Computational Linguistics and Intellectual Technologies:
Proceedings of the International Conference “Dialogue 2016”

Moscow, June 1–4, 2016

DEEP LEARNING AND STRUCTURAL KERNELS FOR SEMANTIC INFERENCE: QUESTION ANSWERING APPLICATIONS TO FORMAL TEXT AND WEB FORUMS

Alessandro Moschitti (amoschitti@gmail.com)

Qatar Computing Research Institute, Hamad Bin Khalifa University, Qatar

In very recent years, there has been a large body of research work dedicated to deep learning: many recent articles report state-of-the-art results obtained by neural networks on many different applications of Natural Language Processing (NLP), e.g., Machine Translation, Speech processing, etc.

However, even the most optimist supporters of Neural Networks acknowledge the fact that dealing with highly complex semantic tasks requires new deep learning solutions. For example, no effective neural network model has been proposed so far for accurately performing discourse and dialog parsing or deep semantic inference, which is, for example, needed in Question Answering, Textual Entailment and Paraphrasing Identification.

The main difficulty in designing effective Neural Networks for such NLP tasks is the automatic learning of embeddings that can encode complex semantic relations, e.g., whose arguments (i) can span more than few consecutive words and (ii) can be located at any arbitrary distance in the target text. In other words, automatically learning semantic structures for language inference is still an open problem.

In this talk, I will elaborate on the claim above by firstly introducing properties and challenges of automatic language inference as well as the latest developments of the related advanced technology. The latter also includes the use of structural kernels applied to syntactic and semantic structures, which can easily encode complex text dependencies in learning algorithms. For this purpose, I will also capitalize on my direct experience with the techniques and models used for engineering the famous IBM Watson system.

Then, I will introduce some simple networks, along with some of their successful applications to simple NLP problems, e.g., sentiment analysis or named entity recognition, highlighting the importance of pre-training the networks with unsupervised models, e.g., using the famous toolkits, word2vec and GloVe.

Finally, I will present deep learning models for solving complex inference on short text, i.e., the one required for solving Question Answering (QA) applications. In particular, I will focus on Community QA, which offers the possibility of (i) utilizing a strict semantic correlation between questions and answers, i.e., user comments; (ii) modeling the user interaction structure, which is implicitly part of the question-comment treads; and (iii) dealing with a real-world application, which also requires the use of robust methods to process noisy text.

The results of our comprehensive comparative study on the above-mentioned machine learning methods suggest that the combination of deep learning and structural kernels applied to syntactic/semantic representations achieves the state of the art in applications requiring highly complex semantic inference. Therefore, although deep learning is greatly contributing to the solution of several NLP tasks, it is still important to combine it with traditional approaches, e.g., kernel methods applied to syntactic/semantic structure.