Neural Network Model Report

Overview: Explain the purpose of this analysis

The purpose of this analysis was to use the provided dataset to create a binary classifier that can predict whether applicants will be successful if funded by Alphabet Soup. I did this by using machine learning and neural networks. The goal was to develop an model with more than 75% accuracy in predicting success. The dataset that was used contains over 34,000 records of all organizations that have received funding from Alphabet Soup.

Results:

* Data Preprocessing
  + **What variable(s) are the target(s) for your model?** The target variable was the IS\_SUCCESSFUL column
  + **What variable(s) are the features for your model?** The remaining 9 columns are considered the features for this model (not including the two that were dropped)
  + **What variable(s) should be removed from the input data because they are neither targets nor features?** The ‘EIN’ and ‘NAME’ columns were dropped since they only contain identification data
* Compiling, Training, and Evaluating the Model
  + **How many neurons, layers, and activation functions did you select for your neural network model, and why?** There were two hidden layers and one output layer, so 3 in total. Then the number of features determined the number of hidden nodes (see picture below) Both hidden layers had 16 neurons each and used the ReLu activation function. This function was chosen since it demonstrated a higher accuracy. While the output layer used the Sigmoid activation function. Graphical user interface, text, email

    Description automatically generated
  + **Were you able to achieve the target model performance?** The goal was to achieve a model performance of 75%, I was unable to achieve this as my final output was only 73%A picture containing letter

    Description automatically generated
  + **What steps did you take in your attempts to increase model performance?** There were a few things that I tried. The first was to add back the ‘NAME’ column, however that only raised the accuracy by .3%. I also tried increasing the number of neurons in the hidden layers however that did not help, in fact it lowered by accuracy by .2%

Summary:

In the end I was not able to hit the target accuracy goal of 75%, the best I could do was 73%. In the end I would recommend keeping the ‘NAME’ column out of the dataset as it did not help in the accuracy and only brought other issues. I would also recommend just changing the type of machine learning applied since tweaking and adjusting the dataset did not seem to work. Maybe something like Random Forest Clustering would be more applicable in this situation.