**Deep Learning Homework – Charity Funding Predictor**

**Marisa Kiger**

**University of Richmond**

**12.18.2021**

**Overview:**

The goal of this project was to create an algorithm to predict the funding success outcome for applicants. Data was binned by application types and government organization classifications. The categorical data was converted to numerical data to test and train the data.

**Results:**

* **Data Processing**
* The target for the model was the data in the IS\_SUCCESSFUL column.
* The features were the data in the APPLICATION\_TYPE, INCOME, and SPECIAL\_CONSIDERATIONS columns.
* The NAME and EIN columns were dropped from the data.
* **Compiling, Training, and Evaluating the Model**
* The original neural network model included two input layers with 80 and 30 units along with Relu activation, and one output layer of 1 unit with Sigmoid activation. This model was used to match the sample in the assignment. The model was fit using Epochs at 100.
* The original model did not reach the target goal of >75% accuracy. The model returned an accuracy of 72.96%. The model appears to be overfitting the data.
* To better optimize the model and improve the accuracy, I adjusted the model in four attempts.
* Optimization Attempt 1 - In my first attempt to optimize the model, I added extra layers and decreased the units to reduce the complexity of the model. I also changed the activation to Tanh which shows less variability to neuron weights. This yielded a slightly lower accuracy at 72.89%.
* Optimization Attempt 2 – The INCOME\_AMT, SPECIAL\_CONSIDERATIONS, STATUS, and APPLICATION\_TYPE\_Other columns were dropped from the features category. The intent was to reduce any data that may be unnecessary to improve the input. The new data was trained and ran through the original neural network model. This returned an accuracy further reduced at 72.25%.
* Optimization Attempt 3 – I used the reduced features, but added more input layers and changed the activation to Tanh. This was an attempt to create a less complex model by reducing the input units and distributing the data through the hidden layers. The return on this attempt was an accuracy of 72.12%.
* Optimization Attempt 4 – While the assignment only required 3 attempts, I wanted to see if extending the Epochs would yield an accuracy closer to the goal. I used the same model as attempt 2 and increased the Epochs to 250. The accuracy went from 72.25% at Epochs of 50 to 72.89% at Epochs of 250. While this showed improvement the rate was minimal and came with a reduced speed of results.

**Summary:**

All attempts failed to yield an accuracy above 75%. My recommendation for further steps would be to examine changes based on the original data binning. Also using a feature selection model in preprocessing could help isolate the features of importance in the data.