CMPS242 Homework #5 – Neural Network Tweet Classification

Benjamin Sherman

&

Zayd Hammoudeh

November 9, 2017

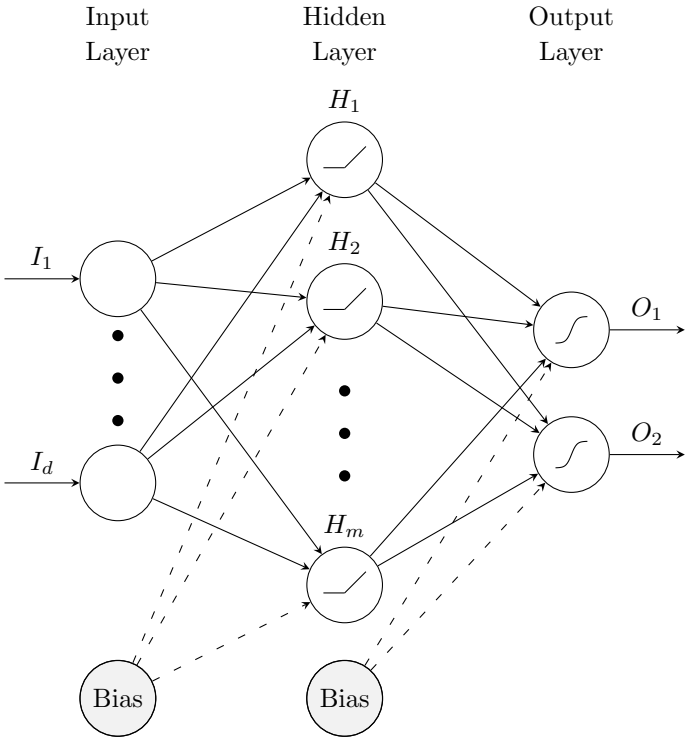

Figure 1: Base Structure of Our Feed-Forward Network

1 Homework Objective

2 Feed-Forward Neural Network Structure

The base structure of our feed forward network is shown in Figure 1. The number of input nodes is dictated by the number of rows in the embedding matrix; as mentioned previously, $d = 25$ in our experiments. Likewise, the number of neurons in the hidden layer was $m = 256$. There were two output nodes (e.g., one for "Donald Trump" and the other for "Hillary Clinton). All nodes had their own bias input to improve performance.

Our final feed-forward network design used the rectified linear and sigmoid activation functions for the hidden and output layers respectively.

3 Extra Credit #1: Bag of Words Model

In the "bag of words" model, each textual input, i.e., tweet, is transformed into an unordered set of words. Hence, the contextual and word ordering information

is discarded. This approach removes any sequential relation in the data; hence, the LSTM added no specific value for training. Hence, we removed the LSTM when performing this experiment and instead trained with just the embedding matrix and the feed-forward network.

Using the previously described neural-network structure, we were able to get 100% accuracy on the complete training set. Likewise, we get an best-case test set of **0.20070** using this approach.

4 Extra Credit #2: Neural Network Experiments

Below are additional experiments we tried beyond the base requirements.

4.1 Extra Credit #2a: Hidden Layer Activation Functions

We experimented with three activation configurations for the hidden layer. In addition to rectified linear, we also tried a "pass-through" activation where the

4.2 Extra Credit #2b: Additional Feed-Forward Hidden Layers