HW0: Classification

Noah Golowich and Jesse Zhang

January 31, 2018

1 Introduction

In this assignment we experiment with different techniques for sentence sentiment classification. We use the Stanford Sentiment Treebank dataset Socher et al. (2013), which consists of sentences ,each of which is labeled either "positive" or "negative" sentiment. We use a Naive Bayes unigram classifier, a logistic regression model, a continuous bag of words model (using the FastText embeddings Bojanowski et al. (2016)), and a convolutional neural network.

[TODO: extensions]

2 Problem Description

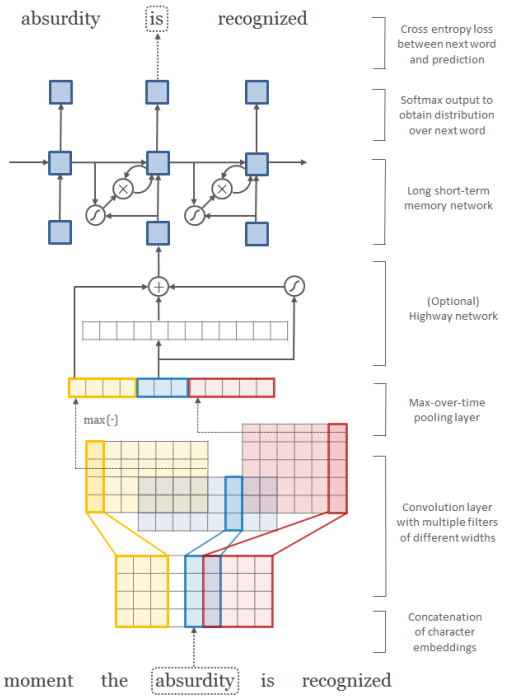
Let $\mathcal{V}$ denote the vocabulary and $\mathcal{T} = \{\text{pos}, \text{neg}\}$ denote the possible labels that can be given to a sentence. We are given a collection of sentences $\mathcal{S} = \{S^1, \ldots, S^N\}$, where each S^j consists of an ordered tuple $(\boldsymbol{x}_1^j, \ldots, \boldsymbol{x}_{\ell(j)}^j)$ of words. Here $\boldsymbol{x}_i^j \in \{0,1\}^{|\mathcal{V}|}$ is a one-hot encoding of the i-th word in the j-th sentence. Corresponding to each S^j is a label $y^j \in \mathcal{T}$. $\ell(j)$ denotes the length of sentence j, $1 \leq j \leq N$; sentences can have varying lengths. Our goal is to predict y^j given S^j.

3 Model and Algorithms

Here you specify the model itself. This section should formally describe the model used to solve the task proposed in the previous section. This section should try to avoid introducing new vocabulary or notation, when possible use the notation from the previous section. Feel free to use the notation from class, but try to make the note understandable as a standalone piece of text.

This section is also a great place to include other material that describes the underlying structure and choices of your model, for instance here are some example tables and algorithms from full research papers:

- diagrams of your model,



- feature tables,

Mention Features	
Feature	Value Set
Mention Head	$\mathcal{V}$
Mention First Word	$\mathcal{V}$
Mention Last Word	$\mathcal{V}$
Word Preceding Mention	$\mathcal{V}$
Word Following Mention	$\mathcal{V}$
# Words in Mention	$\{1, 2, \ldots\}$
Mention Type	$\mathcal{T}$

- pseudo-code,

procedure LINEARIZE($x_1 \ldots x_N$, K, g)
  $B_0 \leftarrow \langle(\langle\rangle, \{1, \ldots, N\}, 0, \boldsymbol{h}_0, \mathbf{0})\rangle$
  for $m = 0, \ldots, M-1$ **do**
    for $k = 1, \ldots, |B_m|$ **do**
      for $i \in \mathcal{R}$ **do**
        $(y, \mathcal{R}, s, \boldsymbol{h}) \leftarrow \text{copy}(B_m^{(k)})$
        for word w in phrase x_i **do**
          $y \leftarrow y$ append w
          $s \leftarrow s + \log q(w, \boldsymbol{h})$
          $\boldsymbol{h} \leftarrow \delta(w, \boldsymbol{h})$
        $B_{m+|w_i|} \leftarrow B_{m+|w_i|} + (y, \mathcal{R} - i, s, \boldsymbol{h})$

Figure 1: Sample qualitative chart.

12: keep top-K of $B_{m+|w_i|}$ by $f(x, y) + g(\mathcal{R})$

13: **return** $B_M^{(k)}$

4 Experiments

Finally we end with the experimental section. Each assignment will make clear the main experiments and baselines that you should run. For these experiments you should present a main results table. Here we give a sample Table 1. In addition to these results you should describe in words what the table shows and the relative performance of the models.

Besides the main results we will also ask you to present other results comparing particular aspects of the models. For instance, for word embedding experiments, we may ask you to show a chart of the projected word vectors. This experiment will lead to something like Figure 1. This should also be described within the body of the text itself.

Model	Acc.
BASELINE 1	0.45
BASELINE 2	2.59
MODEL 1	10.59
MODEL 2	13.42
MODEL 3	7.49

Table 1: Table with the main results.

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.

References

Bojanowski, P., Grave, E., Joulin, A., and Mikolov, T. (2016). Enriching Word Vectors with Subword Information. *arXiv:1607.04606 [cs]*. arXiv: 1607.04606.

Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., and Potts, C. (2013). *Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank*.