

Summaries of Related work

Alex Baier
abaier@uni-koblenz.de

December 1, 2016

Contents

1	Ontology learning	1
1.1	Taxonomy learning – factoring the structure of a taxonomy into a semantic classification decision	1
1.2	Ontology Enrichment by Discovering Multi-Relational Association Rules from Ontological Knowledge Bases	2
1.3	Ontology Learning for the Semantic Web	2
2	Neural networks	2
2.1	Using Recurrent Neural Network for Learning Expressive Ontologies	2
2.2	Recurrent neural network based language model	3
2.3	Efficient Estimation of Word Representations in Vector Space	3
2.4	Deep Neural Networks for Learning Graph Representations	4
2.5	Supervised neural networks for the classification of structures	4
2.6	Neural Networks for Classification: A Survey	4
2.7	A Convolutional Neural Network for Modelling Sentences	4
2.8	DRAW: A Recurrent Neural Network For Image Generation	5

1 Ontology learning

1.1 Taxonomy learning – factoring the structure of a taxonomy into a semantic classification decision

[7] defines a tree-ascending algorithm, which calculates the taxonomic similarity between nearest neighbors. Evaluation shows that the algorithm is good at choosing the correct class for a new word. A combination of tree-ascending and k-nearest-neighbors was developed, which slightly improved the results in finding the correct class. Another algorithm, tree-descending, was shown to be less efficient than standard classifiers.

The paper solves a similar problem to the problem of my work. The exploitation of the taxonomy for classification proves to be effective and should therefore be incorporated into the solution of the thesis' problem.

1.2 Ontology Enrichment by Discovering Multi-Relational Association Rules from Ontological Knowledge Bases

The algorithm presented in [2] discovers relational association rules expressed SWRL, which allows the discovered rules to be directly integrated into the ontology and increase its expressiveness. This is done by detecting hidden patterns in a populated ontological knowledge base, and then obtaining relational association rules from these pattern.

Not relevant for the thesis, because the work is concerned with finding new rules in ontological KBs, while the task of the thesis is to find new connections in the taxonomy of a KB. And other papers solve problems with a higher similarity to the thesis' problem.

1.3 Ontology Learning for the Semantic Web

The task of ontology learning is analyzed by [5]. A semi-automated approach is considered, which consists of 5 steps: import/reuse existing ontologies, extract major parts of target ontology, prune to adjust ontology for primary purpose, refine ontology to complete ontology at fine granularity, and apply ontology on target applications, which validates the results. These steps may be repeated as often as necessary.

The problem of the thesis belongs to the task of ontology refinement, because an ontology is already provided.

2 Neural networks

2.1 Using Recurrent Neural Network for Learning Expressive Ontologies

TODO 1: Some parts of this summary should be moved to the comparison of neural networks. [8] describes a recurrent neural network with short-term memory capabilities through Gated Recursive Units for usage in ontology learning. It is argued that recurrent neural networks can handle this task, because they proved to do well in handling natural languages, and therefore should also be able to "handle the typical syntactic structures of encyclopedic text".

The ontology learning process is described as transduction task. The typical structure of encyclopedic text is exploited by translating the text into a logical formula using a pipeline, which creates a formula template and a tagged sentence. These are combined to create a formula.

This pipeline is assembled by a combination of neural networks.

- Sentence tagging: single Recurrent Neural Network with Gated Recursive Units.
 - Input: sentence in natural language, represented as $n+1$ symbol sequence. First n symbols are words. The last symbol is $\langle \text{EOS} \rangle$, which indicates end of sequence. Each word has a mapping to an integer, which represents its position in the vocabulary. The input is then transformed into a sequence of context windows.
 - Output: Estimation of each tag to be the right one for k -th word. Predicted tag for k -th word is the one with the highest probability.
- Sentence transduction: Recurrent Encoder-Decoder with Gated Recursive Units.
 - Input: The same as in sentence tagging. But the context windowing is not executed.
 - Output: Set of all possible formula terms. Input and output sequences don't have the same length.
- Gated Recursive Unit provides both networks with a short-term memory effect.

2.2 Recurrent neural network based language model

[6] describes a simple recurrent network language based model and tests its performance for the task of speech recognition. It outperforms the typical n -gram model, but has a higher computational performance.

Because speech recognition has only very few similarities to the concerned data of my work, this paper has no further relevance for problem of this work. But it should be noted, that this paper shows the power of recurrent neural networks in predicting contextual data and not having a limited size of context.

2.3 Efficient Estimation of Word Representations in Vector Space

[?] proposes two new neural network architectures, which compute continuous vector representations of words. It is possible to use algebraic operations on such word vectors. An example, which is mentioned in the paper $\text{vector}(\text{"King"}) - \text{vector}(\text{"Man"}) + \text{vector}(\text{"Woman"})$ results into a vector that is very close to the representation of the word *Queen*. The paper tries to maximize the accuracy of such operations by developing their architectures to "preserve the linear regularities among words". The Continuous Bag-of-Words Model (CBOW) and the Continuous Skip-gram Model were developed. CBOW is similar to the Feedforward Neural Net Language Model. It predicts the current word based on its context. Continuous Skip-Gram in contrast predicts the words before and after the current word.

The results of different word vectors versions were compared by applying them to a similarity task. The proposed models are able to create word vectors with a higher syntactic and semantic accuracies than other existing models.

2.4 Deep Neural Networks for Learning Graph Representations

[1] describes a deep neural network, which creates deep graph representation model, in which a low dimensional vector is created for each vertex in the graph. The method was tested on real datasets and outperformed some state-of-the-art systems. It is able to handle noisy data, which may be important for the task of this work. The output of this network, the graph representations, can be used as input for other methods, e.g. supervised classification. Therefore this network could be used for transforming the taxonomy of Wikidata into a small-dimensional input format.

2.5 Supervised neural networks for the classification of structures

[9] describes neural networks using generalized recursive neurons for the classification of structured patterns (e.g. concept graphs). The a-priori encoding of structures for use in the neural network has drawbacks. The authors propose the use a second neural network for encoding the structure, which is trained alongside the classification network. To solve this problem the generalized recursive neuron is introduced. An encoding network is constructed using this neurons. It creates a vector, which encodes the graph, and is finally fed into a feed-forward neural network for classification. **TODO 2: Read this paper again, I think misunderstood, what the network actually does.**

2.6 Neural Networks for Classification: A Survey

[10] reviews the topic neural networks for classification. Key issues and features of neural networks are discussed. Relevant for the thesis, the topic of feature variable selection is discussed. It is important to "find the smallest set of features that can result in satisfactory predictive performance". Different approaches are discussed. It is noted that linear limitations of popular methods like principle component analysis can be overcome by using neural networks to perform networks, because they are able to perform nonlinear PCA.

2.7 A Convolutional Neural Network for Modelling Sentences

[4] defines a convolutional neural network architecture called Dynamic Convolutional Neural Network. The network models sentences using combinations of pooling and convolution layers, which internally induce something akin to parse tree on the input

sentences. The specific output of the network is based on its supervised training. For example classification of sentiments was tested on movie reviews and tweets. The network proved to be effective at these tasks.

2.8 DRAW: A Recurrent Neural Network For Image Generation

Deep Recurrent Attentive Writer (DRAW) is introduced by [3]. It is a neural network architecture, which is able to generate images, by iteratively drawing and refining an image. The architecture consists of an encoder network, used for compressing images during learning, and a decoder network that recreates the image after receiving a code. This type of architecture is called variational auto-encoders. IN DRAW both of these networks are recurrent, which means a sequence of code samples is exchanged between the networks. The image generation works in multiple steps, in which different parts of the image are generated. The decision on where to write and what code samples to read is made by the neural network. After a number of steps an image is produced, which cannot be distinguished from real images.

References

- [1] Shaosheng Cao, Wei Lu, and Qionghai Xu. Deep neural networks for learning graph representations. In Dale Schuurmans and Michael P. Wellman, editors, AAAI, pages 1145–1152. AAAI Press, 2016.
- [2] Claudia d’Amato, Steffen Staab, Andrea G. B. Tettamanzi, Tran Duc Minh, and Fabien L. Gandon. Ontology enrichment by discovering multi-relational association rules from ontological knowledge bases. In Sascha Ossowski, editor, SAC, pages 333–338. ACM, 2016.
- [3] Karol Gregor, Ivo Danihelka, Alex Graves, and Daan Wierstra. DRAW: A recurrent neural network for image generation. CoRR, abs/1502.04623, 2015.
- [4] Nal Kalchbrenner, Edward Grefenstette, and Phil Blunsom. A convolutional neural network for modelling sentences. CoRR, abs/1404.2188, 2014.
- [5] Alexander Maedche and Steffen Staab. Ontology learning for the semantic web. IEEE Intelligent Systems, 16(2):72–79, March 2001.
- [6] Tomas Mikolov, Martin Karafiát, Lukáš Burget, Jan Cernocký, and Sanjeev Khudanpur. Recurrent neural network based language model. In INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association, Makuhari, Chiba, Japan, September 26-30, 2010, pages 1045–1048, 2010.
- [7] Viktor Pekar and Steffen Staab. Taxonomy learning - factoring the structure of a taxonomy into a semantic classification decision. In Proceedings of the 19th

Conference on Computational Linguistics, COLING-2002, August 24 - September 1, 2002 , Taipei, Taiwan, 2002, 2002.

- [8] Giulio Petrucci, Chiara Ghidini, and Marco Rospocher. Using recurrent neural network for learning expressive ontologies. CoRR, abs/1607.04110, 2016.
- [9] Alessandro Sperduti and Antonina Starita. Supervised neural networks for the classification of structures. IEEE Transactions on Neural Networks, 8(3):714–735, may 1997.
- [10] Guoqiang Peter Zhang. Neural networks for classification: a survey. In and Cybernetics - Part C: Applications and Reviews, 2000.