Sentiment Analysis on Movie Reviews

Xuwen Shen, Xinzhe Yang, Katie Yang

February 27, 2018

1 Research Context

We are going to investigate on sentiment analysis using the Cornell Movie Review Dataset. Sentiment analysis is a core task of NLP and we wish to generate new ideas and methods while tackling it. It can also be easily applied to Pizza to fit the CDS team project by informing students the sentiment of answers.

2 Research Objective

We are going to first try different embeddings and algorithms to get a decent 4-class classification result on the movie dataset and hopefully during the process generate new ideas, for example, an embedding using DL.

3 Research Road Map

3.1 Data

We are going to use the Moview Review Dataset by Bo Pang and Lillian Lee. We will focus on the 4-class task:
0: rating <=.3
1: .4 <= rating <=.5
2: .6 <= rating <= .7
3: .8 <= rating
It contains review files in their entireties before passing through tokenization, sentence separation, and subjectivity extraction along with rating scores.

3.2 Experiment Design

Our experimental design is broken up into 5 stages:
1) Pre-processing,
2) Applying existing models and evaluating their performance and choose a proper baseline
3) Improve by generating new embeddings and classifiers
4. Training and fine-tuning
5. Evaluation and conclusions

For pre-processing, we will first apply standard preprocessing methods such as tokenization, stemming, tf-idf as suggested in *The Role of Text Pre-processing in Sentiment Analysis*.

For the baseline, we can start with fitting tf-idf model and Count Vectorizer model in Sciket-Learn to get word embedding and use SVM or other classifiers such as Naive Bayes and LSTM.

In the second stage, we will try to use customized embeddings that produce dense vectors. And meanwhile, play around with LSTM or other neural networks

In the third stage, we will finish constructing the network and fine-tuning the data.

In the last stage, we apply the model and evaluate the results and conduct hypothesis testing to get a conclusion between the baseline method and our own method.

3.3 Experiment Validation

For each review, predict the sentiment out of 4 classes: very negative, negative, positive, very positive.

3.4 Research Timeline

Mar 15: Pre-processing

Mar 30: Try both existing sparse and dense embeddings. Try SVM and LSTM. Figure out and evaulate the baseline

April 15: Try non-traditional embeddings and build a network for classier

April 30: Train the dataset and apply the model to evaluate

May 15: Preliminary conclusion

4 Resources

"Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales.", Proceedings of the ACL, 2005.

5 References

1. "Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales.", Proceedings of the ACL, 2005.

2. Deep Learning for Sentiment Analysis: A Survey

3. The Role of Text Pre-processing in Sentiment Analysis

4. Sentiment Analysis of Review Datasets using Naïve Bayes' and K-NN Classifier