Aalborg Universitet

A New Model to Identify the Reliability and Trust of Internet Banking Users Using Fuzzy Theory and Data-Mining

Bekamiri, Hamid; Ghasempour Ganji , Seyedeh Fatemeh ; Simonetti , Biagio ; Hosseini Seno, Seyed Amin

Article

A New Model to Identify the Reliability and Trust of Internet Banking Users Using Fuzzy Theory and Data-Mining

Hamid Bekamiri [1], Seyedeh Fatemeh Ghasempour Ganji [2], Biagio Simonetti [3,4,5,*] and Seyed Amin Hosseini Seno [6]

1   Business School Department, Aalborg University, Fredrik Bajers Vej 7K, 9220 Aalborg East, Denmark; Hamidb@business.aau.dk
2   Department of Management, Ferdowsi University of Mashhad, Azadi Square 9177948974, Iran; fa.ghasempour@mail.um.ac.ir
3   Department of Law, Economics, Management and Quantitative Methods, University of Sannio, 82100 Benevento, Italy
4   WSB University, 80266 Gdansk, Poland
5   National Institute of Geophysics and Volcanology, INGV, 80124 Rome, Italy
6   Department of Computer Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Azadi Square 9177948974, Iran; hosseini@um.ac.ir
*   Correspondence: simonetti@unisannio.it

Abstract: As a result of changes in approach from traditional to virtual banking system, security in data exchange has become more important; thus, it seems essentially necessary to present a pattern based on smart models in order to reduce fraud in this field. A new algorithm has been provided in this article to improve security and to specify the limits of giving special services to Internet banking users in order to pave appropriate ground for virtual banking. In addition to identifying behavioral models of customers, this algorithm compares the behaviors of any customer with this model and finally computes the rate of trust in customer's behavior. The hybrid data-mining and knowledge based structure has been adapted in this algorithm according to fuzzy systems. In this research, qualitative data was gathered from interviews with banking experts, analyzed by Expert Choice to identify the most important variables of customer behavior analysis, and to analyze customer behavior and customer bank Internet transaction data for a period of one year by MATLAB and Clementine. The results of this survey indicate that the potential of the given structure to recognize the rate of trust in Internet bank user's behavior might be at reasonable level for experts in this area.

Keywords: security; data-mining; fuzzy system; behavior of users

1. Introduction

Due to the rapid pace of Internet accessibility, the use of e-banking has also boosted [1]. Furthermore, the progress in electronic banking and mobile banking leads to the rapid development of real-time banking information [2]. Rising attractiveness of Internet banking for the customers caused ever-increasing growth in doing online transactions. For example, the number of online-shopping by using debit and credit cards has almost doubled in the previous 10 years, from 6.7 billion in 2006, to 16.4 billion purchases in 2016 [3]. In UK, debit cards are the most common payment method increased by 12 percent reaching 17.0 billion in 2019 [4]. In Iran, due to Central Bank of Iran's report, 25% growth in the number of e-banking transactions (20.4 billion transactions) and 27% increase in value terms (177 billion dollars) is shown in operations of key electronic banking platforms during March–September 2019 in comparison with the same period in the previous year [5].

This progress has excessively increased the rate of occurrence of financial crimes in e-banking so that the growing trend of such crimes shows approximately 8 to 9% annual increase. In this regard, it can be referred to the resulted financial losses incurred by British banks which was over GPB 22 million in 2007, although the published figures by UK Cards

Association indicates the growing trend in the given figures of financial loss in these banks has actually increased 14% in 2010 compared to the previous year. Of course, if we add the undeclared statistics from financial institutions to these figures, the rate of this growth will be further [3]. Thus, nowadays online fraud has become more complicated in Internet banking and the security and trust has been more seriously addressed in e-banking [6]. Due to European Union Agency for Fundamental Rights [7] a one per four Europeans (24%) feels a huge amount of worries regards to misuse of their online bank account or payment card details by fraudsters or criminals. This rate is approximately 57 percent in Spain. These statistics indicate the highly importance of e-banking fraud detection systems.

Fraud in online banking sector is always growing and it is also challenging to investigate and detect since the fraudulent behavior is dynamic, vary through diverse customer profiles and present in extremely large and dynamic datasets [8]. Research investigating the customer behavior and evaluating the e-banking users' reliability is scare in the context of Iranian financial institutions [9,10]. Although employing of fuzzy system as one of the supervised learning tool and K-means as the unsupervised learning tool for fraud detection shown to have high fraud coverage and accuracy, there are few research in using these tools for fraud detection in literature [8]. Moreover, a major part of these studies use these tools separately, e.g., [11–13], however, we propose the algorithm which combines these two tools for detecting fraud in customers' behavior.

The current research aims to propose a smart system in order to determine the rate of trust in behavior of Internet banking users thereby it would be possible to detect online banking fraud quicker and define constraint for those unreliable customers in giving banking services. Therefore, the major question which may be raised in current research is that how banks can determine the rate of reliability of each customer in terms of future fraud behavior by means of smart algorithm based on the previous behavior of customers. The paper is organized as follows. In Section 1, the literature review is provided based on practical studies in the field of Internet fraud discovery systems in banking industry. In Section 2, we provide some details regards to research methodology and discuss the research implementation process. In Section 3, we provided data analysis in two phases of model preparation and model testing. Next, in Section 4, the finding of research was represented. Next, in Section 5, we provide some discussion relied on findings section. Finally, in Section 6, the research limitations, directions for further research, and the conclusions are discussed.

2. Literature Review

2.1. Internet Fraud Detection Systems

Fraud detection refers to some actions or methods to discover the frauds which have been occurred or are being occurred [1]. Online fraud has imposed financial burdens on the banks during recent decade while the efficient and effective systems for identifying this type of crimes have not become so far applied perfectly [14]. Therefore, this issue has been turned into the major concern for banking directors. The main subject in this regard is to prevent fraud with respect to financial and credit nature of transactions [15–19].

Integrated use of virtual and social information sources has caused creating a new approach in Internet fraud and phishing so that it could add the complexity in these fraud techniques. The international statistics and reports show that some numbers of cybercrimes follow excessive incremental trend in banking field [6,20,21], while analysis on the given models used in fraud discovery systems in Internet banking has indicated this type of systems possesses weak performance with lower accuracy [8,17].

Financial institutions are mostly in search of the required speed to discover the behaviors and operations of fraudsters. This problem is so crucial due to its indirect impact on customer service in financial institutions, declined operational costs due to providing trusted and valid monetary services. The most applicable algorithms employed for online banking fraud detection are regularly executed across the analysis of customer information such as their transactions [1]. Due to the dispersion of information resources of customers

and unstructured nature of this information, limited techniques have been suggested for dynamic identifying of online fraud. This issue has converted presentation of efficient and immediate structures in discovery of fraud into one of the foremost challenges existing in this area [6]. The reason for poor performance of fraud detection systems in Internet banking is related to dynamic nature of customers' behavior and lack of access to the fraud occurrence data in this area. For instance, recognition of fraud in credit and telecommunication cards may usually emphasize in identifying behavioral models of groups of customers while fraud detection is very dynamic in customers of Internet banking and in line with behaviors of actual customers [17]. The created challenges put emphasis on the fact that the classic methods based on real data in Internet banking may not be highly accurate techniques for detection of all types of Internet banking fraud. In addition, research has shown that the use of modern analysis is a major problem due to its very high cost in big data and its very low error detection power [22]. Accordingly, as well as the results of research, the use of this method is not recommended [23].

2.2. Challenges through Online Banking Fraud Detection Systems

There are some characteristics and challenges regards to e-banking fraud detection explained in literature. Firstly, the dataset of e-banking transactions is large and highly imbalanced. Secondly, fraud detection should be real time to generate fraud alert and detect prompt money loss. Thirdly, the fraud behavior is dynamic. Finally, the pattern of customer behavior is diverse, so fraudsters try to behave similar to genuine customer which makes the fraud detection process so complicated [8]. These characteristics change the fraud detection into extremely challenging issues, which motivates researchers to employ different machine learning and data mining techniques to detect fraud in banking system [24].

Many essays and studies have been prepared concerning recognition of Internet banking fraud, e.g., [25]. The structure of most of these studies is based on comments of experts and use of base of rules to recognize suspicious transactions for which one can refer to the studies conducted by Karlsen and Killingberg [17], Bignell [26], as well as Kovach and Ruggiero [27]. Review on performance of these systems shows that the outputs lack high precision. The noticeable point existing in these studies is concerned with the relationship among performance of system and its compliance with rules and the phishing states which highly depend on the adequate expert knowledge. One of the problems and challenges with which these systems may be exposed is low tolerance level of system and non-creation of dynamic structure for immediate compliance with critical state.

It is noteworthy that many given fraud behaviors in Internet banking indicate this fact that the fraud cases are done independently, and sequence of fraud events may be very improbable in this sense. However, the method of identifying operational sequences in behaviors of customers in real world situation has been employed in some studies to discover fraudulent behaviors. For example, in the given study conducted by Fan et al. [28], the behavior of an actual customer with a Trojan (virus) has been explored. In this study, structure of events done by Trojan and actual user has been assessed with different trends.

Another challenge facing the design and implementation of fraud detection systems is the difficulty of accessing labeled data in Internet banking due to the security constraints of organizations. Therefore, one of the common methods for creation of labeled data regarding forgery and fraud in Internet banking is using historical data and adding knowledge from experts to these data. The unsupervised events possess shorter information gain than the supervised approaches and have potential to identify this approach compared to the supervised approaches, e.g., neural networks and decision trees at lower level [29,30].

It has been suggested that a design of offline system for identifying fraud in Internet banking should be designed in order to verify security and health of financial transactions. It has been proposed in another study to install a module on system of Internet banking users so that data exchange can be more secure. The most important challenge of this solution is the development problems of this system [27,29].

Some other articles are based on the fraud discovery systems developed practically in Internet banking based on knowledge-based database of expert rules. As a result, this type of systems has potential for diagnosis at low level and they may not be often capable to adjust to new types of fraud [28].

To meet these challenges, we conduct a survey considering a mixture of different methods to reduce the errors. This study aims to propose a smart system in order to determine rate of trust in behavior of Internet banking users in three general steps including determination of input parameters, optimization of fuzzy rules, and implementation of fuzzy system.

2.3. Previous Literature

New statistical techniques and smart algorithms have been suggested for discovery and identifying of fraud inside databases; for example, in some of studies, e.g., [15,31], neural networks and decision tree were utilized to disclose fraud and in some others, e.g., [32–34], this structure has been implemented according to logistic regression. It has been tried in many studies, e.g., [35,36], to present an expert system based on database of rules. These models have been employed regarding the subjects, e.g., money-laundering, forgery of credit cards, accounting frauds, and cyber tourism [34,36,37].

Recently, some researchers have highlighted the importance of fuzzy based data mining approaches, e.g., [38,39]. Aburrous et al. [40] proposed the model, combining fuzzy logic and data mining algorithms to classifying the phishing website types and categorizing six e-banking phishing website attack strategies with a layer structure. They proposed fuzzy data mining approach as an effective tool in identifying e-banking phishing websites as this technique offers a way for considering quality factors in an ambiguous process of phishing detection. In terms of considering user behaviors, Atta-ur-Rahman et al. [41] proposed a neuro-fuzzy approach to predict user behavior, considering users' temporal logs including local machine, network, and web usage logs. In terms of e-banking, Rezaiee and Karimi [42] proposed a model of determining validity and reliability of e-banking users by using Fuzzy C-Mean model. Eshghi and Kargari [29] employed multi criteria decision methods, intuitionistic fuzzy set, and evidential reasoning to proposed a new algorithm for fraud detection.

3. Methodology

The methodology of this research is a descriptive-exploratory data mining, employing quantitative and qualitative research methodology.

The interviews and datasets were conducted by the cooperation of Agri-Bank central center located in Tehran. Agri-Bank or Bank Keshavarzi of Iran, established in 1993, is one of the major Iranian bank. Agribank known as the only specialized bank serving financial services in the agricultural sector. This bank has 1817 branches in Iran and provide financial services and credits to almost 70% of the agricultural sector to facilities its development. As a bank with more than eight decades of banking experience, Agri-Bank has been a forerunner in applying the advanced technologies like electronic banking terminal, setting up of the core-banking system, and providing a wholly automatic system of inter-branch data transfer.

To analyze data, we have used qualitative and quantitative methodology. Rather than conducting librarian studies to design of expert fuzzy system, some other methods such as interview with the experts as well as data-mining techniques and Analytical-Hierarchy Process (AHP) were used. The process for research implementation has been given in Figure 1. The rate of accuracy of fuzzy system has been increased by using data-mining in this process.
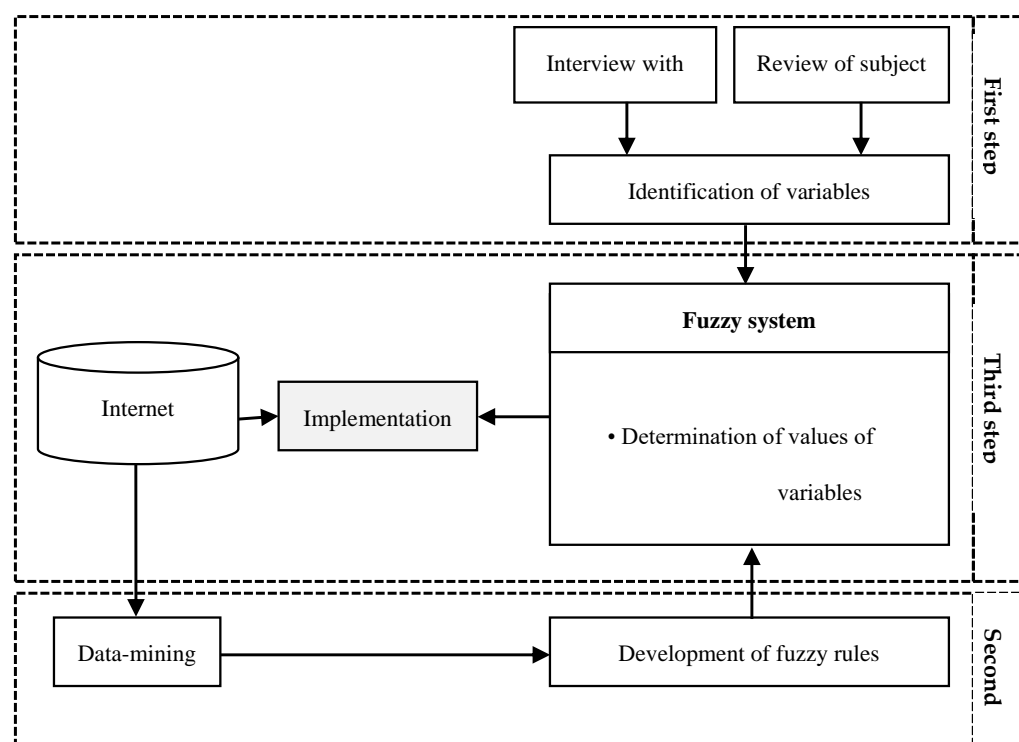
Figure 1. Research implementation process.

The analytical parts of the paper are divided into two steps as follow:

Step 1: Determination of input parameters: On this step, based on analysis of literature, 10 input variables for fraud detection derived from customer behavior were extracted and they were ranked by means of Freidman test and AHP analysis. Data was provided by expert opinion and analyzed by Expert Choice and SPP.

Step 2: Designing the expert fuzzy system: At this stage, according to quantitative method, we use Agricultural Bank annual customer Internet banking data, employing clustering algorism K-Mean by MATLAB. In this phase, data mining and clustering were used to determine the linguistic expressions or linguistic values of each of the variables. It is composed of three main categories, i.e., 'inputs', 'base of rules', and 'outputs' through three phases.

Phase 1: Optimization of fuzzy rules: During this step, the values and intervals of variables in fuzzy system determine based on actual values of variables in Internet banking database and by means of data mining process and K-Means clustering were analyzed.

Phase 2: Architecture of expert system in determination of the rate of trust in behavior of Internet banking users: In this step, the expert fuzzy system was implemented by using MATLAB software and data-mining calculations by means of Clementine software.

Phase 3: Inputs of fuzzy system: Presentation of electronic dynamic behavior algorithm: The expert system deduces each customer's categories based on customers' states in each parameter provided in the previous steps.

Moreover, to test the proposed model, we provided rule based fuzzy system and categorized the customers based on their reliability. Each of these steps were fully described in Analysis section (Section 4).

4. Data Analysis

4.1. Model Preparation

First step: Determination of input parameters According to qualitative method, we interviewed 30 experts in E-banking from Agri-Bank to identify the most important variables of customer behavior analysis. Interviews were analyzed by Analytical-Hierarchy Process (AHP) using Expert choice software. In this phase, AHP was used to rank the vari-

ables and select the most important ones. Among respondents, 23% have BA degree, 70% have MA, and 7% have PhD degree. Following to interview analysis, 10 input variables were extracted.

To select the most important variables, we have designed a questionnaire filled by 100 banking experts. Thus, these data were ranked by means of Freidman test proportional. The conducted analysis is given in Table 1.

Table 1. Ranking of extracted variables (source: research analysis).

Variable	Mean Rank	Variance Coefficient
Number of the Token Errors	6.67	0.923
Number of Card Code Errors to Log in Given System	6.48	0.983
Number of Used Browsers by User.	6.22	0.964
Time Use of Internet Errors	6.08	1.074
Number of Transfers Within Specific Time Interval	4.97	1.159
No Change in User's Code Within Long Time Interval	4.35	1.196
The Amount of Transfers Within Specific Time Interval	4.98	1.269
Number of Different Ips Used via Internet Banking	5.43	1.192
Period of Using Internet Banking as an Active Member	4.75	1.324
Number of Account Deficit Errors	5.07	1.269

After ranking of efficient variables in determination of rate of trust in behavior of Internet banking by means of Freidman test, 6 main variables were selected using fuzzy system (Table 2). The variables were separated into two general classes proportional to nature of these variables and comments of experts.

Table 2. Values and intervals of main variables in fuzzy system (source: research analysis).

Row	Type of Variable	Title of Variable	Time Interval	Symbol	Description
1	User Behavior	Change in Geographical Situation	Last Login	IP	This Variable is Assessed Proportional to the Usual and the Last Geographical Situations Entered in to The System During Recent 24 h
2		Change in Geographical Situation	Throughout the Period	Browser	The Most Frequently Used Browser by User To Login
3		Change in Current Browser	Throughout the Period	Time	A Day and Night is Divided into Three Intervals
4	User Security Information	Card Code Error	At Any Transaction	Card Code Error	The Number of Times User Made Card Code Errors
5		Account Deficit Error	At Any Transaction	Account Deficit Error	Number of Times User Has Made Account Deficit Errors
6		Token Error	When Login	Token Error	Number of Ties User Has Made Token Errors
7	User's Behavior		Monthly	Output	Status of Trust in User's Behavior

1. Variables for change of behavior in use of electronic services:
   - Variations in geographical situation based on IP.
   - The used browsers in Internet banking.
   - The used times for Internet banking.

2. The variables concerning confidential account information:
   - Number of token errors.
   - Number of account deficit errors.
   - Number of card code errors.

Six selected variables were ranked again by AHP method based on the 30 experts' opinions. This is because the AHP method does not require a statistically significant (large) sample size to get statistically robust results [43,44], The given analysis was conducted using Expert Choice software and the results are presented in Figure 2.
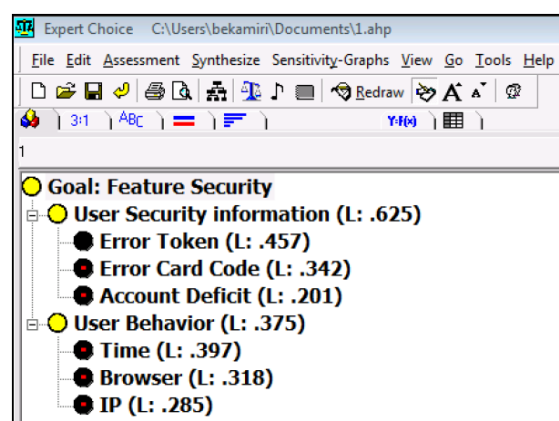

Figure 2. Expert Choice results (source: research analysis).

According to Figure 2, the weight of user security information is estimated 0.625, compared with the weight of user behavior which is 0.397.

Second step: Designing the expert fuzzy system. This step is composed of three main categories, i.e., 'inputs', 'base of rules', and 'outputs' through three phases. Phase 1: Optimization of fuzzy rules. The lingual terms were identified for any variable within intervals of 3–5 terms after recognizing the efficient variables in fuzzy system and then the values and ranges of these terms were determined by questionnaire and based on comments from experts.

With respect to specific relationship among accuracy of fuzzy system and the values and intervals of variables, the absolute reliance on experts' comment may be usually followed by human errors in determination of intervals and their values. Therefore, given the importance of appropriate selection of intervals for variables in implementation of more accurate fuzzy system, it has been suggested in this study to select the intervals based on actual values of variables in Internet banking database and by means of data mining process and K-Means clustering analysis. Thus, employing of clustering technique for determination of lingual intervals in fuzzy system is assumed as one of the innovations in the aforesaid study. It should be noted that with respect to nature of variables of confidential account information, these variables were extracted for all customers by query in SQL-Server in Internet banking database.

The outputs resulting from clustering of confidential account information account are given by means of operation using Clementine software as follows (Tables 3–8).

Table 3. Clustering of variable of token error (source: research analysis).

	Cluster No		
	1	2	3
Card code error	5	1	3

Table 4. Number of customers at any cluster as variable of second error (source: research analysis).

		Number of Customers
	1	4
Cluster no	2	237
	3	36
Number of Valid Data		277
Number of Invalid Data		0

Table 5. Clustering of account deficit error (source: research analysis).

	Number of Cluster			
	1	2	3	4
Number of Account Deficit Error	1	10	3	5

Table 6. Number of customers at any cluster, account deficit error (source: research analysis).

		Number of Customers
	1	59
Cluster no	2	1
	3	2
	4	1
Number of Valid Data		63
Number of Invalid Data		0

Table 7. Clustering variable of card code error (source: research analysis).

	Number of Cluster			
	1	2	3	4
Number of Token Errors	3	10	5	70

Table 8. Number of customers at any cluster, card code (source: research analysis).

		Number of Customers
	1	16,163
Cluster no	2	16
	3	1115
	4	1
Number of Valid Data		17,295
Number of Invalid Data		0

After collection of the related data to any variable, the given output was determined in form of triangular fuzzy numbers as the system input values where the results are displayed in Figure 3.
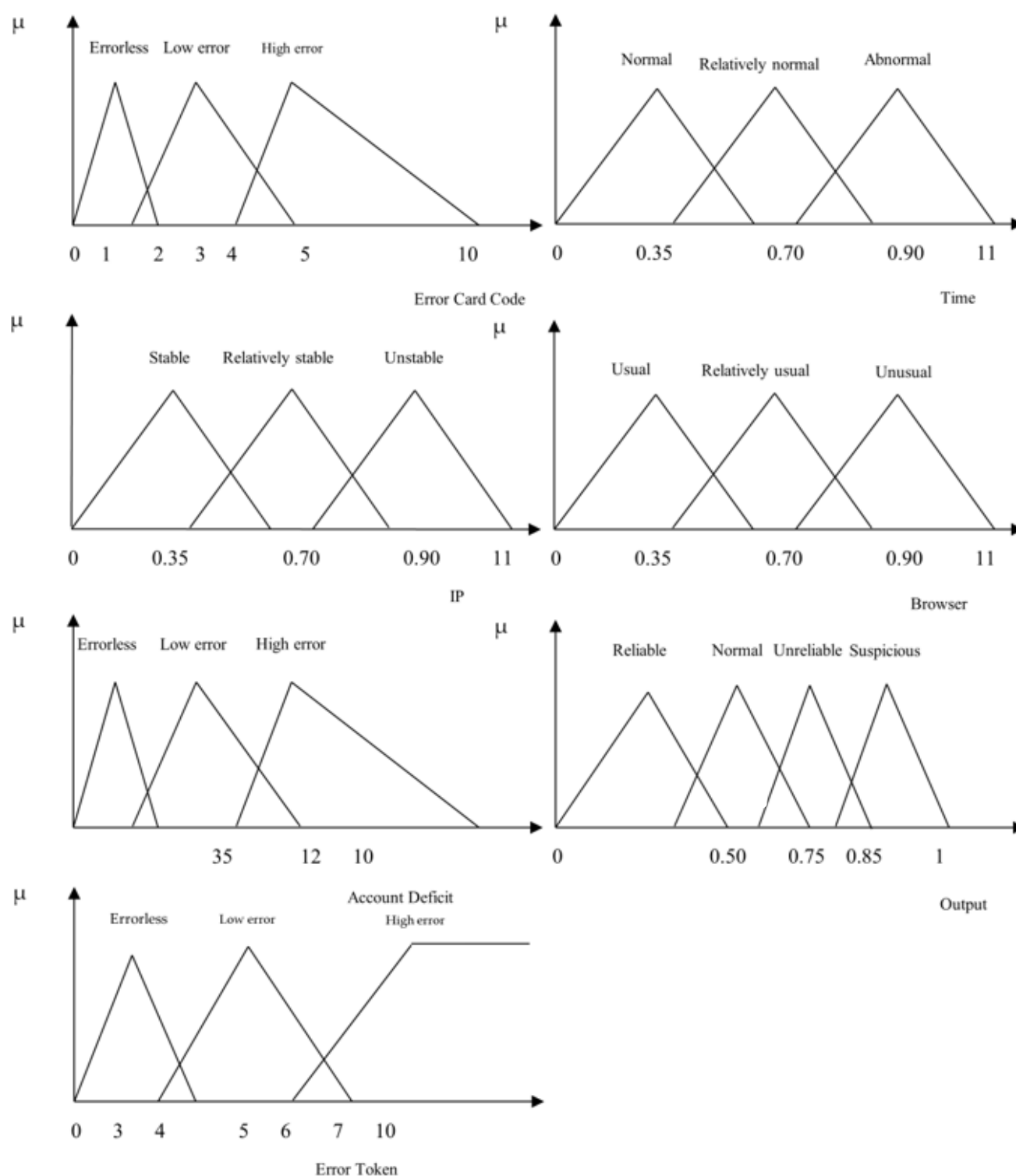
Figure 3. Triangular fuzzy numbers for input and output (source: research analysis).

Phase 2: Architecture of expert system in determination of rate of trust in behavior of Internet banking users. The presented structure in this system is based on identifying behavior of Internet banking users within monthly periods. Specific services may be given without security limit by determination of rate of trust in secure behavior of Internet banking users.

The expert fuzzy system is composed of three main categories, i.e., 'inputs', 'base of rules', and 'outputs'. The expert fuzzy system has been implemented using MATLAB software and data-mining calculations by means of Clementine software.

Phase 3: Inputs of fuzzy system: Presentation of electronic dynamic behavior algorithm.

The input parameters and lingual values of these parameters were determined at first and second phases of study. The expert system deduces each customer's categories based on customers' states in each parameter provided in the previous steps. Then the information of variables of electronic behavior was extracted according to Table 9 for Internet banking users. In order to propose an algorithm with potential for dynamic learning of electronic behavior and personalization of this structure for any customer, score status of any variable was calculated based on Formula 1 after collection of data listed in Table 9.

$$N(x_i) = 1 - \frac{Feature's\ Detail_i}{Total_i} \tag{1}$$

Table 9. Profile of user 1 (source: research analysis).

	User 1			**Recent Login**
Browser	Type	IE7	IE8	IE7
	Quantity	15	28	13
Time of Use	Time	7–15	15–24	7–24
	Number of Login	38	5	0
Geographical Situation	City	Tehran	Tehran	Tehran
	Quantity	43	43	7 October 2018

Feature's detail $_i$: number of occurrence of a type of feature *i Total $_i$*: total number of occurrence of feature *i n*: score for user's electronic behavior in feature *i x $_i$*: Status of feature *i* in user upon login in Internet banking. The given formula is important because the output of this formula may vary from 0 to 10 based on status of any customer. We have the maximum variation in customer's electronic behavior near to 1 and the minimum change in customer's electronic behavior near to 0. For example, if a user login in Internet banking via IE7, the user's score is calculated according to Formula 1 as follows.

$$N(x_i) = 1 - \frac{Feature's\ Detail_i}{Total_i} = 1 - \frac{15}{43} = 0.65 \tag{2}$$

With respect to the given explanations, user's status is placed in this feature at relatively usual level. It is noteworthy that the quantity of this formula may vary proportional to user's status in the long run.

Following to the former example, one can calculate user's score in all variables. The user's score is given in login (input) as described in Table 10.

Table 10. Status of variables for user 1 (source: research analysis).

User	**Variable**	**Status of Variable**	**Input Fuzzy Numbers**	**Lingual Value of Fuzzy System**	**Output of Fuzzy System**	
	IP	Tehran	0	Stable		
	Browser	IE7	0.65	Relatively Usual		
User 1	Time	15–24	0.88	Abnormal	0.521	Normal
	Card Code Error	2	2	Low Error		
	Account Deficit	0	0	Errorless		
	Token Error	1	1	Errorless		

4.2. Model Testing

After determination of triangular values of input and output of fuzzy system (Table 10), the laws were produced for base of fuzzy rules including 720 rules. This base of rules was implemented by Python software. It should be noted that with respect to tree structure of rules, the great number of variable may not be problematic for system processing speed. Generally, if status of a variable at low, medium, and high level, the total number of rules will be reduced to 237 and by analysis on status of the second variable at second step, this structure is reduced to 78 rules and this trend will be continued until recognition of the appropriate fuzzy law. Therefore, the great number of rules may not lead to reduced speed in system processing, and it can tangibly increase accuracy of system.

At this phase, base of fuzzy rules was extracted by means of lingual input variables and comment from experts using 729 rules of '*if and then*'. Some of these rules are presented in Table 11.

Table 11. Examples of description of fuzzy rules (source: research analysis).

Row	Rule Description
1	If the user logins the system with correct card code and within normal time and without making account deficit error and used usual browser and stable geographical situation without error into the Internet banking website, then the behavior of user is reliable.
2	If the user logins the system with card code low error and within normal time and without making account deficit error and used usual browser and stable geographical situation without error into the Internet banking website, then the behavior of user is reliable.
3	If the user logins the system with card code high error and within normal time and without making account deficit error and used usual browser and stable geographical situation without error into the Internet banking website, then the behavior of user is normal.
4	If the user logins the system without card code error and within relatively normal time and in without making account deficit error and uses normal bowser and errorless and stable geographical situation into Internet banking system, then the user's behavior is assumed as reliable.
5	If the user logins the system with card code low error and within relatively normal time and in without making account deficit error and uses normal bowser and errorless and stable geographical situation into Internet banking system, then the user's behavior is assumed as reliable.
6	If the user logins the system with card code high error and within relatively normal time and in without making account deficit error and uses normal bowser and errorless and stable geographical situation into Internet banking system, then the user's behavior is assumed as normal.
7	If the user logins the system without card code error and within abnormal time and in without making account deficit error and uses normal bowser and errorless and stable geographical situation into Internet banking system, then the user's behavior is assumed as reliable.
8	If the user logins the system with card code low error and within abnormal time and in without making account deficit error and uses normal bowser and errorless and stable geographical situation into Internet banking system, then the user's behavior is assumed as normal.
9	If the user logins the system with card code high error and within abnormal time and in without making account deficit error and uses normal bowser and errorless and stable geographical situation into Internet banking system, then the user's behavior is assumed as normal.
10	If the user logins the system without card code error and within normal time and in without making account deficit error and uses normal bowser and errorless and stable geographical situation into Internet banking system, then the user's behavior is assumed as reliable.

5. Findings

The rate of trust in Internet banking customers has been classified in the given system at four levels. In this structure, score of trust in customers' behavior is determined in monthly periods and proportional to the acquired scores by customer, specific level of services is given, e.g., higher level for fund transfer and opening of new account and so forth.

System output is implemented within framework of four general classes of customers. As a customer login in this system, the system determines at what level customer's behavior was placed at former period based on the given fuzzy model and in which group of customers it is classified. The levels of trust in customers' behavior have been classified in four following groups.

Reliable: It includes users who enjoy risk-free and uniform electronic behavior. This group of customers may use special electronic services without the usual constraints.

Normal: This category comprises of users with electronic behavior along with card code error and their electronic behavior is followed by periodic changes. This group of customers may use electronic services with the usual constraints.

Unreliable: It includes that class of Internet banking users who made a great deal of card code errors and also a lot of changes in electronic behavior in different periods. It necessitates for the limited use of this group of customers along with analysis on them.

Suspicious: It comprises of the users who make a lot of card code errors and in process of fund transfer. This group possesses different and abnormal electronic behavior compared to their behavior in the past. It is required not to give any electronic services to this group of customers via Internet banking.

At this step, expert fuzzy system was implemented by means of real situation data. To evaluate system, a few samples of the relevant information items to various users entered in the system of Internet banking database existing in Agri-bank. We will discuss two examples in more details.

The given analysis on user-1 has been proposed with respect to the extracted data in profile of this user. Given the output of fuzzy system, variable of trust in user's behavior is placed at normal level (Table 12). Accordingly, one can take appropriate strategies regarding rate of giving special services proportional to user's electronic behavior.

Table 12. Status of variables in user-1 (source: research analysis).

User	Variable	Status of Variable	Input Fuzzy Numbers	Lingual Value of Fuzzy System	Output of Fuzzy System	Assessment of Experts
User 1	IP	Tehran	0	Stable	0.521	Normal
	Browser	IE7	0.65	Relatively Usual		
	Time	15–24	0.88	Abnormal		
	Card Code Error	2	2	Low error		
	Account Deficit	0	0	Errorless		
	Token Error	1	1	Errorless		

In order to analyze more on performance of the given fuzzy system, another sample of Internet banking user's status entered into fuzzy system while output of system was assessed. The status of electronic behavior of User-2 was collected in Table 13 proportional to the existing data in Internet banking database.

Then status of User-2 is given in variables of fuzzy system described in Table 14.

As the output of fuzzy system indicates, with respect to the data in user's profile and laws in base of rules, the behavior of User-2 is at unreliable status. Thus it is not recommended to give special services to this user.

In this section, to evaluate the model, data related to 100 customers were examined. In this regard, the data was evaluated by the system and experts, using Chi Square test by R software. The results failed to reject H0, indicated the dependency of expert opinion and system results. The coding and results of Chi Square test is provided in Table 15.

Based on Table 15, It can be concluded that the evaluation results of the presented model are accepted by the experts.

Table 13. Profile of User 2 (source: research analysis).

		User 2			
Browser	Type	IE7	IE8	Safari5.5	Firefox32
	Quantity	1	3	49	18
Time of Use	Time	7–15	15–24	7–24	15–30
	Number of Login	62	62	9	0
Geographical Situation	City	Tehran	Manassa	Torrence	Torrence
	Quantity	56	4	11	10
Recent Date		13 November 2019			
Recent City		Torrence			
Recent Browser		Safari5.5			

Table 14. Status of variables for User 2 (source: research analysis).

User	Variable	Status of Variable	Input Fuzzy Numbers	Lingual Value of Fuzzy System	Output of Fuzzy System	Assessment by Experts
User 2	IP	Manassa	0.944	Stable	0.752	Unreliable
	Browser	IE7	0.986	Abnormal		
	Time	15–24	0.874	Unusual		
	Card Code Error	5	5	High Error		
	Account Deficit	0	0	Errorless		
	Token Error	4	4	Errorless		

Table 15. Chi Square test results (source: research analysis).

	Results				Row Totals
	Normal	Reliable	Unreliable	Suspicious	
The Model Result	34 (33.00) (0.03)	51 (53.00) (0.08)	14 (12.50) (0.18)	1 (1.50) (0.17)	100
The Mean Expert Result	32 (33.00) (0.03)	55 (53.00) (0.08)	11 (12.50) (0.18)	2 (1.50) (0.17)	100
Column Totals	66	106	25	3	200 (Grand Total)

The p-value is 0.824249. The result is not significant at $p < 0.5$.

6. Discussion

The results show that four categories can be considered for reliability of Agri-Bank customers based on their behavior. Reliable customers are those who enjoy risk-free and uniform electronic behavior. This group of customers may use special electronic services without the usual constraints. Normal customers are those with card code error followed by periodic changes. This group of customers may use electronic services with the usual constraints. Unreliable users are those who made a great deal of card code errors and also a lot of changes in electronic behavior in different periods. It is required to trace and control the behavior of these customers through Internet banking more accurately, and lastly, suspicious customers' online banking services are those customers who make a lot of card code errors and in process of fund transfer. This group possesses different and abnormal electronic behavior compared to their behavior in the past. It necessitates for the limited use of this group of customers along with analysis on them. Moreover, results of this research show that the proposed model which employed data-oriented approaches based

on data-mining in combination with data-oriented approaches based on fuzzy systems is reliable. Moreover, due to the acceptance of dependency hypothesis in Chi-square test, we can conclude the evaluation results of the proposed model are accepted by the experts.

This research contributes to the current literature by proposing the online banking fraud detection algorithm by both considering the expert opinion and dataset of bank. Such a model can provide the opportunity to use the contextual opinion of experts as well as different characteristics of customer behavior. We also proposed the model based on fuzzy system and K-means tools as two of the accurate tools with high coverage of fraud detection [8]. The foremost given innovation in this study is related to use of data-oriented approaches based on data-mining in combination with data-oriented approaches based on fuzzy systems and also use of clustering to determine lingual intervals in fuzzy system. Moreover, one can consider implementation of this system in real world situation in one of the greatest banks in Iran as a special advantage of the aforesaid research. Moreover, data mining and clustering were used to determine the linguistic expressions or linguistic values of each of the variables, which was one of the most important research innovations. It is noteworthy that the suggested structure has been proposed for improving security of Internet banking system in which the given services depends on rate of trust in electronic behavior of the user.

7. Conclusions and Suggestions

There are some limitations in this study, which provide a direction for further research. Due to time and accessibility limitations, the experts selected from the same bank, thus, it is recommended for future research to benefit from experts' opinion from several banks domain. It was tried in this study to design a dynamic algorithm based on the identified factors in the previous researches. With respect to importance of full identification of the effective factors in determination of credit of users in Internet banking, it is suggested identifying the effective factors in credit by means of content analysis technique in the future investigations and using open-ended questionnaire and direct interview with banking experts.

Banks can use suggested algorithm to increase the security of their Internet banking, to induce a sense of intelligence by observing the behavior of each customer. The extensive changes in IT and emerging of new concepts in this area, e.g., cloud-computing have turned the process of implementation of smart techniques as necessary and vital for improving security in data exchange to various organizations. Therefore, it has been tried in the given study to propose a smart structure for identifying behavior of Internet banking users. Adaption of this system may lead to creating new approach toward giving special services without the usual constraints such as the transferrable fund via Internet banking and opening of new account and so forth to the customers with reliable behavior. Due to the fact that the number of researches in this field is very limited, the implementation of this model in the banking industry and its application can be one of the dimensions of research innovation, while the use of clustering in calculating language values of each of the variables is another innovation of this research.

Data Availability Statement: Data will not share. For having this data, please contact Hamid Bekamiri at Hamidb@business.aau.dk.

Acknowledgments: At the end, we are required expressing our outmost gratitude for the cooperation made by the directors and experts in the field of E-banking of Agri-bank that paved the way duly for conducting the aforesaid study.

Conflicts of Interest: The authors declare that they have no competing interests.

Ethical Consideration: This work is confirmed by ethical committee NO.4217 from Vice president for Research & Technology. The ethical protocols were approved by the Ethics Committee of Ferdowsi university of Mashhad. We analyzed data admitted to Ferdowsi university of Mashhad between 1 December 2018 and 20 February 2020 after approval of ethical committee of Human Research (NO. 4217), the study was carried out based on Agri-bank Internet banking dataset.

Declarations: The authors have no relevant financial or non-financial interests to disclose.

References

1. Pouramirarsalani, A.; Khalilian, M.; Nikravanshalmani, A. Fraud detection in E-banking by using the hybrid feature selection and evolutionary algorithms. *Int. J. Comput. Sci. Netw.* **2017**, *17*, 271–279.
2. Hassani, H.; Huang, X.; Silva, E. Digitalisation and big data mining in banking. *Big Data Cogn. Comput.* **2018**, *2*, 18. [CrossRef]
3. UK Cards Association. Is the Trade Body for the Card Payments Industry in the UK, Representing Financial Institutions which Act as Card Issuers and Acquirers. 2015. Available online: http://www.theukcardsassociation.org.uk (accessed on 5 February 2015).
4. UK Finance. UK Payment Markets Summary 2020. 2020. Available online: https://www.ukfinance.org.uk/system/files/UK-Payment-Markets-Report-2020-SUMMARY.pdf (accessed on 16 April 2021).
5. Eghtesad Online. Growth in Iran's Electronic Banking. 2019. Available online: https://www.en.eghtesadonline.com/Section-economy-4/31265-growth-in-iran-electronic-banking (accessed on 16 April 2021).
6. Huang, Y.P.; Lu, C.C.; Chang, T.W. An Intelligent Approach to Detecting the Bad Credit Card Accounts. In Proceedings of the 25th IASTED International Multi-Conference Artificial Intelligence and Applications, Innsbruck, Austria, 12–14 February 2007; pp. 1–6.
7. European Union Agency for Fundamental Rights. Your Rights Matter: Security Concerns and Experiences. 2020. Available online: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-fundamental-rights-survey-security_en.pdf (accessed on 16 April 2021).
8. Minastireanu, E.A.; Mesnita, G. An Analysis of the Most Used Machine Learning Algorithms for Online Fraud Detection. *Inform. Econ.* **2019**, *23*, 5–16. [CrossRef]
9. Bekamiri, H.; Lagzian, M.; Pooya, A.; Sharif, H. The Banking Industry Foresight Using the Scenario planning Approach and the cross-effects matrix. *IT Manag. Stud.* **2020**. [CrossRef]
10. Bekamiri, H.; Mehraeen, M.; Pooya, A.; Sharif, H. A Stochastic Approach for Valuing Customers in Banking Industry: A Case Study. *Ind. Eng. Manag. Syst.* **2020**, *19*, 744–757.
11. Jain RGour, B.; Dubey, S. A Hybrid Approach for Credit Card Fraud Detection using Rough Set and Decision Tree Technique. *Int. J. Comput. Appl.* **2016**, *139*, 1–16.
12. Choi, D.; Lee, K. Machine Learning based Approach to Financial Fraud Detection Process in Mobile Payment System. *IT Converg. Pract. (INPRA)* **2018**, *5*, 12–24.
13. Sudha, C.; Raj, T.N. Credit Card Fraud Detection in Internet Using K-Nearest Neighbor Algorithm. *Int. J. Comput. Sci.* **2017**, *5*, 22–30.
14. Syniavska, O.; Dekhtyar, N.; Deyneka, O.; Zhukova, T.; Syniavska, O. Security of e-banking systems: Modelling the process of counteracting e-banking fraud. In *SHS Web of Conferences*; EDP Sciences: Ulis, France, 2019; Volume 65, pp. 1–5. [CrossRef]
15. Daliri, S. Using Harmony Search Algorithm in Neural Networks to Improve Fraud Detection in Banking System. *Comput. Intell. Neurosci.* **2020**. [CrossRef]
16. Utami, W.; Nugroho, L.; Mappanyuki, R.; Yelvionita, V. Early warning fraud determinants in banking industries. *Asian Econ. Financ. Rev.* **2020**, *10*, 604. [CrossRef]
17. Karlsen, K.N.; Killingberg, T. Profile Based Intrusion Detection for Internet Banking Systems. Master's Thesis, Norwegian University of Science and Technology, Trondheim, Norway, 2008.
18. Mannan, M.; van Oorschot, P.C. Security and usability: The gap in real-world online banking. In Proceedings of the Workshop on New Security Paradigms (NSPW '07), North Conway, NH, USA, 18–21 September 2007; Association for Computing Machinery: New York, NY, USA, 2008; pp. 1–14.
19. Neill, D.B.; Moore, A.W. Rapid detection of significant spatial clusters. In Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Seattle, WA, USA, 22–25 August 2004; pp. 256–265.
20. Datta, P.; Tanwar, S.; Panda, S.N.; Rana, A. Security and Issues of M-Banking: A Technical Report. In Proceedings of the 2020 8th International Conference on Reliability, Infocom Technologies and Optimization, Trends and Future Directions, ICRITO, IEEE, Noida, India, 4–5 June 2020; pp. 1115–1118.

21. Ilker, K.A.R.A.; Aydos, M. Cyber Fraud: Detection and Analysis of the Crypto-Ransomware. In Proceedings of the 2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference UEMCON, IEEE, Virtual Conference, New York, NY, USA, 28–31 October 2020; pp. 0764–0769.

22. Kasabov, N.K. *Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering*; The MIT Press: Cambridge, MA, USA, 1998.

23. Mendel, J.M. *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*; Prentice Hall PTR: Hoboken, NJ, USA, 2001.

24. Russell, S.J.; Norvig, P. *Artificial Intelligence: A Modern Approach*, 3rd ed.; Prentice Hall: Hoboken, NJ, USA, 2010.

25. Olusolade Aribake, F.; Mat Aji, Z. Modelling the Phishing Avoidance Behaviour Among Internet Banking Users in Nigeria: The Initial Investigation. *J. Comput. Eng. Technol.* **2020**, *4*, 1–17. [CrossRef]

26. Bignell, K.B. Authentication in an Internet banking environment: Towards developing a strategy for fraud detection. In Proceedings of the International Conference on Internet Surveillance and Protection (ICISP), Cote d'Azur, France, 26–29 August 2006; pp. 23–30.

27. Kovach, S.; Ruggiero, W.V. Online banking fraud detection based on local and global behavior. In Proceedings of the Fifth International Conference on Digital Society, Guadeloupe, France, 23–28 February 2011; pp. 166–171.

28. Fan, W.; Miller, M.; Stolfo, S.; Lee, W.; Chan, P. Using artificial anomalies to detect unknown and known network intrusions. *Knowl. Inf. Syst.* **2004**, *6*, 507–527. [CrossRef]

29. Eshghi, A.; Kargari, M. Introducing a new method for the fusion of fraud evidence in banking transactions with regards to uncertainty. *Expert Syst. Appl.* **2019**, *121*, 382–392. [CrossRef]

30. Leung, A.; Yan, Z.; Fong, S. On designing a flexible e-payment system with fraud detection capability. In Proceedings of the IEEE International Conference on e-Commerce Technology, Taipei, Taiwan, 28–31 March 2004; pp. 236–243.

31. Johnson, J.M.; Khoshgoftaar, T.M. Medicare fraud detection using neural networks. *J. Big Data* **2019**, *6*, 1–35. [CrossRef]

32. Itoo, F.; Singh, S. Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection. *Int. J. Inf. Technol.* **2020**, 1–9. [CrossRef]

33. Awoyemi, J.O.; Adetunmbi, A.O.; Oluwadare, S.A. Credit card fraud detection using machine learning techniques: A comparative analysis. In Proceedings of the 2017 International Conference on Computing Networking and Informatics ICCNI, IEEE, Lagos, Nigeria, 29–31 October 2017; pp. 1–9.

34. Rushin, G.; Stancil, C.; Sun, M.; Adams, S.; Beling, P. Horse race analysis in credit card fraud—Deep learning, logistic regression, and Gradient Boosted Tree. In Proceedings of the 2017 Systems and Information Engineering Design Symposium SIEDS, IEEE, Charlottesville, VA, USA, 28 April 2017; pp. 117–121.

35. Baumann, M. Improving a Rule-based Fraud Detection System with Classification Based on Association Rule Mining. 2021. Available online: https://www.researchgate.net/profile/Michaela_Baumann/publication/349244021_Improving_a_Rule-based_Fraud_Detection_System_with_Classification_Based_on_Association_Rule_Mining/links/6026804fa6fdcc37a81df08f/Improving-a-Rule-based-Fraud-Detection-System-with-Classification-Based-on-Association-Rule-Mining.pdf (accessed on 17 February 2021).

36. Öztürk, M.S.; Usul, H. Detection of Accounting Frauds Using the Rule-Based Expert Systems within the Scope of Forensic Accounting. In *Contemporary Issues in Audit Management and Forensic Accounting*; Emerald Publishing Limited: Bingley, UK, 2020.

37. Edge, K.; Raines, R.; Grimaila, M.; Baldwin, R.; Bennington, R.; Reuter, C. The use of attack and protection trees to analyze security for an online banking system. In Proceedings of the 40thAnnual Hawaii International Conference on System Sciences (HICSS), Waikoloa, HI, USA, 3–6 January 2007.

38. Behera, T.K.; Panigrahi, S. Credit card fraud detection using a neuro-fuzzy expert system. In *Computational Intelligence in Data Mining*; Springer: Singapore, 2017; pp. 835–843.

39. Subudhi, S.; Panigrahi, S. Use of optimized Fuzzy C-Means clustering and supervised classifiers for automobile insurance fraud detection. *J. King Saud Univ.—Comput. Inf. Sci.* **2020**, *32*, 568–575. [CrossRef]

40. Aburrous, M.; Hossain, M.A.; Dahal, K.; Thabtah, F. Intelligent phishing detection system for e-banking using fuzzy data mining. *Expert Syst. Appl.* **2010**, *37*, 7913–7921. [CrossRef]

41. Ur Rahman, A.; Dash, S.; Luhach, A.K.; Chilamkurti, N.; Baek, S.; Nam, Y. A Neuro-fuzzy approach for user behavior classification and prediction. *J. Cloud Comput.* **2019**, *8*, 17. [CrossRef]

42. Rezaiee, A.M.; Karimi, A. A New Dynamic Intelligent Model to Determine Reliability and Trust of Online Banking by Using Fuzzy C-Mean. *Indones. J. Electr. Eng. Comput. Sci.* **2016**, *4*, 605–610. [CrossRef]

43. Darko, A.; Chan, A.P.C.; Ameyaw, E.E.; Owusu, E.K.; Pärn, E.; Edwards, D.J. Review of application of analytic hierarchy process (AHP) in construction. *Int. J. Constr. Manag.* **2019**, *19*, 436–452. [CrossRef]

44. Doloi, H. Application of AHP in improving construction productivity from a management perspective. *Constr. Manag. Econ.* **2008**, *26*, 841–854. [CrossRef]