

Analysis

This analysis will explain how a deep learning model was used to predict whether applicants will be successful if funded by Alphabet Soup. This will help Alphabet Soup select applicants with the best chance of success in their ventures. Data was collected in the form of a CSV file with 34,000 various organizations that have received funding from Alphabet Soup. This data was then preprocessed. Deep learning was then used to create a binary classification model, which split the data into two groups (successful and not successful). Finally, the model was compiled, trained, and evaluated to calculate the loss and accuracy.

Results

Data Preprocessing

- Target: "IS_SUCCESSFUL"
 - This column states if the money was used effectively
- Features: "NAME", "APPLICATION_TYPE", "AFFILIATION", "CLASSIFICATION", "USE_CASE", "ORGANIZATION", "STATUS", "INCOME_AMT", "SPECIAL_CONSIDERATION", "ASK_AMT"
 - Used `pd.get_dummies()` to encode categorical variables so had 71 features total
- Removed: "EIN"
 - Used "NAME" identifier instead of "EIN"

Compiling, Training, and Evaluating the Model

- Layers, Neurons, Activation
 - Both the original model and optimized model had 2 hidden layers
 - Both models utilized 65 neurons for the first layer and 40 neurons for the second layer
 - Experimented with different number of layers and neurons
 - Using 2 to 3 times the number of input features resulted in overfitting, no accuracy increase, and a greater loss, so a lower number of neurons was used
 - Both models used ReLU activation for the layers and Sigmoid function for the output layer
- The first model achieved 72.8% accuracy and 55.9% loss, and the optimized model achieved 78% accuracy and 47% loss

- To increase model performance, the “NAME” feature was added as an identifier
 - Increasing / decreasing the number of layers as well as the number of neurons did not influence the performance of the model
 - Increasing / decreasing the number of epochs did not affect model accuracy

Summary

The optimized model was 78% accurate at predicting whether applicants will be successful if funded by Alphabet Soup. While this is greater than the original 72.8% accuracy, it is not accurate enough to rely on when it comes to distributing funding. The model proved to be most effective when utilizing deep learning with multiple layers. More data could be collected to potentially increase this model's accuracy. However, since deep learning is supervised learning, perhaps another supervised model such as SVM, k-nearest neighbors, or decision tree could provide a higher accuracy.