**Performance Report for Deep Learning Homework: Charity Funding Predictor**

**Overview:**

The purpose of this analysis is to overall analyze and report on the performance of the deep learning model that was created for AlphabetSoup. In analyzing and reporting on the functionality of the model, this analysis, also, serves to provide recommendations on how a different model or model(s) could solve this type of a problem.

**Results:**

* **Data Preprocessing** 
  + The variable considered to be the target for my model is if the money was used effectively or not.
  + Originally, variables considered to be the features for my model are the following: application type, affiliation, classification, use case, organization, status, income amount, special considerations, and ask amount. But after compiling, training, and evaluating the model, in order to increase performance, the name of the organization was added as an additional feature.
  + Initially, the variables that were considered neither targets nor features, and were removed from the input data were both the identification columns: the EIN and name of the organization. After training and evaluating the model, the EIN was the only one dropped from the dataset.
* **Compiling, Training, and Evaluating the Model**
  + For my neural network model, I used two hidden layers and one output layer for my initial model and later for the models that were optimize. Because the model is being created to predict the probability of an output, in this case, where the applicants would be successful if funded, the output layer uses the sigmoid activation function. The two hidden layers use the relu activation function, because the values in the dataset aren’t negative and allows for only certain neurons to be activated at a time allowing for a more computational model. The number of neurons used for each of the hidden layers was 16. For the output layer only one neuron was used because the output layer is a binary classification.

**Graphical user interface, text

Description automatically generated with medium confidence**

* + I was able to achieve the target model performance.
  + Steps I took to try and increase the model’s performance was first to add the name of the organizations as another feature, because it allows for more of a distinction increasing accuracy. A cutoff value was designated for the number of names and all the other names were put under other. The number of neurons were changed to the first hidden layer to have 8 and second 16 to prevent the model from overfitting.

**Graphical user interface, text, application, email

Description automatically generated**

**Graphical user interface, text, application

Description automatically generated**

**Summary:**

The overall results of the model were that accuracy was at 78.85% and loss was at 46% concluding that the accuracy of the model was optimized or benefited greatly from the inclusion of the name column in the dataset.

Based on the above, a another type of model that could be used to solve this classification problem is a Random Forest classifier because it reduces the risk of overfitting and provides flexibility, especially with the dataset provided that has more than two features.