Call for Master Thesis

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Concept-level Sentiment Analysis with SenticNet and Word Embeddings

Sentiment Analysis is an extremely difficult task as it requires the understanding and interpretation of data. Humans can do it, machines cannot. To close this gap, word statistics to compute a positive/negative polarity of texts, or word embedding representations which can somehow embed emotional information have been recently investigated. These representations are used to feed Machine Learning algorithms with the hope that they learn patterns and predicts correct sentiment labels.

However, these approaches lack concept-level semantics which allows us to go beyond a mere word analysis. Concept-level sentiment analysis steps away from blind us of words, but rather tries to adopt features associated with natural language concepts. The goal of the thesis is to study how SenticNet [1] resources, which contain emotional information about concepts, can be used in combination with word embedding representations (e.g., word2vec, fasttext, BERT) for sentiment analysis tasks. The approach will be applied for score prediction on learners' reviews of the Coursera dataset [2].

[1] https://sentic.net/downloads/
[2] https://www.kaggle.com/septa97/100k-courseras-course-reviews-dataset

The thesis will be supervised by **Prof. Dr. Harald Sack,** Information Service Engineering at Institute AIFB, KIT, in collaboration with FIZ Karlsruhe.

Pre-requisites:
Good programming skills in Python, creativity, curiosity, and willingness to learn

Contact person:
Danilo Dessì
danilo.dessi@kit.edu
danilo.dessi@fiz-karlsruhe.de

aifb
Institute of
Applied Informatics and
Formal Description Methods
http://www.aifb.kit.edu/

KIT
Karlsruhe Institute of Technology