NEURAL NETWORK REPORT

OVERVIEW

The purpose of this analysis is to predict whether an applicant for grant funding is likely to achieve their stated project goals based on elements of their applications, characteristics of their organizations, and other factors. The results proved to be inconclusive as a result of failing to achieve an accuracy score of 75% or greater

RESULTS

Data Processing

- *Target Variable*
  - IS_SUCCESSFUL
- *Features*
  - APPLICATION_TYPE
  - AFFILIATION
  - CLASSIFICATION
  - USE_CASE
  - ORGANIZATION
  - STATUS
  - INCOME_AMT
  - SPECIAL_CONSIDERATIONS
  - ASK_AMT

Compiling, Training, Evaluating Model

- First Optimization Attempt
  - For the first attempt of the Sequential model, I opted for two hidden layers both with 6 neurons each and both using the relu activation function in order to avoid overfitting the model at the start. I removed the SPECIAL_CONSIDERATIONS column. Considering there were only 27 yes values in the column with over 34,000 rows of data, I felt that this feature was not critical to the analysis. Additionally, I re-binned the APPLICATION_TYPE and CLASSIFICATION bins son the threshold for Other was lowered from 500 to 100 and 1000 to 500 respectively. This first attempt failed with an accuracy score of 0.7264.

- Second Optimization Attempt
  - In the second attempt, I kept all the same parameters from preprocessing as the first attempt, but I changed the activation function in one of the hidden layers from relu to tanh (leaving the output layer activation function as sigmoid). This attempt also failed with an identical accuracy score as the first, 0.7264

- Third Optimization Attempt

- In the third attempt, I kept all the same parameters from preprocessing as the second attempt but I added an additional hidden layer, increased the number of neurons in hidden layer from 6 to 10, and switched all the activation functions to relu (save for the sigmoid function in the output layer). While these modifications failed to achieve the target accuracy, there was a very small improvement in the model performance, resulting in an accuracy score of 0.7296.

SUMMARY

The model failed to produce actionable insights due to the below standard accuracy score. The extant model could potentially produce more accurate results if the SPECIAL_CONSIDERATIONS were added back to the training and test data – it is possible this feature was more important to the analysis than I originally anticipated. Furthermore, additional hidden layers could be built and additional neurons added to each layer to improve accuracy, given that the only (marginal) improvements in accuracy were observed when increasing layers and neurons. In future research, I would recommend the use of a linear regression based model given that previous research has demonstrated that the relationships this analysis hoped to uncover tend to be linear in nature