# **Neural Network Report**

## Overview

The aim of this analysis is to use the features in the provided dataset to create a binary classifier that can predict whether applicants will be successful if funded by Alphabet Soup.

## Results

Data Pre-processing:

* The target for this model is the **IS\_SUCCESSFUL column of the provided dataset. This column indicates whether the money was used effectively.**
* **The features for this model were:**
  + **APPLICATION\_TYPE**—Alphabet Soup application type
  + **AFFILIATION**—Affiliated sector of industry
  + **CLASSIFICATION**—Government organisation classification
  + **USE\_CASE**—Use case for funding
  + **ORGANIZATION**—Organisation type
  + **STATUS**—Active status
  + **INCOME\_AMT**—Income classification
  + **SPECIAL\_CONSIDERATIONS**—Special considerations for application
  + **ASK\_AMT**—Funding amount requested
* Two variables were removed as they were neither targets nor features:
  + **EIN** and **NAME** – Identification columns

### Compiling, Training, and Evaluating the Model

* The initial setup of the neural network consisted of the input and output layers and two hidden layers using the relu activation function. The input and both of the hidden layers had 50 neurons each. This setup was chosen as a simple starting point to see how the model ran.
* I was unable to achieve the 75% target All my attempts were in the range of 72.5-73.5%.
* In trying to increase the performance I introduced dropout layers, extra hidden layers, used the tanh activation function and varied the number of neurons up to as high as 1000 neurons in a layer. I also dropped a couple of the features.

## Summary

Numerous changes were made to this neural network model but none of these changes made any appreciable change to the performance. All results had an accuracy between 72.5 and 73.5%. Quite a lot of time was spent making changes and running them to try and get above the 75% target.

If we use a different binary classifier model we might be able to get better accuracy. A logistic regression model is one option we could use. It’s a supervised learning model so it will be suitable to use with our data set as we have data on which funding applicants were successful.

Another supervised learning model that we could use is a support vector model. This model has the advantage that it can also do non-linear classification which is good for high dimensional features.