Analysis of Deep Learning Model Performance of Alphabet Soup Charity

Purpose of the analysis

The purpose of this analysis is to evaluate the performance of a deep learning model created using TensorFlow/Keras for Alphabet Soup Charity. The model aims to predict the success of funding applications helping the organization make informed decisions about allocating resources effectively.

Data preprocessing

Target variable: The target variable for this model 'IS SUCCESSFUL', indicating whether an applicant was successful after receiving funding (1 for successful 0 for unsuccessful).

Features: All columns except 'IS SUCCESSFUL', 'EIN', and 'NAME' are considered features for the model .

Variables to Remove: 'EIN' and 'NAME' we are removed as they are identifiers and don't contribute to the predictive power of the model.

Compiling Training and Evaluating the Model

Model Architecture: The model consists of an input layer with 128 neurons, a hidden layer with 64 neurons, and an output layer with a sigmoid activation function. ReLU Activation was chosen for hidden layers due to its ability to handle nonlinearities effectively.

Achievement of Target Performance: The model achieved an accuracy of approximately 72.40% on the test set.

Steps to Increase Performance: Strategies include binning low-frequency categories, scaling the data, and experimenting with different neural network architectures.

Results

What's variables at the targets for your model?

- The target variable is 'IS SUCCESSFUL'.

What variables are the features for your model?

- All columns except 'IS SUCCESSFUL, 'EIN', and 'NAME'.

What variables should be removed from the input data becausr they are neither targets nor features?

- ‘EIN’, and ‘NAME’, were removed.

How many neurons layers and activation functions did you select for your neural network model and why?

- The model has an input layer with 128 neurons, a hidden layer with 64 neurons and a sigmoid activation function in the output layer. ReLU activation was chosen for hidden layers due to its effectiveness in handling nonlinearities.

Were you able to achieve the target model performance?

- The model achieved an accuracy of approximately 72.40%

What steps did you take in your attempts to increase model performance?

- Binning low-frequency categories, scaling data, and experimenting with different architectures were undertaken to improve performance.

Overall Model Summary

The deep learning model achieved a reasonable level of accuracy in predicting funding application success. Further improvements could be made through feature engineering, hyperparameter tuning, and exploring different model architectures.

Using a Different Model Approach

One alternative model that could be explored is a Gradient Boosting Machine or XGBoost. GBM models often perform well in tabular data classification tasks and can handle categorical variables effectively. They are robust against overfitting and can capture complex relationships in the data. Using GBM could provide a different perspective on the problem and potentially improve predictive performance.