

CSC411: Assignment #3

Due on Monday, March 19, 2018

Lukas Zhornyak

March 21, 2018

Environment

Parts 1-6 were created with Python 2.7.14 with numpy 1.14.0, scipy 1.0.0, scikit-image 0.13.1, and matplotlib 2.1.1, as well as all associated dependencies.

Part 1

1 (a)

To improve on the performance of the models previously tested, a convolutional neural network is used, based closely on the model presented in [1]. In summary, this network has multiple one dimensional convolution layers of different filter sizes acting on the input vectors. The maximum activation of these filters across the entire sentence are then combined in a final fully-connected layer. The data is presented to the neural network in a slightly different manner than it was to the previous models, consisting of several vectors each representing one word rather than one vector representing all the words. In this way, the network is able to additional, positional information relating to the headlines.

The various hyperparameters used in the model were those found to be most effective in the original study. Three sizes of filters were used (three, four and five words), each with one hundred filters. Early stopping was employed, so training ended after performance on the validation set did not improve for two epochs. Additionally, the length of a sentence was cut off after forty words owing to memory and performance constraints. As the large majority of headlines have fewer than forty words, this is not expected to reduce performance significantly. The model was trained using stochastic gradient descent with a learning rate of 10^{-4} .

The model was trained on the original dataset presented. The learning curve associated with this training is shown in fig. 1. The final accuracy achieved was 96.24% on the training set, but only 69.38% on the validation set and 68.57%. While the accuracy on the training set is comparable to those of other models, the performance on the other sets is much worse, suggesting a large degree of overfitting and that the relations that the model uses to classify a sentence are not effective for the task of fake news detection.

Additional tests were performed by training the model on the "Liar" dataset [2]. In this set, statements labelled "true", "mostly-true", or "half-true" were considered to represent real news, while other labels signified fake news. The performance in this dataset is similar, albeit worse, to that observed in the previous, achieving 92.38% on the training set, 58.41% on the validation set, and 59.51% on the testing set. However, the model trained on this set performed worse than chance on the original data set, achieving 44.66% accuracy. This suggests that, on top of not being very effective in the first place, this framework in this context does not produce easily transferable models. For this reason, the model trained on the original data set is used for all further purposes.

1 (b)

Since this model examines sequences of words rather than individual words, finding the top words in this context does not have much meaning. Instead, the sequence of words that most activate the filter that has the largest corresponding weight in the fully connected layer are generated. These are listed in These sequences do not seem to form coherent phrases, nor do they share very many words with those were identified for previous model.

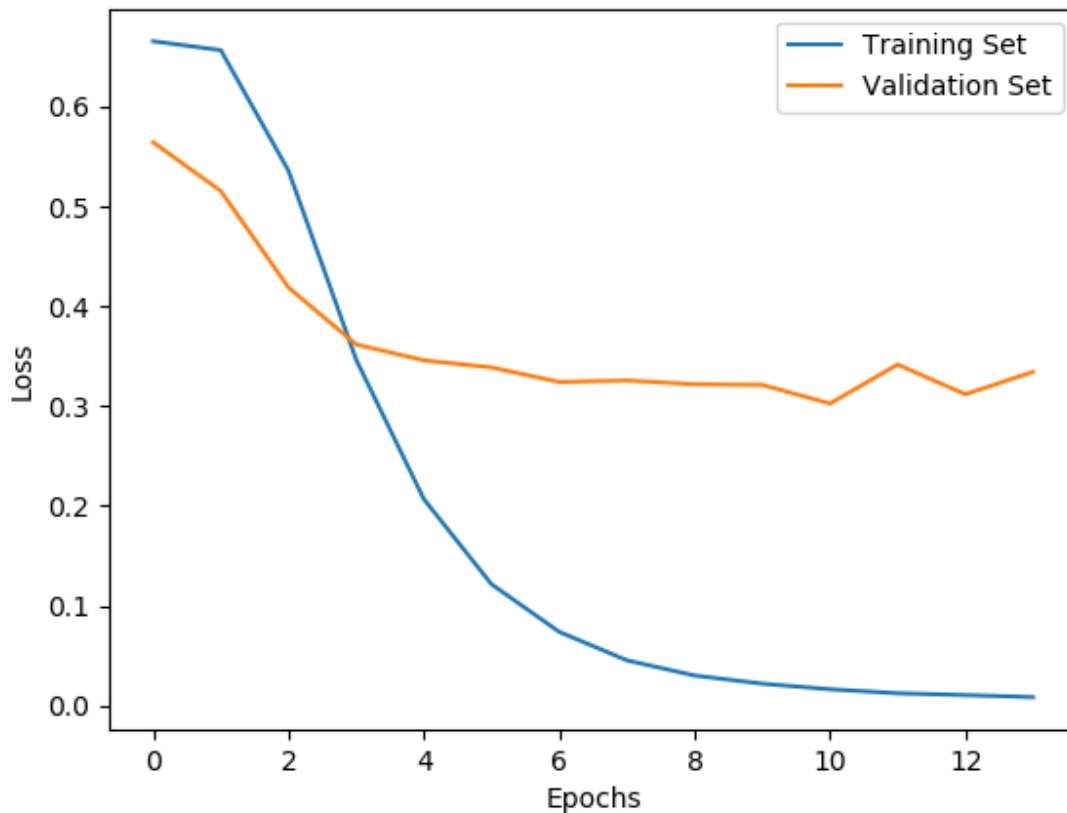


Figure 1: Learning curve for the training and validation sets across epochs. Early stopping was employed, so training ended after performance on the validation set did not improve for two epochs.

Real Sequences

cia chickens tumultuous stir debuts
 analysis contradicts fashion opinion
 tradition load liberals cares
 trumps australia talks
 labels democrats website coy
 trumps australia ban
 killed adjusted responds promoting
 hallways judges with gore
 thrilled exporters softens antique ability
 preemptive guidelines spending repartis refugee

Fake sequences

suddenly silicon message everywhere
 pena sound follows includes
 watch that m
 shredded revolt slams thinks mired
 homeless hillary america
 just an trump
 road henderson military annouces hurricanes
 voting just that
 painter punish off inevitable
 sack duke aware request marblenecltr

References

- [1] Y. Kim, “Convolutional neural networks for sentence classification,” *CoRR*, vol. abs/1408.5882, 2014. arXiv: 1408.5882. [Online]. Available: <http://arxiv.org/abs/1408.5882>.
- [2] W. Y. Wang, ““liar, liar pants on fire”: A new benchmark dataset for fake news detection,” *CoRR*, vol. abs/1705.00648, 2017. arXiv: 1705.00648. [Online]. Available: <http://arxiv.org/abs/1705.00648>.