**Charity Funding Analysis**

**Overview**

A non-profit, Alphabet Soup, wants to be able to predict how successful their applicants will be if they choose to fund them. There is a csv file with around 34,000 organizations but it has not been analyzed for the purpose they are looking for. A neural network was created to try to if an organization will be successful or not.

**Results**

**Preprocessing**

* The target variable for this model is IS\_SUCCESSFUL.
* The feature variables are: APPLICATION\_TYPE, AFFILIATION, CLASSIFICATION, USE\_CASE, ORGANIZATION, STATUS, INCOME\_AMT, SPECIAL\_CONSIDERATIONS, ASK\_AMT, and IS\_SUCCESSFUL.
* The feature variables that were removed: EIN and NAME were removed for the first and second attempts. For the third attempt USE\_CASE was removed as well.

**Compiling, Training, and Evaluating**

* The first attempt had 2 inner layers. It used the relu activation since it is more efficient in regards to computing and converging than sigmoid. Attempt 1 had 20 neurons in both inner layers and used 1 neuron with sigmoid activation in the outer layer.
* Attempt 1 had a loss of 0.5681890845298767 with an accuracy of 0.7197667360305786, not quite meeting the goal of 75% accuracy.
* Attempt 2 added more neurons to both inner layers from 20 to 60. This resulted in a loss of 0.5754483938217163 with an accuracy of 0.7194169163703918. In the final attempt the USE\_CASE column was dropped, values for CLASS\_COUNT and APP\_COUNT were slightly increased, inner layer 1 was changed to 80 neurons, inner layer 2 changed to 70 neurons, and a third layer was added with 60 neurons with the sigmoid activation. The loss of attempt 3 was 0.5903772115707397 with an accuracy of 0.7188338041305542.

**Summary**

After 3 attempts the accuracy stayed just below 72%. Given more time the data will need to undergo further preprocessing such as removing features or adjusting the values for CLASS\_COUNT and APP\_COUNT. Another model could also be used, possibly a random forest classifier. A random forest classifier could potentially out perform a logistic regression model as this data is not linear and is more complex than originally thought.