21. Deep Learning

Presenter: Prathima. Polavarpu

Charity Funding Predictor

**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 three official attempts using machine learning and neural networks. They all resulted in the same accuracy rate – right around 72%, so a little short of 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 72.5%. This was the highest accuracy score of the three models. This means that 72.5% of the model’s predicted values align with the dataset’s true values.

The hyperparameters used were:

• layers = 2

o layer1 = 9 neurons and ‘relu’ activation function

o layer2 = 18 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 72.89%. This means that 72.6% of the model’s predicted values align with the dataset’s true values.

The hyperparameters used were:

• layers = 3

o layer1 = 9 neurons : activation function = ‘relu’

o layer2 = 18 neurons : activation function = ‘relu’

o layer3 = 27 neurons : activation function = ‘relu’

epochs = 100

**ATTEMPT #3**

For my third and final attempt (Resources/AlphabetSoupCharity3.h5) I kept the third layer and changed the activation function for layers 2 and 3. This attempt resulted in an accuracy score of 76.45%. This means that 76.45% of the model’s predicted values align with the dataset’s true values.

The hyperparameters used were:

• layers = 3

o layer1 = 9 neurons : activation function = ‘relu’

o layer2 = 18 neurons : activation function = ‘tanh’

o layer3 = 27 neurons : activation function = ‘tanh’

• epochs = 100

**Summary**

In the three attempts I made, the last model was able to achieve a target predictive accuracy higher than 76%. Hyper tuning resulted in improvement.