasons for it was that SMEs focus more of their knowledge and resources on easier-to-use technologies such as IoT and cloud solutions [28]. It has been identified by Watney and Auer [140] that with AI, SMEs can get benefits from, e.g. predictive maintenance and quality control. Nevertheless, once again, the lack of real-world integration in SMEs is mainly due to the lack of knowledge and resources [141]. To overcome this, pilot projects are suitable starting methods for SMEs. These pilot projects would preferably be with business partners, consultancy companies or research institutes. Afterwards, the companies can start to be independent with the gained knowledge and set up their own AI group [142].

In a survey from Brosset et al. [143], different use cases of AI in manufacturing companies were described. For example, Bridgestone used AI to achieve 15% more uniformity in their tire production. Nokia installed camera surveillance at the production and used AI to alert if any irregularities occurred. A McKinsey article also described how a cement factory used AI to control different processes and thus relied less on operator expertise. A study of using AI in an SME suggested using open alliances with non-competing SMEs to set up a test-driven environment to enhance the knowledge of AI. The study also suggested to start using solvable problems and low-cost areas [144].

8.2 Implications

The book chapter presented both different use cases of AI in manufacturing and the challenges SMEs is faced with. The book chapter works in conjunction with [Paper A | [1]] as a background paper for SMEs. Mainly, the book chapter presented the challenge of lack of knowledge and expertise within the field of AI and machine learning. The main suggestion in the book chapter was to collaborate with companies and research institutions. Moreover, the suggestion is to start pilot projects on low-cost operations and gain knowledge from them. The contributions of this paper can be summarised as follows:

1. The paper gives an overview of AI in manufacturing.
2. Gives an insight that SMEs lack knowledge and resources to implement AI in their production.

Chapter 9

Anonymising and Sharing Production Data

Paper F has the title *On the topic of anonymising production data for machine learning*. The paper has been submitted to Journal of Intelligent Manufacturing, Springer. The paper relates to **RQ2** and answers it through the following sub-question:

> **RQ2.1:** What are the challenges of using and sharing data for SMEs and why should they do it?

This paper discusses the topic of why it is important for companies, including SMEs, to share data. It gives perspectives from other sectors and viewpoints from manufacturing companies. Therefrom, the paper discusses how and why data could be anonymised to preserve intellectual properties (IP) and presents a six-step general guideline for anonymising data. Lastly, experiments with the guidelines are conducted to compare the results before and after anonymisation. As this chapter is an extended abstract of the mentioned paper, [Paper F | [6]], repetition of context, phrasing, results, figures and tables are to be expected and is from that source.

9.1 Extended Abstract

Introduction and Company Perspective

With AI and NN, large datasets are often needed to train the models. As Industry 4.0 has increased focus, the use of AI in manufacturing continues to grow in interest. However, there exists only a few relevant datasets for the manufacturing industry [3, 131, 132, 145] and thus a void of manufacturing dataset is present. To understand the lack of manufacturing dataset, it is required to know the thought process of companies. It has been shown that companies are willing to share their data with business partners, and this mentality is expected to grow [146, 147]. When it comes to share data with third parties and thus not knowing who sees and use the data, the companies are more reluctant. As of this paper, no study on the topic has been conducted. Based on experiment with realising the AICD dataset [Paper C | [3]] and collecting data for [Paper G | [7]], some general thought processes can be described. The main fear is that company IP gets unintentionally distributed to the competition. On the other hand, they also said that they want others to see the data as they expect it could benefit them.

Data anonymisation is not a new practice as it has been conducted for years in, e.g. the health sector. Currently, companies with a business within the European Union should focus on general data protection regulation (GDRP) [148]. The GDPR rules are paramount for companies to apply even though they have shown to be challenging for sharing and anonymising data [149, 150]. As stated, different types of methods and studies in anonymising data within the health sector have been conducted [151–153]. Nonetheless, the study found no prior studies addressing sharing and anonymising manufacturing data regarding machine learning. The research gap for missing datasets from the manufacturing industry has already been established [134, 154].

Anonymisation Guidelines

Before any collection and anonymisation of data for machine learning, it is beneficial to define the problem which is to be solved. This would also make it easier for the companies to select what type to collect and anonymise it. As the guidelines in this study are general, it is impossible to set specific guidelines. Instead, different data categories are discussed.

Personal data

As mentioned, the GDPR rules are paramount, especially when dealing with personal data. If the company deems personal data crucial for the dataset, the correct measures must be assured. This can, e.g. by specifying unique IDs per operator. This ID should not be used elsewhere, so it should not be, e.g. the salary number. Moreover, if the person's sex, age or education is required, it can be necessary to split it into groups if the dataset is not large enough.

Product data

Information regarding the product could be relevant to the dataset. Therefore, the study advises including it. If the company does not want the real product name and information to be present, it should be anonymised. This can be done by changing the product's name and its variations. It is advised to make the anonymised product name so that it is known what the product is and its functionality. If measurements are necessary and the company does not want the real measurements to be disclosed, they can be altered with, e.g. multiplying the same constant to them all, thus still keeping the integrity of the data.

Machine configuration data

Information regarding machine configuration could also be relevant to the dataset. If the machine has different configurations during production, this information is relevant for the produced product. As the configuration settings are often discrete values, they can be anonymised by getting it a name which represents the setting. If the data is continuous data, it can again be anonymised by multiplying a constant to it. If the dataset contains PLC tags, they should also be changed such that they indicate what they do. An example of a dataset where the PLC tag names were changed is [Paper C | [3]].

Time series sensor data

If the dataset contains values from sensor readings, it is probably time series data. One important characteristic of time series is temporal information. This information is paramount to preserve in a dataset meant for machine learning. Prior research has argued it is not possible to create a data-driven anonymisation algorithm for time series data [155]. Thus the study argues that the time series itself should not be altered, but the information surrounding it should, e.g. PLC tag name. If the company is not satisfied with that amount of anonymisation, it is possible to anonymise the time series data by including only statistical features such as standard deviation, kurtosis and skewness.

Image data

Image classification and object detection can be useful for companies. Companies could be more reluctant to share image data as there could be a lot of information and possible IPs. The study does not recommend altering the images as this could lead to unknown behaviour. Instead, the images should be screened for personal data and IPs and removed accordingly. Moreover, it is recommended to alter the file names of the images and remove/alter the embedded meta-data.

Sound data

Sound recording can be useful in manufacturing, e.g., detecting a machine failure. The recordings can be stored in files such as FLAC and MP3. If the data is stored in these types of files, it is important to remove/alter the file names and embedded meta-data. Moreover, the recordings should be screened for personal data, such as a person saying some personal information or IPs in the recordings. If the company does not want to share the raw data, they could share the derived features, such as the spectrogram or fast Fourier transform. It should be noted that it can be possible to reconstruct the sound from these features.

Video data

As video data contain images and sound, the guidelines described there also apply to video. The companies should ensure that no personal data is in the footage, along with removing/altering the embedded meta-data. Depending on the objective, only footage of the process or product of interest is recommended to avoid other irrelevant information.

This study presents a six-step general guideline based on the guidelines for the different data categories and what is done in other sectors. The highlighted six-step guidelines below is quoted from this study, *On the topic of anonymising production data for machine learning*, [Paper F | [6]].

1. **Define the problem**
   Define the problem which the dataset should solve.
2. **Inspect for sensitive personal information**
   Investigate if the dataset contains any personal information. If that is the case, investigate relevant legislation (e.g. GDPR) and act accordingly. If personal information is present, get approval to include it.
3. **Remove sensitive personal information**
   Remove the information that is in contradictions against the legislation and does not have personal approval.
4. **Inspect for IPs**
   Inspect the datasets for company IPs. This can be classified documents, pictures etc.
5. **Anonymise**
   Anonymise the data which is not to be shared openly. Only anonymise the features that contain IP information regarding the suggestions of the specific data types. This includes changing PLC tag names and removing samples if necessary.
6. **Re-check**
   Perform steps 2-3 again to make sure nothing was missing. Moreover, check for hidden information such as meta-data in files such as images and check file names for unwanted information.

Experiments and Results

The study conducted two experiments on public datasets to demonstrate the six-step guidelines. Each experiment was conducted in two trials, one without anonymisation and one with. All of the data preprocessing conducted in both experiments were the same for the original and anonymised datasets.

The first experiment was concerning *time series and anomaly detection*. The dataset used was NASA Bearing [145]. Only data from the first test was used, which was preprocessed down to 2156 samples from 8 sensors. The study performed one of the anonymisation guidelines of time series data, specifically to calculate the statistical features. The statistical features were calculated on a per-sample basis. The statistical features calculated were: mean value, median value, min value, max value, skewness, kurtosis and standard deviation. The datasets were split into training and test data and standardised according to Equation 6.1. A DAE was chosen as the anomaly detection model with the architecture shown in Table 9.1. The loss function was MAE with Adam optimiser, and the experiments were trained for 100 epochs with 32 as the batch size. Since this is an anomaly detection problem, the test data consists of data where the bearing starts to fail. Thus, the MAE (reconstruction loss) is expected to rise towards the end. In Figure 9.1, the MAE can be seen for both the original and anonymised data. It can be seen that both of them spike at the error in the end, and for the first 600 samples, the two trials follow each other.

Table 9.1: The architecture of the DAE model. **Batch** in the *output size* column specifies the batch size of the training. ***8**** in the *output size* column is the 8 features from the original dataset. For the anonymised dataset, it would be 7 instead [Paper F | [6]].

Operation layer	Activation	Output size
Dense	Leaky ReLu	(Batch, ***8****)
Batch Normalisation	-	(Batch, ***8****)
Dense	Leaky ReLu	(Batch, 8)
Dropout of rate 0.2	-	(Batch, 8)
Dense	Leaky ReLu	(Batch, 4)
Dropout of rate 0.2	-	(Batch, 4)
Dense	Leaky ReLu	(Batch, 8)
Batch Normalisation	-	(Batch, 8)
Dense	-	(Batch, ***8****)

The second experiment concerns *structured data and classification* and uses the data from the Vehicle dataset[1]. The Vehicle dataset contains data from used vehicles sold along with different characteristics of said vehicles, e.g.

[1] https://www.kaggle.com/datasets/nehalbirla/vehicle-dataset-from-cardekho

(a) Snippet of first 600 samples.

(b) Complete test set.

Fig. 9.1: The MAE of test data from both the original and the anonymised dataset [Paper F | [6]].

transmission type, fuel type and the number of owners. This experiment's objective was to classify if the vehicle had one owner or more. The column *Name* was removed from the dataset and *Fuel* and *Seller type* were changed into one-hot encoding. Lastly, *Transmission* was changed to Boolean data type describing if the vehicle had a manual transmission or not. The anonymisation was done by changing the column names such that selling_price is price and km_driver and so on. Moreover, the values in *price* and *year* a constant was added to them. Specifically, a value of 10,000 was added to the price and two years to the year column. The datasets were randomly shuffled and split into training and test sets and standardised according to Equation 6.1. The study used a fully connected NN as the model, which can be seen in Table 9.2. The model was trained with the Adam optimiser for 100 epochs with a batch size of 32. The loss function was categorical cross-entropy.

The test showed an accuracy of 74.7% on the original dataset and 74.8% for the anonymised. Thus the models achieved equal performance on the two datasets.

Conclusion

As presented by this study, the topic of anonymising production data is sparse. The first thing companies should do is to figure out what problem the dataset should solve. With this knowledge, they are better equipped to determine what kind of data is needed and how to anonymise it. The three main things they should focus on when anonymising the data are keeping the data's integrity, IPs and GDPR for personal data. This study encourages researchers and companies to share the production data to fill the gap of missing datasets. Moreover, research should focus on finding the drivers behind companies sharing their data, as this is a research gap.

Table 9.2: The architecture of the NN model. The batch value in the output size column specifies the batch size of the training [Paper F | [6]].

Operation layer	Activation	Output size
Dense	Leaky ReLu	(Batch, 16)
Dropout of rate 0.4	-	(Batch, 16)
Dense	Leaky ReLu	(Batch, 32)
Dropout of rate 0.4	-	(Batch, 32)
Dense	Leaky ReLu	(Batch, 32)
Dropout of rate 0.4	-	(Batch, 32)
Dense	Leaky ReLu	(Batch, 16)
Batch Normalisation	-	(Batch, 16)
Dense	Softmax	(Batch, 2)

9.2 Implications

This paper presented a six-step general guideline for anonymising production data to be shared. The paper identified the lack of research within a couple of areas. Firstly, the lack of research on how production data should be anonymised. Secondly, the lack of research related to the drivers from the companies in regards to why they would want to share their data. The paper dealt with manufacturing in general, but since the lack of research within the area is lacking, the same can be said for SMEs. As prior studies have shown, the need to cooperate through, e.g. open innovation is a benefit. Therefore SMEs sharing their data could benefit them as well as the research. In summary, the contributions of this paper can be summarised as follows:

1. Identified the lack of research on how to anonymise manufacturing data.
2. Identified the research gap of the drivers behind companies sharing their data.
3. Proposed six-step general guidelines for anonymising manufacturing data for machine learning.

Chapter 10

Field Study of Machine Learning and IoT in an SME

The final paper of this PhD thesis, Paper G, has the title: *An in-depth investigation of machine learning and IoT adoption at a manufacturing SME: a field study*. The paper was submitted in 2022 to the 32nd International Conference on Flexible Automation and Intelligent Manufacturing (FAIM), Elsevier. The final paper relates to **RQ1** and **RQ3** and answers them through the following sub-questions:

> **RQ1.2:** What are the challenges the SMEs face when integrating IoT and AI solutions?
>
> **RQ3.1:** What is the preferred starting methodology for SMEs beginning their utilisation of IoT and AI?
>
> **RQ3.2:** How can an IoT and AI setup be at an SME, both hardware and software-wise?

This paper is a field study investigating how an IoT setup can look at a manufacturing SME. The IoT setup is shown and discussed with regards to both hardware and software. Furthermore, the data captured through the IoT setup was used to train an NN model to control crucial process parameters. Lastly, the perspective from the company is discussed regarding the added benefits and the change in working methodology. As this chapter is an extended abstract of the mentioned paper, [Paper G | [7]], repetition of context, phrasing, results, figures and tables are to be expected and is from that source.

10.1 Extended Abstract

Introduction and Production Setup

It is known that manufacturing SMEs are often faced with challenges due to globalisation [156–158]. To overcome some of these, SMEs have started backshoring their business to increase quality and development control [159, 160]. For SMEs to enhance the quality and control their processes IoT is a viable method, even though the adoption is low [Paper A | [1]], [28].

This field study was conducted at a manufacturing SME producing the latex teat for baby pacifiers. The SME produce the teat in two sizes named 1 and 2 through the latex dipping method [161]. The latex dipping method involves three main steps: preheat of pacifier forms, dipping the forms in latex, and vulcanising the latex in an oven. An illustration of how the oven works at the SME can be seen in Figure 10.1. The pacifier forms are preheated in two water tubs and are afterwards submerged into a latex tub. The latex then attaches itself to the form, which is then vulcanised in the oven. Finally, the vulcanised pacifiers are ejected from the forms. One problem is that, since latex is a natural product, there are material variations between each barrel, and throughout each batch, the properties also change. To compensate for these changes, the operators currently control the following parameters:

- Water tubs temperature.
- Latex cooler temperature.
- Dipping time.

The two different pacifier sizes each have their target weight which is currently measured just after the latex dipping, i.e. the raw weight. The optimal weight target would be the finished product, but since the latex dipping method takes $\sim$1 hour, it is not feasibly to use. To summarise the production and in collaboration with the company, the following challenges and limitations were identified:

1. There is no general information displayed to the operators in regard to the machine's performance and stability.
2. Handover between shifts is limited as it is based on memory.
3. There is no historical IT data collected from the production, which limits, e.g. batch traceability.
4. Operator experience and general guidelines are used to control the control parameters.
5. The control parameters are chosen based on the raw weight.

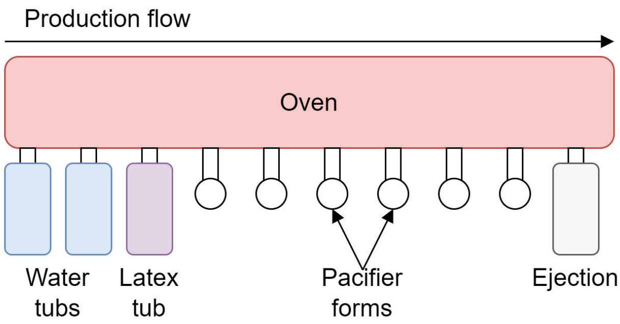


Fig. 10.1: An overview of how the latex dipping production at the SME [Paper G | [7]].

IoT Infrastructure

For the IoT infrastructure, factors such as cost, security, stability, compatibility and what cloud solution to use were considered. As it was unknown exactly what kind of benefits could be expected, the cost was the primary focus. An overview of the IoT setup can be seen in Figure 10.2.

A local computer was chosen as *the cloud system*, mainly because of the low cost and added freedom of the hardware and software used. The computer was an Intel NUC mini PC with the Linux distribution Ubuntu installed as the operating system. The main database was InfluxDB because majority of the data is time series. Moreover, InfluxDB integrates with the dashboard program Grafana, which was chosen to be the platform where information would be displayed to the operators. Lastly, the main communication platform was selected to be MQTT, with Mosquitto as the MQTT broker. The data transferred over MQTT were designed to be in the InfluxDB scheme. Thus the data can be written into the database. All programs were installed as Docker containers to enable easy handling, version control and backup.

To monitor and control the latex machine, *sensors* are needed. The latex machine already had attached sensors, and more required to be retrofitted. For example, additional PT100 temperature sensors needed to be attached to the oven. The PT100 sensors were connected to two ESP32 microcontrollers, which transmit the readings over MQTT. Moreover, an ESP32 with BME280 climate sensors were also placed close to the latex machine. The PLC on the latex machine could send its PLC tag values over MQTT directly. Therefore, the PLC sent data such as latex tub temperature readings and operation mode. Lastly, to control the process, three tablets were placed to measure the weights throughout the production cycle. Manual measurements were made once every hour when the latex machine was running normally.

The two water tubs for preheating the pacifier forms were controlled manually on each device. Therefore a *temperature control system* was created. The temperature control system was created with Python and Pyside6 as the GUI. It communicated with the two water tub heaters through serial communication. The operators then see and change the temperature in this program, and the program sends the data over MQTT to the cloud. When a new barrel of latex is received at the production, an operator tests its temperature, viscosity, and pH value. An overview of all the sensor measurements is shown in the Table 10.1.

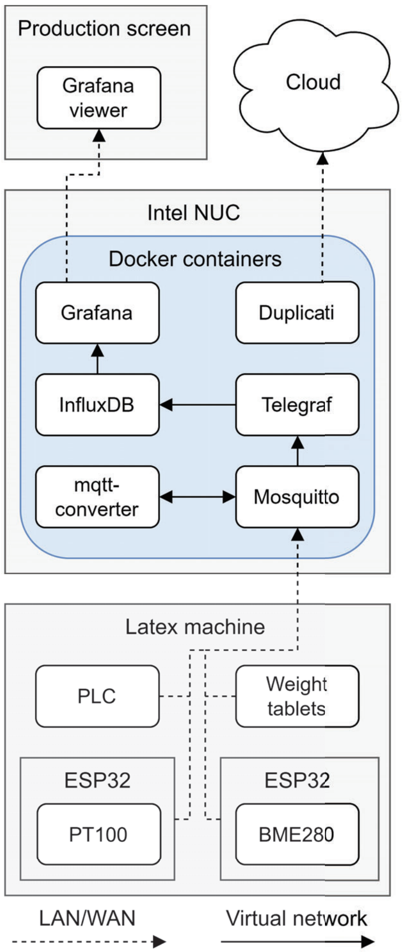


Fig. 10.2: The IoT infrastructure deployed at the SME. LAN/WAN communication is both over WiFi and Ethernet, depending on the device. The virtual network is deployed in the Docker environment enabling the Docker containers of secured inter-communication [Paper G | [7]].

For the IoT system to benefit the operators, business intelligence (BI) solutions were implemented. Specifically, a central screen was placed in the

Table 10.1: The different sensor values and measurements used in this project along with the total number of each measurement [Paper G | [7]].

Measurement	Unit	Amount	Interval (s)
Climate - humidity	%	1	60
Climate - temperature	°C	1	60
Climate - pressure	hPa	1	60
Oven temperature	°C	6	60
Pacifier size	-	1	10
Latex tub temperature	°C	3	10
Latex cooler - temperature	°C	1	1
Latex cooler - set temperature	°C	1	1
Water tubs - temperature	°C	2	1
Water tubs - set temperature	°C	2	1
Water tubs - power	%	2	1
Batch - viscosity	cP	1	-
Batch - pH	pH	1	-
Batch - temperature	°C	1	-

production for the operators. This screen showed a Grafana dashboard for the last 24 hours of production with relevant information such as batch number, incidents etc. This removed the problem between shifts as historical events were now on the screen.

Experiments and Results

To further enhance the production, it would be beneficial if the three main control parameters were not operated by experience but based on data. Unfortunately, the dipping time data was not present in the database; thus, only the water tubs and latex cooler could be estimated. The sensor data, weight measurement and batch specifications were collected over six months. The data was extracted within a window of 10 minutes and only when the latex machine was running. Moreover, now that a data-driven model controls the control parameters, the target weight is chosen to be the finished vulcanised teat. All of the data up to each measurement of the vulcanised teat was mean aggregated to get a snapshot of what the settings and measurements were leading to that exact weight. The final training and test data were then the mean aggregated data with the specified final weight and the target temperature settings for the water and latex tubs. This gave in total a dataset of 2213 samples before the data was split into train and test. Afterwards, the data were standardised according to Equation 6.1.

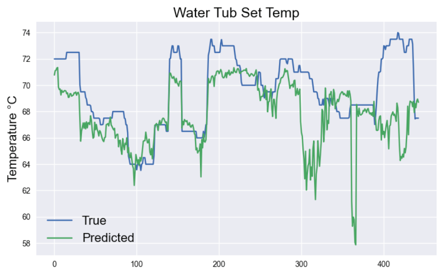
The NN model chosen was a 1D CNN, designed to be able to run in a Docker container in the local cloud. The CNN model's architecture can be seen in Table 10.2. The Adam optimiser was chosen as the optimiser function, and the mean squared error (MSE) function as the loss function. The training was conducted with a batch size of 32, and the model would first stop training when no improvement happened to the loss function in 40 epochs. Two experiments were conducted. The first experiment was with all the data in sequential order, e.i. the same order it was collected in. The second experiment was with the data randomly shuffled across all batches. Both experiments had a training and test data split of 80% and 20%.

Table 10.2: The architecture of our CNN model. The two convolutional layers also use L2 kernel regularisers and no padding. The con The Batch value in the output size column specifies the batch size of the training. The abbreviations are **Activ.** is Activation, **L.ReLu** is Leaky ReLu and **Norm.** is Normalisation [Paper G | [7]].

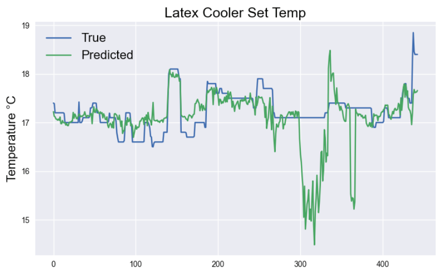
Layer	Activ.	Filters	Kernel	Stride	Output
Input layer	-	-	-	-	(Batch, 22, 1)
Conv1D	L.ReLu	32	5	2	(Batch, 9, 32)
Dropout (0.2)	-	-	-	-	(Batch, 9, 32)
Conv1D	L.ReLu	64	5	2	(Batch, 3, 64)
Dropout (0.2)	-	-	-	-	(Batch, 3, 64)
Dense	L.ReLu	32	-	-	(Batch, 3, 32)
Batch Norm.	-	-	-	-	(Batch, 3, 32)
Flatten	-	-	-	-	(Batch, 96)
Dense	-	2	-	-	(Batch, 2)

The results of the *sequential data* experiment can be seen in Figure 10.3. The training was completed after 183 epochs and had a test MSE loss of 5.158. It can be seen in the figures that the prediction follows the true value at the beginning of the test, but towards the end, it starts to deviate. This could be that a new batch was started at that point, which has new properties that the model has not seen before.

The results of the *shuffled data* experiment can be seen in Figure 10.4. The training was completed after 231 epochs and had a test MSE loss of 0.563. It can be seen in the figures that the predicted value follows the true closely throughout the test data, which the comparably low MSE also indicates. These results suggest that more measurements are still needed for a stable system since the sequential data experience problems when a new batch was introduced.

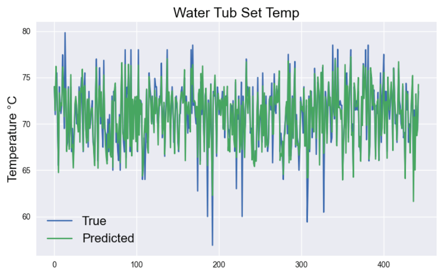


(a) Water temperature.

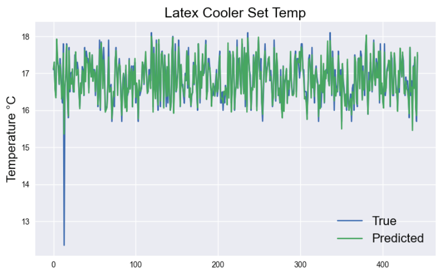


(b) Latex temperature.

Fig. 10.3: The sequential experiment for the water tub and latex tub set temperature. The predicted set temperature is plotted on top of the true values. Both figures are from [Paper G | [7]].



(a) Water temperature.



(b) Latex temperature.

Fig. 10.4: The shuffled experiment for the water tub and latex tub set temperature. The predicted set temperature is plotted on top of the true values. Both figures are from [Paper G | [7]].

Discussion and Conclusion

From the identified challenges and limitations, the first three were solved with the implementation of the IoT infrastructure, the cloud system and the BI implementation. The last two regarding the control parameters are meant to be solved with the NN model. As the experiments showed, currently, not enough data is available to take over this control from the operator fully. The dipping time is missing from the dataset, which could also lead to the control system's instability. Lastly, as the data collected is only from the winter season, the change in climate during the summer is not present in the dataset and could affect the performance. However, when implemented, the NN model would make it possible for the operators to specify a target weight for the pacifiers, and thus the SME is not dependent on the operators' experience.

Though the IoT system is functional as of the implementation, there are some drawbacks. The hardware itself is not enterprise-level, which could prematurely fail. The Intel NUC is not running with redundancy, meaning the entire operating system is compromised if the SSD fails. Nonetheless, all of the Docker containers configuration and the database is currently backup every night of the site. As InfluxDB was chosen as the main database, e.g. batch numbers are not stored by best practices as it is not a time series value. The same is true for the manually typed comments the operators enter in the system.

Such an IoT implementation in any company can be hard to calculate the return on investment. Listening to the experience of the company gives an insight into their viewpoint. They especially say it has transformed their business throughout the organisation. On the production floor, the handover is now done more knowledgeable with the added screen. The operators can now see what happened before their shift and discuss any events. The administration now has insight into their entire production, where they can see a detailed overview throughout all of the batches produced. If any events happen, they can go back in time to see what happened during the production of that specific batch. The company states that they are now making data-driven decisions, and the added transparency and traceability the IoT system provides has enhanced their production. They also expect the NN model to improve the quality of the pacifiers and free up operating hours when implemented.

10.2 Implications

The final paper presented a field study on how IoT and AI can be implemented in a manufacturing SME. The paper showed that a comprehensive and stable IoT infrastructure could be achieved with low-cost hardware and mainly open-source/free software programs. The downside of this approach is that a skilled person is required to set up all of the hardware and software, with the hardest part is making sure everything communications correctly. Moreover, it also showed that companies need to brace themselves on retrofitting sensors and change, e.g. PLC codes, to get high benefits. Lastly, the paper showed that relative simple NN models could potentially control crucial control parameters with enough data collected. While it is hard to estimate how valuable an IoT infrastructure is, the company said it brought a lot of value in the form of available information. The company now has a data-driven decision mentality and is actively seeking new ways of connecting and making parts of the production "smarter". They agree that the most challenging part was how and where to get started, but new ideas and possibilities quickly appear once done. In summary, the contributions of this paper can be summarised as follows:

1. Low-cost hardware and software can be used to build a beneficial IoT infrastructure.

2. Skilled personnel is still needed to make such a system work.
3. Companies should be open to retrofit sensors and change, e.g. PLC codes.
4. AI methods are doable and can bring value to an SME production even with a simple IoT infrastructure.

Chapter 11

Conclusion

This chapter summarises and concludes the research questions presented in Section 2.1 along with presenting future work. Based on the conclusions from the research questions, some general concluding remarks and is presented.

11.1 Research Contribution Summary

This section summarises the three research questions from Chapter 2 and answers them by the sub-questions.

RQ1: What is the current state of IoT and AI adoption in SMEs?

RQ1.1 On what level of integration is the current IoT and AI solutions in SMEs?
It was shown in [Paper A | [1]], IoT and AI are less utilised compared to other Industry 4.0 technologies, such as cloud solutions. From the perspective of IoT and AI, IoT is the more utilised one. The way IoT is utilised is, in general, on a machine-wise level. With the few instances of AI, it was only a basic control scheme for lighting and detecting of machine status. The same status of AI in SMEs was identified by [Paper E | [5]], which also showed some possibilities of AI for SMEs, such as predictive maintenance.

RQ1.2 What are the challenges the SMEs face when integrating IoT and AI solutions?
With the expected benefits of both IoT and AI, the low level of adoption indicates that something must be missing. Both in [Paper A | [1]] and [Paper E | [5]] it was shown that it is too hard and complicated

for SMEs in general to implement IoT and AI. SMEs often lack the resources and knowledge of how and why to implement these technologies in their production or product. In the field study [Paper G | [7]], the challenges were further highlighted with the effort needed to implement said technologies. These challenges were, e.g. retrofitting sensors, managing the infrastructure, and directly changing the PLC code of existing equipment.

RQ1.3 Why should the SMEs adapt the use of IoT and AI and what are the benefits?
The expected benefits of using AI and IoT in SMEs are many. Mainly it can be split into two categories where the first is increasing profit, and the other one is increasing production knowledge. Both [Paper A | [1]] and [Paper E | [5]] identified that the possibilities of AI and IoT, such as predictive maintenance and better quality control, is achievable and can reduce e.g. costs for the SMEs. Moreover, they also showed that production transparency is achievable, which would increase the SMEs' knowledge of their production.

RQ2: How can modern digital solution be designed to aid the adoption within SMEs?

RQ2.1 What are the challenges of using and sharing data for SMEs and why should they do it?
Identified in [Paper C | [3]], there is a gap of manufacturing datasets to train and develop machine learning algorithms. This is a challenge as it makes it harder to train relevant manufacturing models. A contribution to fill this gap was made with the AICD dataset. However, for more companies to share their data, general guidelines are needed to help this process. One of the needed guidelines is anonymising a dataset for machine learning collected at a production. In [Paper F | [6]] a contribution was presented in a six-step general guidelines for anonymising production data. It was also identified that more research is needed to understand the companies' desire to share their data.

RQ2.2 How can an algorithm be designed to overcome the challenges of the SME and thus be used without expert knowledge?
As shown, the topic of AI involves a lot of different sub-areas. This thesis proposed two algorithms in [Paper B | [2]] and [Paper D | [4]] to overcome the challenges of SMEs. [Paper B | [2]] proposed a generic AI-Box architecture. It showed it is possible to design and use a system which hides all the expert parts of training and deploying NN

to the operator. Moreover, it showed it could be designed modularly so that the specific model and sensors can be chosen at creation with little knowledge from the operator. In [Paper D | [4]], an architecture focused on constructing HI for an arbitrary machine was presented. The presented algorithm ModularHI is designed such expert knowledge is not required, nor is historical data. It is modular in the sense it is not fixed to a certain number of inputs, nor is it fixed to a specific sensor type. Both of the two contributions address the major challenges for SMEs, they are designed to be easy to use.

RQ3: How can IoT and AI solutions be integrated in an SME to enhance the production?

RQ3.1 What is the preferred starting methodology for SMEs beginning their utilisation of IoT and AI?
It was found in [Paper A | [1]] that the push for changes in SMEs often comes from within the company itself. It was also found that companies should collaborate with SMEs and research institutes to gather knowledge and experience within the field. The paper also states that starting with a machine-wise IoT implementation is a viable option as it both adds benefits for that specific machine and the gained experience. The field study from [Paper G | [7]] showed an example of a low-cost implementation of an IoT centred around local cloud storage. From the perspective of the company and the stability observed, it is a viable solution to start with.

RQ3.2 How can an IoT and AI setup be at an SME, both hardware and software-wise?
The proposed AI-Box from [Paper B | [2]] addresses one way an AI solution could be built and set up at an SME. It consists of a mini-PC with dedicated GPU and thus accelerated AI capabilities, and with its designed web interface, it can be versatile deployed in production. Its software architecture allows it to be re-configurable to many different objectives without the need for AI expert knowledge. In [Paper G | [7]] an example of an IoT infrastructure is presented. The low-cost architecture shows that achieving a stable and reliable IoT infrastructure with standard consumer sensors and ESP32 microcontrollers is possible. The communication medium with MQTT was able to handle all of the sensor traffic. The local cloud system with the programs installed in Docker containers made it easier to manage and back up the different programs. Lastly, using InfluxDB and Grafana added significant benefits for the SME with easily available information.

11.2 Concluding Remarks and Future Work

This PhD study was centred around the conjugation of two broad technologies (AI and IoT) and a certain type of companies (SMEs). This PhD showed that the use of IoT and AI is limited in SMEs, though the interest in the subjects are still on the rise. The main reason it is limited, is that SMEs generally lack knowledge of the field and have resource constraints. This was also observed during interaction with the IFN project companies. It is recommended both by this PhD study and other researchers that SMEs seek out the possibility for open innovation and collaborate more with other SMEs, research institutions and consultancies. This type of collaboration will enhance their knowledge which was found to be one of the SMEs' drawbacks. It was also found that a suitable method for SMEs to start with these technologies is to select a single machine. They could use this machine as their testing environment for new technologies, which would give them knowledge of said technology and enhance the machine. Implementing IoT in such a way is called machine-wise IoT implementation, and it was found to be a common way for SMEs to implement IoT. This PhD found a gap regarding the Industry 4.0 drivers of SMEs. More research is needed to understand SMEs' needs and drivers. This would help future research in developing more suitable SME solutions.

This PhD has proposed two methods of AI designed with SMEs in mind. The first one was the general architecture of the AI-Box and the second one was a data-driven HI architecture called ModularHI built on top of the underlying architecture of the AI-Box. The AI-Box was designed to be a standalone device with NN capabilities. The main concept was to hide the complex part of machine learning by only exposing the needed information, such as data type, classification or regression problem, and sensor input. Moreover, it was built so that new models and sensors could be added without changing the program. The two experiments performed with the AI-Box showed that it was feasible to have such a standalone device and that the company was fond of the idea. ModularHI was a continuation of this mindset of hiding the underlying model for construction of HI for machinery. The experiments with that model also showed it was possible to construct such an HI system on a dataset with low variance in its signal. More research is needed to enhance its capabilities to more complex data with high variance and change. The literature also showed the significant lack of publicly available datasets from the manufacturing system, along with guidelines on sharing and anonymising them. This PhD thesis both provided such a dataset and proposed anonymisation guidelines. While both have their contributions, more datasets and research into anonymisation and sharing manufacturing data are still needed. The two concepts, AI-Box and ModularHI, show the usability of hiding the complexity of general machine learning problems. It enables SMEs to take advantage of these technologies, which are quickly advancing. While this system will not outperform specifically engineered systems, it will bring benefits for SMEs nonetheless.

The PhD study concluded with a field study of IoT and AI implementation, which put the theoretical and practical experience to the test. The field study showed that SMEs could implement a low-cost IoT architecture both from a hardware and software perspective. Moreover, with the collected data, it is possible to control crucial control parameters within the production. From a company perspective, the IoT implementation gave many benefits. One benefit was the gained production knowledge where the company knows when a batch is produced and under which circumstances. Knowledge sharing between shifts was also improved. The company now makes decisions based on the available data instead of experience and gut feelings. Lastly, the most significant change is that the company went from not knowing where and how to start, to seeing many different opportunities with IoT and AI. More field studies should be conducted to better map the most suitable IoT architectures and AI models.

More easy-to-use devices are needed for more SMEs to adopt IoT and AI. As shown in the field study, it is possible to create such an architecture. However, it still requires an expert to set it up and connect all devices. Moreover, in the future, SMEs should brace the idea of altering already existing machinery by, e.g. retrofitting sensors or changing the PLC code. In retrospect, IoT and AI have come a long way since the beginning of the fourth industrial revolution. As these technologies mature, it is expected to be more accessible for SMEs. It should also be expected for SMEs that their operators and general staff would have a higher knowledge of IT systems to take advantage of said technologies.

Bibliography

[1] E. B. Hansen and S. Bøgh, "Artificial intelligence and internet of things in small and medium-sized enterprises: A survey," *Journal of Manufacturing Systems*, vol. 58, pp. 362–372, 2021, Digital Twin towards Smart Manufacturing and Industry 4.0, ISSN: 0278-6125. DOI: 10.1016/j.jmsy.2020.08.009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0278612520301424.

[2] E. B. Hansen, N. Iftikhar, and S. Bøgh, "Concept of easy-to-use versatile artificial intelligence in industrial small & medium-sized enterprises," *Procedia Manufacturing*, vol. 51, pp. 1146–1152, 2020, 30th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2021), ISSN: 2351-9789. DOI: 10.1016/j.promfg.2020.10.161. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2351978920320187.

[3] E. B. Hansen, E. R. van der Bijl, M. B. Nielsen, M. Svangren Bodilsen, S. V. Berg, J. Kristensen, N. Iftikhar, and S. Bøgh, "A new authentic cloud dataset from a production facility for anomaly detection," in *Towards Sustainable Customization: Bridging Smart Products and Manufacturing Systems*, Cham: Springer International Publishing, 2022, pp. 415–422, ISBN: 978-3-030-90700-6. DOI: 10.1007/978-3-030-90700-6_47.

[4] E. B. Hansen, H. Langseth, N. Iftikhar, and S. Bøgh, "A data-driven modular architecture with denoising autoencoders for health indicator construction in a manufacturing process," unpublished, 2022.

[5] E. B. Hansen and S. Bøgh, "Artificial intelligence and machine learning," in *The Future of Smart Production for SMEs: A Methodological and Practical Approach Towards Digitalization in SMEs*. Cham: Springer International Publishing, 2023, pp. 323–326, ISBN: 978-3-031-15428-7. DOI: 10.1007/978-3-031-15428-7_27. [Online]. Available: https://doi.org/10.1007/978-3-031-15428-7_27.

[6] E. B. Hansen, N. Iftikhar, and S. Bøgh, "On the topic of anonymising production data for machine learning," unpublished, 2022.

[7] E. B. Hansen, K. L. Hansen, N. Iftikhar, and S. Bøgh, "An in-depth investigation of machine learning and iot adoption at a manufacturing sme: A field study," unpublished, 2022.

[8] G. Culot, G. Nassimbeni, G. Orzes, and M. Sartor, "Behind the definition of industry 4.0: Analysis and open questions," *International Journal of Production Economics*, vol. 226, p. 107 617, 2020, ISSN: 0925-5273. DOI: 10.1016/j.ijpe.2020.107617. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925527320300050.

[9] H. Lasi, P. Fettke, H.-G. Kemper, T. Feld, and M. Hoffmann, "Industry 4.0," *Business & Information Systems Engineering*, vol. 6, no. 4, pp. 239–242, Aug. 2014, ISSN: 1867-0202. DOI: 10.1007/s12599-014-0334-4. [Online]. Available: https://doi.org/10.1007/s12599-014-0334-4.

[10] A. Gilchrist, "Introducing industry 4.0," in *Industry 4.0: The Industrial Internet of Things*. Berkeley, CA: Apress, 2016, pp. 195–215, ISBN: 978-1-4842-2047-4. DOI: 10.1007/978-1-4842-2047-4_13. [Online]. Available: https://doi.org/10.1007/978-1-4842-2047-4_13.

[11] W. AULBUR, A. CJ, and R. BIGGHE, "Skill development for industry 4.0," Tech. Rep., 2016. [Online]. Available: https://www.globalskillsummit.com/whitepaper-summary.pdf.

[12] J. Lee, B. Bagheri, and H.-A. Kao, "A cyber-physical systems architecture for industry 4.0-based manufacturing systems," *Manufacturing Letters*, vol. 3, pp. 18–23, 2015, ISSN: 2213-8463. DOI: 10.1016/j.mfglet.2014.12.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S221384631400025X.

[13] M. Wollschlaeger, T. Sauter, and J. Jasperneite, "The future of industrial communication: Automation networks in the era of the internet of things and industry 4.0," *IEEE Industrial Electronics Magazine*, vol. 11, no. 1, pp. 17–27, 2017. DOI: 10.1109/MIE.2017.2649104. [Online]. Available: https://ieeexplore.ieee.org/document/7883994.

[14] B. Chen, J. Wan, L. Shu, P. Li, M. Mukherjee, and B. Yin, "Smart factory of industry 4.0: Key technologies, application case, and challenges," *IEEE Access*, vol. 6, pp. 6505–6519, 2018. DOI: 10.1109/ACCESS.2017.2783682. [Online]. Available: https://ieeexplore.ieee.org/document/8207346.

[15] A. G. Frank, L. S. Dalenogare, and N. F. Ayala, "Industry 4.0 technologies: Implementation patterns in manufacturing companies," *International Journal of Production Economics*, vol. 210, pp. 15–26, 2019, ISSN: 0925-5273. DOI: 10.1016/j.ijpe.2019.01.004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925527319300040.

[16] C. Schröder, "The challenges of industry 4.0 for small and medium-sized enterprises," Tech. Rep., 2016.

[17] I. Colotla, A. Fæste, A. Heidemann, A. Winther, P. H. Andersen, T. Duvold, and M. Hansen, "Winning the industry 4.0 race," *Innovation Fund Denmark*, 2016. [Online]. Available: https://innovationsfonden.dk/sites/default/files/2018-07/bcg-winning-the-industry-40-race-dec-2016.pdf.

[18] Joint Standing Committee on Foreign Affairs, Defence and Trade, *From little things big things grow: Supporting australian smes go global*, 2019. [Online]. Available: https://apo.org.au/node/220426.

[19] S. L. Aranoff, D. R. Pearson, D. T. Okun, C. R. Lane, I. A. Williamson, and D. A. Pinkert, "Small and medium-sized enterprises: Overview of participation in us exports," Jan. 2010. [Online]. Available: https://mdgs.un.org/unsd/trade/s_geneva2011/refdocs/CDs/USA%20-%20ITC%20-%20Small%20and%20medium%20sized%20enterprises%20(Jan%202010).pdf.

[20] EC, *What is an sme?* 2009. [Online]. Available: https://ec.europa.eu/growth/smes/sme-definition_en.

[21] W. Rosen, *The most powerful idea in the world*. Random House, Jun. 2010, ISBN: 978-0-679-60361-0.

[22] IEC 62264-1:2013, "Enterprise-control system integration," International Organization for Standardization, Geneva, CH, Standard, May 2013.

[23] A. Rojko, "Industry 4.0 concept: Background and overview," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 11, no. 5, pp. 77–90, Jul. 2017. DOI: 10.3991/ijim.v11i5.7072. [Online]. Available: https://online-journals.org/index.php/i-jim/article/view/7072.

[24] M.-F. Körner, D. Bauer, R. Keller, M. Rösch, A. Schlereth, P. Simon, T. Bauernhansl, G. Fridgen, and G. Reinhart, "Extending the automation pyramid for industrial demand response," *Procedia CIRP*, vol. 81, pp. 998–1003, 2019, 52nd CIRP Conference on Manufacturing Systems (CMS), Ljubljana, Slovenia, June 12-14, 2019, ISSN: 2212-8271. DOI: 10.1016/j.procir.2019.03.241. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212827119305463.

[25] M. Saturno, V. M. Pertel, F. Deschamps, and E. Loures, "Proposal of an automation solutions architecture for industry 4.0," in *24th international conference on production research*, 2017, pp. 2–3. DOI: 10.12783/dtetr/icpr2017/17675.

[26] M. Rahman, A. D. Fentaye, V. Zaccaria, I. Aslanidou, E. Dahlquist, and K. Kyprianidis, "A framework for learning system for complex industrial processes," in *AI and Learning Systems*, K. Kyprianidis and E. Dahlquist, Eds., Rijeka: IntechOpen, 2021, ch. 2. DOI: 10.5772/intechopen.92899.

[27] A. Khan and K. Turowski, "A survey of current challenges in manufacturing industry and preparation for industry 4.0," in *Proceedings of the First International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'16)*, A. Abraham, S. Kovalev, V. Tarassov, and V. Snášel, Eds., Cham: Springer International Publishing, 2016, pp. 15–26, ISBN: 978-3-319-33609-1. DOI: 10.1007/978-3-319-33609-1_2.

[28] A. Moeuf, S. Lamouri, R. Pellerin, S. Tamayo-Giraldo, E. Tobon-Valencia, and R. Eburdy, "Identification of critical success factors, risks and opportunities of industry 4.0 in smes," *International Journal of Production Research*, vol. 58, no. 5, pp. 1384–1400, 2020. DOI: 10.1080/00207543.2019.1636323.

[29] S. Quinton, A. Canhoto, S. Molinillo, R. Pera, and T. Budhathoki, "Conceptualising a digital orientation: Antecedents of supporting sme performance in the digital economy," *Journal of Strategic Marketing*, vol. 26, no. 5, pp. 427–439, 2018. DOI: 10.1080/0965254X.2016.1258004.

[30] S. Laforet and J. Tann, "Innovative characteristics of small manufacturing firms," *Journal of Small Business and Enterprise Development*, 2006. DOI: 10.1108/14626000610680253.

[31] B. Schönfuß, D. McFarlane, G. Hawkridge, L. Salter, N. Athanassopoulou, and L. de Silva, "A catalogue of digital solution areas for prioritising the needs of manufacturing smes," *Computers in Industry*, vol. 133, p. 103 532, 2021, ISSN: 0166-3615. DOI: 10.1016/j.compind.2021.103532. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166361521001391.

[32] A. H. Lassen, B. V. Wæhren, C. Møller, and O. Madsen, *Innovation factory north*, Danish, Project, 2019. [Online]. Available: https://www.ifn.aau.dk/.

[33] S. Li, L. D. Xu, and S. Zhao, "The internet of things: A survey," *Information Systems Frontiers*, vol. 17, no. 2, pp. 243–259, Apr. 2015, ISSN: 1572-9419. DOI: 10.1007/s10796-014-9492-7.

[34] N. L. Fantana, T. Riedel, J. Schlick, S. Ferber, J. Hupp, S. Miles, F. Michahelles, and S. Svensson, "Iot applications—value creation for industry," *Internet of Things: Converging technologies for smart environments and integrated ecosystems*, pp. 153–206, 2013.

[35] J. Manyika, M. Chui, P. Bisson, J. Woetzel, R. Dobbs, J. Bughin, and D. Aharon, "The internet of things: Mapping the value beyond the hype," *McKinsey Global Institute*, 2015.

[36] M. Haenlein and A. Kaplan, "A brief history of artificial intelligence: On the past, present, and future of artificial intelligence," *California management review*, vol. 61, no. 4, pp. 5–14, 2019. DOI: 10.1177/0008125619864925.

[37] H. Shevlin, K. Vold, M. Crosby, and M. Halina, "The limits of machine intelligence," *EMBO reports*, vol. 20, no. 10, e49177, 2019. DOI: 10.15252/embr.201949177. [Online]. Available: https://www.embopress.org/doi/abs/10.15252/embr.201949177.

[38] M. Miškuf and I. Zolotová, "Comparison between multi-class classifiers and deep learning with focus on industry 4.0," in *2016 Cybernetics & Informatics (K&I)*, 2016, pp. 1–5. DOI: 10.1109/CYBERI.2016.7438633.

[39] M. Sewak, S. K. Sahay, and H. Rathore, "Comparison of deep learning and the classical machine learning algorithm for the malware detection," in *2018 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, 2018, pp. 293–296. DOI: 10.1109/SNPD.2018.8441123.

[40] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS Quarterly*, vol. 28, no. 1, pp. 75–105, 2004, ISSN: 02767783. DOI: 10.2307/25148625. [Online]. Available: http://www.jstor.org/stable/25148625.

[41] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45–77, 2007. DOI: 10.2753/MIS0742-1222240302.

[42] B. Bejczy, R. Bozyil, E. Vaičekauskas, S. B. Krogh Petersen, S. Bøgh, S. S. Hjorth, and E. B. Hansen, "Mixed reality interface for improving mobile manipulator teleoperation in contamination critical applications," *Procedia Manufacturing*, vol. 51, pp. 620–626, 2020, 30th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2021), ISSN: 2351-9789. DOI: 10.1016/j.promfg.2020.10.087. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2351978920319442.

[43] A. S. Staal, C. G. Salvatierra, D. D. Albertsen, M. Mahendran, R. Ravichandran, R. F. Thomsen, E. B. Hansen, and S. Bøgh, "Towards a collaborative omnidirectional mobile robot in a smart cyber-physical environment," *Procedia Manufacturing*, vol. 51, pp. 193–200, 2020, 30th

International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2021), ISSN: 2351-9789. DOI: 10.1016/j.promfg.2020.10.028. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2351978920318837.

[44] S. Kumar, P. Tiwari, and M. Zymbler, "Internet of things is a revolutionary approach for future technology enhancement: A review," *Journal of Big Data*, vol. 6, no. 1, p. 111, Dec. 2019, ISSN: 2196-1115. DOI: 10.1186/s40537-019-0268-2.

[45] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks*, vol. 52, no. 12, pp. 2292–2330, 2008, ISSN: 1389-1286. DOI: 10.1016/j.comnet.2008.04.002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128608001254.

[46] W. Li and S. Kara, "Methodology for monitoring manufacturing environment by using wireless sensor networks (wsn) and the internet of things (iot)," *Procedia CIRP*, vol. 61, pp. 323–328, 2017, The 24th CIRP Conference on Life Cycle Engineering, ISSN: 2212-8271. DOI: 10.1016/j.procir.2016.11.182. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212827116313427.

[47] D. Kandris, C. Nakas, D. Vomvas, and G. Koulouras, "Applications of wireless sensor networks: An up-to-date survey," *Applied System Innovation*, vol. 3, no. 1, 2020, ISSN: 2571-5577. DOI: 10.3390/asi3010014. [Online]. Available: https://www.mdpi.com/2571-5577/3/1/14.

[48] D. Jaspert, M. Ebel, A. Eckhardt, and J. Poeppelbuss, "Smart retrofitting in manufacturing: A systematic review," *Journal of Cleaner Production*, vol. 312, p. 127 555, 2021, ISSN: 0959-6526. DOI: 10.1016/j.jclepro.2021.127555. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S095965262101773X.

[49] T. Stock and G. Seliger, "Opportunities of sustainable manufacturing in industry 4.0," *Procedia CIRP*, vol. 40, pp. 536–541, 2016, 13th Global Conference on Sustainable Manufacturing – Decoupling Growth from Resource Use, ISSN: 2212-8271. DOI: 10.1016/j.procir.2016.01.129. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S221282711600144X.

[50] Á. García, A. Bregon, and M. A. Martínez-Prieto, "A non-intrusive industry 4.0 retrofitting approach for collaborative maintenance in traditional manufacturing," *Computers & Industrial Engineering*, vol. 164, p. 107 896, 2022, ISSN: 0360-8352. DOI: https://doi.org/10.1016/j.cie.2021.107896. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360835221008007.

[51] P. Strauß, M. Schmitz, R. Wöstmann, and J. Deuse, “Enabling of predictive maintenance in the brownfield through low-cost sensors, an iiot-architecture and machine learning,” in *2018 IEEE International Conference on Big Data (Big Data)*, 2018, pp. 1474–1483. DOI: 10.1109/BigData.2018.8622076.

[52] F. Lima, A. A. Massote, and R. F. Maia, “Iot energy retrofit and the connection of legacy machines inside the industry 4.0 concept,” in *IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society*, vol. 1, 2019, pp. 5499–5504. DOI: 10.1109/IECON.2019.8927799.

[53] A. Flammini, P. Ferrari, D. Marioli, E. Sisinni, and A. Taroni, “Wired and wireless sensor networks for industrial applications,” *Microelectronics Journal*, vol. 40, no. 9, pp. 1322–1336, 2009, Quality in Electronic Design 2nd IEEE International Workshop on Advances in Sensors and Interfaces Thermal Investigations of ICs and Systems, ISSN: 0026-2692. DOI: 10.1016/j.mejo.2008.08.012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0026269208004618.

[54] T. R. Kurfess, C. Saldana, K. Saleeby, and M. P. Dezfouli, “A Review of Modern Communication Technologies for Digital Manufacturing Processes in Industry 4.0,” *Journal of Manufacturing Science and Engineering*, vol. 142, no. 11, Sep. 2020, ISSN: 1087-1357. DOI: 10.1115/1.4048206.

[55] S. K. Rao and R. Prasad, “Impact of 5g technologies on industry 4.0,” *Wireless Personal Communications*, vol. 100, no. 1, pp. 145–159, May 2018, ISSN: 1572-834X. DOI: 10.1007/s11277-018-5615-7.

[56] M. Rizzi, P. Ferrari, A. Flammini, E. Sisinni, and M. Gidlund, “Using lora for industrial wireless networks,” in *2017 IEEE 13th International Workshop on Factory Communication Systems (WFCS)*, 2017, pp. 1–4. DOI: 10.1109/WFCS.2017.7991972.

[57] S. Profanter, A. Tekat, K. Dorofeev, M. Rickert, and A. Knoll, “Opc ua versus ros, dds, and mqtt: Performance evaluation of industry 4.0 protocols,” in *2019 IEEE International Conference on Industrial Technology (ICIT)*, 2019, pp. 955–962. DOI: 10.1109/ICIT.2019.8755050.

[58] K. L. Berg, T. Seymour, and R. Goel, “History of databases,” *International Journal of Management & Information Systems (IJMIS)*, vol. 17, no. 1, pp. 29–36, 2012. DOI: 10.19030/ijmis.v17i1.7587.

[59] V. F. de Oliveira, M. A. d. O. Pessoa, F. Junqueira, and P. E. Miyagi, “Sql and nosql databases in the context of industry 4.0,” *Machines*, vol. 10, no. 1, 2022, ISSN: 2075-1702. DOI: 10.3390/machines10010020. [Online]. Available: https://www.mdpi.com/2075-1702/10/1/20.

[60] S. N. Z. Naqvi, S. Yfantidou, and E. Zimányi, "Time series databases and influxdb," *Studienarbeit, Université Libre de Bruxelles*, vol. 12, 2017.

[61] L. Thames and D. Schaefer, "Software-defined cloud manufacturing for industry 4.0," *Procedia CIRP*, vol. 52, pp. 12–17, 2016, The Sixth International Conference on Changeable, Agile, Reconfigurable and Virtual Production (CARV2016), ISSN: 2212-8271. DOI: 10.1016/j.procir.2016.07.041. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212827116307910.

[62] N. Niknejad, W. Ismail, I. Ghani, B. Nazari, M. Bahari, and A. R. B. C. Hussin, "Understanding service-oriented architecture (soa): A systematic literature review and directions for further investigation," *Information Systems*, vol. 91, p. 101 491, 2020, ISSN: 0306-4379. DOI: 10.1016/j.is.2020.101491. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306437920300028.

[63] J. Puttonen, A. Lobov, M. A. C. Soto, and J. L. M. Lastra, "Cloud computing as a facilitator for web service composition in factory automation," *Journal of Intelligent Manufacturing*, vol. 30, no. 2, pp. 687–700, Feb. 2019, ISSN: 1572-8145. DOI: 10.1007/s10845-016-1277-z.

[64] M. Saraswat and R. Tripathi, "Cloud computing: Analysis of top 5 csps in saas, paas and iaas platforms," in *2020 9th International Conference System Modeling and Advancement in Research Trends (SMART)*, 2020, pp. 300–305. DOI: 10.1109/SMART50582.2020.9337157.

[65] S. Yi, C. Li, and Q. Li, "A survey of fog computing: Concepts, applications and issues," in *Proceedings of the 2015 Workshop on Mobile Big Data*, ser. Mobidata '15, Hangzhou, China: Association for Computing Machinery, 2015, pp. 37–42, ISBN: 9781450335249. DOI: 10.1145/2757384.2757397.

[66] K. H. Abdulkareem, M. A. Mohammed, S. S. Gunasekaran, M. N. Al-Mhiqani, A. A. Mutlag, S. A. Mostafa, N. S. Ali, and D. A. Ibrahim, "A review of fog computing and machine learning: Concepts, applications, challenges, and open issues," *IEEE Access*, vol. 7, pp. 153 123–153 140, 2019. DOI: 10.1109/ACCESS.2019.2947542.

[67] E. Hechler, M. Oberhofer, and T. Schaeck, *Deploying AI in the enterprise*, 1st ed. Berlin, Germany: APress, Sep. 2020. DOI: 10.1007/978-1-4842-6206-1.

[68] J. Balzategui, L. Eciolaza, N. Arana-Arexolaleiba, J. Altube, J.-P. Aguerre, I. Legarda-Ereño, and A. Apraiz, "Semi-automatic quality inspection of solar cell based on convolutional neural networks," in *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2019, pp. 529–535. DOI: 10.1109/ETFA.2019.8869359.

[69] M. Etxeberria-Garcia, M. Labayen, M. Zamalloa, and N. Arana-Arexolaleiba, "Application of computer vision and deep learning in the railway domain for autonomous train stop operation," in *2020 IEEE/SICE International Symposium on System Integration (SII)*, 2020, pp. 943–948. DOI: 10.1109/SII46433.2020.9026246.

[70] T. Maruyama, N. Hayashi, Y. Sato, S. Hyuga, Y. Wakayama, H. Watanabe, A. Ogura, and T. Ogura, "Comparison of medical image classification accuracy among three machine learning methods," *Journal of X-Ray Science and Technology*, vol. 26, pp. 885–893, 2018, 6, ISSN: 1095-9114. DOI: 10.3233/XST-18386.

[71] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, Curran Associates, Inc., 2012, pp. 1097–1105. [Online]. Available: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf.

[72] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, *Squeezenet: Alexnet-level accuracy with 50x fewer parameters and <0.5mb model size*, 2016. DOI: 10.48550/ARXIV.1602.07360. [Online]. Available: https://arxiv.org/abs/1602.07360.

[73] K. He, X. Zhang, S. Ren, and J. Sun, *Deep residual learning for image recognition*, 2015. DOI: 10.48550/ARXIV.1512.03385. [Online]. Available: https://arxiv.org/abs/1512.03385.

[74] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 5987–5995. DOI: 10.1109/CVPR.2017.634.

[75] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, *An image is worth 16x16 words: Transformers for image recognition at scale*, 2020. DOI: 10.48550/ARXIV.2010.11929. [Online]. Available: https://arxiv.org/abs/2010.11929.

[76] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv*, 2018.

[77] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot MultiBox detector," in *Computer Vision – ECCV 2016*, Springer International Publishing, 2016, pp. 21–37. DOI: 10.1007/978-3-319-46448-0_2.

[78] K. He, G. Gkioxari, P. Dollár, and R. Girshick, *Mask r-cnn*, 2017. DOI: 10.48550/ARXIV.1703.06870. [Online]. Available: https://arxiv.org/abs/1703.06870.

[79] M. Uemura, M. Tomikawa, T. Miao, R. Souzaki, S. Ieiri, T. Akahoshi, A. K. Lefor, and M. Hashizume, "Feasibility of an ai-based measure of the hand motions of expert and novice surgeons," *Computational and Mathematical Methods in Medicine*, Mar. 2018.

[80] T.-Y. Kim and S.-B. Cho, "Particle swarm optimization-based cnn-lstm networks for forecasting energy consumption," in *2019 IEEE Congress on Evolutionary Computation (CEC)*, 2019, pp. 1510–1516. DOI: 10.1109/CEC.2019.8789968.

[81] B. Zhao, H. Lu, S. Chen, J. Liu, and D. Wu, "Convolutional neural networks for time series classification," *Journal of Systems Engineering and Electronics*, vol. 28, no. 1, pp. 162–169, 2017. DOI: 10.21629/JSEE.2017.01.18.

[82] H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, "Transfer learning for time series classification," in *2018 IEEE International Conference on Big Data (Big Data)*, IEEE, Dec. 2018. DOI: 10.1109/bigdata.2018.8621990.

[83] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997. DOI: 10.1162/neco.1997.9.8.1735.

[84] A. Sagheer and M. Kotb, "Time series forecasting of petroleum production using deep lstm recurrent networks," *Neurocomputing*, vol. 323, pp. 203–213, 2019, ISSN: 0925-2312. DOI: 10.1016/j.neucom.2018.09.082.

[85] A. Essien and C. Giannetti, "A deep learning model for smart manufacturing using convolutional lstm neural network autoencoders," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 9, pp. 6069–6078, 2020. DOI: 10.1109/TII.2020.2967556.

[86] K.-L. Li, H.-K. Huang, S.-F. Tian, and W. Xu, "Improving one-class svm for anomaly detection," in *Proceedings of the 2003 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.03EX693)*, vol. 5, 2003, 3077–3081 Vol.5. DOI: 10.1109/ICMLC.2003.1260106.

[87] M. Sakurada and T. Yairi, "Anomaly detection using autoencoders with nonlinear dimensionality reduction," in *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis*, ser. MLSDA'14, Gold Coast, Australia QLD, Australia: Association for Computing Machinery, 2014, pp. 4–11, ISBN: 9781450331593. DOI: 10.1145/2689746.2689747.

[88] J. Chen, J. Li, W. Chen, Y. Wang, and T. Jiang, "Anomaly detection for wind turbines based on the reconstruction of condition parameters using stacked denoising autoencoders," *Renewable Energy*, vol. 147, pp. 1469–1480, 2020, ISSN: 0960-1481. DOI: 10.1016/j.renene.2019.09.041. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960148119313710.

[89] J. An and S. Cho, "Variational autoencoder based anomaly detection using reconstruction probability," *Special Lecture on IE*, vol. 2, no. 1, pp. 1–18, 2015.

[90] P. S. Deshpande, S. C. Sharma, and S. K. Peddoju, "Predictive and prescriptive analytics in big-data era," in *Security and Data Storage Aspect in Cloud Computing*. Singapore: Springer Singapore, 2019, pp. 71–81. DOI: 10.1007/978-981-13-6089-3_5.

[91] M. Haarman, M. Mulders, and C. Vassiliadis, *Predictive maintenance 4.0: Predict the unpredictable*, 2017.

[92] T. Zonta, C. A. da Costa, R. da Rosa Righi, M. J. de Lima, E. S. da Trindade, and G. P. Li, "Predictive maintenance in the industry 4.0: A systematic literature review," *Computers & Industrial Engineering*, vol. 150, p. 106 889, 2020, ISSN: 0360-8352. DOI: 10.1016/j.cie.2020.106889. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360835220305787.

[93] Z. M. Çınar, A. Abdussalam Nuhu, Q. Zeeshan, O. Korhan, M. Asmael, and B. Safaei, "Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0," *Sustainability*, vol. 12, no. 19, 2020, ISSN: 2071-1050. DOI: 10.3390/su12198211. [Online]. Available: https://www.mdpi.com/2071-1050/12/19/8211.

[94] L. Guo, N. Li, F. Jia, Y. Lei, and J. Lin, "A recurrent neural network based health indicator for remaining useful life prediction of bearings," *Neurocomputing*, vol. 240, pp. 98–109, 2017, ISSN: 0925-2312. DOI: 10.1016/j.neucom.2017.02.045. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231217303363.

[95] L. Guo, Y. Lei, N. Li, T. Yan, and N. Li, "Machinery health indicator construction based on convolutional neural networks considering trend burr," *Neurocomputing*, vol. 292, pp. 142–150, 2018, ISSN: 0925-2312. DOI: 10.1016/j.neucom.2018.02.083. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231218302583.

[96] L. Chen, G. Xu, S. Zhang, W. Yan, and Q. Wu, "Health indicator construction of machinery based on end-to-end trainable convolution recurrent neural networks," *Journal of Manufacturing Systems*, vol. 54, pp. 1–11, 2020, ISSN: 0278-6125. DOI: 10.1016/j.jmsy.2019.11.008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0278612519300974.

[97] A. Rai and J.-M. Kim, "A novel health indicator based on information theory features for assessing rotating machinery performance degradation," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6982–6994, 2020. DOI: 10.1109/TIM.2020.2978966.

[98] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to rul prediction," *Mechanical Systems and Signal Processing*, vol. 104, pp. 799–834, 2018, ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2017.11.016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888327017305988.

[99] Y. Hu, P. Baraldi, F. Di Maio, and E. Zio, "Online performance assessment method for a model-based prognostic approach," *IEEE Transactions on Reliability*, vol. 65, no. 2, pp. 718–735, 2016. DOI: 10.1109/TR.2015.2500681.

[100] O. Gassmann, E. Enkel, and H. Chesbrough, "The future of open innovation," *R&D Management*, vol. 40, no. 3, pp. 213–221, 2010. DOI: 10.1111/j.1467-9310.2010.00605.x. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9310.2010.00605.x.

[101] V. van de Vrande, J. P. de Jong, W. Vanhaverbeke, and M. de Rochemont, "Open innovation in smes: Trends, motives and management challenges," *Technovation*, vol. 29, no. 6, pp. 423–437, 2009, ISSN: 0166-4972. DOI: 10.1016/j.technovation.2008.10.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166497208001314.

[102] A. Leckel, S. Veilleux, and L. P. Dana, "Local open innovation: A means for public policy to increase collaboration for innovation in smes," *Technological Forecasting and Social Change*, vol. 153, p. 119 891, 2020, ISSN: 0040-1625. DOI: 10.1016/j.techfore.2019.119891. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0040162518315403.

[103] C. Schröder, *The Challenges of Industry 4.0 for Small and Medium-sized Enterprises*. Jul. 2016, ISBN: 978-3-95861-543-4.

[104] J. Lenz, T. Wuest, and E. Westkämper, "Holistic approach to machine tool data analytics," *Journal of Manufacturing Systems*, vol. 48, pp. 180–191, 2018, Special Issue on Smart Manufacturing, ISSN: 0278-6125. DOI: 10.1016/j.jmsy.2018.03.003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0278612518300360.

[105] H. Kagermann, W.-D. Lukas, and W. Wahlster, "Industry 4.0: Using the internet of things on the road to the 4th industrial revolution," *VDI news*, vol. 13, no. 1, 2011, German. [Online]. Available: https://www.dfki.de/fileadmin/user_upload/DFKI/Medien/News_Media/Presse/Presse-Highlights/vdinach2011a13-ind4.0-Internet-Dinge.pdf.

[106] K. T. Park, S. J. Im, Y.-S. Kang, S. D. Noh, Y. T. Kang, and S. G. Yang, "Service-oriented platform for smart operation of dyeing and finishing industry," *International Journal of Computer Integrated Manufacturing*, vol. 32, no. 3, pp. 307–326, 2019. DOI: 10.1080/0951192X.2019.1572225.

[107] Y.-C. Chen, K.-C. Ting, Y.-M. Chen, D.-L. Yang, H.-M. Chen, and J. J.-C. Ying, "A low-cost add-on sensor and algorithm to help small- and medium-sized enterprises monitor machinery and schedule processes," *Applied Sciences*, vol. 9, no. 8, 2019, ISSN: 2076-3417. DOI: 10.3390/app9081549.

[108] S. Hyun-Jun, K. Sung-Jin, J. Min-Ho, and O. Chang-Heon, "Development of automatic mold shot measurement and management system for smart factory," English, *International Journal of Electrical and Computer Engineering*, vol. 6, no. 6, pp. 3142–3147, Dec. 2016. [Online]. Available: https://search.proquest.com/docview/1868300470?accountid=8144.

[109] M. Ushada, T. Okayama, N. Khuriyati, and A. Suyantohadi, "Affective temperature control in food smes using artificial neural network," *Applied Artificial Intelligence*, vol. 31, no. 7-8, pp. 555–567, 2017. DOI: 10.1080/08839514.2017.1390327.

[110] J.-u. Jung and K.-h. Jin, "Case studies for the establishment of the optimized smart factory with small and medium-sized enterprises," in *Proceedings of the 2Nd International Symposium on Computer Science and Intelligent Control*, ser. ISCSIC '18, Stockholm, Sweden: ACM, 2018, 27:1–27:5, ISBN: 978-1-4503-6628-1. DOI: 10.1145/3284557.3284692.

[111] D. Mourtzis, N. Milas, and A. Vlachou, "An internet of things-based monitoring system for shop-floor control," *Journal of Computing and Information Science in Engineering*, vol. 18, no. 2, Mar. 2018, ISSN: 1530-9827. DOI: 10.1115/1.4039429.

[112] P. E. Ramírez-Correa, E. E. Grandón, J. Arenas-Gaitán, F. J. Rondán-Cataluña, and A. Aravena, "Explaining performance expectancy of iot in chilean smes," in *Business Information Systems*, W. Abramowicz and R. Corchuelo, Eds., Cham: Springer International Publishing, 2019, pp. 475–486, ISBN: 978-3-030-20485-3.

[113] J. M. Müller and K.-I. Voigt, "Sustainable industrial value creation in smes: A comparison between industry 4.0 and made in china 2025," *International Journal of Precision Engineering and Manufacturing-Green Technology*, vol. 5, no. 5, pp. 659–670, Oct. 2018, ISSN: 2198-0810. DOI: 10.1007/s40684-018-0056-z.

[114] L. Kaňovská and E. Tomášková, "Drivers for smart servitization in manufacturing companies," English, *AGRIS On-line Papers in Economics and Informatics*, vol. 10, no. 3, pp. 57–68, 2018. [Online]. Available: https://search.proquest.com/docview/2121520304?accountid=8144.

[115] K. S. Potapov, P. V. Terentyeva, B. T. Drapikovskaya, I. G. Fomina, and O. A. Shvetsova, "About erp and dss system's introduction and using prospects at small and medium-sized enterprises in russia," in *2016 XIX IEEE International Conference on Soft Computing and Measurements (SCM)*, May 2016, pp. 438–440. DOI: 10.1109/SCM.2016.7519806.

[116] X. Song, H. Tian, and X. Wu, "Study on smes-oriented strategic decision support system based on competitive intelligence," in *2010 International Conference on Management and Service Science*, Aug. 2010, pp. 1–4. DOI: 10.1109/ICMSS.2010.5576728.

[117] H.-K. Kim, W.-H. So, and S.-M. Je, "A big data framework for network security of small and medium enterprises for future computing," *The Journal of Supercomputing*, vol. 75, no. 6, pp. 3334–3367, Jun. 2019, ISSN: 1573-0484. DOI: https://doi.org/10.1007/s11227-019-02815-8.

[118] S. Paul, "Reinforcing your sme against cyberthreats," *Computer Fraud & Security*, vol. 2017, no. 10, pp. 13–15, 2017, ISSN: 1361-3723. DOI: 10.1016/S1361-3723(17)30091-X.

[119] M. Carcary, E. Doherty, G. Conway, and S. McLaughlin, "Cloud computing adoption readiness and benefit realization in irish smes—an exploratory study," *Information Systems Management*, vol. 31, no. 4, pp. 313–327, 2014. DOI: 10.1080/10580530.2014.958028.

[120] V. Varaniūtė, E. Vitkauskaitė, and A. Tarutė, "Peculiarities of iot-based business model transformations in smes," 2018.

[121] D.-I. Shin, "An exploratory study of innovation strategies of the internet of things smes in south korea," *Asia Pacific Journal of Innovation and Entrepreneurship*, vol. 11, no. 2, pp. 171–189, 2017.

[122] M. Wieland, F. Steimle, B. Mitschang, D. Lucke, P. Einberger, D. Schel, M. Luckert, and T. Bauernhansl, "Rule-based integration of smart services using the manufacturing service bus," in *2017 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computed, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, Aug. 2017, pp. 1–8. DOI: 10.1109/UIC-ATC.2017.8397529.

[123] J. Gao, *Machine learning applications for data center optimization*, 2014.

[124] N. Lazic, C. Boutilier, T. Lu, E. Wong, B. Roy, M. Ryu, and G. Imwalle, "Data center cooling using model-predictive control," in *Advances in Neural Information Processing Systems 31*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds., Curran Associates, Inc., 2018, pp. 3814–3823. [Online]. Available: http://papers.nips.cc/paper/7638-data-center-cooling-using-model-predictive-control.pdf.

[125] S. Mittal, M. A. Khan, D. Romero, and T. Wuest, "A critical review of smart manufacturing & industry 4.0 maturity models: Implications for small and medium-sized enterprises (smes)," *Journal of Manufacturing Systems*, vol. 49, pp. 194–214, 2018, ISSN: 0278-6125. DOI: 10.1016/j.jmsy.2018.10.005.

[126] G. E. Krasner, S. T. Pope, *et al.*, "A description of the model-view-controller user interface paradigm in the smalltalk-80 system," *Journal of object oriented programming*, vol. 1, no. 3, pp. 26–49, 1988.

[127] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336–359, Oct. 2019, ISSN: 1573-1405. DOI: 10.1007/s11263-019-01228-7.

[128] T. Linjordet and K. Balog, "Impact of training dataset size on neural answer selection models," in *Advances in Information Retrieval*, Cham: Springer International Publishing, 2019, pp. 828–835, ISBN: 978-3-030-15712-8.

[129] J. Prusa, T. M. Khoshgoftaar, and N. Seliya, "The effect of dataset size on training tweet sentiment classifiers," in *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, 2015, pp. 96–102. DOI: 10.1109/ICMLA.2015.22.

[130] H. Purohit, R. Tanabe, K. Ichige, T. Endo, Y. Nikaido, K. Suefusa, and Y. Kawaguchi, *Mimii dataset: Sound dataset for malfunctioning industrial machine investigation and inspection*, 2019. arXiv: 1909.09347 [cs.SD].

[131] A. Saxena and K. Goebel, "Turbofan engine degradation simulation data set," *NASA Ames Prognostics Data Repository*, pp. 1551–3203, 2008.

[132] M. Arias Chao, C. Kulkarni, K. Goebel, and O. Fink, "Aircraft engine run-to-failure dataset under real flight conditions for prognostics and diagnostics," *Data*, vol. 6, no. 1, 2021, ISSN: 2306-5729. DOI: 10.3390/data6010005. [Online]. Available: https://www.mdpi.com/2306-5729/6/1/5.

[133] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, 2017. arXiv: 1412.6980 [cs.LG].

[134] T. Wuest, D. Weimer, C. Irgens, and K.-D. Thoben, "Machine learning in manufacturing: Advantages, challenges, and applications," *Production & Manufacturing Research*, vol. 4, no. 1, pp. 23–45, 2016. DOI: 10.1080/21693277.2016.1192517.

[135] J. Abrhám, W. Strielkowski, M. Vošta, and J. Šlajs, "Factors that influence the competitiveness of Czech rural SMEs," English, *Agricultural Economics*, vol. 61, no. 10, pp. 450–460, 2015. [Online]. Available: https://www.proquest.com/scholarly-journals/factors-that-influence-competitiveness-czech/docview/2507285194/se-2.

[136] V. Atamuradov, K. Medjaher, P. Dersin, B. Lamoureux, and N. Zerhouni, "Prognostics and health management for maintenance practitioners - review, implementation and tools evaluation," *International Journal of Prognostics and Health Management*, vol. 8, no. 060, pp. 1–31, Dec. 2017. [Online]. Available: https://oatao.univ-toulouse.fr/19521/.

[137] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mechanical Systems and Signal Processing*, vol. 115, pp. 213–237, 2019, ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2018.05.050. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888327018303108.

[138] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2021. DOI: 10.1109/JPROC.2020.3004555.

[139] A. Reuther, P. Michaleas, M. Jones, V. Gadepally, S. Samsi, and J. Kepner, "Survey and benchmarking of machine learning accelerators," in *2019 IEEE High Performance Extreme Computing Conference (HPEC)*, 2019, pp. 1–9. DOI: 10.1109/HPEC.2019.8916327.

[140] C. Watney and D. Auer, *Encouraging ai adoption by eu smes*, 2021. [Online]. Available: https://www.progressivepolicy.org/wp-content/uploads/2021/03/PPI_Encouraging-AI-adoption-by-EU-SMEs-3.24.21-2.pdf.

[141] R. Welte, M. Estler, and D. Lucke, "A method for implementation of machine learning solutions for predictive maintenance in small and medium sized enterprises," *Procedia CIRP*, vol. 93, pp. 909–914, 2020, 53rd CIRP Conference on Manufacturing Systems 2020, ISSN: 2212-8271. DOI: 10.1016/j.procir.2020.04.052. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212827120306223.

[142] A. Ng, *Ai transformation playbook*, 2020. [Online]. Available: https://landing.ai/wp-content/uploads/2020/05/LandingAI_Transformation_Playbook_11-19.pdf.

[143] P. Brosset, A. Jain, Y. Khemka, J. Buvat, A. Thieullent, A. Khadikar, and S. Patsko, "Scaling ai in manufacturing operations: A practitioners' perspective," *Capgemini Research Institute*, vol. 10, 2019.

[144] N. Iftikhar and F. E. Nordbjerg, "Implementing machine learning in small and medium-sized manufacturing enterprises," in *Towards Sustainable Customization: Bridging Smart Products and Manufacturing Systems*, A.-L. Andersen, R. Andersen, T. D. Brunoe, M. S. S. Larsen, K. Nielsen, A. Napoleone, and S. Kjeldgaard, Eds., Cham: Springer International Publishing, 2022, pp. 448–456, ISBN: 978-3-030-90700-6.

[145] J. Lee, H. Qiu, G. Yu, and J. Lin, "Bearing data set," *NASA Ames Prognostics Data Repository*, 2007. [Online]. Available: http://ti.arc.nasa.gov/project/prognostic-data-repository.

[146] J. Björkdahl, "Strategies for digitalization in manufacturing firms," *California Management Review*, vol. 62, no. 4, pp. 17–36, 2020. DOI: 10.1177/0008125620920349.

[147] E. Commission, C. Directorate-General for Communications Networks, Technology, E. Scaria, A. Berghmans, M. Pont, C. Arnaut, and S. Leconte, *Study on data sharing between companies in Europe : final report*. Publications Office, 2018. DOI: doi/10.2759/354943.

[148] Council of European Union, *European data protection directive 95/46/ec*, 2016. [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng.

[149] P. Quinn, "The anonymisation of research data — a pyric victory for privacy that should not be pushed too hard by the eu data protection framework?" *European Journal of Health Law*, vol. 24, no. 4, pp. 347–367, 2017. DOI: 10.1163/15718093-12341416. [Online]. Available: https://brill.com/view/journals/ejhl/24/4/article-p347_347.xml.

[150] J. M. M. Rumbold and B. K. Pierscionek, "A critique of the regulation of data science in healthcare research in the european union," *BMC Medical Ethics*, vol. 18, no. 1, p. 27, Apr. 2017, ISSN: 1472-6939. DOI: 10.1186/s12910-017-0184-y.

[151] M. Gentili, S. Hajian, and C. Castillo, "A case study of anonymization of medical surveys," in *Proceedings of the 2017 International Conference on Digital Health*, ser. DH '17, London, United Kingdom: Association for Computing Machinery, 2017, pp. 77–81, ISBN: 9781450352499. DOI: 10.1145/3079452.3079490.

[152] S. G. Tsiafoulis, V. C. Zorkadis, and E. Pimenidis, "Maximum entropy oriented anonymization algorithm for privacy preserving data mining," in *Global Security, Safety and Sustainability & e-Democracy*, C. K. Georgiadis, H. Jahankhani, E. Pimenidis, R. Bashroush, and A. Al-Nemrat, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 9–16, ISBN: 978-3-642-33448-1.

[153] C. Keerie, C. Tuck, G. Milne, S. Eldridge, N. Wright, and S. C. Lewis, "Data sharing in clinical trials – practical guidance on anonymising trial datasets," *Trials*, vol. 19, no. 1, p. 25, Jan. 2018, ISSN: 1745-6215. DOI: 10.1186/s13063-017-2382-9.

[154] R. S. Peres, X. Jia, J. Lee, K. Sun, A. W. Colombo, and J. Barata, "Industrial artificial intelligence in industry 4.0 - systematic review, challenges and outlook," *IEEE Access*, vol. 8, pp. 220 121–220 139, 2020. DOI: 10.1109/ACCESS.2020.3042874.

[155] V. Thouvenot, D. Nogues, and C. Gouttas, "Data-driven anonymization process applied to time series.," in *SIMBig*, 2017, pp. 80–90.

[156] B. Anand, "Reverse globalization by internationalization of sme's: Opportunities and challenges ahead," *Procedia - Social and Behavioral Sciences*, vol. 195, pp. 1003–1011, 2015, World Conference on Technology, Innovation and Entrepreneurship, ISSN: 1877-0428. DOI: 10.1016/j.sbspro.2015.06.359. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877042815038380.

[157] D. K. Mutalemwa, "Does globalisation impact sme development in africa?" *African Journal of Economic and Management Studies*, vol. 6, no. 2, P. Felix Moses Edoho, Ed., pp. 164–182, Jan. 2015, ISSN: 2040-0705. DOI: 10.1108/AJEMS-01-2015-0012.

[158] K. Das, "Smes in india: Issues and possibilities in times of globalisation," *Asian SMEs and Globalization, ERIA Research Project Report*, vol. 5, pp. 69–97, 2007. [Online]. Available: https://www.eria.org/SMEs%20in%20India%5C_Issues%20and%20Possibilities%20in%20Times%20of%20Globalisation.pdf.

[159] J. Heikkilä, S. Nenonen, J. Olhager, and J. Stentoft, "Manufacturing relocation abroad and back: Empirical evidence from the nordic countries," *World Review of Intermodal Transportation Research*, vol. 7, no. 3, pp. 221–240, 2018. DOI: 10.1504/WRITR.2018.093563. [Online]. Available: https://www.inderscienceonline.com/doi/abs/10.1504/WRITR.2018.093563.

[160] M. Johansson, J. Olhager, J. Heikkilä, and J. Stentoft, "Offshoring versus backshoring: Empirically derived bundles of relocation drivers, and their relationship with benefits," *Journal of Purchasing and Supply Management*, vol. 25, no. 3, p. 100 509, 2019, ISSN: 1478-4092. DOI: 10.1016/j.pursup.2018.07.003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1478409218300359.

[161] D. C. Blackley, "Latex-dipping processes," in *Polymer Latices: Science and Technology Volume 3: Applications of latices*. Dordrecht: Springer Netherlands, 1997, pp. 155–228, ISBN: 978-94-011-5848-0. DOI: 10.1007/978-94-011-5848-0_3.

Glossary

Autoencoder	A machine learning algorithm which tries to recreate its input. It can be used as both a compression tool and a recreation tool. In this thesis, it is mainly used to detect anomalies.
Artificial intelligence	An ambiguous term describing a collection of different methods and technologies which enables computers to perform objectives normally requiring human intelligence.
Cyber-physical system	A system where the physical aspect is highly integrated with software for, e.g. controlling and monitoring.
Data silo	When data is collected and stored at a machine without an infrastructure to send the data to other relevant parties in the company.
Deep learning	Part of machine learning where the mathematical model imitates how humans learn using neural networks.
Digitisation	Changing from physical information into a digital one.
Digitalisation	Changing the digitisation into a more inter-connected system to improve, e.g. a business.
Industry 4.0	The fourth industrial revolution consists of many new IT technologies and a digitalisation transition.
Industrial internet of things	The industrial term for the internet of things is commonly used in industrial settings such as a manufacturing site.
Industrial revolution	A rapid shift within the manufacturing sector based on new available technologies or methods.

Internet of things	Physical objects or devices which can connect to each other and share information. The devices can, e.g. be equipped with sensors or have computing power.
Fog computing	A computer which acts as a middleman between the cloud and the production to provide faster computation and reduce the data sent to the cloud. AI models can also be deployed here to reduce latency and increase security.
Machine learning	A subset of AI which allows a program to learn a specific outcome without being explicitly programmed to do so.
Process	A manufacturing process of changing a product either as the physical change or the associated equipment conducting the change.
Product	The product produced at the process.
Retrofit	The process of attaching new equipment such as sensors to existing machinery.
Smart factory	A manufacturing factory where different digitalisation methods are used to enhance the factory.
Temporal information	The underlying information present in time series data.

Acronyms

AAU	Aalborg University
AE	autoencoder
AI	artificial intelligence
AICD	authentic industrial cloud data
ANN	artificial neural network
BCG	Boston Consultancy Group
CNN	convolutional neural network
DAE	denoising autoencoder
DBMS	database management system
DSR	design science research
DSS	decision support systems
GDPR	general data protection regulation
GUI	graphical user interface
IaaS	infrastructure as a service
IIoT	industrial internet of things
IoT	internet of things
IP	intellectual properties
IT	information technology

LSTM	long short-term memory
MAE	mean absolute error
MSE	mean squared error
MVC	model-view-controller
NLP	natural language processing
NN	neural network
PaaS	platform as a service
PCA	principal component analysis
PHM	prognostics and health management
PLC	programmable logic controller
SaaS	software as a service
SOA	service-oriented architecture
SVM	support vector machines
UCN	University College of Northern Denmark
WSN	wireless sensor networks

Part III

Papers

Paper A

Artificial intelligence and internet of things in small and medium-sized enterprises: A survey

Emil Blixt Hansen and Simon Bøgh

The paper has been published in the
Journal of Manufacturing Systems Vol. 58, pp. 362–372, 2021.

Journal of Manufacturing Systems 58 (2021) 362–372

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

journal homepage: www.elsevier.com/locate/jmansys

Technical Paper

Artificial intelligence and internet of things in small and medium-sized enterprises: A survey

Emil Blixt Hansen [a,b,*], Simon Bøgh [a]

[a] Robotics & Automation Group, Department of Materials and Production, Aalborg University, Denmark
[b] Technology Education & Sustainable Growth, University College of Northern Denmark, Aalborg, Denmark

A R T I C L E I N F O

Keywords:
Artificial intelligence
SME
IoT
Machine learning
Industry 4.0
Predictive analytics

A B S T R A C T

Internet of things (IoT) and artificial intelligence (AI) are popular topics of Industry 4.0. Many publications regarding these topics have been published, but they are primarily focused on larger enterprises. However, small and medium-sized enterprises (SMEs) are considered the economic backbone of many countries, which is why it is increasingly important that these kinds of companies also have easy access to these technologies and can make them operational. This paper presents a comprehensive survey and investigation of how widespread AI and IoT are among manufacturing SMEs, and discusses the current limitations and opportunities towards enabling predictive analytics. Firstly, an overview of the enablers for AI and IoT is provided along with the four analytics capabilities. Hereafter a comprehensive literature review is conducted and its findings showcased. Finally, emerging topics of research and development, making AI and IoT accessible technologies to SMEs, and the associated future trends and challenges are summarised.

1. Introduction

The internet of things (IoT) and artificial intelligence (AI) are a few of the technologies which together form the fourth industrial revolution, often referred to as Industry 4.0. Both research and industry have explored Industry 4.0 for years, and different companies and research institutions have tried to categorise the technologies and methods of the fourth industrial revolution. A popular example is Boston Consulting Group (BCG), which has categorised nine technologies referred to as the nine pillars of Industry 4.0 [1]. These pillars include *the industrial internet of things* and *big data and analytics*, where the latter can include AI and machine learning methods. The term Industry 4.0 is primarily a European term whereas *Smart Manufacturing* is used in USA and *Smart Factory* in Asia [2,3]. The three terms' definitions have some minor differences; however, since they are a result of the same technical advancement in the industry, they are used interchangeably throughout this paper. Industry 4.0 includes advanced and modern technologies and methods where small and medium-sized enterprises (SMEs) lack the resources and knowledge to utilise and set up a dedicated strategy for the transformation [4].

BCG and Innovation Fund Denmark specify in a white paper concerning Industry 4.0 in Danish manufacturing SMEs, that IoT was one of the main drivers of Industry 4.0, enabling real-time communication and decision making between processes. Moreover, they also specify that big data and analytics will enable new insight into the production, and support the real-time decision making [5,6]. A recent study from 2020 by Moeuf et al. [7] found that SMEs do not need to exploit all nine pillars of Industry 4.0. Furthermore, they found IoT to be a key area to generate data which can be exploited by big data and analytics, i.e. AI methods. According to Chiang and Lee [8], IoT is an opportunity for SMEs to seek new collaborations and Moeuf et al. showed that 90% of the experts agree in IoT is key for SMEs' industrial performance. Moreover, over 55% agree that big data is also key to improve their performance, but none of the experts agree that SMEs have the knowledge and expertise to utilise AI. From the expert group, 75% agree that research teams should promote the implementations of Industry 4.0 in SMEs.

In this paper, our focus will therefore be on the two technologies of IoT and AI, and how they are used in manufacturing SMEs. However, before describing these two technologies further, it is necessary to understand what an SME is and its characteristics.

1.1. Small and medium-sized enterprises

The European Commission defines small and medium-sized enter-

* Corresponding author at: Fibigerstræde 16, 9220 Aalborg East, Denmark.
E-mail addresses: ebh@mp.aau.dk (E.B. Hansen), sb@mp.aau.dk (S. Bøgh).

https://doi.org/10.1016/j.jmsy.2020.08.009
Received 30 October 2019; Received in revised form 7 August 2020; Accepted 7 August 2020
Available online 13 August 2020

prises (SMEs) as having less than 250 employees and less than € 50 million in turnover or € 43 million in revenue [9]. A more detailed list is shown in Table 1 and even though the three categories (micro, small, medium) can be distinguished, throughout this paper we will describe them all under the collective category "SMEs".

The European Commission states that 99% of all those employed in the EU are in an SME and thus the European Commission says that they are the backbone of the European economy. Therefore, they want to promote entrepreneurship and improve the business environment of the SMEs [10]. Moreover, the USA's government also states the importance of their SMEs, which are responsible for two-thirds of all new jobs created [11].

Besides the economic aspect of SMEs, they have the advantage of being more fluid in their internal structure and can thus benefit more from methods such as open innovation [12]. Previous literature reviews have also shown that the technologies of Industry 4.0 are under-exploited by the industrial SMEs and are sometimes outright neglected [13]. Different studies have found and described the characteristics of SMEs. Quinton et al. [14] described specific SME characteristics which are an advantage in a digital transformation. The identified characteristics were entrepreneurship, innovation, learning orientation, and power centralisation (the leadership of one leader). Moeuf et al. [7] identified four managerial characteristics of SMEs: short-term strategy, importance of the SME manager, lack of expert support function, and short hierarchical line. Moreover, Laforet and Tann [15] found four key characteristics: culture, process, leadership, and company strategy, as the important factors. Regarding the manufacturing aspect, they discovered that there is a correlation between the internal culture of a company and process innovation and that the flexible and informal environment along with customer relationship are also vital factors. For this paper we assemble these characteristics in the following way:

- Culture and leadership
- Process innovation
- Company strategy
- Customer relationship
- Flexible and informal environment

In contrast, larger enterprises have a different set of characteristics. According to Nicholas et al. [16], large enterprises have the characteristics of having several layers of management and have a slow response time to changes. They have a rigid and formal work environment rather than a highly flexible environment compared to SMEs. Moreover, their personnel have less individual authority. They are also more reluctant to process innovation but excel in product innovation according to Laforet and Tann [15], which is contrary to SMEs.

A study from before Industry 4.0 also found the key areas for SMEs to stay competitive; they need to invest and improve their processes, systems, and technologies [17]. With Industry 4.0 these investments and continued improvements are still paramount for SMEs to stay competitive and especially cloud solutions have been found to be of high adoption partly because of its simplicity compared to other Industry 4.0 technologies [7] and the potential impact of utilising it [18]. Moreover, they also found that other technologies of Industry 4.0 (autonomous robots, cyber-physical systems, machine-to-machine communication, etc.) is not feasibly for SMEs based on the relative high cost. With regards to the described characteristics of SMEs and the importance of IoT and AI in SMEs, according to Moeuf et al. [7], we will in the rest of this section describe IoT and AI along with their possibilities before describing the literature review in Section 4.

Table 1
The European Commission definition of SMEs. The columns denoted with * only requires one of them to be fulfilled and m denotes millions [9].

Category	Staff	Turnover*	Total balance*
Medium	<250	≤€ 50 m	≤€ 43 m
Small	<50	≤€ 10 m	≤€ 10 m
Micro	<10	≤€ 2 m	≤€ 2 m

1.2. Internet of things

The term internet of things (IoT) refers to interconnected devices on the internet or local network. These devices encompass computers, refrigerators, and small sensors. The industry uses the term *industrial internet of things* (IIoT) for their version of IoT. However, the two terms are often used interchangeably in an industrial context, which is also the case for this paper.

IoT in the industry is commonly used with a *cloud* solution in mind, which harvests all of the data from the sensors, machines, MES, etc. [19, 20]. The recent advances in communications standards further enabling the capabilities of IoT systems include Wi-Fi 6 (also called IEEE 802.11ax), which aims to increase throughput, improve power efficiency, and the efficient use of spectral resources [21]. Along with Wi-Fi 6, the next generation of cellular network, 5G, is also expanding the possibilities of IoT with significant speeds and lower latency [22]. Moreover, the new internet protocol, IPv6, gives sufficient uniquely IP addresses for the foreseeable future [23]. Beside the mentioned consumer-based communications standards, dedicated standards for IoT have also emerged, such as NB-IoT, LoRa, and Sigfox. These standards are designed with IoT in mind, and, e.g. NB-IoT has a small package size but a wide range which could help small IoT devices in gaining access to the internet [24].

1.3. Artificial intelligence

Analytics methods such as machine learning can be applied to data from IoT devices. As a part of Industry 4.0, machine learning, or AI, is a growing trend in many companies, especially software companies such as Google and Facebook [25,26]. Deep learning has become a popular method with the increased computational power and larger datasets made available, and they have in, e.g. image classification outperformed previous methods [27,28]. Deep learning methods are continuously getting better and outperforming their predecessors [29,30], which indicates the high research activity within the field. In 1997 a study discussed that AI methods were an emerging technology to be used in the industry for maintenance [31]. The industry is also moving towards cloud computing and deep learning with, e.g. Microsoft Azure cloud platform which incorporates different Industry 4.0 technologies such as big data and analytics [32]. It has been shown that the industry is moving towards machine learning methods where, in particular, neural networks (deep learning) and support vector machines (SVM) have an increased interest [33]. Moreover, for deep learning convolutional neural network (CNN), restricted Boltzmann machine, autoencoder (AE), and recurrent neural network (RNN) have all attracted considerable interest and are used in topics of surface integration inspection, machinery fault diagnosis, and predictive analytics and defect prognosis [34].

In this paper, we investigate the use of AI and IoT within SMEs. The use of these two technologies in SMEs can result in multiple different sub-domain technologies, an example of such can be seen in Fig. 1. The explanation of these sub-domain examples are:

Manufacturing data: When IoT is used at manufacturing SMEs, the subsequent is IIoT and thus manufacturing data. These data can be used to give clarity to the production.

Models: In between AI and IoT, mathematical models of the topic of interest can be created in, e.g. a cloud solution.

Intelligent decision system: Using AI within SMEs can enable the company to use technologies such as decision support system (DSS) to combat tacit knowledge.

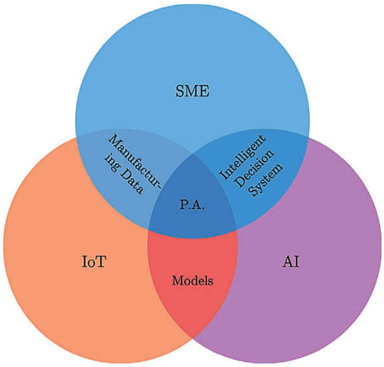


Fig. 1. An example of how the two technologies IoT and AI, and SME interconnect with each other and the possible technology interconnections. P.A.: Predictive analytics.

[Predictive analytic (P.A.): In the middle is predictive analytics, which with a combination of the three topics and their sub-domains will enable SMEs to perform predictive analytics on, e.g. maintenance.

The remainder of the paper is constructed as follows. Firstly, a more detailed overview of the enablers of AI and IoT is presented in Section 2 and in Section 3 a potential outcome of these technologies is presented with the four analytic capabilities. Then in Section 4 the methodology used for a systematic review is outlined, and its result is shown in Section 5. Finally, in Section 6 the topics of AI and IoT in manufacturing SMEs are discussed along with the potential future of AI and IoT in SMEs and in Section 7 conclusions are drawn on the current state along with future work for the topic.

2. Enablers of AI and IoT

To further understand the different areas of AI and IoT, this section will provide a better overview of the technologies and methods within the aforementioned technologies.

2.1. AI enablers

Artificial intelligence is a category that encompasses different methods. Throughout this paper, AI will be used as a synonym for machine learning and deep learning, which is a way of describing how a computer can learn patterns based on the inputs. AI also encompasses natural language processing and machine vision, to name a few [35].

2.1.1. Paradigms

In the machine learning and deep learning topic of AI, three paradigms exist [36]:

Supervised learning is the method to use when *labels* of a particular dataset are present. It is normally used to separate different classes from each other to classify a new object.
Unsupervised learning is the method to use if unknown patterns in a dataset are to be found. The data have no label, and this method can be used to, e.g. find outliers in a fault detection system.
Reinforcement learning is used when it is not straightforward to use supervised learning (no distinct classes are available). The algorithm learns the objective by the trial-and-error principle, where it gets rewards and punishments based on its decisions.

2.1.2. Methods

Depending on the required machine learning paradigm approach, different methods are available. Table 2 provides an overview of different methods for the different paradigms.

2.2. IoT enablers

The internet of things, or industrial internet of things, is an interconnected communication architecture. An IoT architecture has been constructed by Xu et al. [48] and consists of four layers: sensing, networking, service, and interface. These layers can be compared with the OSI model [49], where the bottom layer (sensing) is the hardware layer and the top layer (interface) is where the connection with, for instance, the cloud takes place.

2.2.1. Sensors

The sensing layer is where the sensors are placed. When choosing a production area of interest, it is necessary to carefully select the best sensors for the job. The wrong choice of sensors will yield bad results [50]. The choice of sensor could vary from accelerates, temperature, magnetic and sound sensors (microphone) depending on the use case.

2.2.2. Communication

The second layer, networking, is the layer that is responsible for the communication of the IoT devices. This includes protocols such as QoS (quality of service) to ensure a stable and minimal data loss connection between critical devices. Communication platforms could be the open standard OPC-UA which is designed for industrial communication.

2.2.3. Databases and the cloud

The service and interface layer contains the databases which store the data along with different GUI's and API interfaces for seeing and utilising the IoT network. Databases could be of various kinds from simple CSV files to more structured SQL or MS Access databases. Whereas the interface could be a cloud interface where the operator could monitor the process in real-time or the manager could receive reports of the production.

Table 2
Examples of different machine learning methods and their paradigms. Paradigms in italics indicate that it is commonly used for it but not exclusively. SVM: Support vector machine, FNN: fully connected neural network, CNN: convolutional neural network, RNN: recurrent neural network, ACN: actor-critic network, AE: autoencoder.

Name	Deep learning	Paradigm	Usage	Reference
SVM	No	Supervised	Based on user feature extraction, effective class separation	[37,38]
FNN	Yes	Unsupervised supervised reinforcement	Normal fully connected neural network used in all NN tasks	[39,40]
CNN	Yes	*Supervised*	Highly effective on image classification problems	[41,29]
RNN	Yes	*Supervised*	Good performance in time series problems such as language processing	[42,43]
ACN	Yes	Reinforcement	Two NN which make each other perform better	[44,45]
AE	Yes	Unsupervised	Good at compressing data, without losing too much performance	[46,47]

3. The four analytic capabilities for SMEs

Predictive analytics is the third step of *the four analytic capabilities* presented by Gartner Inc. [51]. They described steps from manual labour with no data collection to where employees determine what went wrong in the production up to *predictive* and *prescriptive analytics*. The latter two are where a production facility reaches a point where the machines have more control and information and will aid the employees in making the correct decision if they are needed at all. This is currently a utopia, especially for SMEs; however, there are already examples of predictive analytics [52,53], but we were unable to find any relevant examples of predictive analytics in SMEs besides [54]. For SMEs to reach the point of predictive analytics, they need to achieve all the steps of Gartner's four analytic capabilities:

1. Descriptive analytics: Most SMEs do not start at the *descriptive* level since they have no information from the system if something goes wrong. Instead, they rely on their expertise to understand and solve the problem. To get to the first step, SMEs could investigate the subdomain of IoT and SMEs from Fig. 1, namely *manufacturing data*. The data from the machinery will enable the operators and managers to investigate incidents and make a factually based decision. This step has been used in supply-chain management to manage the extensive network of suppliers and buyers [55]. Convolutional neural networks have also been used to detect defects in steel production [56]. Moreover, CNN has also been used to extract surface defects such as dirt and scratches [57].
2. Diagnostic analytics: For SMEs to take the next step, they need to look at intelligent systems such as AI (or machine learning). With *diagnostics* the system is now able to provide information on why it happened instead of only what happened. This step can be a subdomain from Fig. 1 of AI and IoT (*models*) or AI and SME (*intelligent decision system*). Depending on the use case, either of these will drive the SME to the diagnostic level. Subsequently, this reduces the time for the employee to diagnose, e.g. the machinery. The use of diagnostic analytic extends to applications such as Wang et al. [58] where they can perform diagnostics of featureless data such as vibration data. A cloud platform with the use of big data has also been used for fault diagnostics in cloud-based manufacturing systems and tested on a steel plate manufacturer use case [59]. A more dated example is one from Zhou et al. [60], where in the year 2000, they used different sensors and in a total of 72 features to make fault diagnoses in manufacturing systems.
3. Predictive analytics: To reach the step of *predictive* analytics in e.g. maintenance, both AI and IoT needs to be utilised at the SME. The system will then be able to predict if a motor is failing or if a production part will be faulty. However, this requires that the right sensors are available to obtain this information, and it also requires the AI-IoT model for prediction. With such a system, SMEs can reduce the down-time of machinery and increase part production. The semiconductor industry has, e.g. researched in the use of predictive analytics for fault detection using decision tree and Naïve Bayes methods [61]. A machine learning approach for solving the predictive maintenance is used in a use case from the semiconductor industry where the task was replacing tungsten filaments used in ion implantation [62]. A low-cost example of predictive maintenance in uninterruptible power supplies has also been conducted [63]. Sahal et al. [64] proposed a set of requirements to enable predictive maintenance with big data.
4. Prescriptive analytics: The ultimate goal of SMEs (and most enterprises) is the *prescriptive* step, where the system can make changes according to how e.g. the part should be produced. This setup will enable smart factories to run close to autonomously, and the system will be able to make changes such as adjusting valves and pressures to achieve the desired part output. This kind of setup is enabled in the same way as the *predictive* step, however; much better models and sensors are needed, and a system should be able to make adjustments at the production hardware level, which is unlikely for today's standard. Use cases for prescriptive analytics are lacking. Even though publications on how the use of prescriptive maintenance in a modern cyber-physical system have been performed [65]. A framework for prescriptive maintenance solutions using auditable methods to complex systems in different industries has been proposed [66]. Moreover, a data mining concept of recommendation for business process optimisation has also been proposed [67].

These steps encompass the technologies of IoT and AI models to reach the predictable analytics state. Therefore, drivers are both the different enablers of IoT and AI already described in Section 2 and the careful selection of the right data and construction of a sufficient model to perform the prediction. Technologies such as the cloud are not a prerequisite but could be beneficial to reach the step as they can give a unified platform where everything is implemented.

4. Literature survey

To fill the research gap of AI and IoT in manufacturing SMEs, a systematic review of the state-of-the-art of IoT and AI for industrial SMEs is conducted. The search queries were used in Web of Science and Scopus and were the following:

> "Small and Medium Sized Enterprise" and "Internet of Things"; "Small and Medium Sized Enterprise" and "The Cloud"; "Small and Medium Sized Enterprise" and "Machine Learning"; "Small and Medium Sized Enterprise" and "Deep Learning"; "Small and Medium Sized Enterprise" and "Neural Networks"; "Small and Medium Sized Enterprise" and "Artificial Intelligence"; "Small and Medium Sized Enterprise" and "Digital Twin".

These search queries were carefully selected to emphasise the SME aspect of IoT and AI, along with synonyms where, e.g. *deep learning* is used instead of *machine learning*. Moreover, broader terms such as *digital twin* could be used in the context of machine learning and IoT, without a mention of them explicitly and are therefore included. The broad selection of search queries was chosen, to catch publications which do not have IoT and AI as an explicit focus but still utilise them. The search results were then initially screened (title and abstract) to exclude misfits. The following list of publications was fully screened to validate the relevance of this paper. The relevance was validated against the criteria shown in Table 3. Publications relevant for the manufacturing and production industry is concerned; these include management of the production, but not financial risk assessment and other publications specific to SMEs in, e.g. the service industry.

5. Literature survey results

Because of this area's novelty, the broadly used term *artificial*

Table 3
Criteria of inclusion. DT: Digital twin.

No.	Criteria	Reason for inclusion
1	SME	The publication has to be relevant for SMEs.
2	IoT/AI/DT	Industry 4.0 is a collection of different technologies and thus different terminologies are often used; however, the publications should fall within at least one of these technologies.
3	2010–present	Industry 4.0 as a terminology was first used by the German government in 2011 [68] and since these technologies also first started to be relevant at that time, the publications should be post-2010.
4	Manufacturing	The publications should be relevant to the manufacturing industry.
5	English material	Because of the global aspect of Industry 4.0 and a way to avoid national biases, only English material is considered.

intelligence was also included as a search query. This resulted in many misfits where AI was used as a brief example. After the initial screening, 69 papers were identified as relevant out of 155 results, 7 of them were not accessible, and thus a full screening was conducted on the remaining 62. The full screening showed that 25 articles were not deemed relevant. In the end, 37 papers fulfilled the criteria and were accepted for this paper (Fig. 2 illustrates the review process).

The novelty was further underlined by the distribution of the publication year of the accepted publications. From 2010 to 2016 the number of accepted publications per year ranged from 0 to 3, while from 2017 to 2019 it ranged from 6 to 10. This also showed an increase in the topics related to IoT and machine learning in SMEs. The noted increase in the topics is further emphasised with an investigation of the Google Trend Index from Fig. 3, which shows an increase in search queries for machine learning and deep learning around 2016. The demographics of this literature review showed that Europe and Asia are the continents with the most activity with, 11 and 9 publications specific to the region, where, respectively, Germany and South Korea are the countries with the most publications. Africa, South America, and Australia had one publication each, and there were none from North America. Finally, there were 14 publications which were not country- nor region-specific.

5.1. Focus area findings

Through the literature survey, different categories of AI and IoT in SMEs were revealed. In Table 4 the different categories are shown, where it can be seen that, in particular, the cloud category has been exploited along with IoT. Moeuf et al. [13] also found this in their literature review of SMEs and Industry 4.0 from 2018.

5.1.1. IoT focused

Throughout this literature review, we discovered that IoT was broadly used in aspects ranging from being a synonym for Industry 4.0 to implementing IoT use cases. Park et al. [72] implemented a service-orientated platform through IIoT and machine learning in a textile manufacture to reduce energy consumption. In Chile, a questionnaire showed that the adoption level of IoT is still low [73]. On the contrary, Kaňovská and Tomášková [77] showed in a survey, that SMEs started initialisation of smart services in eight Czech Republic agricultural SMEs on their own. This conforms with a survey by Müller and Voigt [75] which concluded that SMEs should increase their focus on IoT. Multiple publications had a focus on both the sensor and the network part of IoT. Hyun-Jun et al. [85] measured machine utilisation, Chen et al. [50] showcased monitoring of machinery and Jung and Jin [76] showed a low-cost IoT solution and Mourtzis et al. [79] utilised OPC-UA and Zigbee to implement low-cost monitoring of machinery whereas [78] showcased a rule-based manufacturing integration assistant. Mohammed et al. [80] used the networking aspect to connect and extract information from the ERP system in a company. Ushada et al. [95] used IoT in the monitoring and controlling of room heat in an SME, and Uhlemann et al. [103] showed an example of IoT in a digital twin. Whereas many of the other IoT-related publications were related to the sub-domain for sensors and networking, only one concerned the topic of machine-to-machine communication. Nonetheless, the paper only concerned the topic of innovation and was not a real use case, nor was it technical [70].

5.1.2. AI focused

On the topic of AI, only a few publications mentioned the use of it explicitly. Only one did not use a variation of neural networks, namely Roitberg et al. [98]. They demonstrated a collaborative robot manipulator tracking human movement to act accordingly. Kim et al. [94] used a fully connected neural network to conduct a regression to determine cyber security threats. Chen et al. [50] used a neural network to detect the status of machinery, and Ushada et al. [95] also used a neural network to change the heat and light in a food SME. Chen et al. [54] showcased a quality of service platform for the better utilisation of SME clusters for CNN.

5.1.3. Cloud solutions

The cloud solutions attracted the most interest among the majority publications in this literature survey, and covers multiple sub-domains of the cloud. Huang et al. [71] proposed a cloud-based manufacturing service platform specifically for SMEs. Moreover, Soika et al. [100] used an ant colony optimiser for production planning, and it has also been found that manufacturing-focused SMEs should consider an ERP decision support system (DSS) [97]. Song et al. [102] also designed a DSS to be used in SMEs with the help of expert knowledge. Mohammed et al.

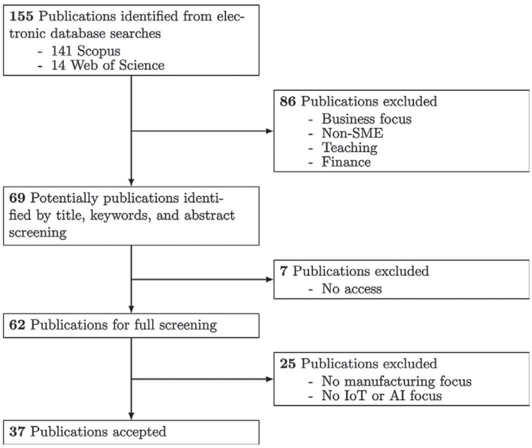


Fig. 2. The systematic literature review approach.

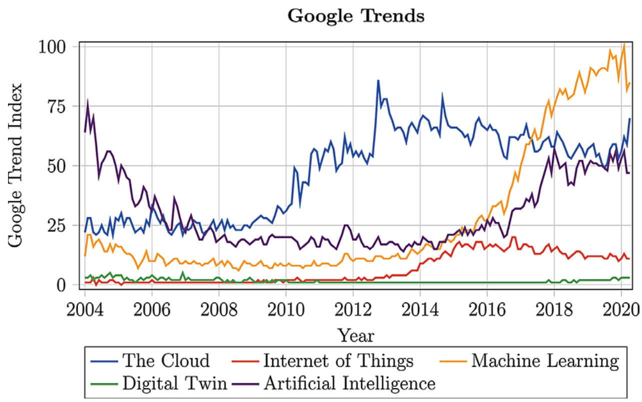


Fig. 3. The Google Trend Index for digital twin, AI, machine learning, the cloud, and IoT [69].

Table 4
Examples of different machine learning methods and their paradigms. The literature review results and which category each publication was focused upon. The abbreviations are: Sens: sensor, Net: network, M2M: machine-to-machine, No-NN: classical machine learning methods, Reg: regression, Comp: computing, SECTY: security, DT: digital twin, Innov: innovation.

	IoT			AI				Cloud						Business	
	Sens	Net	M2M	No-NN	Reg	FNN	CNN	Comp	DSS	ERP	MES	SECTY	DT	Model	Innov
[70]		X	X												X
[71]								X							
[72]	X	X						X					X		
[73]		X													X
[74]														X	
[75]														X	X
[76]	X	X													
[77]	X													X	X
[78]	X	X						X							
[79]	X	X						X							
[80]		X								X					
[81]														X	X
[82]											X				
[83]															X
[84]															X
[85]	X	X													
[86]								X						X	
[87]								X				X			
[88]								X				X			
[89]								X							
[90]								X			X	X			
[91]												X			
[92]								X							
[93]								X							
[94]					X	X		X				X			
[54]							X								
[50]	X					X									
[95]	X					X									
[96]								X	X						
[97]									X	X					
[98]				X											
[99]										X					
[100]										X					
[101]														X	X
[102]									X						
[103]	X											X	X		
[104]													X		

[80] also focused on ERP systems and how data can be extracted from them into an Excel database using multi-agents. An interest in a smart MES using the cloud was also found [82] and how SMEs could use business intelligence systems to enhance their resource utilisation [96]. The use of a service-oriented platform as a cloud solution was also described and utilised in the textile dyeing industry SME [72], and Ding et al. [93] describe a study of a better utilisation of SMEs' hardware as a cluster for training CNNs. A literature review by Schäfer et al. [99] found that it could be beneficial for SMEs to adopt cloud-based enterprise systems. However, they will not gain as much benefit as larger enterprises; and it is unknown how big the gap is. An IoT cloud solution was proposed by Mourtzis et al. [79] showed that with the use of communication standards such as Zigbee and OPC-UA sensors could communicate with a cloud solution in SMEs. A helpful integration assistant for IoT devices has also been proposed [78]. Kloch et al. [86] discussed the need for internet access and IT capabilities in order for SMEs to gain leverage over the competition. However, this brings the concern of cybersecurity which many publications found in is this review focused upon. Copie et al. [87] proposed a dedicated cloud governance and how to handle security threats. A self-protecting cloud has been described to combat the security threats for SMEs [88]. It has been found that SMEs are vulnerable to attacks [94], and a study showed that 91% of attacks on manufacturing industries are from outside the company [91]. Malik et al. [89] showed that in Africa and Asia, cloud solution has challenges, but it is still beneficial. Moreover, a survey based on SMEs in Ireland showed that the current cloud solutions are not suitable for SMEs and thus justifies the development of SME focused cloud solutions [92]. Wang et al. [90] gave an in-depth overview of these needed capabilities for SMEs.

5.1.4. Digital twin

A few of the accepted publications concerned the topic of digital twins. Two of them focused on a learning factory to teach SMEs the fundamentals of Industry 4.0 and to realise a digital twin in SMEs, where Uhlemann et al. [103] focus on the concept of realising the digital twin and their follow-up article [104] focused on the learning factory itself. Only one use case was found by Park et al. [72], where they used it in the textile industry.

5.1.5. Innovation and business model

The last found category of interest regards business models and innovation. Shin [70] found through an exploratory study that SMEs within IoT should embrace open innovation and the open market, and they should have a close linkage throughout the supply chain. Kaňovská and Tomášková [77] found that these innovation processes started from within the company. A literature review found that the topic of IoT in SMEs is still novel, and SMEs should adopt new business models to stay competitive [81]. Müller and Voigt [75] found that SMEs should focus more on adopting the technologies to enhance the sustainable focus and Reuter et al. [84] showed a focus area of data acquisition to enable a sustainable production. Müller and Däschle [74] found that business models are more tailored to the manufacturing industry rather than the service industry and Müller and Hopf [83] showcase a competence centre to demonstrate Industry 4.0 for SMEs. Kloch et al. [86] showed that general IT functionality and competences are required for the SMEs to survive. The competence centre could be backed up by a MADBE software program by Lurgi and Estanyol [101], which helps SMEs to collaborate based on their profile. Even though there is a trend showing that SMEs are adopting new business model and innovation practises a study from 2018 by Ramírez-Correa et al. [73] showed a low adoption of IoT in Chile, and it is dependent on personal innovation attitude.

5.2. SME characteristics findings

As mentioned, the manufacturing SMEs have a different set of characteristics compared to larger enterprises. Throughout this literature review, these characteristics were mapped to understand what were the drivers behind the usage of AI and IoT. Many of the accepted publication were survey-, review- and technology-specific and did not have a clear indication of why SMEs adopted (or should adopt) IoT or AI in their company. Therefore, these are excluded from the overview, shown in Table 5.

5.2.1. Process innovation

The first characteristic found that drives SMEs was process innovation. This underlying reason as to why they chose process optimisation could be a reduction of cost, clarity in the production or shorter changeover time. Park et al. [72] did it to reduce energy consumption, and examples of bringing down the cost of after-sales were also shown [76]. Wieland et al. [78] demonstrated a tool for the quicker implementation of new IoT devices. In contrast, Mourtzis et al. [79] tested the implementation of IoT monitoring devices at different processes (e.g. moulding machine as in [85]) at three different SMEs. Wang et al. [90] showcased a cloud solution for SMEs for multiple processes such as a cutting tool. Moreover, Chen et al. [50] developed a low-cost sensor to monitor, in this case, a spring manufacturing process.

5.2.2. Company strategy

The second characteristic found was company strategy. These fundamental changes in a company's strategy were seen in Shin [70] where a Korean SME changed their strategy to the overseas export of products. Kaňovská and Tomášková [77] found in an exploratory study that six of eight interviewed companies changed their company strategy by themselves and the other two were changed by the competition and customers. A change in the company strategy to provide a better work environment for the employees has also been found by Ushada et al. [95].

6. Discussion

The focus of this survey was the use of AI and IoT in SMEs in publications. The literature survey showed that it has a high novelty factor and because of the search queries the accepted publications also included some which did not have an explicit focus on the topic in question.

6.1. Survey findings

The distribution of the different focus areas found can be seen in Fig. 4. It can be seen that a majority of the publications focused on the cloud and IoT.

The cloud aspect of the publication was where most of the publications were related to cloud computing and as a service platform. Different publication discussed and showcased the extraction and utilisation of data from an ERP system, and the use of DSS should help SMEs in enhancing their cloud capabilities. However, the different literature reviews found in this survey indicates that the current cloud solutions are too focused on larger enterprises and are thus not suitable for SMEs in general. This was found both in reviews from 2013 in a global context [99], 2019 in an African and Asian context [89] and in Ireland from 2014 [92]. However, all of them agrees that the potential for SMEs is there and thus justifies the development of SMEs specific versions. An aspect of cloud solutions that cannot be ignored is cybersecurity. It was found that SMEs are most vulnerable to attacks and the majority of attacks are from outside the company [91,94]. These threats are not only

Table 5
The different characteristics and the found related publications.

	Publications
Process innovation	[50,72,76,78,79,85,90]
Company strategy	[70,77,95]

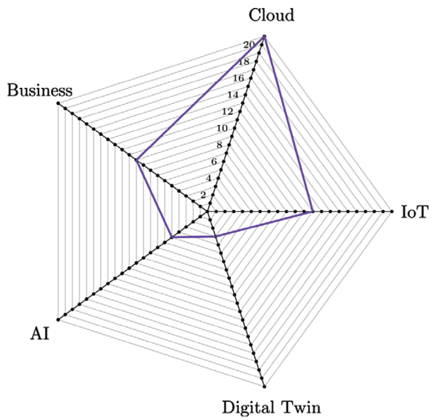


Fig. 4. The different sub-domains that were covered by the accepted publications.

to cloud solutions but also, according to Tuptuk and Hailes [105], their IT in general, and the general precautions of IT security should be made.

The AI topic was shown to have been used at both cloud solutions and more close to production. However, the lack of IT knowledge in SMEs can deter them from adopting machine learning projects. SMEs might not know where to start or even know that the technology exists and can be used to their benefits, as it was the case in [72,95]. These examples give an optimistic mindset and should encourage SMEs to dive further into AI and machine learning techniques. The machine learning methods have also become more accessible where commercial cloud solutions already have built-in easy-to-use machine learning tools such as Microsoft Azure [106]. Furthermore, if an SME still lacks digitisation, the focus should be there instead.

Our literature review showed that it does not require much from the SMEs to get started with small beneficial implementations of IoT. Nonetheless, the general adoption of IoT is still low, and it could be because of a lack of knowledge and expertise. The drivers of adoption of IoT also showed that they primarily come from inside the company and not outside factors (customers, suppliers); however, as the IoT technologies mature, this will shift to outside factors. SMEs should start to investigate the IoT possibilities since it has been shown that, with little effort, results can be achieved such as in Jung and Jin [76] and Hyun-Jun et al. [85].

To understand what the drivers of the companies involved in the accepted literature, a particular set of distinct characteristics was described. As seen in Fig. 5, we only found two of the five characteristics to be the driver of the SMEs, which were *process innovation* and *company strategy*. From the 37 accepted publications, 25 of them did not specify the drivers behind them, or they were surveys and did not determine which characteristics were the driver. The manufacturing SMEs' equipment and production is often acquired over time, and it is not uncommon to have old machinery for many years. It was seen that behind the *process innovation* characterised the SMEs primarily worked to improve these old aspects of the machinery instead of investing in new machinery. Process optimisation was sometimes performed to have more transparency in the production, for instance, machinery utilisation. Instances of the company strategy changing were also noticed. It was found that most undertook this shift in company strategy by themselves, while some were market-/customer-driven. However, in

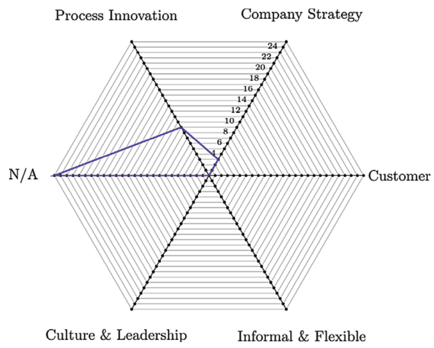


Fig. 5. The different characteristics of SMEs and number of publications concerning them. N/A: Not applicable.

cases of where the company strategy led the way there might have been other drivers that were not directly reported, such as customers or the culture. The drivers behind the company strategy should incorporate the topic of open innovation since it has shown to be beneficial to work with other SMEs who are also undertaking the Industry 4.0 transformation.

6.2. *The future of IoT and AI in SMEs*

The current development of machine learning techniques is taking place at a rapid rate, where new methods are outperforming their predecessors. Taking a look at machine vision in the industry, only one accepted publication used machine learning and vision together [98]. However, machine vision has reached a state where more than acceptable results can be achieved [28,107]. The future for SMEs in manufacturing could enforce these robust machine learning methods to ensure better quality control and monitoring of the production, or at least as an assistive tool. AI and IoT together would also enable the utilisation of predictive analytics of Gartner's four analytic steps. However, predictive analytics for a complete manufacturing line could be a tedious and costly affair, especially for SMEs. Still, smart IoT devices (such as Blackbird devices [108]) that could be attached to a single machine could enable predictive analytics machine-wise. This will make it less costly to adopt predictive analytics which would be beneficial for SMEs; moreover, it could also be used for dynamically value stream mapping [109]. Besides a local predictive analytics system, these data would be uploaded to a cloud service where the data will be stored and could be used for management reports and increase the accuracy of the prediction models. This future version of an SME IoT company is illustrated in Fig. 6. In the figure, IoT data from the production will stream to the cloud and the local operator monitor station. This local monitor station will show the operator and team leader the current state of the machinery. The manager can obtain a manager report from the cloud with information regarding the utilisation of the production machinery along with predictive maintenance reports. The technologies enabling machine learning and IoT is also described in [110] to be enablers of a digital twin, however, as they also describe, and as this survey showed, there are still problems to overcome.

7. Conclusion and future work

Through this literature survey, research gaps were discovered regarding Industry 4.0 and SMEs concerning the sub-domains of AI and IoT. Examples of IoT and its use were found in the review, where it was,

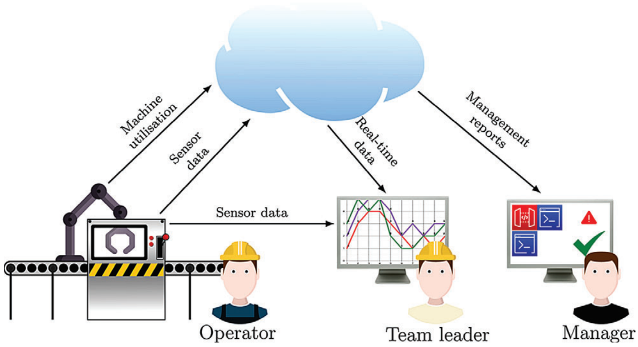


Fig. 6. IoT devices at the production sending data to a local monitor station as well as the cloud.

in general, machine-wise and involved the utilisation of one machine that was being recorded and used. Machine-wise implementation is cheaper compared to full production-wise implementation, and therefore SMEs should initially pursue this implementation. Nonetheless, many of the cases were machine utilisation recording. While it is an excellent way to gain insight into the production, it is not the full spectrum of IoT in Industry 4.0. SMEs should also focus on using IoT for other scenarios such as machine-wise predictive analytics. The absence of machine vision applications was also noted, and that good results can be achieved with a low-cost solution, as long as the knowledge of the area is present. The survey also showed that SMEs need to be at the forefront of this new industrial revolution if they want to stay competitive. Moreover, they should have an open mind regarding new business models and embrace knowledge sharing methods such as open innovation.

We also discovered that many of the successful use cases were the low costs and simple implementation of IoT and cloud solutions. This conforms with the findings of Moeuf et al. [7], stating the reason behind cloud solution being the most commonly utilised in SMEs is the simplicity of it compared to the rest of Industry 4.0. Future research should, therefore, focus on making other parts of Industry 4.0, such as IoT and AI, more simple to use and implement. This could be in the form of off-the-shelves product that is easy to implement at a given production aspect without the need for computer scientists to set up.

As stated, this survey showed the novelty of the subject of IoT and AI in SMEs, and therefore more research is needed to be conducted on the topic to unlock the potentials of Industry 4.0 to SMEs. Because of the novelty, search queries were relatively broad to catch publications that were not explicitly focused on the topic. This led to a high amount of misfits which were not caught in the initial screening; however, articles which would not have shown up otherwise were retrieved. Moreover, for future work it would be beneficial to perform more research in the characteristics and drivers of SMEs, since out of the 37 accepted publications, 25 of them did not disclose what the key drivers were. Therefore, future research should find out more precisely what characteristics drove the SMEs since our results, in that regard, are inadequate to conclude other than the fact that process innovation and company strategy are areas of focus. Although this correlates with the findings of many process optimisations papers in this survey. The reason behind the findings of few AI focused papers, is down to the complexity of AI in its current state, where SMEs lacks the knowledge and resources to utilise the technology, even though it deemed to be beneficial for them. Therefore, future research should focus on simplifying AI solutions for SMEs and thus make them more directly applicable to them.

Acknowledgements

The work was conducted as a part of the Danish strategy for Industry 4.0 in SMEs called Innovation Factory North.

Conflict of interest: The authors declare that there is no conflict of interest.

References

[1] Rüßmann M, Lorenz M, Gerbert P, Waldner M, Justus J, Engel P, et al. Industry 4.0: the future of productivity and growth in manufacturing industries. Boston Consult Group 2015;9(1):54–89.

[2] Thoben K-D, Wiesner S, Wuest T. "Industrie 4.0" and smart manufacturing – a review of research issues and application examples. Int J Autom Technol 2017;11(1):4–16. https://doi.org/10.20965/ijat.2017.p0004.

[3] Mittal S, Khan MA, Romero D, Wuest T. A critical review of smart manufacturing & Industry 4.0 maturity models: implications for small and medium-sized enterprises (SMEs). J Manuf Syst 2018;49:194–214. https://doi.org/10.1016/j.jmsy.2018.10.005.

[4] Schröder C. The challenges of Industry 4.0 for small and medium-sized enterprises. 2016.

[5] Colotla I, Fæste A, Heidemann A, Winther A, Andersen PH, Duvold T, et al. Winning the Industry 4.0 race, Innovation Fund Denmark. 2016. Available from: https://innovationsfonden.dk/sites/default/files/2018-07/bcg-winning-the-industry-40-race-dec-2016.pdf. [Accessed 22 April 2020].

[6] Lenz J, Wuest T, Westkämper E. Holistic approach to machine tool data analytics. J Manuf Syst 2018;48:180–91. https://doi.org/10.1016/j.jmsy.2018.03.003.

[7] Moeuf A, Lamouri S, Pellerin R, Tamayo-Giraldo S, Tobon-Valencia E, Eburdy R. Identification of critical success factors, risks and opportunities of Industry 4.0 in SMEs. Int J Prod Res 2020;58(5):1384–400. https://doi.org/10.1080/00207543.2019.1636323.

[8] Chiang Y, Lee D. Smart manufacturing with the internet of makers. J Chin Inst Eng 2017;40(7):585–92. https://doi.org/10.1080/02533839.2017.1362324.

[9] EC. What is an SME?. 2009. Available from: https://ec.europa.eu/growth/smes/business-friendly-environment/sme-definition. [Accessed 1 October 2019].

[10] EC. Statistics on small and medium-sized enterprises. 2019. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php/Statistics_on_small_and_medium-sized_enterprises#Context. [Accessed 1 October 2019].

[11] USTR. Small business. 2018. Available from: https://ustr.gov/issue-areas/small-business. [Accessed 1 October 2019].

[12] Spithoven A, Vanhaverbeke W, Roijakkers N. Open innovation practices in SMEs and large enterprises. Small Bus Econ 2013;41(3):537–62. https://doi.org/10.1007/s11187-012-9453-9.

[13] Moeuf A, Pellerin R, Lamouri S, Tamayo-Giraldo S, Barbaray R. The industrial management of SMEs in the era of Industry 4.0. Int J Prod Res 2018;56(3):1118–36. https://doi.org/10.1080/00207543.2017.1372647.

[14] Quinton S, Canhoto A, Molinillo S, Pera R, Budhathoki T. Conceptualising a digital orientation: antecedents of supporting sme performance in the digital economy. J Strateg Mark 2018;26(5):427–39. https://doi.org/10.1080/0965254X.2016.1258004.

[15] Laforet S, Tann J. Innovative characteristics of small manufacturing firms. J Small Bus Enterp Dev 2006. https://doi.org/10.1108/14626000610680253.

[16] Nicholas J, Ledwith A, Perks H. New product development best practice in SME and large organisations: theory vs practice. Eur J Innov Manage 2011;14(2): 227–51. https://doi.org/10.1108/14601061111124902.
[17] Kennedy J, Hyland P. A comparison of manufacturing technology adoption in SMEs and large companies. Proceedings of 16th annual conference of Small Enterprise Association of Australia and New Zealand 2003:1–10.
[18] Wu D, Greer MJ, Rosen DW, Schaefer D. Cloud manufacturing: strategic vision and state-of-the-art. J Manuf Syst 2013;32(4):564–79. https://doi.org/10.1016/j.jmsy.2013.04.008.
[19] Biswas AR, Giaffreda R. IoT and cloud convergence: opportunities and challenges. 2014 IEEE World Forum on Internet of Things (WF-IoT) 2014:375–6. https://doi.org/10.1109/WF-IoT.2014.6803194.
[20] Tao F, Qi Q, Liu A, Kusiak A. Data-driven smart manufacturing. J Manuf Syst 2018;48:157–69. https://doi.org/10.1016/j.jmsy.2018.01.006.
[21] Afaqui MS, Garcia-Villegas E, Lopez-Aguilera E. IEEE 802.11 ax: challenges and requirements for future high efficiency WiFi. IEEE Wirel Commun 2017;24(3): 130–7. https://doi.org/10.1109/MWC.2016.1600089WC.
[22] Li S, Xu LD, Zhao S. 5G internet of things: a survey. J Ind Inf Integr 2018;10:1–9. https://doi.org/10.1016/j.jii.2018.01.005.
[23] Madakam S, Ramaswamy R, Tripathi S. Internet of things (IoT): a literature review. J Comput Commun 2015;3(5):164.
[24] Ray B. NB-IoT vs. LoRa vs. Sigfox. 2018 06. Available from: https://www.link-labs.com/blog/nb-iot-vs-lora-vs-sigfox. [Accessed 28 October 2019].
[25] Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, et al. Tensorflow: a system for large-scale machine learning. 12th USENIX symposium on operating systems design and implementation (OSDI 16). Savannah, GA: USENIX Association; 2016. p. 265–83. Available from: https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi. [Accessed 1 October 2019].
[26] Paszke A, Gross S, Chintala S, Chanan G, Yang E, DeVito Z, et al. Automatic differentiation in pytorch. NIPS-W 2017.
[27] Schmidhuber J. Deep learning in neural networks: an overview. Neural Netw 2015;61:85–117. https://doi.org/10.1016/j.neunet.2014.09.003.
[28] Redmon J, Farhadi A. Yolov3: an incremental improvement. 2018. arXiv: 1804.02767.
[29] Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, et al. Human-level control through deep reinforcement learning. Nature 2015;518 (7540):529. https://doi.org/10.1038/nature14236.
[30] Hessel M, Modayil J, Van Hasselt H, Schaul T, Ostrovski G, Dabney W, et al. Rainbow: combining improvements in deep reinforcement learning. Thirty-second AAAI conference on artificial intelligence 2018.
[31] Luxhøj JT, Riis JO, Thorsteinsson U. Trends and perspectives in industrial maintenance management. J Manuf Syst 1997;16(6):437–53. https://doi.org/10.1016/S0278-6125(97)81701-3.
[32] Barga R, Fontama V, Tok WH, Cabrera-Cordon L. Predictive analytics with Microsoft Azure machine learning. Springer; 2015.
[33] Sharp M, Ak R, Hedberg T. A survey of the advancing use and development of machine learning in smart manufacturing. J Manuf Syst 2018;48:170–9. https://doi.org/10.1016/j.jmsy.2018.02.004.
[34] Wang J, Ma Y, Zhang L, Gao RX, Wu D. Deep learning for smart manufacturing: methods and applications. J Manuf Syst 2018;48:144–56. https://doi.org/10.1016/j.jmsy.2018.01.003.
[35] Rouse M. AI (artificial intelligence). 2018. Available from: https://searchenterpriseai.techtarget.com/definition/AI-Artificial-Intelligence. [Accessed 29 October 2019].
[36] Parikh D. Learning paradigms in machine learning. 2018. Available from: https://medium.com/datadriveninvestor/learning-paradigms-in-machine-learning-146ebf8b5943. [Accessed 29 October 2019].
[37] Joachims T. Text categorization with support vector machines: learning with many relevant features. In: Nédellec C, Rouveirol C, editors. Machine learning: ECML-98. Berlin, Heidelberg: Springer Berlin Heidelberg; 1998. p. 137–42.
[38] Hearst MA, Dumais ST, Osuna E, Platt J, Scholkopf B. Support vector machines. IEEE Intell Syst Appl 1998;13(4):18–28. https://doi.org/10.1109/5254.708428.
[39] Hsu K, Li H, Psaltis D. Holographic implementation of a fully connected neural network. Proc IEEE 1990;78(10):1637–45. https://doi.org/10.1109/5.58357.
[40] Lewenstein M, Nowak A. Fully connected neural networks with self-control of noise levels. Phys Rev Lett 1989;62:225–8. https://doi.org/10.1103/PhysRevLett.62.225.
[41] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ, editors. Advances in neural information processing systems 25. Curran Associates, Inc.; 2012. p. 1097–105. Available from: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf. [Accessed 29 October 2019].
[42] Mikolov T, Karafiát M, Burget L, Černocký J, Khudanpur S. Recurrent neural network based language model. Eleventh annual conference of the International Speech and Communication Association 2010.
[43] Zaremba W, Sutskever I, Vinyals O. Recurrent neural network regularization. 2014. arXiv:1409.2329.
[44] Vamvoudakis KG, Lewis FL. Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem. Automatica 2010;46(5):878–88. https://doi.org/10.1016/j.automatica.2010.02.018.
[45] Bahdanau D, Brakel P, Xu K, Goyal A, Lowe R, Pineau J, et al. An actor-critic algorithm for sequence prediction. 2016. arXiv:1607.07086.
[46] Ng A. Sparse autoencoder, CS294A lecture notes 72. 2011. p. 1–19.
[47] Deng L, Seltzer ML, Yu D, Acero A, Mohamed A-r, Hinton G. Binary coding of speech spectrograms using a deep auto-encoder. Eleventh annual conference of the International Speech and Communication Association 2010.
[48] Xu LD, He W, Li S. Internet of things in industries: a survey. IEEE Trans Ind Inform 2014;10(4):2233–43. https://doi.org/10.1109/TII.2014.2300753.
[49] Zimmermann H. OSI reference model – the ISO model of architecture for open systems interconnection. IEEE Trans Commun 1980;28(4):425–32. https://doi.org/10.1109/TCOM.1980.1094702.
[50] Chen Y-C, Ting K-C, Chen Y-M, Yang D-L, Chen H-M, Ying JJ-C. A low-cost add-on sensor and algorithm to help small- and medium-sized enterprises monitor machinery and schedule processes. Appl Sci 2019;9(8). https://doi.org/10.3390/app9081549.
[51] Hagerty J. 2017 planning guide for data and analytics. 2016. Available from: https://www.gartner.com/binaries/content/assets/events/keywords/catalyst/catus8/2017_planning_guide_for_data_analytics.pdf. [Accessed 24 October 2019].
[52] Schmidt B, Gandhi K, Wang L, Galar D. Context preparation for predictive analytics – a case from manufacturing industry. J Qual Maint Eng 2017;23(3): 341–54. https://doi.org/10.1108/JQME-10-2016-0050.
[53] Butte S, Patil S. Big data and predictive analytics methods for modeling and analysis of semiconductor manufacturing processes. 2016 IEEE workshop on microelectronics and electron devices (WMED) 2016:1–5. https://doi.org/10.1109/WMED.2016.7458273.
[54] Chen Z, Luo L, Yang H, Yu J, Wen M, Zhang C. GENIE: QoS-guided dynamic scheduling for CNN-based tasks on SME clusters. 2019 design, automation test in Europe conference exhibition (DATE) 2019:1599–602. https://doi.org/10.23919/DATE.2019.8715279.
[55] Cox A, Sanderson J, Watson G. Supply chains and power regimes: toward an analytic framework for managing extended networks of buyer and supplier relationships. J Supply Chain Manage 2001;37(1):28–35. https://doi.org/10.1111/j.1745-493X.2001.tb00097.x.
[56] Masci J, Meier U, Ciresan D, Schmidhuber J, Fricout G. Steel defect classification with max-pooling convolutional neural networks. The 2012 international joint conference on neural networks (IJCNN) 2012:1–6. https://doi.org/10.1109/IJCNN.2012.6252468.
[57] Park J-K, Kwon B-K, Park J-H, Kang D-J. Machine learning-based imaging system for surface defect inspection. Int J Precis Eng Manuf Green Technol 2016;3(3): 303–10. https://doi.org/10.1007/s40684-016-0039-x.
[58] Wang J, Zhuang J, Duan L, Cheng W. A multi-scale convolution neural network for featureless fault diagnosis. 2016 international symposium on flexible automation (ISFA) 2016:65–70. https://doi.org/10.1109/ISFA.2016.7790137.
[59] Kumar A, Shankar R, Choudhary A, Thakur LS. A big data MapReduce framework for fault diagnosis in cloud-based manufacturing. Int J Prod Res 2016;54(23): 7060–73. https://doi.org/10.1080/00207543.2016.1153166.
[60] Zhou Z, Chen Y, Fuh J, Nee A. Integrated condition monitoring and fault diagnosis for modern manufacturing systems. CIRP Ann 2000;49(1):387–90. https://doi.org/10.1016/S0007-8506(07)62971-0.
[61] Munirathinam S, Ramadoss B. Predictive models for equipment fault detection in the semiconductor manufacturing process. IACSIT Int J Eng Technol 2016;8(4): 273–85. https://doi.org/10.7763/IJET.2016.V8.898.
[62] Susto GA, Schirru A, Pampuri S, McLoone S, Beghi A. Machine learning for predictive maintenance: a multiple classifier approach. IEEE Trans Ind Inform 2015;11(3):812–20. https://doi.org/10.1109/TII.2014.2349359.
[63] Abdennadher K, Venet P, Rojat G, Rétif J, Rosset C. A real-time predictive-maintenance system of aluminum electrolytic capacitors used in uninterrupted power supplies. IEEE Trans Ind Appl 2010;46(4):1644–52. https://doi.org/10.1109/TIA.2010.2049972.
[64] Sahal R, Breslin JG, Ali MI. Big data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case. J Manuf Syst 2020;54:138–51. https://doi.org/10.1016/j.jmsy.2019.11.004.
[65] Ansari F, Glawar R, Nemeth T. PriMa: a prescriptive maintenance model for cyber-physical production systems. Int J Comput Integr Manuf 2019;32(4-5): 482–503. https://doi.org/10.1080/0951192X.2019.1571236.
[66] Marques H, Giacotto A. Prescriptive maintenance: building alternative plans for smart operations. FT2019. Proceedings of the 10th aerospace technology congress, October 8–9, 2019, Stockholm, Sweden, no. 162 2019:231–6. https://doi.org/10.3384/ecp19162027.
[67] Gröger C, Schwarz H, Mitschang B. Prescriptive analytics for recommendation-based business process optimization. In: Abramowicz W, Kokkinaki A, editors. Business information systems. Cham: Springer International Publishing; 2014. p. 25–37. https://doi.org/10.1007/978-3-319-06695-0_3.
[68] Kagermann H, Lukas W-D, Wahlster W. Industry 4.0: using the internet of things on the road to the 4th industrial revolution. VDI News 2011;13(1). German. Available from: https://www.dfki.de/fileadmin/user_upload/DFKI/Medien/News_Media/Presse/Presse-Highlights/vdinach2011a13-ind4.0-Internet-Dinge.pdf. [Accessed 8 October 2019].
[69] Google Trend. 2019. Available from: https://trends.google.com/. [Accessed 24 October 2019].
[70] Shin D-I. An exploratory study of innovation strategies of the internet of things SMEs in south korea. Asia Pac J Innov Entrep 2017;11(2):171–89.
[71] Huang B, Li C, Yin C, Zhao X. Cloud manufacturing service platform for small- and medium-sized enterprises. Int J Adv Manuf Technol 2013;65(9):1261–72. https://doi.org/10.1007/s00170-012-4255-4.
[72] Park KT, Im SJ, Kang Y-S, Noh SD, Kang YT, Yang SG. Service-oriented platform for smart operation of dyeing and finishing industry. Int J Comput Integr Manuf 2019;32(3):307–26. https://doi.org/10.1080/0951192X.2019.1572225.

[73] Ramírez-Correa PE, Grandón EE, Arenas-Gaitán J, Rondán-Catalu na FJ, Aravena A. Explaining performance expectancy of IoT in Chilean SMEs. In: Abramowicz W, Corchuelo R, editors. Business information systems. Cham: Springer International Publishing; 2019. p. 475–86.
[74] Müller JM, Däschle S. Business model innovation of Industry 4.0 solution providers towards customer process innovation. Processes 2018;6(12):260.
[75] Müller JM, Voigt K-I. Sustainable industrial value creation in SMEs: a comparison between Industry 4.0 and made in china 2025. Int J Precis Eng Manuf Green Technol 2018;5(5):659–70. https://doi.org/10.1007/s40684-018-0056-z.
[76] Jung J-u, Jin K-h. Case studies for the establishment of the optimized smart factory with small and medium-sized enterprises. In: Proceedings of the 2nd international symposium on computer science and intelligent control, ISCSIC'18; 2018. https://doi.org/10.1145/3284557.3284692. pp. 27:1–27:5.
[77] Kaňovská L, Tomášková E. Drivers for smart servitization in manufacturing companies. AGRIS On-line Pap Econ Inform 2018;10(3):57–68.
[78] Wieland M, Steimle F, Mitschang B, Lucke D, Einberger P, Schel D, et al. Rule-based integration of smart services using the manufacturing service bus. 2017 IEEE smartworld, ubiquitous intelligence computing, advanced trusted computed, scalable computing communications, cloud big data computing, internet of people and smart city innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI) 2017:1–8. https://doi.org/10.1109/UIC-ATC.2017.8397529.
[79] Mourtzis D, Milas N, Vlachou A. An internet of things-based monitoring system for shop-floor control. J Comput Inf Sci Eng 2018 03;18(2):021005. https://doi.org/10.1115/1.4039429.
[80] Mohammed WM, Ferrer BR, Martinez JL, Sanchis R, Andres B, Agostinho C. A multi-agent approach for processing industrial enterprise data. 2017 international conference on engineering, technology and innovation (ICE/ITMC) 2017:1209–15. https://doi.org/10.1109/ICE.2017.8280018.
[81] Varaniūtė V, Vitkauskaitė E, Tarutė A. Peculiarities of IoT-based business model transformations in SMEs. 2018.
[82] Menezes S, Creado S, Zhong RY. Smart manufacturing execution systems for small and medium-sized enterprises. Procedia CIRP 2018;72:1009–14. https://doi.org/10.1016/j.procir.2018.03.272.
[83] Müller E, Hopf H. Competence center for the digital transformation in small and medium-sized enterprises. Procedia Manuf 2017;11:1495–500. https://doi.org/10.1016/j.promfg.2017.07.281.
[84] Reuter C, Brambring F, Hempel T, Kopp P. Benefit oriented production data acquisition for the production planning and control. Procedia CIRP 2017;61:487–92. https://doi.org/10.1016/j.procir.2016.11.142.
[85] Hyun-Jun S, Sung-Jin K, Min-Ho J, Chang-Heon O. Development of automatic mold shot measurement and management system for smart factory. Int J Electr Comput Eng 2016;6(6):3142–7.
[86] Kloch C, Petersen EB, Madsen OB. Cloud based infrastructure, the new business possibilities and barriers. Wirel Pers Commun 2011;58(1):17–30. https://doi.org/10.1007/s11277-011-0286-7.
[87] Copie A, Fortiş T-F, Munteanu VI, et al. Data security perspectives in the framework of cloud governance. In: Caragiannis I, Alexander M, Badia RM, Cannataro M, Costan A, Danelutto M, editors. Euro-Par 2012: parallel processing workshops. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 24–33.
[88] Sulistio A, Reich C, et al. Towards a self-protecting cloud. In: Meersman R, Panetto H, Dillon T, Eder J, Bellahsene Z, Ritter N, editors. On the move to meaningful internet systems: OTM 2013 conferences. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 395–402.
[89] Malik BH, Asad J, Kousar S, Nawaz F, Zainab, Hayder F, et al. Cloud computing adoption in small and medium-sized enterprises (SMEs) of asia and africa a cross-continent overview of advantages and challenges. 2019.
[90] Wang XV, Wang L, Gördes R. Interoperability in cloud manufacturing: a case study on private cloud structure for SMEs. Int J Comput Integr Manuf 2018;31(7):653–63. https://doi.org/10.1080/0951192X.2017.1407962.
[91] Paul S. Reinforcing your SME against cyberthreats. Comput Fraud Secur 2017;2017(10):13–5. https://doi.org/10.1016/S1361-3723(17)30091-X.
[92] Carcary M, Doherty E, Conway G, McLaughlin S. Cloud computing adoption readiness and benefit realization in irish SMEs – an exploratory study. Inf Syst Manage 2014;31(4):313–27. https://doi.org/10.1080/10580530.2014.958028.
[93] Ding F, an Mey D, Wienke S, Zhang R, Li L. A study on today's cloud environments for hpc applications. In: Helfert M, Desprez F, Ferguson D, Leymann F, editors. Cloud computing and services science. Cham: Springer International Publishing; 2014. p. 114–27.
[94] Kim H-K, So W-H, Je S-M. A big data framework for network security of small and medium enterprises for future computing. J Supercomput 2019;75(6):3334–67. https://doi.org/10.1007/s11227-019-02815-8.
[95] Ushada M, Okayama T, Khuriyati N, Suyantohadi A. Affective temperature control in food SMEs using artificial neural network. Appl Artif Intell 2017;31(7–8):555–67. https://doi.org/10.1080/08839514.2017.1390327.
[96] Ayoubi E, Aljawarneh S. Challenges and opportunities of adopting business intelligence in SMEs: collaborative model. In: Proceedings of the first international conference on data science, e-learning and information systems, DATA'18; 2018. https://doi.org/10.1145/3279996.3280038. pp. 42:1–42:5.
[97] Potapov KS, Terentyeva PV, Drapikovskaya BT, Fomina IG, Shvetsova OA. About ERP and DSS system's introduction and using prospects at small and medium-sized enterprises in Russia. 2016 XIX IEEE international conference on soft computing and measurements (SCM) 2016:438–40. https://doi.org/10.1109/SCM.2016.7519806.
[98] Roitberg A, Perzylo A, Somani N, Giuliani M, Rickert M, Knoll A. Human activity recognition in the context of industrial human-robot interaction. Asia-Pacific signal and information processing association annual summit and conference (APSIPA), 2014 2014:1–10. https://doi.org/10.1109/APSIPA.2014.7041588.
[99] Schäfer G, Schulze M, Yusuf Y, Musa A. The benefits of saas-based enterprise systems for SMEs – a literature review. In: Hernández JE, Liu S, Delibašić B, Zaraté P, Dargam F, Ribeiro R, editors. Decision support systems II – recent developments applied to DSS network environments. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 34–45.
[100] Soika C, Teich T, Militzer J, Kretz D. Generation of process variants in automated production planning by using ant colony optimization. 2011 IEEE 3rd international conference on communication software and networks 2011:52–6. https://doi.org/10.1109/ICCSN.2011.6014847.
[101] Lurgi M, Estanyol F. MADBE: a multi-agent digital business ecosystem. 4th IEEE international conference on digital ecosystems and technologies 2010:262–7. https://doi.org/10.1109/DEST.2010.5610637.
[102] Song X, Tian H, Wu X. Study on SMEs-oriented strategic decision support system based on competitive intelligence. 2010 international conference on management and service science 2010:1–4. https://doi.org/10.1109/ICMSS.2010.5576728.
[103] Uhlemann TH-J, Lehmann C, Steinhilper R. The digital twin: realizing the cyber-physical production system for Industry 4.0. Procedia CIRP 2017;61:335–40. https://doi.org/10.1016/j.procir.2016.11.152.
[104] Uhlemann TH-J, Schock C, Lehmann C, Freiberger S, Steinhilper R. The digital twin: demonstrating the potential of real time data acquisition in production systems. Procedia Manuf 2017;9:113–20. https://doi.org/10.1016/j.promfg.2017.04.043.
[105] Tuptuk N, Hailes S. Security of smart manufacturing systems. J Manuf Syst 2018;47:93–106. https://doi.org/10.1016/j.jmsy.2018.04.007.
[106] Azure machine learning services. 2019. Available from: https://azure.microsoft.com/services/machine-learning-service/. [Accessed 24 October 2019].
[107] Raghavendra U, Fujita H, Bhandary SV, Gudigar A, Tan JH, Acharya UR. Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images. Inf Sci 2018;441:41–9. https://doi.org/10.1016/j.ins.2018.01.051.
[108] Blackbird. 2019. Available from: https://www.blackbird.online/. [Accessed 24 October 2019].
[109] Huang Z, Kim J, Sadri A, Dowey S, Dargusch MS. Industry 4.0: development of a multi-agent system for dynamic value stream mapping in SMEs. J Manuf Syst 2019;52:1–12. https://doi.org/10.1016/j.jmsy.2019.05.001.
[110] Qi Q, Tao F, Hu T, Anwer N, Liu A, Wei Y, et al. Enabling technologies and tools for digital twin. J Manuf Syst 2019. https://doi.org/10.1016/j.jmsy.2019.10.001.

Paper B

Concept of easy-to-use versatile artificial intelligence in industrial small & medium-sized enterprises

Emil Blixt Hansen, Nadeem Iftikhar, and Simon Bøgh

The paper has been published in the
Procedia Manufacturing Vol. 51, pp. 1146–1152, 2020.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 51 (2020) 1146–1152

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

30th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2021)
15-18 June 2021, Athens, Greece.

Concept of easy-to-use versatile artificial intelligence in industrial small & medium-sized enterprises

Emil Blixt Hansen[a,c,*], Nadeem Iftikhar[b], Simon Bøgh[c]

[a]Technology Education and Sustainable Growth, University College of Northern Denmark, Aalborg, Denmark
[b]Department of Computer Science, University College of Northern Denmark, Aalborg, Denmark
[c]Department of Materials and Production, Aalborg University, Denmark

Abstract

In this paper, the concept of what we call AI-Box is presented. This concept is targeting small and medium-sized enterprises within the manufacturing industry sector. The AI-Box aims to bring technologies from Industry 4.0 to them, with the use of easy-to-use and versatile implementation. Preliminary experiments have been conducted at Aalborg University and at an industrial partner to solve vision tasks, which would be too expensive with conventional vision techniques. Moreover, three different convolutional neural networks were tested to find the best-suited architecture. The three networks tested were the simple AlexNet, the complex ResNeXt, and small and complex SqueezeNet. Our results show that it is possible to solve the classification problem in a few epochs. Furthermore, with the use of augmented data, the performance can be improved. Our preliminary results also showed that the simpler convolutional neural network architecture from AlexNet yields a better result when classifying simple data.

Keywords: Artificial Intelligence; Machine Learning; Concept; Small and Medium Sized Enterprises; Versatile

1. Introduction

The topic of *artificial intelligence*, or its commonly used abbreviation, AI, is used in more and more sectors and is only expected to grow [1, 2]. More specifically, the use of the AI sub-category *machine learning* is showing tremendous potential with the different breakthroughs throughout the 2010s in, e.g. image data [3], and can now reach and outperform humans in games [4]. With quickly advances in AI and machine learning, the manufacturing industry is also looking towards these technologies. AI, along with other new technologies such as Big Data and 3D printing, is categorised under the name Industry 4.0 [5]. Large enterprises have already begun using these technologies within their products and production [6, 7]. Moreover, the use of machine learning has also been exploited in welding and robotic context [8, 9]. Most of the research and developed solution on the market is concerning big enterprises or research, and there is a lack of focus on the small- and medium-sized enterprises (SMEs). The lack of focus on SMEs is also evident in a study from 2018, which showed that the current maturity assessment for Industry 4.0 is not suitable for SMEs [10]. Gartner's steps of analytics can be used to reflect on how advanced analytic capabilities a company has [11]. It includes four steps: descriptive, diagnostic, predictive, and prescriptive, where SMEs generally is placed on the first step *descriptive*. The descriptive step specifies that the system can describe what the problem is but nothing else. The step of *predictive* analytics is a step that would benefit the industry; hence they would be able to predict equipment failure and maintenance. With Industry 4.0, different companies have created *smart devices* that enables manufactures to monitor their production, OEE, and live data. Examples of such devices are the Factbird by Blackbird [12] and M-Box by Monitor-Box [13]. Within research, the Danish Institute of Technology has created what they call Vision Box, which brings 2D and 3D quality inspections to manufactures with the use of deep learning [14]. Research exists in the context of visual inspection of production to find faults [15]

* Corresponding author. Tel.: +45 42494707
E-mail address: ebh@mp.aau.dk (Emil Blixt Hansen).

10.1016/j.promfg.2020.10.161

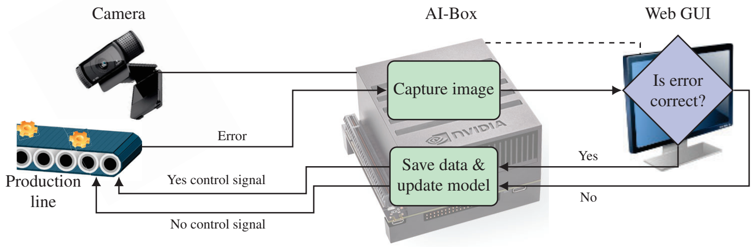


Fig. 1: A use case example of the AI-Box. The line between the camera and the AI-Box represent a physical USB connection between the two. The dashed line between the monitor and the AI-Box illustrates a wireless connection between the GUI and AI-Box.

and distribute the computation in between other edge nodes, also known as fog computing [16].

In this paper, we present the concept of what is called AI-Box. The AI-Box enables manufacturing SMEs to utilise the Industry 4.0 technologies of machine learning and IoT without them needing software engineers for setup and maintenance. Firstly in Section 2 three examples of use cases are described. In Section 3, the concept of the AI-Box is described and in Section 4 preliminary experiments are described. Finally, in Section 5, the conclusion and future work for the project are described.

2. Use case description for SMEs

Before introducing the underlying architecture, it is essential to have an understanding of what kinds of problems the AI-Box aims to solve. This section contains three hypothetical cases describing different use cases of the AI-Box at industrial SMEs.

Use case 1: False alarm

A company is producing plastic gears, and after they come out of the mould, sometimes there is leftover material on the cut teeth. This leftover material is not crucial to be removed at this stage. Still, the production triggers a false alarm, hence the production is temporarily stopped until an operator acknowledges the false error. Here the AI-Box will be set up with a camera pointing at the place from where the alarm was triggered and is connected to the relevant PLC out- and inputs. Each time the alarm occurs a picture is taken and is labelled depending on the input from the operator. When enough data is collected, the AI-Box starts to train on the data, and when it has trained enough, it can take over control and operate the PLC inputs by itself. This type of operation is illustrated in Fig. 1.

Use case 2: An audible error

At a production line, an experienced operator can hear when a tool should be changed but is unable to see it. The AI-Box is connected to a microphone that would monitor the utilisation. The operator will record the tool sound under normal operation and under a situation where the tool is in a condition where it should be changed. The AI-Box will learn this signature and will create an alert when the tool is about to be worn down. This will help inexperienced operators with performing tool change and helper monitor the machine.

Use case 3: Unknown error occurs

At a production line sometimes an operation fails, but no alarms are created. However, it is believed that it can be measured through vibration, and thus an accelerator sensor is attached to the place where the operation fails. The AI-Box is then set into an *outlier* detection mode, and here it will then create an alarm if the vibration is at an abnormal level.

These three use cases are examples of the different aspects of the AI-Box, but not limited to. An input could also be temperature, magnetic or video feeds. The AI-Box can, therefore, be summed up to the following key features:

- Simple to deploy and use
- Various built-in machine learning models
- Read and write PLC signals
- Handle different types of measurement data

Together, these features will enable SMEs to find and solve problems in there production, which they are unable to solve with traditional means.

3. Concept framework & architecture

The AI-Box concept consists of a complex architecture involving hardware and software. To give a better overview, the section is split up in hardware, software, and the internal system architecture.

3.1. Hardware

The main hardware for the AI-Box is the computer that trains and deploys the network for any given task. It should be balanced between cost, performance, and deployability, where the latter means it should be small, lightweight and not depending on a vast amount of peripherals. An NVIDIA Jetson AGX Xavier was chosen as the main hardware component [17] since it is small, not too expensive and it has a dedicated GPU with CUDA cores and thus enabling GPU hardware to accelerate deep learning performance. Even though the recommendation is that the Xavier is meant for deployment only, it can still be used to train on. Webcam and microphone are connected directly through a USB port while sensors such as accelerometer and temperature sensor, is connected through the GPIO pins on the Xavier.

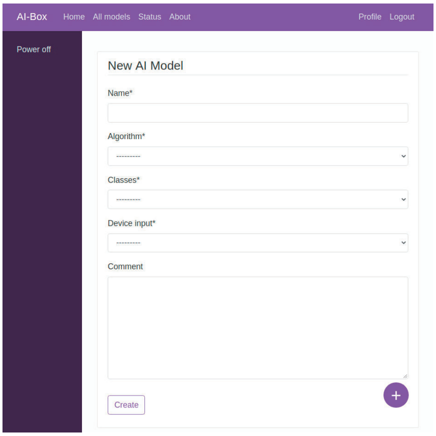


Fig. 2: The setup screen of a new model, where the algorithm is chosen, the number of classes there are, and the input device are chosen too.

3.2. Software

Rapid development was deemed a paramount aspect of creating the AI-Box. Therefore, the primary programming language used was Python 3.6 with Tensorflow 2.0 as the deep learning framework. Moreover, as one of the requirement was the reduced use of peripherals, a web-based interface was chosen as the best solution, here Django 3.0 was chosen. Through the web-interface of the AI-Box, the operator will make the initial setup for the problem at hand, which include specifying if, e.g. it is a camera that should be used (see Fig. 2). This setup screen enables the operator to set up the right deep learning architecture without he/she knowing it. Moreover, the classes that can be selected is a yes/no scenario or an outlier instance. When the model has been created, a running view is displayed

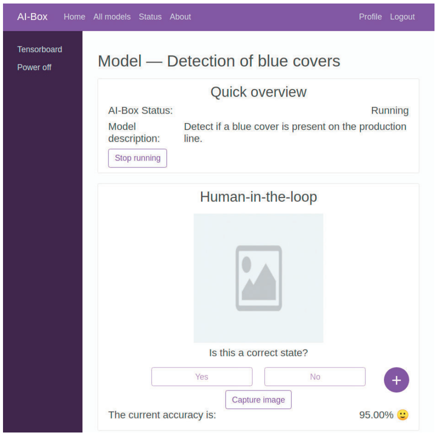


Fig. 3: The running screen shows the status and current accuracy. Moreover, when a new sample is captured, it will ask the operator to label it.

(see Fig. 3) to the operator. Here, the operator can see relevant information, such as current accuracy. Also, when a new sample is taken, and the AI-Box has not trained enough, or the model is not confident enough of the class, the operator is asked. The operator's decision then labels the sample and is saved in the sample database. The database storing all of the data has chosen to be a Zarr database [18]. Moreover, the web-interface also contains a view of all created models along with details view, such that the operator can view and change parameters of the model.

3.3. Architecture

The internal architecture of the model is firstly based on the web interface architecture *model-view-controller* (MVC) [19]. The MVC architecture is native to Django and is thus automatically implemented. The complexity of the AI-Box requires an additional architecture design to handle the state of the device, the deep learning models, and its associated data. This additional architecture is loosely based on the *layered pattern* architecture, where each layer has a different level of abstraction and serves the layers above and below it. In Fig. 4 the system can be seen. The top layer is the *GUI* where the operator interacts with the AI-Box, depending on what is displayed and interacted with, the information is pulled from the SQLite3 database from Django. When the operator changes the state of the device (e.g. starting a model) the *global state handler* ensures to initialise the correct parameters depending on the model information stored in the SQLite3 database. The *model handler* is a singleton python class that handles initialisation of the *loop controller* and the *sensor handler*, moreover, it also handles their intercommunication. The loop controller is where the deep learning is handled. The dataset is read from the Zarr database and is handling the learning of the model and when to stop to avoid overfitting. It also initialises the actual deep learning model by calling the respected *ML model* class. The *sensor handler* controls the setup, retrieves the data from the sensor. It also starts the *sensor* driver for the corresponding sensor. This driver performs the low-level data collection and preprocessing before it is sent upstream to the sensor handler. With this architecture, it is a simple procedure to add a new deep learning model without breaking the system. Moreover, potentially new sensor inputs can also be added to the system.

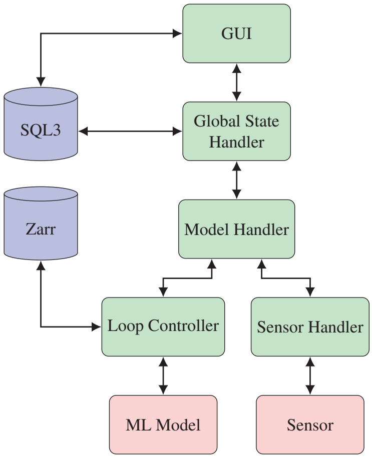


Fig. 4: The system architecture of the AI-Box. The *green* squares indicate that they are static models, the *blue* is databases, and the *red* is non-static model depended layers.

4. Experiment

To validate that it was possible to train and utilise the AI-Box and it would be feasible in an SME, two experiments were conducted. Firstly it was tested at an industrial partner, and secondly, locally at Aalborg University.

4.1. Experiment at industrial partner

The experiment was conducted at an industrial partner specialising in palletizing solutions. A known problem at their palletizer machine is that slip-sheets gets stuck under the picked up layer and no alarm is activated. The AI-Box was placed beside the palletizer, and a webcam was mounted in the corner of the palletizer pointing towards the surface area beneath the picked up layer. The palletizer picked up a layer of cardboard boxes with no slip-sheet below it. Different types and looks of slip-sheets were then placed under the picked-up layer, and images were taken of them. Moreover, images were also captured with no slip-sheets present. The deep learning model used was a simple convolutional neural network (CNN) based on AlexNet [3]. In Fig. 5, the resulting train and validation accuracy is shown. It can be seen that the model converge quickly, with the first 100% test accuracy at epoch 36, which took 37 seconds. After 2 minutes, the test accuracy started to converge. The training first starts after 30 unique samples have been acquired, moreover, all training data was randomly augmented on the fly to increase the variations in the relatively small dataset. There is, in total, nine possible types of augmentation to be performed on each sample, and they are brightness, contrast, flipping, hue, saturation, quality, rotation, blurring, and cropping. On each sample, there is a probability of 0.6 that a random augmentation is applied which is evaluated after each performed augmentation. In Fig. 6, the probability for the number of augmentation per sample is shown. Because of the small dataset, the validation accuracy is the same data as the training just without augmentation applied. Throughout this test, a total of 87 samples was captured.

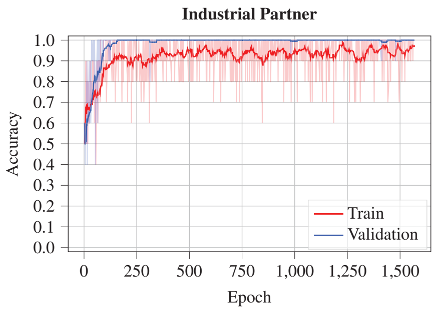


Fig. 5: The train and validation results of the slip-sheet detection at the industrial partner. The train and validation lines are averaged over 20 runs.

4.2. Experiment at FESTO CP Factory

Besides the test at the industrial partner, a test at the FESTO CP Factory line located at Aalborg University was also conducted. The FESTO CP Factory line serves as a learning factory of Industry 4.0 for students and researches at Aalborg University, and it produces smartphones mock-ups. The test had two purposes:

1. Test the applicability of the AI-Box
2. Test the implemented model on a different environment

The objective was to classify whether a blue case or a black case was present at the conveyor. In Fig. 7, the two classes, is shown. The AI-Box was connected to a power outlet and connected to

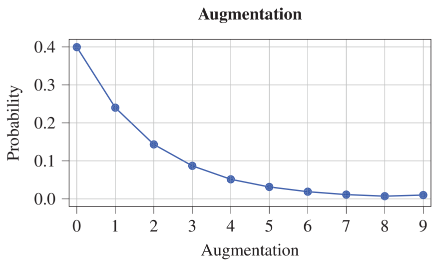


Fig. 6: The probability for the number of augmentations applied to each sample. The values are calculated from historical data from the slip-sheet test.

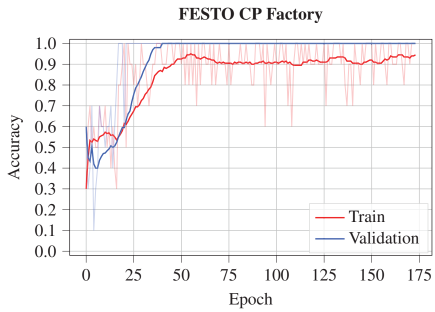


Fig. 8: The train and validation results of the cover detection at FESTO CP Factory. The train and validation lines are averaged over 20 runs.

the LAN of the FESTO CP Factory. Then a webcam was placed at the line and connected to the AI-Box. The AI-Box was turned on, and the URL of the AI-Box was entered in a browser on a laptop. Here the interface of the AI-Box was shown. A new model in the interface was created, and the AI-Box was started and waiting for new image data. In total, 41 samples were col-

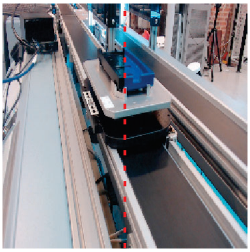


Fig. 7: The view from the AI-Box webcam of the classification object, with a size of 256x256 pixels. The left side is a black case, and the right is a blue case.

lected at the line and, as with the slip-sheet experiment, the training first started after 30 collected samples. The same deep learning model architecture was used, and the result can be seen in Fig. 8. In this experiment, the same premises applied as with the slip-sheet detection, because of the shallow sample size the training data had randomly performed augmentation to it, and the validation data was the same dataset just with no augmentation applied. The results showed that after just 17 epochs, the validation started to converge, which only took 35 seconds after the training began.

4.3. Comparison of CNNs

There exist many high performing CNNs [20]. To select the right one for the AI-Box can be a challenge. Our requirements for the model was high general performance and low complexity. Bianco et al. [20] compared the top of the line CNNs on both a desktop GPU and an Nvidia Jetson TX1. Their research showed that "*there is not a linear relationship between model complexity and accuracy*". For the AI-Box three architectures was chosen to be tested, AlexNet [3], ResNeXt [21], and SqueezeNet [22]. They all three represent diferent apspects of the CNN research:

AlexNet: Is a now classical CNN architecture with convolutional layers, max-pooling for feature extraction and fully connected layers for classification. Compared to more modern architectures, AlexNet has a simple layout though with many trainable parameters.

ResNeXt: Is a more modern CNN built on the idea of residual blocks from He et al. [23]. It contains multi-stacked residual blocks which counteract vanishing gradients. Compared to AlexNet, its architecture is more complicated, and it also has more trainable parameters.

SqueezeNet: Is the last CNN considered in this paper. SqeuzeNet is meant to be a lightweight model where its fire modules which reduce and expand the image helps with keeping the number of parameters down. It has no fully connected layers in the end for classification but relies on the last convolutional layer to do the classification and extract them with a last global average-pooling layer.

These three architectures have all been altered slightly, mainly to reduce memory footprint, and thus all are a custom implementation in Tensorflow 2.0 and Python 3.6. The changes for AlexNet is that there is no *local response normalization* and reduced convolutional and fully connected layers. ResNeXt changes are mainly the filters for the residual block is halved and some removed. The SqueezeNet is already a rather small model (memory-wise), and thus the only alterations are the changes of the last convolution to match the output classes.

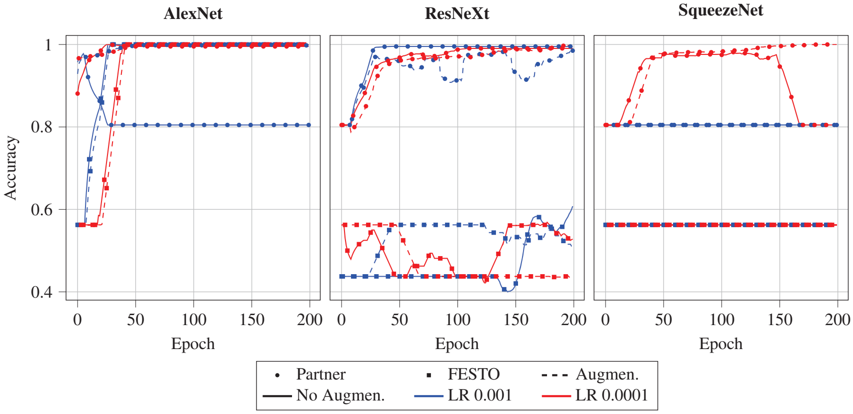


Fig. 9: The validation plot for the different models. The lines are averaged over 20 samples.

In Tab. 1 the architecture of the implemented models are presented. The models were verified with the dataset of the covers from the FESTO CP Factory with a total of 41 samples and 53/47 class distribution. Eight samples were taken from the 41 samples as validation data. Secondly, the models were also tested on a larger dataset from the industrial partner with a total of 1050 images and a class split of 80-20, where 210 samples were selected as validation data. It should be noted that the validation data was only used in inference mode and was not used to change hyper-parameters. The comparison test was conducted to test the three models and to test if data augmentation increases the performance, and lastly how the learning rate affected the model. The test was only allowed to run for 200 epochs to validate the needed quick learning. In Fig. 9 the result is shown. From the results, we made several observations:

Table 1: The different specifications for the implemented architectures. For AlexNet a dropout is placed after all convolutional operations. In ResNeXt nobias is used in the convolutional operations. All models use Adam optimiser with default hyperparameters.

AlexNet	ResNeXt	SqueezeNet
2 x Conv(32,3,1)	ZeroPad(3)	Conv(96,7,2)
MaxPool(2)	Conv(64,7,2)	MaxPool(3)
Conv(64,3,1)	BatNorm(1.001e-5)	2 x F.Module(16,64,64)
MaxPool(2)	ZeroPad(1)	F.Module(64,128,128)
Conv(128,3,1)	MaxPool(3)	MaxPool(3)
MaxPool(2)	3 x R.Block(64,1,32)	F.Module(32,128,128)
Dense(128)	3 x R.Block(128,2,32)	2 x F.Module(48,192,192)
Dense(68)	3 x R.Block(256,2,32)	F.Module(64,256,256)
Dense(2)	2 x R.Block(512,2,32)	MaxPool(3)
	GlobalAvgPool	F.Module(64,256,256)
	Dense(2)	Conv(2,1,1)
		GlobalAvgPool
Parameters	**Parameters**	**Parameters**
16,888,226	22,576,706	736,450

- AlexNet is the overall best performer, with the only time it did not learn was slip-sheet detection with a high learning rate and no data augmentation.
- ResNeXt is able to learn the larger dataset in a few epochs, whereas it did not learn the smaller dataset within the 200 epochs.
- SqueezeNet was only able to learn the larger dataset with a low learning rate and with the use of augmented data.
- A lower learning rate, as expected, halts the learning, but it is also observed that the higher learning rate does not find a sufficient minimum in time.
- Performing random augmentation to the data also decreased the learning rate slightly but also enabled some of the models to learn the dataset because of the added variation to the data.
- AlexNet is located in between the two others regarding the number of trainable parameters, it is the less complex of the three, with only convolutional, max pooling and fully connected layers. Our results show that a less complicated CNN architecture performs better on a less complicated problem (few classes).

5. Conclusion and future work

In this paper, our focus was to describe the current work being done on what is named the AI-Box. The AI-Box serves as a versatile easy to use deep learning device which can aid industrial SMEs to enhance part of their production that will otherwise be too costly/difficult. In its current form, only images are possible to be classified, and experiments were conducted to gain a better understanding of what type of deep learning model is required. The conclusion was that a less complicated CNN showed to be beneficial for small dataset along with an additional augmentation of the sample data. Moreover, the two experiments conducted showed that the AI-Box is versatile and functional in an expected manner regarding setup and performance.

For future work, regarding the image models, it could be of interest to investigate the use of methods such as Grad-CAM [24] to visualise what in the image is learned. This will improve the easy to use for operators and ensure that the model learns the correct features of the image. As of now, all the models are trained from scratch, so for better performance and faster learning transfer learning should be implemented. Besides the image models, deep learning models that should classify sound, vibration, and temperature should also be implemented and tested where recurrent neural networks should be tested. Moreover, outlier detection for all the different sensor inputs should also be exploited for use cases which do not have specific classes. A crucial task to solve is the preprocessing of data. As of now, the image data is reasonably easy to preprocess:

1. Reduce the size and convert to a three-dimensional matrix.
2. Add to the Zarr database along with the label.
3. Normalise the image between 0-1 before use.

For time-series data, this is a more complicated matter. Here different techniques as slicing the time-series data into short frames for classification can be used. Moreover, the various features such as the spectrogram, MFCC spectrogram and the time domain waveform should be investigated. These aspects need to be addressed in future work, and the core idea of the AI-Box is still the easy-to-use and versatility. Therefore, these feature selection should be made without the knowledge of the user through methods such as principal component analysis. Furthermore, more tests should be performed with different operators to verify the versatility and easy to use of the AI-Box with different data and classification objectives.

Acknowledgements

The work was conducted as a part of the Danish strategy for Industry 4.0 in SMEs called Innovation Factory North.

References

[1] E. Plastino, M. Purdy, Game changing value from artificial intelligence: eight strategies, Strategy & Leadership (2018).
[2] B. W. Wirtz, J. C. Weyerer, C. Geyer, Artificial intelligence and the public sector—applications and challenges, International Journal of Public Administration 42 (7) (2019) 596–615. doi:10.1080/01900692.2018.1498103.
[3] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems 25, Curran Associates, Inc., 2012, pp. 1097–1105. URL http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf
[4] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, et al., Human-level control through deep reinforcement learning, Nature 518 (7540) (2015) 529–533.
[5] M. Rüßmann, M. Lorenz, P. Gerbert, M. Waldner, J. Justus, P. Engel, M. Harnisch, Industry 4.0: The future of productivity and growth in manufacturing industries, Boston Consulting Group 9 (1) (2015) 54–89.
[6] J. Gao, Machine learning applications for data center optimization (2014).
[7] N. Lazic, C. Boutilier, T. Lu, E. Wong, B. Roy, M. Ryu, G. Imwalle, Data center cooling using model-predictive control, in: S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, R. Garnett (Eds.), Advances in Neural Information Processing Systems 31, Curran Associates, Inc., 2018, pp. 3814–3823. URL http://papers.nips.cc/paper/7638-data-center-cooling-using-model-predictive-control.pdf
[8] A. Sumesh, K. Rameshkumar, K. Mohandas, R. S. Babu, Use of machine learning algorithms for weld quality monitoring using acoustic signature, Procedia Computer Science 50 (2015) 316–322.
[9] E. B. Hansen, R. E. Andersen, S. Madsen, S. Bøgh, Transferring human manipulation knowledge to robots with inverse reinforcement learning, in: 2020 IEEE/SICE International Symposium on System Integration (SII), IEEE, 2020, pp. 933–937. doi:10.1109/SII46433.2020.9025873.
[10] S. Mittal, M. A. Khan, D. Romero, T. Wuest, A critical review of smart manufacturing & industry 4.0 maturity models: Implications for small and medium-sized enterprises (smes), Journal of Manufacturing Systems 49 (2018) 194 – 214. doi:doi.org/10.1016/j.jmsy.2018.10.005.
[11] C. Sapp, D. Brabham, J. Antelmi, H. Cook, T. Craig, S. Barot, D. Galli, S. Pal, S. Mohan, G. Gilbert, 2019 planning guide for data and analytics (2018). URL https://www.gartner.com/en/doc/361501-2019-planning-guide-for-data-and-analytics
[12] Blackbird, Blacbird homepage (2020). URL https://www.blackbird.online/
[13] Monitor-Box, M-box homepage (2020). URL https://www.monitor-box.com/m-box/
[14] DTI, Dti vision box: Quality control and inspection in 2d and 3d (2020). URL https://www.dti.dk/specialists/dti-vision-box-quality-control-and-inspection-in-2d-and-3d/35124
[15] Y. Yang, Z.-J. Zha, M. Gao, Z. He, A robust vision inspection system for detecting surface defects of film capacitors, Signal Processing 124 (2016) 54 – 62, big Data Meets Multimedia Analytics. doi:doi.org/10.1016/j.sigpro.2015.10.028.
[16] L. Li, K. Ota, M. Dong, Deep learning for smart industry: Efficient manufacture inspection system with fog computing, IEEE Transactions on Industrial Informatics 14 (10) (2018) 4665–4673. doi:10.1109/TII.2018.2842821.
[17] NVIDIA webpage, Jetson agx xavier developer kit (2020). URL https://developer.nvidia.com/embedded/jetson-agx-xavier-developer-kit
[18] Zarr Developers, Zarr, https://github.com/zarr-developers/zarr-python (2020).
[19] G. E. Krasner, S. T. Pope, et al., A description of the model-view-controller user interface paradigm in the smalltalk-80 system, Journal of object oriented programming 1 (3) (1988) 26–49.
[20] S. Bianco, R. Cadene, L. Celona, P. Napoletano, Benchmark analysis of representative deep neural network architectures, IEEE Access 6 (2018) 64270–64277. doi:10.1109/ACCESS.2018.2877890.
[21] S. Xie, R. Girshick, P. Dollár, Z. Tu, K. He, Aggregated residual transformations for deep neural networks, in: Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017, pp. 5987–5995.
[22] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, K. Keutzer, Squeezenet: Alexnet-level accuracy with 50x fewer parameters and¡ 0.5 mb model size, arXiv preprint arXiv:1602.07360 (2016).
[23] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
[24] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-cam: Visual explanations from deep networks via gradient-based localization, International Journal of Computer Vision 128 (2) (2019) 336–359. doi:10.1007/s11263-019-01228-7.

Paper C

A new authentic cloud dataset from a production facility for anomaly detection

Emil Blixt Hansen, Emil Robenhagen van der Bijl, Mette Busk Nielsen, Morten Svangren Bodilsen, Simon Vestergaard Berg, Jan Kristensen, Nadeem Iftikhar, and Simon Bøgh

The paper has been published in the
Towards Sustainable Customization: Bridging Smart Products and Manufacturing Systems pp. 415–422, 2022.

A New Authentic Cloud Dataset from a Production Facility for Anomaly Detection

Emil Blixt Hansen[1,2], Emil Robenhagen van der Bijl[1], Mette Busk Nielsen[1], Morten Svangren Bodilsen[1], Simon Vestergaard Berg[1], Jan Kristensen[3], Nadeem Iftikhar[4], and Simon Bøgh[1]

[1] Department of Materials and Production, Aalborg University, Aalborg, Denmark
ebh@mp.aau.dk
[2] Technology Education and Sustainable Growth, University College of Northern Denmark, Aalborg, Denmark
[3] Körber Supply Chain DK, Arden, Denmark
[4] Department of Computer Science, University College of Northern Denmark, Aalborg, Denmark

Abstract. As technology advances and modern Industry 4.0 solutions are becoming more widespread, the need for better-suited datasets is rising. The commonly used datasets for training machine learning focus on simple data of often publicly available information. Within the industry, there is only a handful of datasets publicly available to use. In this paper, we present a new authentic industrial cloud data (AICD) dataset collected from an actual operating pick-and-place machine handling items with variations in shape, size, and weight. The AICD dataset contains various analogue sensor values and states of the machine, collected from an existing cloud solution. Within the data, an error is present when the machine fails. Therefore, this dataset is suited for testing and developing predictive maintenance and anomaly detection algorithms to be used in the industry. Moreover, the paper also presents a baseline implementation as a performance indicator for future models.

Keywords: Machine learning · Big data · Dataset · Anomaly detection · Predictive maintenance

1 Introduction

With the materialisation of Industry 4.0 and its subsequent technologies, there is a need for more open datasets focusing on industrial applications. Specifically, to use machine learning and deep learning methods, a large dataset is often required for successful results [1,2]. With the rise of big data, more datasets have been made available. However, many of them are often not in the manufacturing domain, such as MNIST [3], California Housing Prices [4], ImageNet [5], and IMDB Reviews [6]. These datasets are meant for different machine learning topics such as machine vision (MNIST, ImageNet), prediction (California Housing Prices), and natural

A.-L. Andersen et al. (Eds.): CARV 2021/MCPC 2021, LNME, pp. 415–422, 2022.
https://doi.org/10.1007/978-3-030-90700-6_47

language processing (IMDB Reviews). In a production environment, it can be beneficial to discover a malfunction before it happens, such that the production can be stopped and maintenance can be carried out. This is called predictive maintenance, and it has economic potential, e.g. reduce downtime and lifetime extension of old assets [7]. Even with the economic potential, the adoption of the broader scope of artificial intelligence in, e.g. small and medium-sized enterprises (SMEs) is lacking [8].

Currently, there exist datasets that are tailored to the manufacturing environment. Purohit et al. [9] published an extensive sound dataset, with recordings of both functional and malfunctioning valves, fans, sliders, and pumps. Moreover, the C-MAPSS jet engine dataset from NASA is also commonly used [10]. The C-MAPSS dataset contains different sensor measurements and settings for multiple jet engines over a lifetime. In 2021 a new version of the popular dataset was released, which is more extensive compared to the original [11]. The new C-MAPSS dataset contains complete simulated flights of different lengths. Moreover, the initial data is collected at real commercial aeroplanes, and the run-to-failure is simulated in C-MAPSS. These three datasets are suitable for data scientists to explore machine learning methods. However, one of the shortcomings of them is that they are all artificial in some way. The sound dataset [9] contains actual recordings, but the recordings of anomalies are artificially introduced and thus is not a real failure. Moreover, the popular original C-MAPSS [10] is completely simulated with added sensors noise. The newer version [11] addresses this by starting with real recordings and simulating the remaining time to failure. Eduardo Oliveira [12] published in 2017 a dataset from a real mining operation of silica. This dataset does not concern malfunction but the prediction of silica concentration to improve the manufacturing process.

In this paper, we introduce a new authentic industrial cloud data (AICD) public dataset from an actual production process, where the failures are not simulated. In Sect. 2, the production process and the data collection are elaborated, and in Sect. 3, the content of the dataset is explained. In Sect. 4, a baseline model is presented, and in Sect. 5, the conclusion is described. The AICD dataset can be found at: https://www.kaggle.com/emilblixthansen/aicd-dataset

2 Data Collection

2.1 The General Use Case

The use case concerns a pick-and-place operation of large items with variations in shape, size and weight. The operation is carried out by a machine, with a tool containing different components to carry out the operation successfully. The tool is also equipped with sensors to measure the various components and a safety sensor to detect if the picked item is dropped. Items are occasionally dropped since the machine handles a wide variety of products throughout a shift. When an item is dropped, the production line needs to be stopped, and cleanup of the dropped item is required before the machine can continue its operations.

Therefore, if the machine could detect if an item is about to be dropped, it could stop the operation and avoid dropping the item.

2.2 Data Extraction

The data extraction process took place on a machine at a real production located in Europe. The data was collected on and off over the course of two days. The data was collected at an interval of 10 ms through the already existing cloud solution.

2.3 Anonymising Production Data

The machine and its operation contain vital company information. Therefore, the exact information regarding the machine and its whereabouts are not disclosed within the paper. Moreover, information within the dataset has also been altered. This includes dates and the column names within the dataset. However, there have been no alterations to the exact data within the dataset, and thus keeping its integrity.

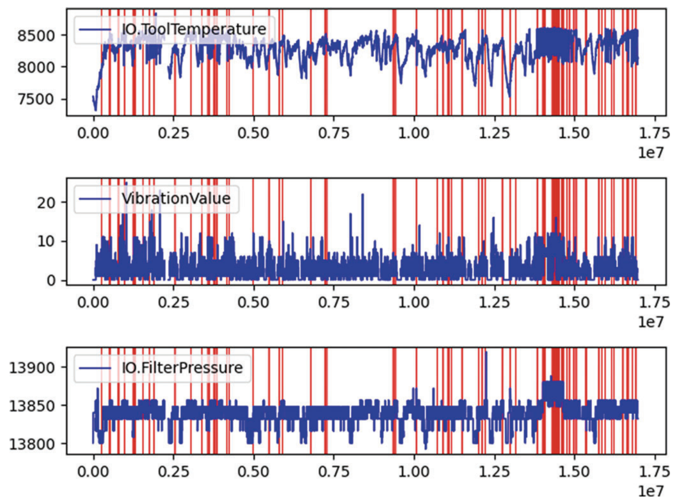


Fig. 1. All 16.9 million sensor measurements from three different sensors, measuring suction, vibration, and pressure. The red lines are when the drop detection sensor has been triggered.

3 Dataset Content

The AICD dataset is distributed in a CSV file format and is split into five CSV files because of the size of the dataset. Moreover, a compressed Python Pickle file containing the same data is also distributed. All recordings are from

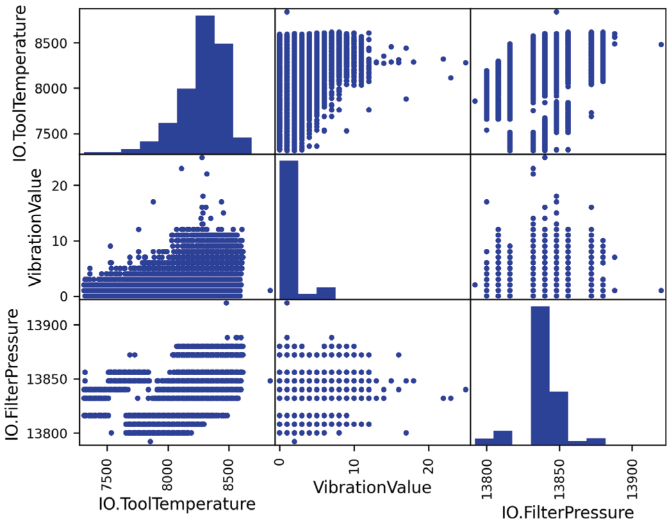


Fig. 2. Scatter matrix of the tree sensor values from Fig. 1. Correlation between the different sensor values is shown along with a histogram of each sensor value along the diagonal.

the same machine. One of the purposes of the presented dataset is to make an authentic dataset open to the community of machine learning engineers within the industry. Henceforth, the different challenges are also present. The challenges include preprocessing the data, where our dataset is not preprocessed; thus, there are no training and test set. The data is not split into drops and no-drops. When a drop has happened, a binary signal goes high (*Alarm.ItemDroppedError*), which should be used to indicate a drop. Moreover, this trigger can also be used to isolate the data before a drop which can be used to train an algorithm to learn when a drop is about to happen.

The AICD dataset contains 96 columns of various data and data types. The first three columns are machine-uptime in nanoseconds, date, and timestamp. The remaining 93 columns are features based on sensors, machine status, and time recordings. Since the dataset is not preprocessed, some features are trivial where either they do not change in value or accumulate to infinity. In Fig. 1, three examples of the dataset are shown, along with indications for when a drop has happened. From the figure, it is not visible that the drop is correlated with these sensors measurements, henceforth, the challenges with this multivariate dataset. Moreover, in Fig. 2 the correlation between the three sensor values is shown. From the figure, it can be seen that there is a correlation between, e.g. *VibrationValue* and *IO.ToolTemperature*, where the vibration is prone to rise with the increased temperature.

The dataset has 16,990,692 rows, and since no preprocessing has been made, some of the rows contain null values. An overview of all the different columns can be found in the *README.md* file along with the dataset.

4 Baseline Experiment

To demonstrate the usage of the dataset and test our initial hypothesis, we present a minimal baseline experiment. The purpose of the experiment is to detect anomalies in the dataset when a drop is happening. Because this dataset consists of multivariate time series data, a neural network LSTM (Long Short Term Memory, [13]) autoencoder is selected as the method. LSTM has been successfully deployed in time series problems [14–16] and autoencoder in anomaly detection [17,18]. Furthermore, together they have been successful in solving time dependant anomaly detection [19,20]. In Table 1, our model architecture is shown. The experiment was conducted in Python 3.6, and the model was created with Tensorflow 2.2. The LSTM layers of the model use the default hyperparameters from Tensorflow, and the complete script can be found in the file *baseline_experiment.py* along with the dataset.

Table 1. The used LSTM autoencoder architecture.

Layer (type)	Output shape	Param #
Input (InputLayer)	(None, 1, 48)	0
Encoder (LSTM)	(None, 1, 16)	4160
Encoder (Dropout)	(None, 1, 16)	0
Encoder (LSTM)	(None, 8)	800
Code (RepeatVector)	(None, 1, 8)	0
Decoder (LSTM)	(None, 1, 8)	544
Decoder (LSTM)	(None, 1, 16)	1600
Output (TimeDist)	(None, 1, 4)	816
Total trainable params:		7920

Before the model could be trained, the data needs to be preprocessed. For this experiment, where the objective is anomaly detection, the data has been split into training and test data, where the training data is from the *Data1.csv* file where drops have been removed from that file. Moreover, all rows two minutes before and after the detected drops were also removed. The test set entails data from the *Data4.csv* file. Only 48 columns were kept for training and testing, and the rest were removed either because they contained no unique values or rose to infinity. The training data was standardised by Eq. 1, where x is the training matrix, μ is the corresponding mean vector, and s is the standard deviation vector. The test data also applied Eq. 1, but used μ and s from the training set.

$$z = \frac{x - \mu}{s} \tag{1}$$

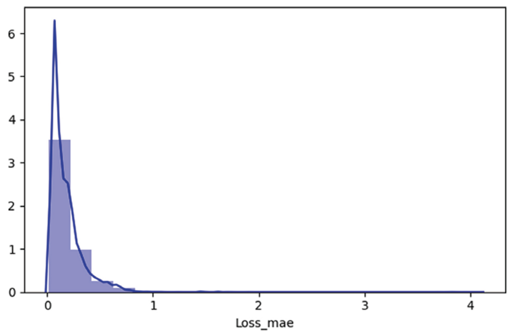
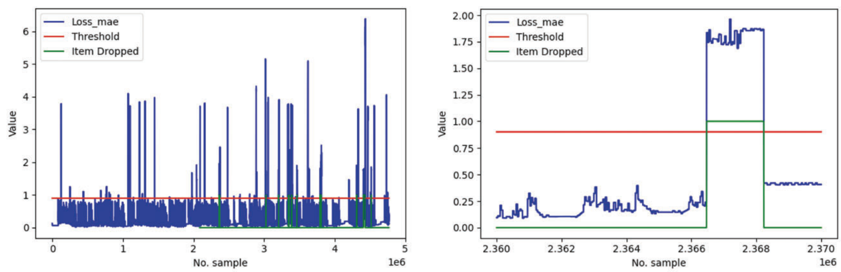


Fig. 3. The MAE training loss distribution.

The training loss function was chosen to be the Mean Absolute Error (MAE). The training was done in 40 epochs with Adam optimiser [21]. The training loss distribution can be seen in Fig. 3. Based upon the training loss distribution, a threshold of 0.9 was chosen to flag an anomaly. Thus, if the loss exceeds 0.9, an anomaly has occurred.

To compare the train and test data, the MAE loss has been combined in Fig. 4a. The test data starts where the green *item dropped* line starts. It can be seen that there are anomalies exceeding the threshold without a drop; this further emphasises the challenges present for data scientists working with production data. To further validate that the baseline model does learn, a closer inspection of the first drop from the test dataset is carried out. In Fig. 4b a snapshot of the drop is shown. Here it is visible that the drop is measurable in the MAE loss. The model has an MAE value of 0.24767 on the complete test set. The MAE score can serve as a baseline score for future anomaly experiments with the dataset.



(a) Combined train and test data.

(b) The snapshot of the first drop in the test data.

Fig. 4. The MAE loss is illustrated in blue, the red line is the threshold of 0.9, and the green line starts with the test data and is the binary item dropped signal.

5 Discussion and Conclusion

In this paper, we presented a new cloud dataset named AICD. The data was collected at an operational running production environment and contained 96 different data features, and a total of 16,990,692 samples was collected. Along with this new public available dataset, we presented a baseline experiment of drop detection to demonstrate the dataset usage and compare futures improvements. The AICD dataset and the baseline experiment can be found at: https://www.kaggle.com/emilblixthansen/aicd-dataset

Wuest et al. [22] identified acquisition and availability of relevant data to be some of the challenges there is in using machine learning in a production setting. This works contribution is a publication of a public relevant dataset within the sparse field op production relevant dataset. As this dataset is collected live at a running production environment and is not preprocessed, we believe it is one of the few production datasets that mostly resemble how data collection in a real production would look like. Therefore, this dataset can bring research from academia closer to the actual challenges from the industry and thus enhance the impact of said research. We encourage researchers to use this dataset to learn the challenges of a real production and develop and research new methods for anomaly detection and forecasting. This can further aid the progression of big data and analytics domain of Industry 4.0.

Acknowledgements. The work was conducted as a part of the Danish strategy for Industry 4.0 in SMEs called Innovation Factory North and in collaboration with Körber Supply Chain. The research was funded by the European Regional Development Fund and University College of Northern Denmark.

References

1. Linjordet, T., Balog, K.: Impact of training dataset size on neural answer selection models. In: Azzopardi, L., Stein, B., Fuhr, N., Mayr, P., Hauff, C., Hiemstra, D. (eds.) ECIR 2019. LNCS, vol. 11437, pp. 828–835. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-15712-8_59
2. Prusa, J., Khoshgoftaar, T.M., Seliya, N.: The effect of dataset size on training tweet sentiment classifiers. In: 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), pp. 96–102 (2015)
3. LeCun, Y., Cortes, C.: MNIST handwritten digit database (2010). http://yann.lecun.com/exdb/mnist/
4. Kelley Pace, R., Barry, R.: Sparse spatial autoregressions. Stat. Prob. Lett. **33**(3), 291–297 (1997). https://www.sciencedirect.com/science/article/pii/S016771529600140X
5. Deng, J., et al.: Imagenet: a large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. IEEE (2009)
6. Maas, A.L., Daly, R.E., Pham, P.T., Huang, D., Ng, A.Y., Potts, C.: Learning word vectors for sentiment analysis. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, pp. 142–150. Association for Computational Linguistics (June 2011)

7. Haarman, M., et al.: Predictive maintenance 4.0 - beyond the hype: Pdm 4.0 delivers results (2018). https://www.pwc.be/en/documents/20180926-pdm40-beyond-the-hype-report.pdf
8. Hansen, E.B., Bøgh, S.: Artificial intelligence and internet of things in small and medium-sized enterprises: a survey. J. Manuf. Syst. **58**, 362–372 (2021). https://www.sciencedirect.com/science/article/pii/S0278612520301424, digital Twin towards Smart Manufacturing and Industry 4.0
9. Purohit, H., et al.: Mimii dataset: Sound dataset for malfunctioning industrial machine investigation and inspection (2019)
10. Saxena, A., Goebel, K.: Turbofan engine degradation simulation data set. NASA Ames Prognostics Data Repository pp. 1551–3203 (2008)
11. Arias Chao, M., Kulkarni, C., Goebel, K., Fink, O.: Aircraft engine run-to-failure dataset under real flight conditions for prognostics and diagnostics. Data **6**(1), 5 (2021). https://www.mdpi.com/2306-5729/6/1/5
12. Oliveira, E.M.: Quality prediction in a mining process (2017). https://www.kaggle.com/edumagalhaes/quality-prediction-in-a-mining-process
13. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997). https://doi.org/10.1162/neco.1997.9.8.1735
14. Pham, V., Bluche, T., Kermorvant, C., Louradour, J.: Dropout improves recurrent neural networks for handwriting recognition (2014)
15. Sønderby, S.K., Winther, O.: Protein secondary structure prediction with long short term memory networks (2015)
16. Zhao, R., Wang, J., Yan, R., Mao, K.: Machine health monitoring with lstm networks. In: 2016 10th International Conference on Sensing Technology (ICST), pp. 1–6 (November 2016)
17. Chen, Z., Yeo, C.K., Lee, B.S., Lau, C.T.: Autoencoder-based network anomaly detection. In: 2018 Wireless Telecommunications Symposium (WTS), pp. 1–5 (2018)
18. Zhao, Y., Deng, B., Shen, C., Liu, Y., Lu, H., Hua, X.S.: Spatio-temporal autoencoder for video anomaly detection. In: Proceedings of the 25th ACM International Conference on Multimedia, pp. 1933-1941. Association for Computing Machinery, New York (2017)
19. Park, D., Hoshi, Y., Kemp, C.C.: A multimodal anomaly detector for robot-assisted feeding using an lstm-based variational autoencoder. IEEE Robot. Autom. Lett. **3**(3), 1544–1551 (2018)
20. Nguyen, H., Tran, K., Thomassey, S., Hamad, M.: Forecasting and anomaly detection approaches using lstm and lstm autoencoder techniques with the applications in supply chain management. Int. J. Inf. Manage. **57**, 102282 (2021)
21. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization (2017)
22. Wuest, T., Weimer, D., Irgens, C., Thoben, K.D.: Machine learning in manufacturing: advantages, challenges, and applications. Prod. Manuf. Res. **4**(1), 23–45 (2016)

Paper D

A data-driven modular architecture with denoising autoencoders for health indicator construction in a manufacturing process

Emil Blixt Hansen, Helge Langseth, Nadeem Iftikhar, and Simon Bøgh

The paper has been submitted to
IEEE Access, submitted August 2022.

In peer review

A data-driven modular architecture with denoising autoencoders for health indicator construction in a manufacturing process

Emil Blixt Hansen[a,b,*], Helge Langseth[c], Nadeem Iftikhar[b], Simon Bøgh[a]

[a]*Department of Materials and Production, Aalborg University, Denmark*
[b]*Industrial Digital Transformation, University College of Northern Denmark, Aalborg, Denmark*
[c]*Department of Computer Science, Norwegian University of Science and Technology, Trondheim, Norway*

Abstract

Within the field of prognostics and health management (PHM), health indicators (HI) can be used to aid the production and, e.g. schedule maintenance and avoid failures. However, HI is often engineered to a specific process and typically requires large amounts of historical data for set-up. This is especially a challenge for SMEs, which often lack sufficient resources and knowledge to benefit from PHM. In this paper, we propose ModularHI, a modular approach in the construction of HI for a system without historical data. With ModularHI, the operator chooses which sensor inputs are available, and then ModularHI will compute a baseline model based on data collected during a burn-in state. This baseline model will then be used to detect if the system starts to degrade over time. We test the ModularHI on two open datasets, CMAPSS and N-CMAPSS. Results from the former dataset showcase our system's ability to detect degradation, while results from the latter point to directions for further research within the area. The results shows that our novel approach is able to detect system degradation without historical data.

Keywords: prognostics and health management, health indicators, health index, machine learning, manufacturing

---

1. Introduction

The manufacturing industry has seen an increased interest in modern IT technologies to enhance their production. These technologies are commonly described as Industry 4.0. Nonetheless, many of the processes use old equipment, which often lacks sensors and communication protocols to be used with technologies of Industry 4.0, such as big data and analytics [1]. A study also found that the use of old equipment negatively impacts the competitiveness and innovation of a small and medium-sized enterprises (SME) [2]. Prognostics and health management (PHM) concerns the topic of monitoring and health indication of machinery. The health indicator (HI) part is concerned with estimating the health of machinery. The HI can be constructed differently depending on the equipment and setting, and an important design choice is whether the HI is to be model-based, data-driven, or hybrid [3]. Model-based models takes the system's underlying information into accounts using mathematical models or other descriptions of the system behaviour. The data-driven approach is where the HI score is not built on top of system information, but rather learned from data. Lastly, the hybrid approach is a combination of the other two. In this work we focus on the *data-driven* approach. Within this category, there are different ways the HI can be constructed, and examples include signal processing and statistical techniques, including time series domains using time-domain or frequency-domain feature [4, 5]. The continued development in deep learning has resulted in several approaches to construct HI through deep learning. A variant of autoencoders (AE) is for instance commonly used [6, 7]. Experiments from Zhao et al. [7] showed that denoising autoencoders (DAE) outperformed other methods for multivariate time series reconstruction problems, whereas [8] constructed a convolutional neural network (CNN) to construct the HI directly.

Within the PHM domain, estimating the remaining useful life (RUL) of a given machinery can be beneficial. RUL estimation is often based on the HI score, but the RUL estimates will then typically rely on historical data of failures to estimate the relationship between HI and RUL. Ensemble learning with time dependant degradation weights has been shown to generate accurate RUL predictions [9], while [10, 11] used an adaptive denoising algorithm to make feature extraction and compute the RUL of aircraft engines.

*Corresponding author, Fibigerstræde 16, 9220 Aalborg East, Denmark
Email address: ebh@mp.aau.dk (Emil Blixt Hansen)

Moreover, it has been shown that utilising semi-supervised learning can improve the performance of RUL prediction [12]. A combination of ARIMA and LSTM models has been proposed to increase performance in anomaly detection and forecasting [13].

The use of HI for PHM is often engineered to the specific machine/task, and furthermore requires that sufficient data of the machine is available to construct the HI. However, SMEs does not in general have this available to them; when it comes to Industry 4.0, they are often behind larger enterprises and do not have the knowledge to use technologies such as artificial intelligence (AI) [14]. Previous work has focused on how to make these technologies more available for SMEs, focusing on ease of use. For instance, Hansen et al. [15] described a framework and conducted tests within image classification problems.

To make HI scoring easy-to-use for SMEs, it is not feasible to assume that the HI scoring system is particularly engineered to each specific machine used by the SME. On the contrary, it should be possible for an operator to set up the framework without understanding the underlying algorithms. Furthermore, since a production process typically consists of multiple different machinery and equipment, it could be beneficial for the SME not to be limited to a specific engineering application. Therefore, ModularHI is not limited to specific sensor input types. Instead it uses a modular divide-and-conquer strategy, which is in-line with current thinking in AI. For instance, DeepMind [16] presented PathNet, a scalable modular network suited for transfer learning between different tasks in 2017. Here small neural network models together form a larger network. The work was later extended by Stepwise PathNet [17]. These works demonstrate that modularity play a central role within deep learning to enhance the overall performance.

In this paper we present ModularHI, a novel approach to construct an HI score for an arbitrary machinery based on the available sensor inputs. ModularHI is aimed at, but not limited to, use cases within the production sector. Our working hypothesis is that while it is infeasible to engineer health monitoring solutions for each machine, it is still beneficial to monitor a machine's health even if the employed system is less accurate than a tailor-made solution would have been. The goal is thereby to provide a low-cost and generally applicable HI monitoring system for SMEs, that can enable them to start benefitting from Industry 4.0 technologies. The rest of the paper is structured as follows. In section 2 the algorithm and components of ModularHI is explained. In section 3 two different tests scenarios are described and in section 4 the results are shown. Furthermore, in section 5 and section 6 the discussion and conclusion is presented, respectively.

2. System architecture

ModularHI aims to support a broad set of use cases, and is not engineered for specific tasks. We have therefore split the execution into three states: *setup*, *burn-in*, and *inference*. The general execution flow of ModularHI's three states is shown in Figure 1. Besides the execution flow, the underlying system comprises two main parts: the *component models* and the *aggregator*. The execution flow and the two main parts are discussed next.

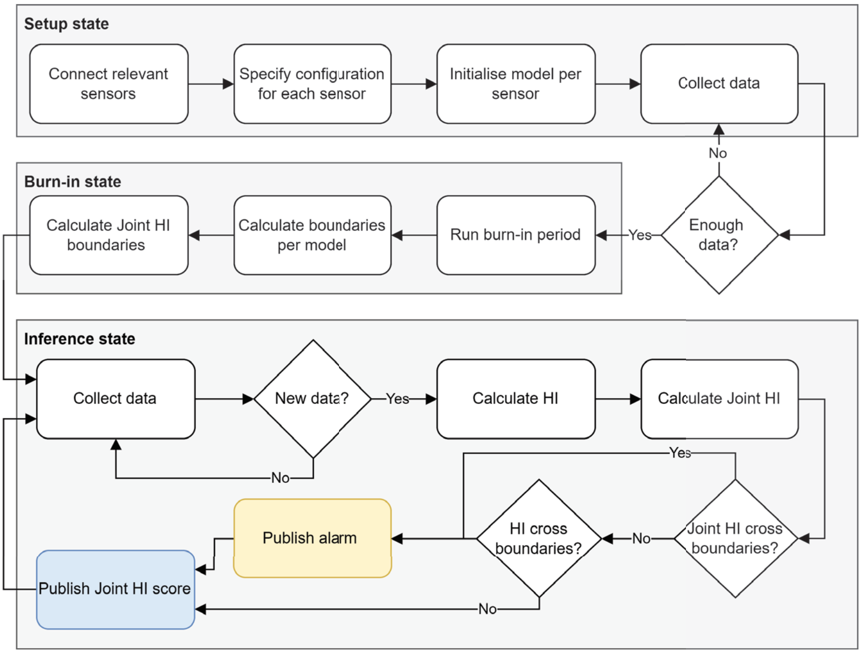


Figure 1: The execution flow of the three states: setup, burn-in, and inference.

2.1. Execution flow

In the first state, *setup state*, the different models and their specific input are configured. The models are called *component models*, and are further described in section 2.2. Furthermore, the initial data collection also takes place in this state. When enough data is collected, the execution goes to the burn-in state.

The *burn-in state* is where the component models are fitted to the collected datasets. We will assume that the data describes the monitored system in its optimal performance state, preferably directly after machine maintenance. When each component model has been sufficiently trained on their individual datasets, the components model can calculate a real-time health indicator (more on this in section 2.2). In the inference state the component model will monitor its own status, and posts an alarm if the HI value is deemed to be outside its acceptable region. First, however, the acceptable region must be established. While several methods can be used to determine this acceptable region, we use a simple strategy in this paper. We monitor the HI during the burn-in state, calculate the standard deviation of the calculated HI during this period, and define the acceptable region for the HI to between zero and nine times the calculated standard deviation. Finally we define the HI at the system level. This is done by the aggregator model, which collects and combines the HI values from the component models. The combination is named the *joint HI* (HI^j), represents the overall system's HI score, and has its own acceptable region. Again, the region is defined by nine standard deviations calculated during the burn-in stage. The combinations of the different component models is further described in section 2.3. Next, the execution enters the third state, the inference state.

The *inference state* is where the ModularHI is executed on new data. Here a continuous collection of data occurs, and every time a new data point is collected, it is executed through the models. This includes calculating the current HI score comparing the new datapoint to data collected during the burn-in state. After that, HI^j is calculated, and all of the boundaries are checked. If at least one boundary is crossed, the system is assumed to either be in a faulty state or sliding towards it, indicating that maintenance or inspection is recommended. This therefore results in an alarm state being published.

2.2. Component models

ModularHI has been designed with simplicity of use in mind. To handle different sensor categories, with different data types and data dimensions, we therefore designed specific models per sensor type. Hence, each sensor has its own model, specifically designed for that measurement type. It follows that component models are specifically trained for a certain type of measurement (e.g., a specific model for temperature data).

As mentioned above, there are different ways of constructing a HI score. One of the requirements for ModularHI is that it should be modular and handle a different mixture of sensor inputs. Therefore, it is infeasible to train a separate model to generate the system's HI score for each possible scenario. Instead, ModularHI consists of individual component models, each outputting a HI score based on a single stream of data. To enable these HI scores to be comparable to each other throughout the different sensor inputs, we have chosen the DAE [7] as our go-to model type. Since the component models are designed to only handle univariate time series, the model size is limited. In the experiments reported in this paper, we have chosen to use a DAE with LSTM layers to capture the time perspective in the data. The DAE architecture is the same across all sensors and all experiments, and can be seen in Table 1. We note that the same core modelling approach could be used to analyse other streaming data like sound or video simply by adapting the DAE to handle that data-type.

Table 1: The LSTM autoencoder.

Layer type	**Specification**
Input	Shape (batch size, window size, 1)
LSTM	8 units, returned sequence
Dropout	Probability 0.5
LSTM	4 units
Repeat vector	8 times
LSTM	4 units, returned sequence
Dropout	Probability 0.5
LSTM	8 units, returned sequence
Output	Time distributed of window size

Currently, this DAE architecture is used for all time series sensor input, such as temperature, vibration, etc. Even though all models have the same architecture, they are independently pre-trained on a relevant dataset. This application of transfer-learning reduces the data-requirements during system setup, reduces the compute required for training the model during burn-in, and improves the model's performance [18].

The HI score is based on the mean absolute error (MAE) during reconstruction of the input-signal:

$$\mathrm{HI}_k^m = \frac{1}{n}\sum_{i=1}^{n} |x_i - \tilde{x}_i| \,. \tag{1}$$

Here HI_k^m refers to the HI score of the k'th component model m (out of the set of M component models for a specific setup case), $\boldsymbol{x}$ is a vector of measurements of window-size n and $\tilde{\boldsymbol{x}}$ is the corresponding reconstructed measurements also of size n. As stated earlier, each model has its own upper-bound calculated by taking the mean of the HI of burn-in and finding standard deviation σ from the sample mean. Specifically, the upper boundary is specified by $9\,\sigma$ and the lower is set to 0 (since MAE is a non-negative real number).

2.3. Aggregator

The objective of the *aggregator* is to combine the information from the component models to give a system-wide HI, which we call the *joint HI* (HI^j). We note that HI^j can be calculated in many different ways, ranging from simple aggregations like maximum or sum to increasingly complex combinations, e.g., represented by a deep neural network. As ModularHI is designed for ease-of-use within a wide range of application areas, to be data-driven, and to not require engineering input to define the aggregation function, we choose to let each sensor contribute equally to HI^j by default. The joint health indicator is now simply the average of the N components' HI scores:

$$\mathrm{HI}^j = \frac{1}{N}\sum_{i=1}^{N} \mathrm{HI}_i^m \tag{2}$$

Nevertheless, if an operator knows that some sensor is more critical to the stability of a system, they can specify weights $\boldsymbol{w}$ for each sensor input. An example of this could be that the operator knows that the system starts to be unstable when the system temperature rises. The operator will then give the temperature sensor input a higher weight the the others. Thus the HI^j will be more acceptable to changes from the temperature sensor. The HI^j is then calculated by a weighted average as seen in Equation 3. By default all the weights are specified as 0.5, hence they are all equal and the HI^j will be calculated as Equation 2.

$$\mathrm{HI}^j = \frac{\sum_{i=1}^{N} w_i \cdot \mathrm{HI}_i^m}{\sum_{i=1}^{N} w_i} \tag{3}$$

ModularHI calculates the upper-bound for HI^j at $9\,\sigma$ during the burn-in state. This is done when all component models have finished their burn-in step. Then Equation 3 is calculated to find all of the HI^j from the burn-in period to calculate the sample mean and thus the corresponding σ.

When the burn-in state is completed, the cycle of the inference state is executed. For a new data point x_t all component models will run a window of the last n samples including the new data point, such that $\boldsymbol{x} = \{x_{t-n+1}, \ldots, x_t\}$. After that, all calculated HI^m values will be used to calculate the current HI^j. With the new HI^j all boundaries are checked such that if any boundary is crossed in their respected HI, an alarm is published. An illustration of the inference state of ModularHI can be seen in Figure 2 where test 7 from section 3.1 is shown.

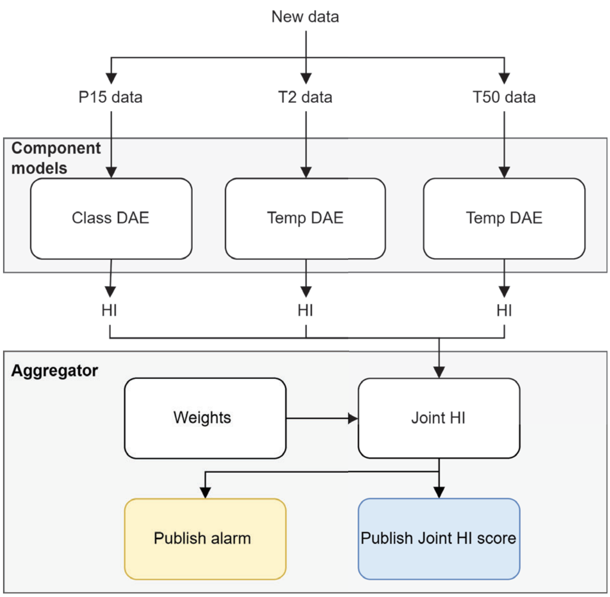


Figure 2: The combination of the different component models and the aggregator from CMAPSS T7.

3. Experiments

To validate the use of ModularHI, we needed to test it on a relevant dataset. This dataset should preferably contain multiple different sensors (to validate the modularity), include recordings from a steady-state (to enable the initial data collection), and have a continuous data steam eventually leading to a faulty state. Finally, it would be beneficial if the dataset was relevant to production applications, either being extracted from one or similar to a process that could exist in one. The MIMII dataset, which contains various recordings of failure of processes [19] is one potential candidate. However, it does not contain continuous recording from steady-state to failure, and does not include information such as vibration. A dataset collected from a real production line containing various different sensors and faulty states was presented in [20]. Unfortunately, this dataset consists mainly of operations failures and not process failures. NASA published the jet engine degradation dataset CMAPSS in 2008 [21]. This dataset contains simulated jet engines monitored until failure, where each data point is an aggregated value from one flight. More recently, they released an updated version, named N-CMAPSS [22]. This dataset contains various sensor values from the same setting. The main difference is that data in N-CMAPSS is collected throughout the flights and includes both ascent and descent. Moreover, the initial states of the flights are collected at actual flights, and only the run to failure is simulated. Hence, N-CMAPSS is a larger and more complex dataset compared to the original CMAPSS. We have chosen to test ModularHI against both CMAPSS and N-CMAPSS. We chose both datasets because both of them have a continuous data stream from steady-state until failure. Moreover, the two datasets symbolise two different levels of complexity. The available datasets also show the lack of available authentic production datasets which can be used in our case. Thus the two CMAPSS datasets are the closest to an actual production dataset with their continuous data stream of various sensors.

3.1. CMAPSS dataset test

The original CMAPSS dataset consists of multiple flights, which are commonly used for RUL predictions [10, 11]. Here, however, we use ModularHI to generate a HI and determine if the equipment/process is in steady operation. Therefore, our test will only consist of a single jet engine (engine number 1 from the dataset FD001). Besides which sensors to use, we also need to decide on two hyperparameters: the duration of the burn-in period and the window size used by the autoencoder. Since engine number 1 only has 220 recordings, a burn-in period of 78 is chosen as it includes a stable amount of data. A window size of 8 was chosen as it should capture enough of the temporal data. We remark that the relatively small dataset used during burn-in is sufficient due to the auto-encoder being pre-trained on a separate dataset. The component model concerning temperature is pre-trained on historical weather data[1] and the component models concerning generic time series data is pre-trained on accelerometer data collected at Aalborg University. The general purpose of pre-training the networks is to train them to reproduce the input; the underlying characteristics of the individual sensor data will then be learned during the burn-in state.

Table 2: The eight test setups for CMAPSS dataset. Each sensor has its own model; if several sensors are mentioned together with only one model, then all sensors have that type of model. The same is true with the assigned weights. The model T is pre-trained temperature data and C is a generic model pre-trained on accelerometer data.

Tests	**Sensors**	**Models**	**Weights**
CMAPSS - T1	T50	T	0.5
CMAPSS - T2	T30, T50	T	0.5
CMAPSS - T3	T2, T30, T50	T	0.5
CMAPSS - T4	T2, T30, T50	T	0.6, 0.2, 0.2
CMAPSS - T5	P15	C	0.5
CMAPSS - T6	P15, T2, T50	C, T, T	0.5, 0.5, 0.5
CMAPSS - T7	P15, T2, T50	C, T, T	0.2, 0.2, 0.6
CMAPSS - T8	P2, P15, epr, farB, Nf_dmd, PCNfR_dmd, T50	T50: T, Rest: C	0.5

We conducted eight tests on the data from Engine number 1, each with a different combination of sensors and weights (refer to Equation 3). The test-cases are described in Table 2. Moreover, the values for the sensors are plotted in Figure 3. The eight tests are conducted to validate various aspects of ModularHI. Tests 1 and 2 examine the usability of ModularHI with inputs where degradation is present. Test 5 considers what happens when degradation is not clearly present. Tests 3, 6 and 8 look at a combination of sensor with "informative" and "uninformative" sensor values with equal weights. Finally, Tests 4 and 7 (inference state visualised in Figure 2) highlight the same when a larger weight has been applied to the "informative" measurements.

[1]https://www.kaggle.com/budincsevity/szeged-weather

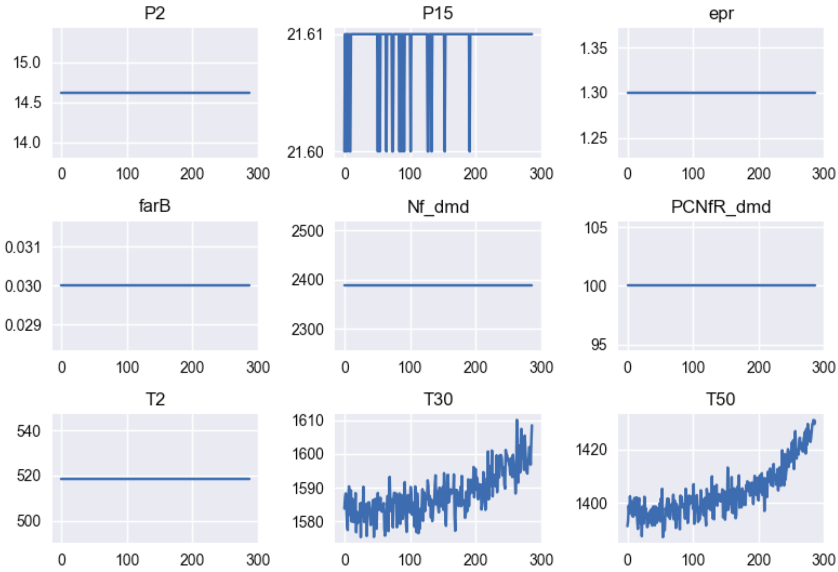


Figure 3: The nine used sensor from CMAPSS engine number 1 used, only in sensor T30 and T50 is the degradation visible with an upward trend.

3.2. N-CMAPSS dataset test

Besides the original CMAPSS data, we also tested ModularHI on the more complex N-CMAPSS dataset. Since N-CMAPSS consists of multiple flights of different lengths, we chose flight number 2 from dataset DS01. These flights comprised of data from flights over 3000 feet. All the flights include ascend from 3000 feet, cruise altitude and descending to 3000 feet. Flight number 2 describe class three flights, meaning all flights is over 5 hours long [22]. The sensor readings and flight settings are subject to high variance during a flight's ascend and descend. Since the focus of ModularHI is to be able to detect deviations from stable operation, we have filtered out the ascending and descending parts of the flights. Moreover, only cruise parts between 25,000 and 30,000 feet were included during the burn-in state. Furthermore, we only included flights where the cruise part had a minimum of 1024 observations (as we used window-size $n = 1024$ for the autoencoders). The cleaned file consists of 4 sensors, and the first 105,876 readings are used for the burn-in state.

Table 3: The five test setups for N-CMAPSS dataset. Each sensor has its own model; if several sensors are mentioned together with only one model, then all sensors have that type of model. The same is true with the assigned weights. The model T is pre-trained temperature data and C is a generic model pre-trained on accelerometer data.

Test	**Sensors**	**Models**	**Weights**
N-CMAPSS - T1	T40	T	0.5
N-CMAPSS - T2	SmLPC	C	0.5
N-CMAPSS - T3	T40, SmLPC, SmHPC	T, C, C	0.5
N-CMAPSS - T4	T2, SmLPC	T, C	0.5
N-CMAPSS - T5	T2, SmLPC, SmHPC	T, C, C	0.6, 0.2, 0.2

We conducted five tests of the cleaned DS01 engine 2 data. All of them had a sliding window size of 1024 and was tested in batches of 256. To combat the abrupt changes in the values when going from a single cruise to another, we only trained with mini-batches where one cruise was involved. The different tests, along with their models and weights are reported in Table 3. A plot of the used sensors after data cleaning can be seen in Figure 4. While some of the sensors in Figure 4 show signs of degradation (see, in particular, SmLPC), the apparently noisy variations in the readings dominate the signals. The dataset is therefore particularly challenging for ModularHI, where we for simplicity assume that all data during the burn-in period is from stable operation of the equipment/process. From just plotting the dataset, one should therefore expect that the monitoring system will be more challenged by this dataset than from the previous one.

As with the CMAPSS dataset, the five tests were conducted to test different aspects of ModularHI. Tests 1 and 2 was to see the effect of using only a single sensor, where one of them (Test 2) had a clear degradation. Tests 3 and 4 use multiple sensor values, and Test 5 employs sensors with different weights.

4. Results

The component models were learning until there was no improvement in the validation loss for 25 consecutive epochs, and at that time the weights from the epoch with the lowest validation loss was used.

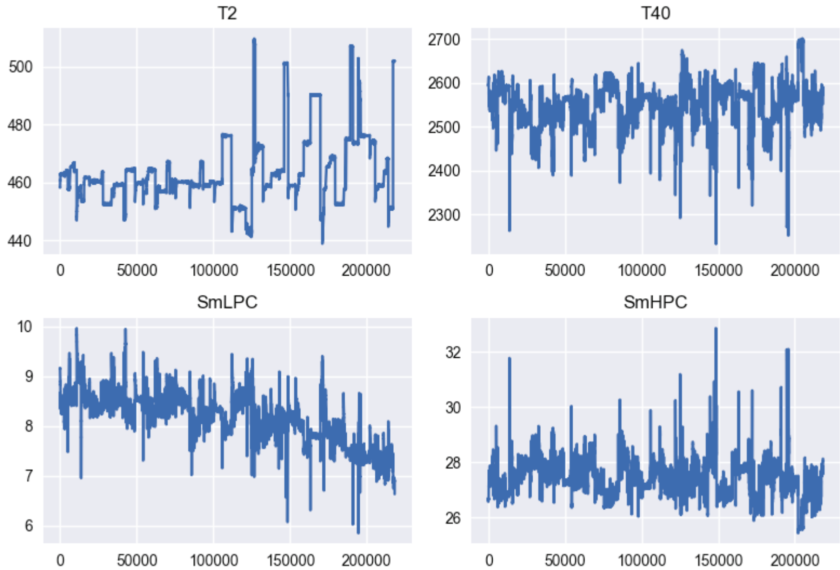


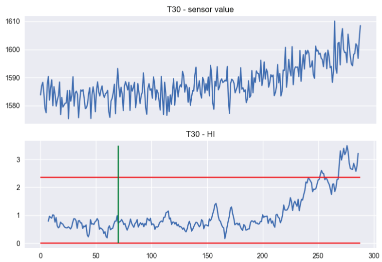
Figure 4: The 4 used sensor from N-CMAPSS engine number 2. Only in sensor SmLPC is the degradation clearly visible with an downwards trend.

4.1. CMAPSS dataset result

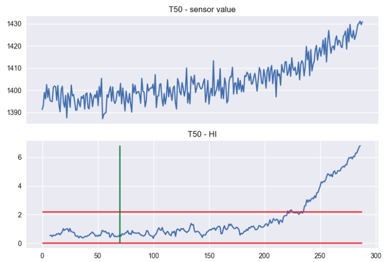
The first two tests, Tests 1 and 2, examined the system when degradation is visually present in the sensor readings. Since *T50* is included in both tests, the results are almost identically. The result from Test 2 reported in Figure 5 show that HI^j reacts at time-step 212. The two component models first publish alarms at step 221 and 252 for sensor T50 and T30, respectively.

In Test 7, we used sensors T50, T2, and P15. Sensor T50 displays a clear degradation in the sensor values, while sensor P15 only has a small alteration in the data, and T2 is static. The results can be seen in Figure 6, and show that sensor T2 and P15 do not indicate any depredations, neither does their HI score. However, since sensor T50 crosses its upper-bound an alarm is published. We also note that HI^j exceeds its upper bound even though two sensors do not indicate a degradation.

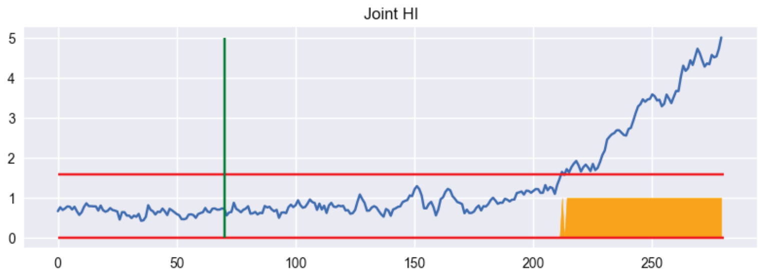
Test 8 was conducted to see if a majority of non-contributing sensor values will prevent HI^j to detect that one sensor measures a degradation. Therefore,



(a) T30 sensor value and HI.



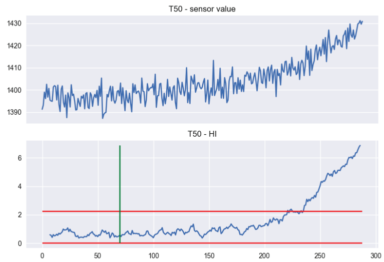
(b) T50 sensor value and HI.



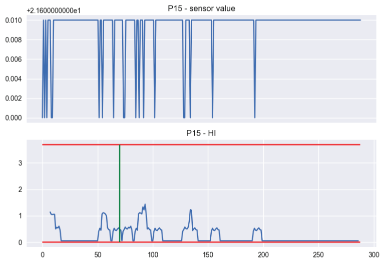
(c) HI^j for T30 and T50.

Figure 5: The result from Test 2. The green vertical line indicates the end of the burn-in state. The horizontal red lines indicate the calculated boundaries. The orange area in the HI^j is when an alarm is published.

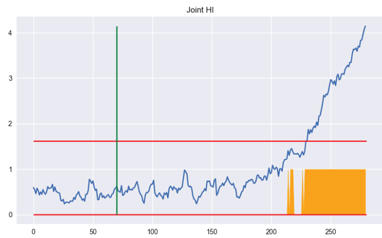
(a) T2 sensor value and HI.



(b) T50 sensor value and HI.



(c) P15 sensor value and HI.



(d) HI^j for T2, T50, and P15.

Figure 6: The result from Test 7. The green vertical line indicates the end of the burn-in state. The horizontal red lines indicates the calculated boundaries. The orange area in the HI^j is when an alarm is published. The HI score from Figure 6a is below the upper boundary line.

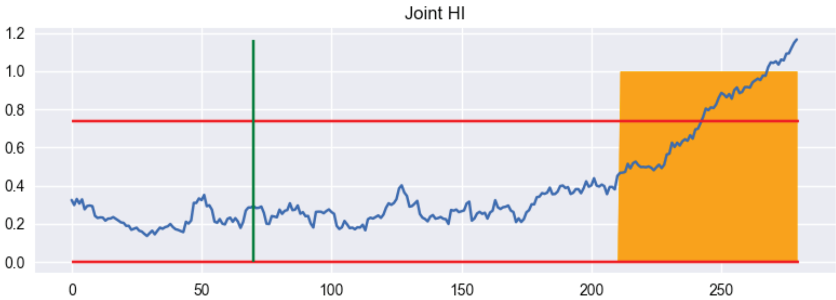


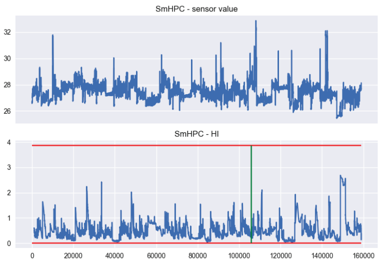
Figure 7: The result from Test 8. The green vertical line indicates the end of the burn-in state. The horizontal red lines indicates the calculated boundaries. The orange area in the HI^j is when an alarm is published.

this test includes one sensor with the degradation evident (T50) and six without (P2, P15, epr, farB, Nfdmd, and PCNfRdmd). The HI^j results in Figure 7 show that an alarm is published before HI^j crosses its boundary. This is initiated by the sensor T50 crosses its upper-bound. Moreover, it can also be seen that HI^j itself crossed its boundary at time 243.

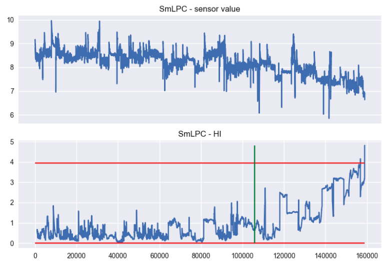
4.2. N-CMAPSS dataset result

We now focus on the more challenging N-CMAPSS dataset. Here, Test 3 was designed to examine ModularHI with multiple sensors. We chose sensors T40, SmLPC, and SmHPC, and gave them equal weight. The results in Figure 8 show that the only component model that indicates a degradation is the SmLPC sensor. After the burn-in state, HI^m for this sensor is rising and exceeds its upper bound in the end. The same trend is also present in the aggregator where HI^j is approaching the upper-bound. The two alarms are published when SmLPC crosses its boundary.

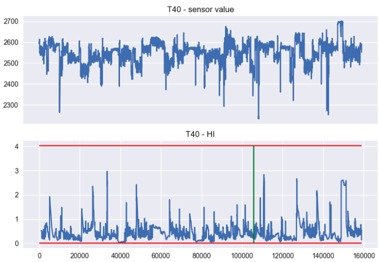
Test 5 was conducted to test ModularHI when a higher weight is applied to one measurement, and sensor T2 was given a higher weight than the other two sensors (SmHPC and SmLPC) in this test. The T2 sensor values and its HI score can be seen in Figure 9 along with the HI^j. Due to its higher weight, T2 highly impacts HI^j, moreover, multiple alarms are published.



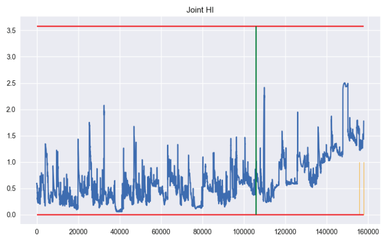
(a) SmHPC sensor value and HI.



(b) SmLPC sensor value and HI.



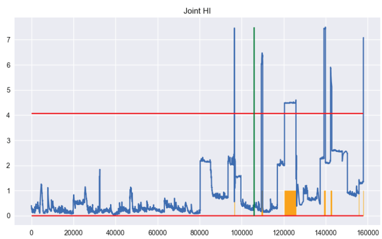
(c) T40 sensor value and HI.



(d) HI^j for SmHPC, SMLPC, and T40.

Figure 8: The result from test 3 of the N-CMAPSS dataset. The green vertical line indicates the end of the burn-in state. The horizontal red lines indicates the calculated boundaries. The orange area in the HI^j is when an alarm is published.

(a) T2 sensor value and HI.



(b) HI^j for SmHPC, SmLPC, and T2.

Figure 9: The result from test 5 of the N-CMAPSS dataset. The green vertical line indicates the end of the burn-in state. The horizontal red lines indicates the calculated boundaries. The orange area in the HI^j is when an alarm is published.

5. Discussion

The tests showed successful results on the original CMAPSS dataset. Each test was able to detect the degradation within the dataset and successfully publish an alarm at an appropriate time. ModularHI succeeded regardless of how many non-degradation sensors were included to complicate the endeavour. This must be partially attributed to the system design, as a single HI^m crossing its boundary is sufficient to publish an alarm. The tests on the more complex N-CMAPSS dataset were not as successful, as alarms only were published when a single sensor value crossed its boundary. Furthermore, sensor T2 triggered an alarm during Test 5 even though no clear degradation trend can be seen from that sensor.

The component models used in this paper are built around the idea that one model should learn the underlying characteristics of a specific sensor. We have pre-trained our models to speed up the training process, i.e., related to the data collection and learning in the burn-in state. All the univariate models are sharing the same DAE architecture (Table 1). A direction for future research is to evaluate the performance of different architectures when applied to univariate data, potentially reflecting what sensor the data originates from. Some sensors output multivariate data streams, and while we hypothesise that a DAE architecture similar to what is currently used for univariate data could also work for multivariate streams, this is so far not thoroughly tested. Similarly, specific DAE models should also be made for images, videos, and sound to handle those types of sensor inputs. Another issue is how to model correlated sensor values. Vibration can, for instance, be measured by both a 3 degree of freedom (DoF) and 6 DoF accelerometers. To detect increasing levels of vibration it could be beneficial to analyse readings from both sensors together. The current version of ModularHI has traded such fidelity in the model for simplicity, but the performance gains of combining correlated signals into component-overarching models should be investigated.

The aggregator handles all HI-values from the component models and decides when an alarm should be published. One specific challenge is that the main objective is that it should be data-driven. Therefore, the aggregator should internally work out its optimal behaviour during the burn-in state. Currently, the aggregator's HI value is a weighted average of those from the component models, and no computational work is done to find a better aggregation formula. Nevertheless, since we can use a *weighted* average, it is possible to get at least some expert knowledge into the system by tweaking the weights. If doing so, we take a step away from an entirely data-driven approach, and move towards a hybrid approach. Another idea is to find, e.g., a neural network that uses the component models' HI^m-values to generate HI^j. This would enable the aggregator to learn the different characteristics and dependency of all the components models and thus understand the connection between sensor values, but would increase the data-requirements to initialise the system.

Currently, the boundaries from both the component models and the aggregator are calculated by the 9'th standard deviation from the HI mean during the burn-in state. While this boundary was successful in the CMAPSS dataset as described in section 4.1, it was less successful for the N-CMAPSS dataset. This was partly because the DAE models used to calculate each component's HI value were not successful in faithfully reproducing the input-signal during the burn-in state. As a consequence, we saw a larger variability in HI^m, leading to rather large standard deviation during burn-in, and therefore to a less responsive behaviour for our monitoring system. A solution could be a locally adaptive boundary that is based on the historical HI scores and changes as each new score is calculated. This would make the boundaries dynamic, but potentially also prevent the detection of slowly developing problem situations.

When working with time-series data in neural networks, the way data is presented to the model is extremely important. We use a rolling window in our implementation, and fix the window-size based on the available data and the speed of the underlying dynamics. This choice, while difficult to make, can influence the results dramatically. If the window-size is too small the model cannot learn the temporal information, and if it is too-large the result can be a prolonged training time and a reduction in available data batches reducing the learning quality. We used a window size of 8 for the CMAPSS test, and 1024 the N-CMAPSS dataset, as the dataset contained more data and high variation throughout the dataset. To reduce the sensitivity wrt. this parameter one could try to define the window size dynamically based on the amount of data collected in the burn-in state along with other statistical characteristics such as standard deviation, skewness, etc.

While other implementations of HI split them up into multiple zones compared to the severity (see, e.g., [23]), we have chosen to have only one type of alarm. This was done to avoid scenarios where a system would fail in, e.g. warning state instead of in a critical state. Other researchers based their zoning on historical data, which is not assumed to be present in our case. Finally, it would be beneficial for the operators to have an RUL estimation, but to the best of our knowledge there is still no method available to produce reliable RUL estimates without historical data available.

To utilise ModularHI in practice, a few things are considered paramount to be in place. Firstly, each component model is assumed to be already pre-trained on relevant datasets. The operator will then place and select the sensor inputs through a GUI where he would define the stable period for the burn-in to occur. When the burn-in period is finished, the system will atomically go into the inference state. While we in this paper performed tests on the publicly available datasets CMAPSS and N-CMAPSS, they are not the ideal use case. ModularHI was designed with more generic manufacturing systems in mind, such as punching machines, drill presses, CNC and bending machines. These types of systems are more comparable to the CMAPSS dataset since the system is not as complex as the data represented in the N-CMAPSS dataset. Nonetheless, as it is now, ModularHI does not distinguish between different part numbers, e.i. ModularHI has no way of knowing if the machine is currently producing a part it was not present in the burn-in state.

6. Conclusion and future work

In this paper we presented ModularHI, a novel modular architecture for constructing a HI in a production process. Our work does not require any historical data, as the DAE is fitted during a stable process on-site. Moreover, the operators choose the sensors available/required, and ModularHI does not require any knowledge of the type of machinery. The system is data-driven, easy-to-use and modular, and these design-choices make ModularHI suitable for SMEs that do not have the resources to invest in PHM solutions that large enterprises do.

We tested our contribution on two open datasets, CMAPSS and N-CMAPSS. All the tests on the CMAPSS dataset successfully published an alarm and detected the degradation in the underlying process. This indicates our novel approach of using a combination multiple of individual trained DAEs aggregated together is able to produce an HI score and degradation detection without any prior historical data. This will benefit SMEs as it has been shown that they often lack the resources and knowledge to invest in large PHM solutions.

The test of the N-CMAPSS was not as successful in detecting the degradation as on the CMAPSS tests. This is because of the higher complexity within that dataset since it contains different flights of different lengths and altitudes; state changes that influence the data substantially. We tried to overcome this problem by limiting the data to cruise flights above 10,000 feet, but found that this was not enough to make the data sufficiently stable. One could try to limit this specific dataset to have different models for the different stages of flights. Nonetheless, this is out of scope for ModularHI since it is focused on a more "stable" system, which does not change extensively throughout its execution.

For future work, ModularHI should be tested on real production data with various sensors to validate it in its intended settings. Moreover, tests should be carried out with both images and sound as the component models. More research should be conducted on each component model to ensure the correct architecture for each model. A more automatic system for choosing the window size should also be researched to make the system more easy-to-use. Lastly, more research within the aggregator should be conducted to see if it can learn more from the component models instead of aggregating the results as it is now.

Acknowledgements

The work was conducted as a part of the Danish strategy for Industry 4.0 in SMEs called Innovation Factory North.

References

[1] I. Gurkov, Dominant strategic archetype of the Russian industrial firm, JEEMS Journal of East European Management Studies 12 (2007) 35–52.

[2] J. Abrhám, W. Strielkowski, M. Vošta, J. Šlajs, Factors that influence the competitiveness of Czech rural SMEs, Agricultural Economics 61 (2015) 450–460.

[3] V. Atamuradov, K. Medjaher, P. Dersin, B. Lamoureux, N. Zerhouni, Prognostics and health management for maintenance practitioners - review, implementation and tools evaluation, International Journal of Prognostics and Health Management 8 (2017) 1–31.

[4] V. Atamuradov, K. Medjaher, F. Camci, N. Zerhouni, P. Dersin, B. Lamoureux, Machine health indicator construction framework for failure diagnostics and prognostics, Journal of Signal Processing Systems 92 (2020) 591–609.

[5] J. Zhu, T. Nostrand, C. Spiegel, B. Morton, Survey of condition indicators for condition monitoring systems, in: Annual Conference of the PHM Society, volume 6, 2014.

[6] B. Luo, H. Wang, H. Liu, B. Li, F. Peng, Early fault detection of machine tools based on deep learning and dynamic identification, IEEE Transactions on Industrial Electronics 66 (2019) 509–518.

[7] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, R. X. Gao, Deep learning and its applications to machine health monitoring, Mechanical Systems and Signal Processing 115 (2019) 213–237.

[8] L. Guo, Y. Lei, N. Li, T. Yan, N. Li, Machinery health indicator construction based on convolutional neural networks considering trend burr, Neurocomputing 292 (2018) 142–150.

[9] Z. Li, D. Wu, C. Hu, J. Terpenny, An ensemble learning-based prognostic approach with degradation-dependent weights for remaining useful life prediction, Reliability Engineering & System Safety 184 (2019) 110 – 122. Impact of Prognostics and Health Management in Systems Reliability and Maintenance Planning.

[10] T. Berghout, L. Mouss, O. Kadri, L. Saïdi, M. Benbouzid, Aircraft engines remaining useful life prediction with an adaptive denoising online sequential extreme learning machine, Engineering Applications of Artificial Intelligence 96 (2020) 103936.

[11] A. Saxena, K. Goebel, D. Simon, N. Eklund, Damage propagation modeling for aircraft engine run-to-failure simulation, in: 2008 International Conference on Prognostics and Health Management, 2008, pp. 1–9. doi:https://doi.org/10.1109/PHM.2008.4711414.

[12] A. L. Ellefsen, E. Bjørlykhaug, V. Æsøy, S. Ushakov, H. Zhang, Remaining useful life predictions for turbofan engine degradation using semi-supervised deep architecture, Reliability Engineering & System Safety 183 (2019) 240–251.

[13] K. Hollingsworth, K. Rouse, J. Cho, A. Harris, M. Sartipi, S. Sozer, B. Enevoldson, Energy anomaly detection with forecasting and deep learning, in: 2018 IEEE International Conference on Big Data (Big Data), 2018, pp. 4921–4925. doi:https://doi.org/10.1109/BigData.2018.8621948.

[14] E. B. Hansen, S. Bøgh, Artificial intelligence and internet of things in small and medium-sized enterprises: A survey, Journal of Manufacturing Systems 58 (2021) 362–372. Digital Twin towards Smart Manufacturing and Industry 4.0.

[15] E. B. Hansen, N. Iftikhar, S. Bøgh, Concept of easy-to-use versatile artificial intelligence in industrial small & medium-sized enterprises, Procedia Manufacturing 51 (2020) 1146–1152. 30th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2021).

[16] C. Fernando, D. Banarse, C. Blundell, Y. Zwols, D. Ha, A. A. Rusu, A. Pritzel, D. Wierstra, PathNet: Evolution channels gradient descent in super neural networks, 2017. arXiv:1701.08734.

[17] S. Imai, S. Kawai, H. Nobuhara, Stepwise PathNet: a layer-by-layer knowledge-selection-based transfer learning algorithm, Scientific Reports 10 (2020) 8132.

[18] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He, A comprehensive survey on transfer learning, Proceedings of the IEEE 109 (2021) 43–76.

[19] H. Purohit, R. Tanabe, K. Ichige, T. Endo, Y. Nikaido, K. Suefusa, Y. Kawaguchi, MIMII Dataset: Sound Dataset for Malfunctioning Industrial Machine Investigation and Inspection, 2019. URL: https://doi.org/10.5281/zenodo.3384388. doi:10.5281/zenodo.3384388.

[20] E. B. Hansen, E. R. van der Bijl, M. B. Nielsen, M. Svangren Bodilsen, S. V. Berg, J. Kristensen, N. Iftikhar, S. Bøgh, A new authentic cloud dataset from a production facility for anomaly detection, in: Towards Sustainable Customization: Bridging Smart Products and Manufacturing Systems, Springer International Publishing, Cham, 2022, pp. 415–422.

[21] A. Saxena, K. Goebel, Turbofan engine degradation simulation data set, NASA Ames Prognostics Data Repository (2008) 1551–3203.

[22] M. Arias Chao, C. Kulkarni, K. Goebel, O. Fink, Aircraft engine run-to-failure dataset under real flight conditions for prognostics and diagnostics, Data 6 (2021).

[23] K. Velmurugan, P. Venkumar, P. R. Sudhakara, SME 4.0: Machine learning framework for real-time machine health monitoring system, Journal of Physics: Conference Series 1911 (2021) 012026.

Paper E

Artificial intelligence and machine learning

Emil Blixt Hansen and Simon Bøgh

The paper has been published in the book
The Future of Smart Production for SMEs pp. 323–326, 2023.

Artificial Intelligence and Machine Learning

Emil Blixt Hansen and Simon Bøgh

Key words: Artificial intelligence, Machine learning, Small and medium sized enterprise

1 Introduction

With the ever-expanding digitalisation, more information or data is generated and is available within the digital ecosystem. The expansion of available data and the increased competition caused by globalisation contribute to why manufacturers are looking for more advanced methods to optimise their production and products. The general usage of artificial intelligence (AI) within different fields is expanding. As part of Industry 4.0, AI is also gaining interest within the industrial sector, where companies are expanding and trying different usages of AI, both within their production and as a product or service. Nonetheless, the term *artificial intelligence* is ambiguous and is a collection of many different methods and fields of statistics, data- and computer science. When the term AI is used as a tool, often it refers to *artificial narrow intelligence* (ANI), which indicates a specific field or problem an AI is applied to. Whereas the term *artificial general intelligence* (AGI) refers to an AI which succeeds in multiple fields and is closer to human beings type of intelligence. Examples of ANI is voice assistants like Google Assistant and Siri and image recognition like Google Lens. Moreover, AI's like DeepMind's AlphaGo is also an ANI. As of this paper, there exist no examples of an AGI yet. Thus, we will only refer to ANI when we use AI from here on.

Emil Blixt Hansen
Material and Production, Aalborg University, Denmark e-mail: ebh@mp.aau.dk

Simon Bøgh
Material and Production, Aalborg University, Denmark e-mail: sb@mp.aau.dk

AI is often utilised fields such as robotics, planning, computer vision, natural language processing, expert system etc. The tool used to solve the challenges within these fields is often machine learning and deep learning in recent years. Deep learning is a subset of machine learning which utilises artificial neural networks to learn a given problem. To apply machine learning to a problem generally, three components is required:

- A decision process
- An error function
- An optimiser

The *decision process* is a set of calculations where the algorithm makes a best guess on what the output should be. The *error function* (also known as loss function) then calculates the error between the guess and the correct answer if present. At last, the *optimiser* minimises this error function to improve the decision process; and then the cycle continues until stopped.

Depending upon the task, machine learning can be split into three different paradigms. The three paradigms are supervised learning, unsupervised learning and reinforcement learning.

Supervised learning:
It is used when the problem has a specific and known class or outcome, often referred to as a label. This can, e.g. be image recognition and classification. Moreover, it can also be used to solve regression problems such as predicting future sales.

Unsupervised learning:
Unsupervised can be used to find unknown patterns in a dataset and to reduce the dimension of a large dataset. Moreover, it can be used to detect faults e.g. a machinery through methods such as autoencoders.

Reinforcement learning:
It is often used where a more dynamic approach is suited, where, e.g. labels can be hard to obtain. The reinforcement agent learns from the action it takes within the environment it is deployed. The training of an agent often requires a simulation environment since the agent learns through the means of trial and error. Reinforcement learning can be beneficial in cases of planning where the environment is more dynamic.

In general, it is required to either have a substantial amount of data or the ability to create/collect it to produce a good result with machine learning. This is especially true when dealing with the best-performing algorithms and models. Nonetheless, since this field is seeing a rapid involvement, new methods focusing on how different ways to optimise the algorithms, both for shorter training time, more energy-efficient and dedicated hardware (Reuther et al., 2019). Companies are also releasing either products with build in AI and machine learning or having it as a part of their toolbox for manufacturers, such as Microsoft Azure.

2 The use of AI in SMEs

In 2021 a survey found a low adoption of AI in SMEs, where only five publications utilised AI in SMEs related to the manufacturing industry (Hansen and Bøgh, 2021). It showed that all though limited, that the SMEs was more focused on the internet of things (IoT), cloud solutions and relevant business opportunities. One of the reasons behind the higher adoption of IoT and cloud solutions is that they are easy to use (Moeuf et al., 2020). This also indicates that SMEs, in general, lack the knowledge and resources to use AI themselves; however, researchers are researching methods to make AI more easy-to-use, applicable and tailored towards manufacturing companies, including SMEs.

Within an SME, there are numerous ways AI can bring value (Watney and Auer, 2021). Some of them are predictive maintenance, resource optimisation, quality control, and logistics. Which subfield of AI is used depends on whether it is predictive maintenance, logistics or one of the other problems AI is applied on. This also leads to one of the shortcomings of AI, e.i. the knowledge required to have a successful integration. Normally AI projects are engineered to the specific task, either in-house or by a consulting company. Which further indicates the challenges and knowledge required for implementation. Therefore, the problem is more present in SMEs where, in general, the expertise and resources is a constraint (Welte et al., 2020). Henceforth, pilot projects are a way for SMEs to get started with AI, either in collaboration with a research institution or an AI consultancy company. Thereafter, the SMEs have gained knowledge of AI in connection with their business. Then can start to build up their own internal AI specialist or group (Ng, 2020).

3 Use cases of AI

As stated in the previous section that the adoption of AI in SMEs is low, it is possible to look at large enterprises to understand how AI can be used. A survey from Brosset et al. (2019) describe how Bridgestone uses an AI to tune its tire production based upon 480 sensor inputs. It resulted in a 15% more uniformity tire production. Moreover, Nokia implemented a camera surveillance system of the production line, which monitors the production, and if any inconsistency occurs, an operator is alerted. It is not only the manufacturing process AI is used within; Kellogg's has an AI which aids consumers in selecting a recipe they want. Also, logistics and supply chain, along with inventory management, have examples of AI use cases. Some of these use AI to optimise the orders and planning, while also examples of communication with partners with the help of natural language processing.

Charalambous et al. (2019) from McKinsey analytics also published an article where a cement manufacture improved their process. The problem was that the operators had a lot of expertise, and when they went on retirement, it was complicated to replace them with new operators. Therefore they applied AI to control the different processes, and thus the expert knowledge is not lost when an operator retires.

A study from 2021 also showed a use case where an SME wanted to have an *remaining useful life* (RUL) estimation of a critical component (Iftikhar and Nordbjerg, 2022). The study came with different suggestions to SMEs and how to adapt machine learning within their company. The suggestions included the need to join an alliance with non-competing SMEs, consultancies and research institutions in a test-driven environment to enhance the knowledge of the field. Moreover, they also suggested starting with the areas where the problem is solvable and where it is most suitable from a cost perspective.

References

Brosset P, Jain A, Khemka Y, Buvat J, Thieullent A, Khadikar A, Patsko S (2019) Scaling ai in manufacturing operations: A practitioners' perspective. Capgemini Research Institute 10

Charalambous E, Feldmann R, Richter G, Schmitz C (2019) Ai in production: A game changer for manufacturers with heavy assets. McKinsey Analytics

Hansen EB, Bøgh S (2021) Artificial intelligence and internet of things in small and medium-sized enterprises: A survey. Journal of Manufacturing Systems 58:362–372, DOI https://doi.org/10.1016/j.jmsy.2020.08.009, URL https://www.sciencedirect.com/science/article/pii/S0278612520301424, digital Twin towards Smart Manufacturing and Industry 4.0

Iftikhar N, Nordbjerg FE (2022) Implementing machine learning in small and medium-sized manufacturing enterprises. In: Andersen AL, Andersen R, Brunoe TD, Larsen MSS, Nielsen K, Napoleone A, Kjeldgaard S (eds) Towards Sustainable Customization: Bridging Smart Products and Manufacturing Systems, Springer International Publishing, Cham, pp 448–456

Moeuf A, Lamouri S, Pellerin R, Tamayo-Giraldo S, Tobon-Valencia E, Eburdy R (2020) Identification of critical success factors, risks and opportunities of industry 4.0 in smes. International Journal of Production Research 58(5):1384–1400, DOI 10.1080/00207543.2019.1636323, URL https://doi.org/10.1080/00207543.2019.1636323, https://doi.org/10.1080/00207543.2019.1636323

Ng A (2020) Ai transformation playbook. URL https://landing.ai/wp-content/uploads/2020/05/LandingAI_Transformation_Playbook_11-19.pdf

Reuther A, Michaleas P, Jones M, Gadepally V, Samsi S, Kepner J (2019) Survey and benchmarking of machine learning accelerators. In: 2019 IEEE High Performance Extreme Computing Conference (HPEC), pp 1–9, DOI 10.1109/HPEC.2019.8916327

Watney C, Auer D (2021) Encouraging ai adoption by eu smes. URL https://www.progressivepolicy.org/wp-content/uploads/2021/03/PPI_Encouraging-AI-adoption-by-EU-SMEs-3.24.21-2.pdf

Welte R, Estler M, Lucke D (2020) A method for implementation of machine learning solutions for predictive maintenance in small and medium sized enterprises. Procedia CIRP 93:909–914, DOI https://doi.org/10.1016/j.procir.2020.04.052, URL https://www.sciencedirect.com/science/article/pii/S2212827120306223, 53rd CIRP Conference on Manufacturing Systems 2020

Paper F

On the topic of anonymising production data for machine learning

Emil Blixt Hansen, Nadeem Iftikhar, and Simon Bøgh

The paper has been submitted to and is in second review in
Journal of Intelligent Manufacturing, submitted April 2022.

In peer review

Paper G

An in-depth investigation of machine learning and IoT adoption at a manufacturing SME: A field study

Emil Blixt Hansen, Kent Lindegaard Hansen, Nadeem Iftikhar and Simon Bøgh

The paper has been submitted to
32nd International Conference on Flexible Automation and Intelligent Manufacturing (FAIM), submitted October 2022.

In peer review

ISSN (online): 2446-1636
ISBN (online): 978-87-7573-793-2

AALBORG UNIVERSITY PRESS