

## Alphabet Soup Charity Success Predictor

### Overview:

The purpose of this analysis is to generate a deep learning neural network capable of predicting if a charitable organization will be successful or not. The end goal would be to have a recommendation accurate enough to give insight to Alphabet Soup regarding which organizations to give money.

### Results:

#### **Data Preprocessing:**

The target variable used in this dataset is the: IS\_SUCCESSFUL column. This column has either a 1 or 0, corresponding to whether the funding project was successful.

The feature variables are the following columns: 'APPLICATION\_TYPE', 'AFFILIATION', 'CLASSIFICATION', 'USE\_CASE', 'ORGANIZATION', 'STATUS', 'INCOME\_AMT', 'SPECIAL\_CONSIDERATIONS', and 'ASK\_AMT'

The columns included in the dataset that are neither features nor targets, and are therefore removed from the analysis, are: 'EIN', 'NAME'

#### **Compiling, Training, and Evaluating the Model:**

The first attempt at constructing the DL model was a 4-layer network with 86, 43, 21, and 1 (output layer) neuron(s). The first layer was chosen to be 2 x input features and the following layers regress by factors of 2 until the final output layer. The output layer is one neuron because this is a binary classification problem. The activation function used on the hidden layers was 'relu' and the output layer's was 'sigmoid'.

| Layer (type)            | Output Shape | Param # |
|-------------------------|--------------|---------|
| dense_12 (Dense)        | (None, 86)   | 3784    |
| dense_13 (Dense)        | (None, 43)   | 3741    |
| dense_14 (Dense)        | (None, 21)   | 924     |
| dense_15 (Dense)        | (None, 1)    | 22      |
| Total params: 8,471     |              |         |
| Trainable params: 8,471 |              |         |
| Non-trainable params: 0 |              |         |

Figure 1: First attempt DL model

The model did not achieve the target model performance of 75% accuracy. The final accuracy of this model was 72.8%.

Additional layers were added, number of neurons increased within the layers was increased/decreased, and the feature tuning was modified to reduce the number of features the model was trained on. Unfortunately, none of these modifications caused the accuracy to increase above 75%.

## Summary:

Unfortunately, the model did not perform as well as expected and did not achieve the 75% accuracy mark. However, other classification models may outperform this DL model. Neural Networks typically require larger datasets to achieve high accuracy. This dataset only has approximately 30,000 rows, so a simpler model may achieve better accuracy. The Random Forest decision classifier may show better results since the dataset contains categorical features and a relatively small number of entries.