Neural Network Model Report

Overview

The purpose of this analysis is to create a tool/ algorithm to help select applicants for funding by predicting their chance of success in the company. Neural Networks, which is a type of machine learning model can be used to create these predictions. Neural network allows you to use predictors/ features from the dataset to create a binary classifier to predict which applicant could be successful if funded.

 The data elements captured in this dataset are the following:
- EIN
- Name
- Application Type
- Affiliation
- Classification
- Use_case
- Organization
- Status
- Income amt
- Special considerations
- Ask amt
- Is successful

Results

Step 1: Data Preprocessing
- The target variable in this neural network model is the "IS_ SUCCESSFUL" column which has the information we are trying to predict.
- The features in the model all the other variables except for 'IS SUCCESSFUL', 'EIN', and 'NAME' columns
- EIN and NAME were removed because they were non-beneficial columns, neither target or features

Step 2 : Compiling, Training, and Evaluating the Mode

```
⯈  Model: "sequential_2"

    Layer (type)              Output Shape           Param #
    ================================================================
    dense_9 (Dense)           (None, 100)            4400

    dense_10 (Dense)          (None, 80)             8080

    dense_11 (Dense)          (None, 80)             6480

    dense_12 (Dense)          (None, 200)            16200

    dense_13 (Dense)          (None, 1)              201


    ================================================================
    Total params: 35361 (138.13 KB)
    Trainable params: 35361 (138.13 KB)
    Non-trainable params: 0 (0.00 Byte)
```

Fitting and training a model consists of completing the features and target variable, training the model on the training set and then testing the model on the test data. Then the model metrics such as the accuracy was calculated. 3 models were created to find the optimal model . The activation for these models were, relu which prevents exponential growth in computation required to operate neural networks and sigmoid which makes a function for modeling decision boundaries in binary classification problems. The original model had an epoch of 50, 2 hidden layers and one output layer. This model had accuracy of .728. The 2nd model had 200 epochs, with an accuracy of .7275. The final model had 5 hidden layers and 50 epochs in attempts to optimize the model. This final model had an accuracy of .7296. The third model was the best model, showing the increasing the hidden layers for this dataset impacts mode performance . However the target model performance of 75% accuracy was not met.

Summary

The best neural network model has an accuracy score of .729. This model had multiple hidden layers with relu and sigmoid as activations. Other binary classification models like random forest, logistic regression and SVM could be used to solve this problem. These binary models rely on features and target variables similar to neural networks to predict the target variable.