Charity Funding Predictor – Report

# Overview

The purpose of this analysis is to use machine learning by way of deep neural networks to find which applicants have the greatest chance of successfully receiving funding.

The dataset comprises of the following fields:

Name  
Application type  
Industry  
Government organisation classification  
Use case for funding  
Organisation type  
Active status  
Income classification  
Special considerations for application  
Funding amount requested   
Organisation was successful or not

# Results

## What variable(s) are the target(s) for your model?

The target variable for the model was the column ‘IS\_SUCCESSFUL’ (Organisation was successful or not)

## What variable(s) are the features for your model?

APPLICATION\_TYPE  
AFFILIATION  
CLASSIFICATION  
USE\_CASE  
ORGANIZATION STATUS   
INCOME\_AMT  
ASK\_AMT

## What variable(s) should be removed from the input data because they are neither targets nor features?

To start with the variables EIN and NAME were removed as they had no bearing on the impact to our model. Follow that, looking into the variables it was uncovered that ‘STATUS’ and ‘SPECIAL\_CONSIDERATION’ both displayed results heavily skewed to one side, roughly 99% of the results. For this reason both these variables were excluded from the model.

## How many neurons, layers, and activation functions did you select for your neural network model, and why?

For the initial Keras model created, two layers were introduced, both using Relu activiation comprising of 50 and 40 nodes respectfully. The decision behind this was it provided a deep enough network for the initial baseline where it would still be efficient to compute a result that was too time consuming.

Baseline – The accuracy derived from this model was 73.08% after 200 epochs. our goal was to achieve 75% accuracy and so began the quest to optimise our model to try achieve this result.

First attempt – increased node across both layers to 80 and 60 respectfully

Accuracy: 0.7316617965698242

Second attempt – Increased layers from 2 to 3 with 80, 30 and 30 nodes respectfully

Accuracy: 0.7320116758346558

Third attempt – Using hyper parameter tuning we were able to get the best model with the following specs:

activation: 'tanh',

first\_units: 9,

num\_layers: 6,

units\_0: 7,

units\_1: 3,

units\_2: 7,

units\_3: 5,

units\_4: 7,

units\_5: 5,

tuner/epochs: 10

Accuracy: 0.7338775396347046

## Summary

From the results we can see the highest achieved accuracy was 0.73378 using hyper parameter tuning, seeing as this approach was restrained in the amount of epochs it ran we can further our modelling by adjusted the best parameters found using this method.

Even though we failed to achieved the desired result of 75% accuracy, what we have learned undertaking this process is that its advisable to carry out hyper parameter tuning early on to then adjust onto it seeing whether we are able to get it closer to our target accuracy.