



**AALBORG UNIVERSITY**  
DENMARK

**Aalborg Universitet**

## **Learning and Interpreting Multi-Multi-Instance Learning Networks**

Tibo, Alessandro

*Publication date:*  
2019

[Link to publication from Aalborg University](#)

*Citation for published version (APA):*

Tibo, A. (2019). *Learning and Interpreting Multi-Multi-Instance Learning Networks*. Poster presented at Kick-off: AI for the people, Aalborg, Denmark.

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

### **Take down policy**

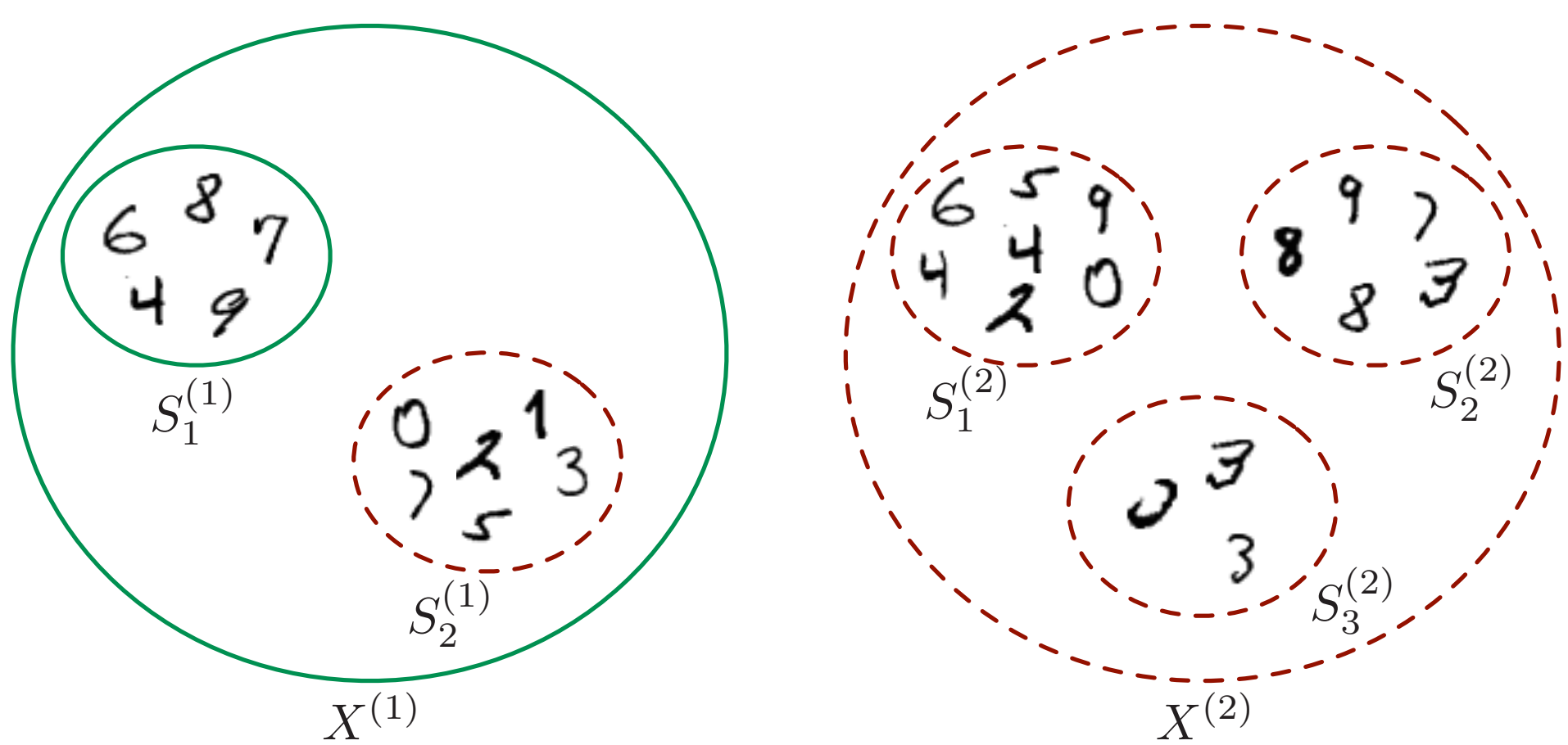
If you believe that this document breaches copyright please contact us at [vbn@aub.aau.dk](mailto:vbn@aub.aau.dk) providing details, and we will remove access to the work immediately and investigate your claim.

## Introduction

- ▶ Dietterich et al., 1997, introduced the multi-instance learning (MIL), perhaps the simplest form of relational learning where data consists of labeled bags of instances
- ▶ We extend the MIL setting by considering examples consisting of labeled nested bags of instances. Labels are observed for top-level bags, while instances and lower level bags have associated latent labels
- ▶ We propose a solution to the MMIL problem based on neural networks with a special layer called **bag-layer**
- ▶ Multi-multi instance learning enables a particular way of **interpreting** the models by reconstructing instance and sub-bag latent variables; this allows to explain the prediction for a particular data point, and to describe the structure of the decision function in terms of symbolic rules

## Framework

- ▶ For any set  $\mathcal{X}$ ,  $\mathcal{M}(\mathcal{X})$  denotes the set of all multisets of  $\mathcal{X}$
- ▶ When the level of nesting is two, an example is a labeled bag-of-bags  $(X, y) \in \mathcal{M}(\mathcal{M}(\mathcal{X})) \times \mathcal{Y}$  drawn from a distribution  $p(X, y)$
- ▶ We assume that instance and sub-bag labels only depend on properties of the respective instances or sub-bags



**Fig. 1:** A positive (left) and a negative (right) top-bag. Solid green lines represent positive (sub-)bags while dashed red lines represent negative (sub-) bags

## Bag Layers

We model the conditional distribution  $p(y|X)$  with a neural network architecture that handles bags-of-bags of variable sizes by aggregating intermediate internal representations. For this purpose, we introduce a new layer called bag-layer. A bag-layer takes as input a bag of  $m$ -dimensional vectors  $\{\phi_1, \dots, \phi_n\}$ , and first computes  $k$ -dimensional representations

$$\rho_i = \alpha(w\phi_i + b), \quad (1)$$

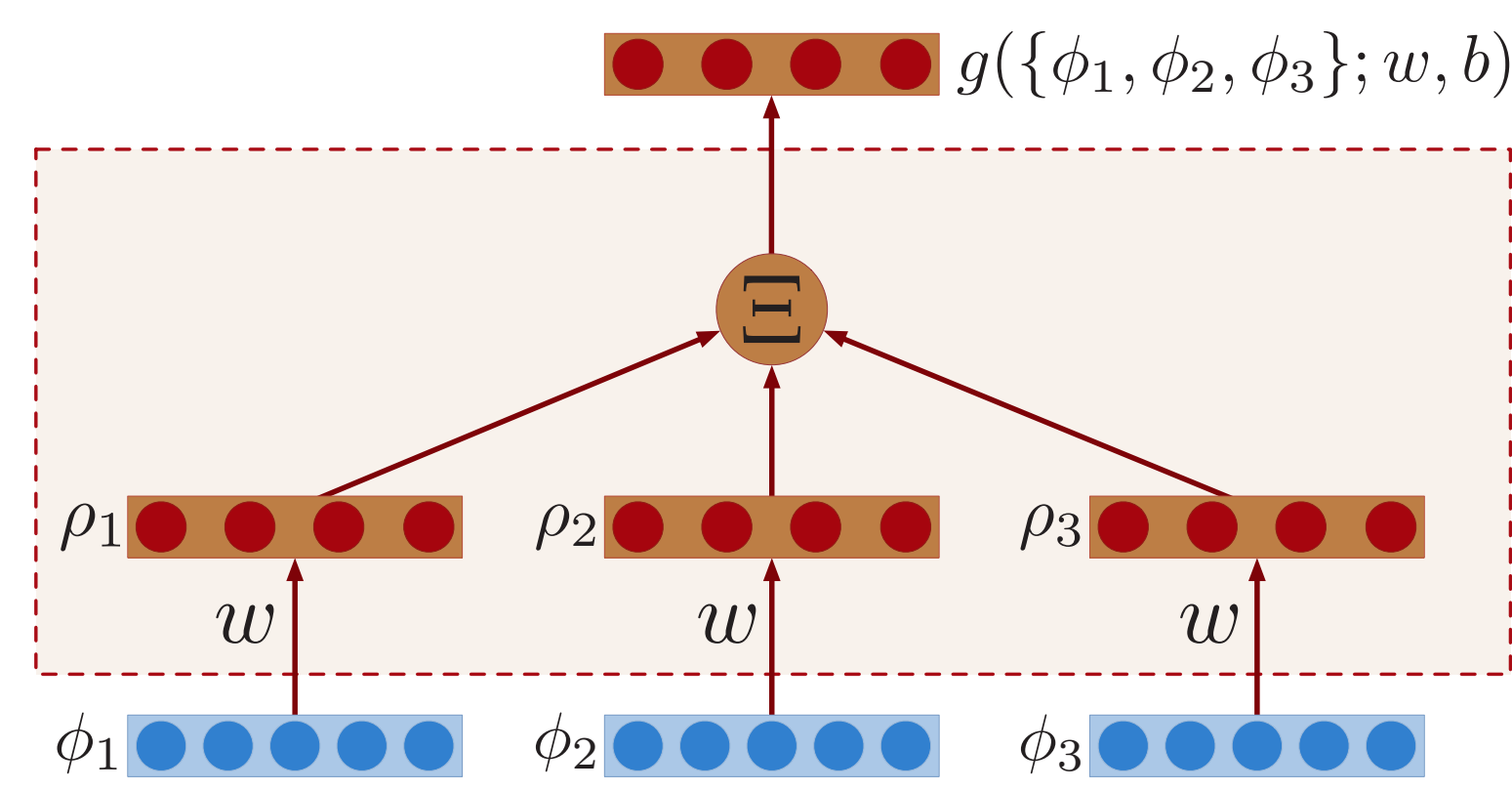
using a weight matrix  $w \in \mathbb{R}^{k \times m}$ , a bias vector  $b \in \mathbb{R}^k$ , and an activation function  $\alpha$  (such as ReLU, tanh, or linear).

The bag layer then computes its output as:

$$g(\{\phi_1, \dots, \phi_n\}; w, b) = \Xi_{i=1}^n \rho_i, \quad (2)$$

where  $\Xi$  is element-wise aggregation operator (such as max or average)

## Bag Layers



**Fig. 2:** A bag-layer receiving a bag of cardinality  $n = 3$ . In this example  $k = 4$  and  $m = 5$ .

## Interpretability

The MMIL settings offers also some advantages for interpreting results. Indeed, if instance or sub-bag labels were observed, they would provide more information about bag-of-bags than mere predictions. Latent variables are indeed associated with each individual “part” of the top-bag, as opposite to the prediction which is associated with the whole.

The idea we propose in the following is based on four steps.

1. We employ clustering at the level of instance and sub-bag representations to construct pseudo-labels as surrogates for hypothesized actual latent labels
2. We provide semantic interpretations of the pseudo-labels for human inspection
3. We apply a transparent learner to extract a human readable representation of the mappings between pseudo-labels at the different levels of a bag-of-bags structure
4. We explain predictions for individual top-bag examples by exhibiting the relevant components and their pseudo-labels which determine the predicted top-bag label

## Interpreting IMDB

$v_1$ (11.37%)	$v_2$ (41.32%)	$v_3$ (15.80%)	$v_4$ (31.51%)
overrated poorly written badly acted	I highly recommend you to NOT waste your time on this movie as I have	I loved this movie and I give it an 8/ 10	It's not a total waste
It is badly written badly directed badly scored badly filmed	This movie is poorly done but that is what makes it great	Overall I give this movie an 8/ 10	horrible god awful
This movie was poorly acted poorly filmed poorly written and overall horribly executed	Although most reviews say that it isn't that bad i think that if you are a true disney fan you shouldn't waste your time with...	final rating for These Girls is an 8/ 10	Awful awful awful

Table 1: Interpreting sentence (sub-bag) pseudo-labels in the MMIL setting.

Story about three eclipse (maybe even Indigo, ha) children beginning their love for murder. Oh, and the people who are “hot” on their trail.  
 $[v_1]$  **Bloody Birthday, a pretty mediocre title**<sup>4</sup> for the film, was a nice lil<sup>1</sup> surprise. I was in no way expecting a film that dealt with blood-thirsty psychopath kids.  
 $[v_3]$  And I may say it's also **one of the best flicks**<sup>1</sup> I've seen with kids as the villains. By the end of the movie I seriously wanted these kids to die in horrible fashion.  
 $[v_3]$  **It's a really solid 80s**<sup>1</sup> horror flick, but how these kids are getting away with all this mayhem and murder is just something that **you can't not**<sup>2</sup> think about. Even the slightest bit of investigation would easily uncover these lil sh!ts as the murderers. But there seems to be only a couple police in town, well by the end, only one, and he seemed like a dimwit, so I suppose they could have gotten away with it. Haha, yeah, and I'm a Chinese jet-pilot. Nevertheless, this movie delivered some evilass kids who were more than entertaining, a lot of premarital sex and a decent amount of boobage. No kiddin! If you're put off by the less than stellar title, dash it from your mind and give this flick a shot.  $[v_3]$  **It's a very recommendable and underrated 80s**<sup>1</sup> horror flick.

**Fig. 3:** A sample positive review with MMIL labeling.

## Experiments on IMDB

We use the IMDB (Maas et al., 2011) dataset, which is a standard benchmark movie review dataset for binary sentiment classification. IMDB consists of 25,000 training reviews, 25,000 test reviews.

Text data exhibits a natural bags-of-bags structure by viewing a text as a bag of sentences, and each sentence as a bag of words. Hence, a MMIL dataset was constructed from the reviews, where then each review (top-bag) is a bag of sentences. However, instead of modeling each sentence (sub-bag) as a bag of words, we represented sentences as bags of trigrams in order to take into account possible negations, e.g. “not very good”, “not so bad”.

**MMIL Structure:** Conv1D layer with 300 filters, ReLU activations and kernel size of 100, two stacked bag-layers (with ReLU activations) with 500 units each (250 max-aggregation, 250 mean-aggregation) and an output layer with sigmoid activation.

We obtained a final accuracy of 92.26%.

## References

- [1] Dietterich, T. G., Lathrop, R. H., and Lozano-Pérez, T. (1997). Solving the multiple instance problem with axis-parallel rectangles. *Artificial intelligence*, 89(1-2), 31-71
- [2] Maas, A. L., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., and Potts, C. (2011, June). Learning word vectors for sentiment analysis. In *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies-volume 1* (pp. 142-150). Association for Computational Linguistics.