

Overview

The purpose of this analysis is to create a model that a non-profit organisation (in this case Alphabet Soup) can use to better select funding candidates. ie select funding applicants that will have a better chance of success. To do this they have provided data on 34000 clients across a number of years, with limited metric and whether they have been successful after funding.

Data Pre-processing information

Target Variable

- IS_SUCCESSFUL column

Feature Variables

- AFFILIATION - Sector of industry
- CLASSIFICATION - Government classification
- USE_CASE - What is their use case
- ORGANIZATION - Organization type
- STATUS - Active status
- INCOME_AMT - Income
- SPECIAL_CONSIDERATIONS - Any special considerations for application
- ASK_AMT - Amount requested

Removed Variables

- EIN
- NAME

Compiling, Training and Evaluation

Many attempts were made to optimise the model. The number of layers, neurons, epochs and activation methods were all varied through models 1 to 3 with no significant changes in accuracy. The final optimised model (model 4) was chosen after running keras-tuner to help decide the parameters. Even this did not yield significant improvements although it did prove to be the most accurate.

The final optimised model achieved an accuracy of only 73% (as below) against a target of 75%.

```
268/268 - 1s - loss: 0.5740 - accuracy: 0.7318 - 509ms/epoch - 2ms/step  
Loss: 0.5740066170692444, Accuracy: 0.7317784428596497
```

There is a lot of data in this set and it may be useful to use more epochs although processing time was a consideration here. Spending time deciding which metrics are truly helpful and if there is any missing data either not collected or omitted by the client could also improve the result.