Jason Chee
Nick Mauthes
Sam Ordonia

Hw3 DNN experiments

1. **What is the effect of training set size on how good the trained model is? How does this compare to the number of weights that your model has?**

Hypothesis: The accuracy of the model will improve as the number of the training examples increases

Description/Approach: Perform training and testing on different size example sets

Results/Conclusion: Accuracy did not improve necessarily with larger training size. It depended on the test examples.

2500 examples -> 94% accuracy
5000 -> 74% accuracy
120000 -> 16% accuracy

We found that as the number of training examples grew very large, the accuracy of the model went decreased. This could be due to the fact as the number of examples increases, a greater number of random mutations occur. This introduces noise which makes it more difficult for the network to train accurately. Since there are 5 densely connected layers of 40 neurons each, there are roughly 40 * 40 * 5 = 8,000 weights in the network, and we can see that as the number of examples increases beyond the number of weights, the accuracy greatly decreases.

2. **What is the difference between training on data chosen completely at random versus on well chosen examples? What is the difference in choosing the testing data at random versus data that is likely to be equally representative of the classes that we are training for?**

Hypothesis: The accuracy will be better when data is well-chosen and not random

Description/Approach: Train on example sets with varying degrees of randomness

Results/Conclusion: Models trained on data where there is a high concentration of one class of palindrome perform better than those trained on random data

When training and testing on mostly NONSTICK examples, the accuracy of the model is quite high. However, when training and testing on other types of labels, the accuracy drops. Despite the presence of six labels, the model seems to only be able to easily distinguish between NONSTICK and non-NONSTICK examples.

Specifically choosing certain datasets does improve the “accuracy” of the testing, particular in the case of the NONSTICK label. When choosing the dataset at random, the accuracy becomes lower on average due to the likelihood of getting a non-NONSTICK label.

3. **What is the difference in accuracy in using cross-validation for testing versus using separate test data**

Hypothesis: Cross-validation will provide a better accuracy than using separate test data.

Description/Approach: Split the dataset between training and testing using k-fold cross-validation. In this experiment, *k* was chosen to be five.

Results/Conclusion: Accuracy Increased with cross-validation. The training size was reduced. From the results of experiment 1, we concluded that accuracy improves with smaller data sets. This result is consistent with the findings of experiment 1.

Cross-validation provides slightly better accuracy results than when testing vs separate test data. Performing cross-validation on the first data folder (data1 created from the shell script provided) gives an average accuracy of 0.01 across the 5 tests. When testing the same training_data against the first test folder (test1 created from the provided shell script), the accuracy is 3.33333e-05.