Semi-Supervised Recursive Autoencoders

Adrian Guthals
aguthals@cs.ucsd.edu

David Larson
dplarson@ucsd.edu

Abstract

We evaluate semi-supervised recursive autoencoders (RAE) as a method for predicting the sentiment of sentences. Using random word initialization, we are able to predict the sentiment of a movie review dataset with a 00.0% accuracy, which is comparable to the 00.0% accuracy reported in the 2011 paper "Semi-Supervised Recursive Autoencoders" by Soch et al.

1 Introduction

Socher et al. presented a semi-supervised method for learning meanings of sentences using recursive autoencoders [1].

The lecture notes state blah [2].

Mention: neural networks, sentence meaning/sentiment

2 Semi-Supervised Recursive Autoencoders

RAE, neural networks, backpropogation, error functions, greedy algorithm, calculating derivatives numerically using finite center-difference

3 Experiments

3.1 Datasets

We use the same movie reviews dataset as in [1], which consists of 10662 snippets from reviews posted to the Rotten Tomatoes website[1]. Each snippet is roughly equivalent to a single sentence and includes a positive/negative label, with the entire dataset containing 5331 positive and 5331 negative labelled snippets. For all experiments we have randomly divided the dataset into a 0000 snippet training set ($\sim 70\%$) and a 0000 snippet ($\sim 30\%$) testing set 1. Care has been taken to prevent any snippets from existing in both datasets, so as to not contaminate the results.

3.2 Experiment 1: RAE

The full method (RAE)

3.3 Experiment 2: RAE without Derivatives

RAE without derivatives to adjust the meaning vector of each word

[1] http://www.rottentomatoes.com

Table 1: Number of total snippets (N_{total}), positive snippets (N_{pos}), and negative snippets (N_{neg}) for the original, training, and testing datasets.

Dataset	N_{total}	N_{pos}	N_{neg}
Original	10662	5331	5331
Training	0000	0000	0000
Testing	0000	0000	0000

3.4 Experiment 3: Bag-of-Words

4 Conclusion

Final remarks

References

[1] R. Socher, J. Pennington, E. H. Huang, A. Y. Ng, and C. D. Manning, "Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions," in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2011.

[2] C. Elkan, "Learning meanings for sentences," February 2013.