OBJECTIVE

The project's objective is to construct a machine learning algorithm with neural networks to forecast the success likelihood of applicants receiving funding from the fictional nonprofit organization, Alphabet Soup.

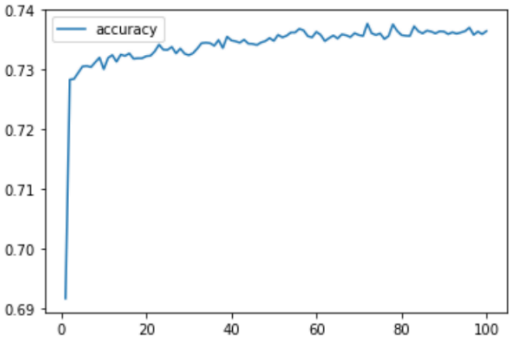
PRE-PROCESSING

1. Removing unnecessary columns,
2. Calculating the count of data points for each unique value in columns with more than 10 unique values, namely APPLICATION_TYPE and CLASSIFICATION,
3. Grouping rare categorical values (below 600 for APPLICATION_TYPE and 300 for CLASSIFICATION) into a new category called "Other",
4. Converting categorical data to numeric using pd.get_dummies(),
5. Splitting the data into target (IS_SUCCESSFUL) and feature arrays,
6. Employing train_test_split to create training and testing datasets,
7. Lastly, applying StandardScaler to standardize the training and testing sets.

COMPILE, TRAIN AND EVALUATE

The objective was to attain a predictive accuracy surpassing 75%. I conducted three formal attempts employing machine learning and neural networks. All three yielded a consistent accuracy rate, hovering at approximately 72%, falling slightly short of the desired target accuracy.

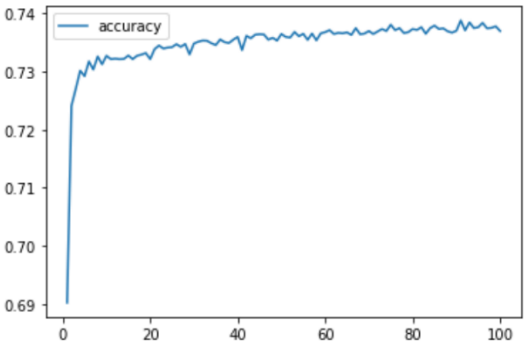
FIRST MODEL



In the initial attempt (using Resources/AlphabetSoupCharity1.h5), an accuracy score of 72.8% was achieved. This marked the highest accuracy among the three models, signifying that 72.8% of the model's predictions matched the actual values in the dataset.

SECOND MODEL

In my second endeavor (utilizing Resources/AlphabetSoupCharity2.h5), I introduced an additional layer to the model. This adjustment led to an accuracy score of 72.6%, indicating that 72.6% of the model's predictions corresponded to the actual dataset values.



THIRD MODEL

In my third and concluding effort (utilizing Resources/AlphabetSoupCharity3.h5), I maintained the third layer while altering the activation functions for layers 2 and 3. This iteration yielded an accuracy score of 72.7%, indicating that 72.7% of the model's predictions were in accordance with the actual dataset values.