HW21 – Report

# Analysis Overview

Per Non-profit foundation Alphabet Soup’s request, a total of four Neural Network Models were built to predict whether applicants to their funding program would be successful.

# Results

## Data Preprocessing

* What variable(s) are the target(s) for your model?

The target variable is IS\_SUCCESSFUL (1=YES and 0=NO)

* What variable(s) are the features for your model?

I used CLASSIFICATION and APPLICATION\_TYPE for binning. Both variables had more than 10 unique values so I opted to combine the ones with less repetitions in a category called “other” to reduce fluctuations.

* What variable(s) should be removed from the input data because they are neither targets nor features?

I removed the variables EIN and NAME because they were not relevant to the model or the analysis that was being performed.

## Compiling, Training and Evaluating the Model

* How many neurons, layers and activation functions did you select for your neural network model and why?

A screenshot of a computer program

Description automatically generatedFor the first Model (nn1) we used 3 layers, the third one using a probability activation function. The first layer had 5, the second one 3, and third one 1. The number of features dictated the number of hidden nodes. The overall accuracy for nn1 was 0.725.

A screen shot of a graph

Description automatically generatedA screen shot of a graph

Description automatically generated

For nn2 through nn4 I played with the nodes and layers. The best model was nn2 with 3 layers with 15, 7 and 5 nodes respectively (accuracy = 0.730). I believe that in order to reach that 0.75 accuracy I would have to bring back other variables like NAME or switch to a different type of model.

* Were you able to achieve the target model performance?

No, target accuracy was 0.75 and the closest I got was 0.73. Even when I brought back the NAME column, I got an accuracy of 0.745 which is closer but not enough.

* What steps did you take in your attempts to increase model performance?

In order to increase model performance I added nodes and layers and even tried bringing the NAME variable back, however I wasn’t successful in reaching the 0.75 accuracy requirement.

# Summary

Deep learning models predict and classify information thanks to the addition of nodes and layers. In this particular case, the best we could do was achieve a 0.73 accuracy with the variables that we initially intended to use, however adding back certain variables proved to be useful in improving accuracy. Even so, I believe that Neural Networks is not the right model for this particular set of data and that a Tree Base Model will probably be more accurate and therefore a better fit for this type of analysis.