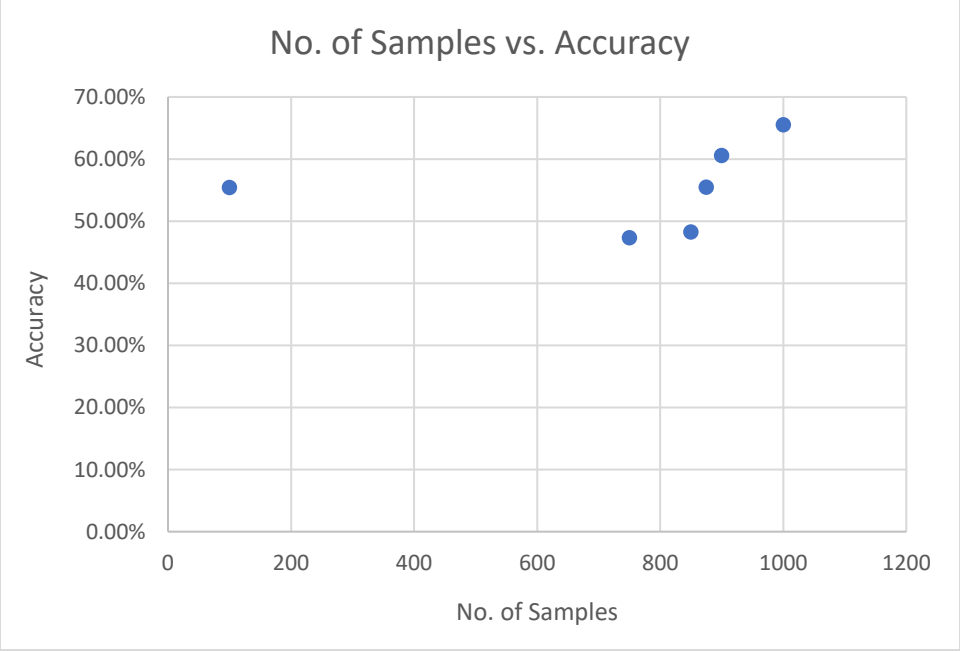
Brandon Palazzo

4/11/21

Advanced Machine Learning Assignment 3

Example	No. of Samples	Accuracy
Embedding Layer Original	25,000	86.36%
Pretrained Layer	100	55.41%
Hypertuned Embedding Layer 1	1000	65.50%
Hypertuned Embedding Layer 2	850	48.24%
Hypertuned Embedding Layer 3	750	47.33%
Hypertuned Embedding Layer 4	900	60.56%
Hypertuned Embedding Layer 5	875	55.43%



After using both an embedding layer and a pretrained layer, I was able to determine that the plain embedding layer was the best approach. Not only was it the simplest method, but it produced the greatest accuracy (86.36% instead of 55.41% for the pretrained layer). After running some hypertuned embedding layers, I was able to determine that the regular embedding layer does better than the pretrained layer once the sample size reaches 875 samples. Anything below 875 samples, it would be best to use a pretrained word embedding layer. I ran the test originally with 1000 samples and

incrementally increased and decreased the number of samples until I reached an accuracy slightly above the pretrained's accuracy of 55.41%. The pretrained technique is best suited for small samples sizes, but once the sample size reaches 875, the best method to use is the regular embedding layer. To model this data, I made a table to show the data values and created a scatter plot to show the relationship between sample size and accuracy. One thing to note for the scatter plot, I left off the original embedding layer with 25,000 samples because it made it too difficult to see the relationship between the other examples on the x-axis.