Deep Learning Challenge Report

Overview

The goal of this project is to create an algorithm using machine learning and neural networks to predict whether applicants will be successful if funded by the fictional non-profit foundation, Alphabet Soup.

Process

I was given a CSV file that I read into Pandas. This file contained more than 34,000 organizations that have received funding from the fictional foundation along with several columns of metadata about each organization.

Preprocessing

I preprocessed the data by:

• dropping non-beneficial columns,

• finding the number of data points for each unique value for each of the columns that had more than 10 unique values - APPLICATION_TYPE and CLASSIFICATION,

• choosing a cutoff point of 600 and 300, respectively, to bin rare categorical values together into a new value called "Other",

• using `pd.get_dummies()` to convert categorical data to numeric,

• dividing the data into a target array (IS_SUCCESSFUL) and features arrays,

• applying the `train_test_split` to create a testing and a training dataset,

• and finally, using `StandardScaler` to scale the training and testing sets.

The resulting data included 44 features. The target variable (y) was IS_SUCCESSFUL. The data was split into training and test subsets.

Compiling, Training and Evaluating the model

The model was required to achieve a target predictive accuracy higher than 75%. I made two official attempts using machine learning and neural networks. The first model did not achieve the target accuracy, but the second model achieved the required target accuracy. Results from each model attempt are detailed below:

ATTEMPT #1

The first attempt (Resources/AlphabetSoupCharity1.h5) resulted in an accuracy score of 72.6%. This was the highest accuracy score of the three models. This means that 72.6% of the model's predicted values align with the dataset's true values.

The hyperparameters used were:

- layers = 2
  - layer1 = 8 neurons and 'relu' activation function
  - layer2 = 5 neurons and 'relu' activation function
- epochs = 100

ATTEMPT #2

For my second attempt (Resources/AlphabetSoupCharity2.h5) I added another layer. This attempt resulted in an accuracy score of 75.9%. This

means that 75.9% of the model's predicted values align with the dataset's true values.

The hyperparameters used were:

• layers = 3

o layer1 = 10 neurons : activation function = 'relu'

o layer2 = 8 neurons : activation function = 'relu'

o layer3 = 6 neurons : activation function = 'relu'

• epochs = 30

Summary

In the two attempts I made, the model and able to achieve a target predictive accuracy higher than 75% in second attempt.