Alphabet Soup Model Analysis

The nonprofit foundation Alphabet Soup wants a tool that can help it select the applicants for funding with the best chance of success in their ventures. From Alphabet Soup's business team, we have received a CSV containing more than 34,000 organizations that have received funding from Alphabet Soup over the years.

With our knowledge of machine learning and neural networks, we used the features in the provided dataset to create a binary classifier that can predict whether applicants will be successful if funded by Alphabet Soup.

- The target variables for this model is the success of the business ventures of past applicants to the funding program, represented in the dataset by a 1 for successful and a 0 for not successful.
- The feature variables for this model are the name of the business venture, application type, affiliation, classification, use case, organization type, active status, income amount, special considerations for the application, and the amount of funding requested.
- To optimize model output, the unnecessary 'EIN' identification number column was removed.
- The model contained three node layers containing eighty, fifty, and thirty neurons, respectively. The first layer contained a ReLU activation function, while the second and third layers contained sigmoid activation functions.

- Unfortunately, model performance peaked during the second attempt of fine-tuning the model, where we achieved an accuracy score of slightly over 73%.

- Most attempts to optimize the model were made manually. First, we made the addition of using the ‘Name’ column’s value counts to further narrow down the dataframe overall to increase the accuracy of our predictions. Second, we tried several variations of node layers and activation functions in order to find the optimal combination. Adding a third layer did indeed seem to boost accuracy, as well as giving sigmoid activation functions to layers two and three rather than ReLU. Third, we manually increased the epoch count for our model training.

- Finally, we attempted to use the keras tool to optimize the model automatically. We created a function that would loop through the data to test different models in order to find the most optimal. This tool yielded an optimal output of 73%.

In summary, the results of this model only achieved an accuracy score of 73% at peak performance. The target score was 75% which would have been a sufficient benchmark to use for decision-making. Our model was not far off, so there is no doubt that with further fine-tuning, this tool could assist in predicting the success of business ventures who apply for funding with Alphabet Soup Charity, but it is not yet ready for deployment.