**#A Written Report on the Neural Network Model**

**# Overview of the analysis**

The purpose of the analysis is to predict if the applicants that will be funded by the Alphabet Soup, a charitable organization will be successful. We have been provided with the data set called charity data with approx. 34K applicants, the data contains name, income range, amount requested and the final outcome of funding successful or not denoted by 1 for success and 0 for failure.

The project encompasses three stages;

1. Preprocessing data for applying neuron network
2. Scaling, training and evaluating model and
3. Optimizing the model by dropping irrelevant columns, increasing neurons to the hidden layers and changing activation function of output layers.

The target data for the project is “IS SUCCESSFUL” column and features are other columns in the file.

**#Results**

1. Preprocessing data

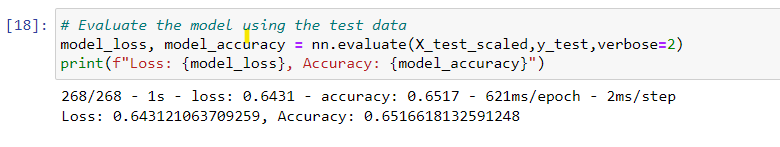
The column named “IS SUCCESSFUL” is the target column (y axis), and I have dropped column names “EIN” and “NAME” from the data frame together with “IS SUCCESSFUL” column from the feature data (X axis)

I have grouped application type count as others where the total count is less than 100.

1. Scaling, training and evaluating model

I have selected two hidden layers, the 1st hidden layer has 80 layers and 2nd hidden layer has 30 layers, I have activation function of sigmoid for the output layer.I have used “relu” activation function for the two hidden layers.

After doing preprocessing of data and activating layers and scalling and training the data set, I ran the model with epochs 30, which gave the accuracy score of 69.67%

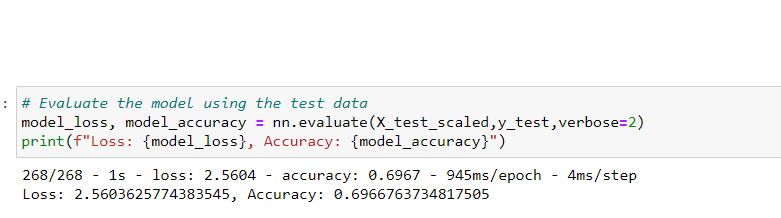


1. Optimizing model in order to achieve 75% accuracy

I have tried three attempts to optimize model in order to get 75% accuracy, I have done following in various attempts

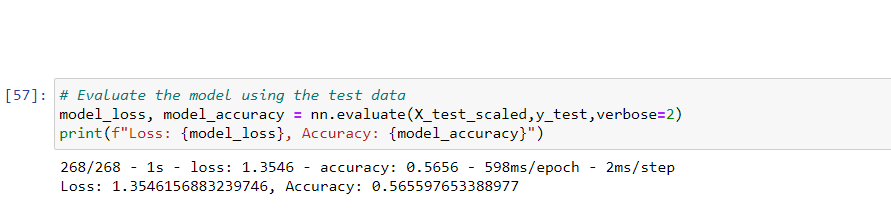
1st attempt- I have removed four features from the data set which seems to me not relevant for the model, those features are 'EIN', 'NAME','USE\_CASE','SPECIAL\_CONSIDERATIONS'

After removing these features and running the neuron network model, I have achieved an accuracy score of 69.67% which is an improvement over initial try i.e 65.17%.



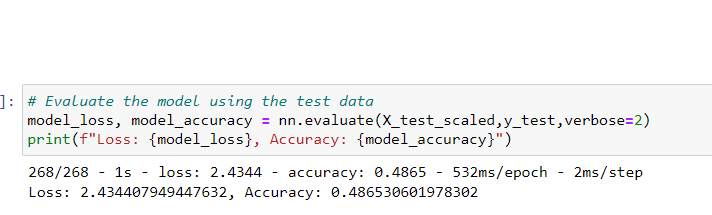
2nd attempt- In the second attempt, I have added one more neuron to the hidden layer, 1st layer had 100 neuron, 2nd layer had 80 neuron and 3rd layer had 30 neurons.

Unfortunately, increasing the neuron had reduced the accuracy of the model to 56.56%



3rd attempt- In my third attempt, keeping the layers and neuron as same as in the 2nd attempt, I have changed the activation function for output layer, the function changed to “tanh” from “sigmoid”.

The accuracy of the model after above change has reduced compared to 2nd attempt, the accuracy score was 48.65%



**#Summary**

From above attempts, it is concluded that, model accuracy will increase only if we drop irrelevant columns and add more data set to the model. Increasing neuron leads to overfitting and hence the accuracy of the model has reduced.

We can try with decision tree model or random forest model to the data set to check if accuracy of the model can be increased to 75%.