HW1: Sentiment Classification

Jonah Philion
jonahphilion@college.harvard.edu
github

February 2, 2018

1 Introduction

In this problem set, we attempt to classify movie reviews as positive or negative. All models investigated take the form of

$$p(y_i) = \sigma(W\phi(x) + b)$$

where $p(y_i)$ is the probability a review x is negative. The models studied are

- naive bayes
- logistic regression with "bag of words" features
- multi-layer perceptron with "continuous bag of words" features
- convolutional neural net

And just for fun

- Fine-tuning ResNet classifier on images of text

2 Problem Description

For all models, a sentence $\boldsymbol{x}_i$ is encoded as a sequence $x_1, ..., x_n$ where each x_j is a one-hot vector of length the vocabulary $\mathcal{V}$. The classification y_i associated with $\boldsymbol{x}_i$ is 0 if $\boldsymbol{x}_i$ is positive and 1 if $\boldsymbol{x}_i$ is negative. Embeddings $\mathcal{E}$ map a one hot vector x_j to a dense vector of size d.

3 Model and Algorithms

All models are trained on the Stanford Sentiment Treebank (SST1). Unless otherwise specified, models requiring gradient descent are trained with an Adam optimizer of learning rate 0.001 and weight decay 0.0005. Training loss and validation loss are recorded in real time with visdom.

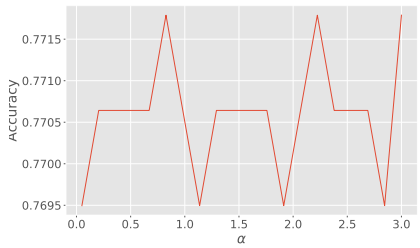

Figure 1: Affect of global smoothing parameter α on validation accuracy.

4 Experiments

4.1 Naive Bayes

4.2 Logistic Regression

4.3 Continuous bag of Words

4.4 CNN

4.5 ResNet

5 Conclusion

End the write-up with a very short recap of the main experiments and the main results. Describe any challenges you may have faced, and what could have been improved in the model.