Report: Alphabet Soup - Applicant Success Classifier

Overview:

I was tasked with creating a neural network machine learning model that is capable of predicting whether or not an applicant would be successful if funded by Alphabet Soup. Using a dataset containing over 34,000 organizations that have received funding from Alphabet Soup, I trained the model to predict the success of a funding applicant based on metadata about each organization.

Results:

Data Preprocessing:

- The features used in the model are:
    - Application Type
    - Classification
    - Use Case
    - Organization (type)
    - Status
    - Income Amount
    - Special Considerations
    - Ask Amount
- The target in this model is the determination of whether or not an applicant was successful after receiving funding, the column 'IS_SUCCESSFUL'.
- The unique identifiers 'NAME' and 'EIN' were removed from the dataset.
- Both Classification and Application Type were binned to reduce outliers.

Compiling, Training, and Evaluating:

- My initial model contained 2 layers with 20 and 15 neurons respectively, and used the ReLU activation function as a starting point.
- The target model performance was an accuracy score of 0.75. The initial model achieved a score of 0.726.

- Even using several different methods, I was not able to achieve the target score. Binning the Income Amount got a score of 0.7273, adding another node was 0.7277, using the sigmoid activation method scored 0.7265, using more epochs achieved 0.7290, and finally, using a combination of all four had an accuracy of 0.7280.

Summary:

All things considered, an accuracy score of ~72% for predicting success is okay, but in a financial context, it may be too risky to gamble the success on higher funding amounts. For further research, a Principal Component Analysis model may give more accurate results by showing which factor or factors are most influential on the success of a grant applicant, and could isolate those variables within the model.