Charity Funding Predictor: Deep Learning Summary

By: Adam Armagost

**Summary**:

The non-profit foundation Alphabet Soup wants to create an algorithm to predict whether applicants for funding will be successful. With your knowledge of machine learning and neural networks, you’ll 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.

From Alphabet Soup’s business team, you have received a CSV containing more than 34,000 organizations that have received funding from Alphabet Soup over the years. Within this dataset are several columns that capture metadata about each organization, such as the following:

**Findings, Outcome & Recommendation:**

* Findings – Overall modeling efforts yielded the highest accuracy with Neural Network models yielding the highest accuracy in version 2 (79%) and a Random Forest Classification Model (77%).
  + This was done by preforming additional observation binning with donor Names in conjunction with Application Type and Classification binning done in version 1.
* Outcome – The initial modeling efforts had a high level of prediction as to weather a donation would be successful or not; allowing Alphabet Soup to determine success with a ~80% of certainty.
* Recommendation- It would be good to test some additional ML modeling techniques to improve accuracy. However, version 2 of the neural network model did find the donor Name to be an important factor in predicting success.

**Process:**

**Data Cleaning:**

* Dropped Observations
  + EIN – dropped as irrelevant
  + Name – dropped in initial modeling but added back in second version of model
  + Status – dropped for second version for lack of observation diversity
  + Special Considerations - dropped for second version for lack of observation diversity
* Binned Observations
  + Name – Names with less than 5 observations were binned into other for the second round of modeling
  + Application Type – App Type with less than 500 observations were binned into other for both rounds of modeling
  + Classification – Classification with less than 1000 observations were binned into other for both rounds of modeling

Input Observations:

* 'NAME',
* 'APPLICATION\_TYPE',
* 'AFFILIATION',
* 'CLASSIFICATION',
* 'USE\_CASE',
* 'ORGANIZATION',
* 'INCOME\_AMT'

Outcome Observation:

* Is\_Successful
  + 1 = ‘Yes’
  + 2 = ‘No’

**Model Outcomes:**

Neural Network version 1 –

**Loss**: 0.56759113073349

**Accuracy**: 0.7282798886299133

Neural Network version 2 –

**Loss**: 0.47051095962524414

**Accuracy**: 0.7867055535316467

**Model Improvements:**

* + - Addition and binning of donor Name
    - Additional hidden layer +1

Random Forest -

**Accuracy**: 0.7713119533527697