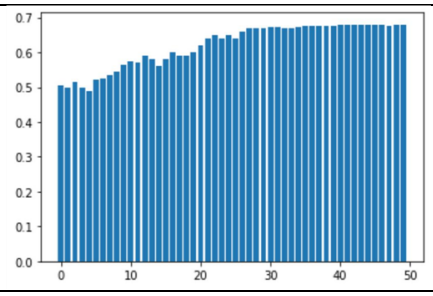
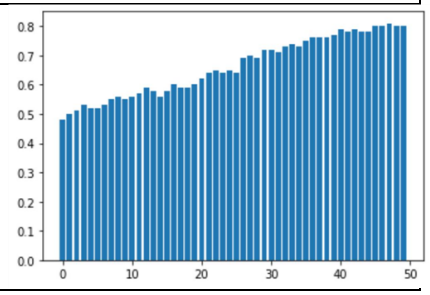
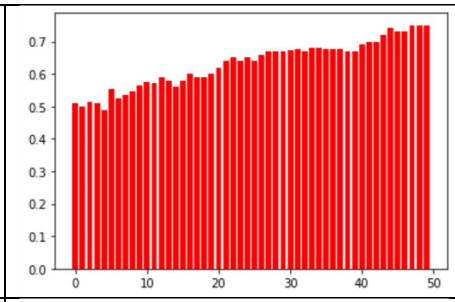
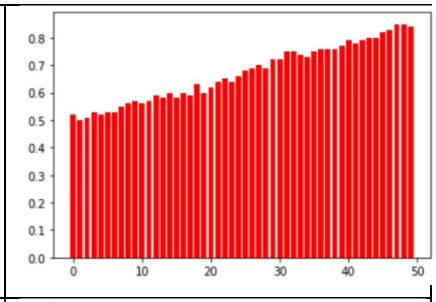
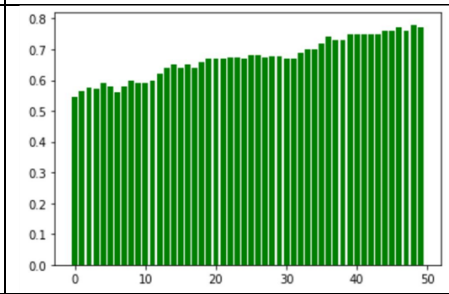
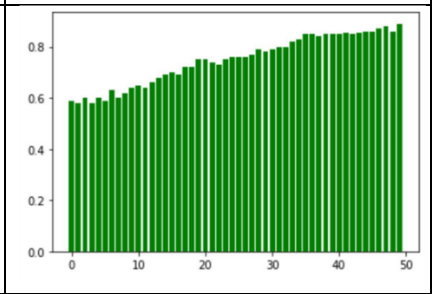
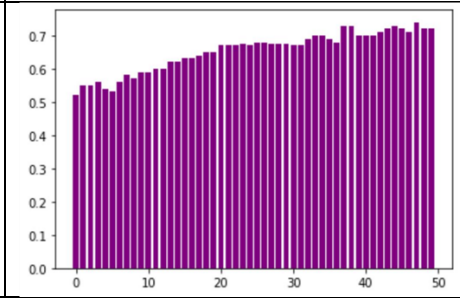
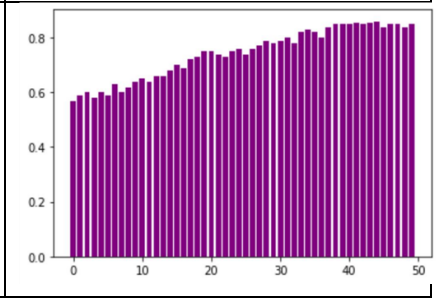
MSDS 422 Assignment #8
James Casey

In order to determine which language modeling method works best for classifying customer service reviews, a recurrent neural network was used to test these different methods. Movie reviews were employed to serve as a guide to how well an RNN model can classify reviews of different language models as either positive or negative. For this study, there where 500 of each type of review that were classified by a thumbs up or thumbs down.

The four languages that were used are Glove.6B.50D, Glove.6B.100D, word2vec.300D, and fastText.300D. Each of these language modeling methods, classify words with vectors in distinct ways, so we will be able to compare different modeling techniques. Furthermore, these methods have a certain number of dimensions per word that are used. By using an RNN model one will be able to determine which one of these methods works best.

For the RNN models, movie reviews of each language models were obtained and put together in lists. For each method, the data was split into training and testing data via a randomized 80:20 split. The RNN model was trained using a random 80% of the data for a total of 50 iterations. Below is a summary of the results of this experiment:

Language Modeling Method Type	**# of Dime-nsions per word**	**Test Accuracy scores (high/ low)**	**Test Accuracy score (mean)**	**Test Iterations**	**Train Accuracy score (high/ low)**	**Train Accuracy score (mean)**	**Train Iterations**
Glove.6B	50	0.695/ 0.505	0.621		0.81/ 0.48	0.661	

Glove.6B	100	0.75/ 0.51	0.632		0.85/ 0.52	0.671	
fastText	300	0.78/ 0.54	0.671		0.88/ 0.57	0.749	
word2vec	300	0.72/ 0.53	0.653		0.86/ 0.56	0.743	

One can recognize the fastText with 300 dimensions per word performs the best out of the 3 language models. It produces the highest accuracy scores out of any model for both test and training sets. One can recognize that the two methods with 300 dimensions performed the best, so a larger number of dimensions may assist in producing more accurate classifications. Furthermore, as more iterations were completed the accuracy score increased. Therefore, more iterations may lead to better scores.

When it comes to recommending a language model method to management, I would recommend the fastText.300D. This method had the most accurate scores out of the four tested. I would recommend for the company to implement a method with a higher dimension per word and to perform more iterations. Consequently this will increase the amount of space and time required, so the business must keep this under consideration. In order to put in place a system

that can help improve the classification of negative reviews, a rating system should be put in place (thumbs up/down, star rating, etc.), employ multiple choice questioning, and have a written review section with a word limit. These along with a language modeling method can help classify negative from positive reviews.