

DEEP LEARNING FOR POLARITY DETECTION TO PREDICT STOCK PRICE MOVEMENT

Ankit Bhutani - Y9227094
Harshvardhan Sharma - 11299

Advisor - Prof. Amitabha Mukerjee

Problem

There has been a lot of recent work in the field of Natural Language Processing to study the financial effects of news on the stock market and on the problem of sentiment analysis and polarity detection of text. The problem of polarity detection has usually been approached by supervised learning methods. Here we will use a generic neural network architecture for polarity detection to predict changes in stock prices by analyzing news articles in an unsupervised manner. This approach uses very minimal domain knowledge and the requirement for labelled data should also be less as compared to other approaches owing to the unsupervised pre-training as proposed by Collobert et. al

Previous work

Collobert et al. have proposed a unified neural network framework for natural language processing tasks that uses deep learning techniques to learn internal representations from unlabelled training data. Their approach uses very minimal language expertise and the builds a generic network capable of performing a lot of labelling tasks like Semantic Role Labelling, Named Entity Recognition, POS tagging etc. with some supervised fine-tuning for separate cases.

Xie et al. have demonstrated the effectiveness of models based on semantic frames in the task of polarity detection. They use the semantic parse generated by semafor (Das and Smith) for the purpose of polarity detection. They also propose a vector space model to represent the data. The model includes features from semantic frames, Dictionary of Affect in Language (Whissel) and bag of words model.

Approach

We propose to discover features relevant to the task by first building a domain specific language model through an approach similar to the one used by Collobert et. al. and then fine-tuning the network based on the polarity labels. Such a weakly supervised approach using neural networks has not been used for stock price prediction in the past to the best of our knowledge.

References

1. R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu and P. Kuksa. **Natural Language Processing (Almost) from Scratch**. *Journal of Machine Learning Research*, 12:2493-2537, 2011.
2. Xie, Boyi; Passonneau, Rebecca J.; Creamer, German; Wu, Leon. 2013. [Semantic Frames to Predict Stock Price Movement](#). *Proceedings of the 2013 Annual Meeting of the Association for Computational Linguistics*. August 4-9, Sofia, Bulgaria
3. Dipanjan Das and Noah A. Smith. Semi-Supervised Frame-Semantic Parsing for Unknown Predicates.. *Proceedings of ACL 2011*
4. Apoorv Agarwal, Fadi Biadisy, and Kathleen Mckeown, 2009. Contextual phrase level polarity analysis using lexical affect scoring and syntactic N-grams. *Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)*