THESIS PROPOSAL

Embarrassingly Stupid Parsing with Neural Networks

Background

Parsing is the task of mapping a natural language sentence to a formal representation of its syntax or semantics. Previous work in the area has developed highly powerful but rather complex parsing algorithms that cast parsing as combinatorial optimisation over feature-rich linear models. However, recent work in the area suggests that we can throw much of the complexity of these algorithms overboard when we switch to neural networks. The purpose of this thesis is to explore how close to the state of the art we can get with this 'stupid' approach.

Project description

Your task is to design, implement and evaluate an 'embarrassingly stupid' parser for natural language. The basic parsing algorithm takes only a few lines of code; what you will experiment with is how close this approach can be pushed towards the state of the art by using different machine learning models based on neural networks. This includes recurrent neural networks such as Long Short-Time Memory networks (LSTMs). You will evaluate your parsing using standard measures such as accuracy, precision, and recall, and compare your results to the research literature.

Customer

NLPLAB (internal)

Contact

Marco Kuhlmann, marco.kuhlmann@liu.se

Student profile

Interest in neural networks and software libraries for neural networks such as Keras, PyTorch, Theano or TensorFlow. Background knowledge in Natural Language Processing, for instance via courses such as TDDE09.