

Purpose:

The purpose of this analysis is to accurately predict whether or not a project will be successful after being funded by AlphabetSoup. If AlphabetSoup is more able to more accurately predict if a project is going to be more successful, then they will be able to allocate more funds to such projects. This will then make a bigger impact on the causes AlphabetSoup is trying to support.

Preprocessing:

- The target variable for this model is whether the funded project is successful. This was represented by the IS_SUCCESSFUL column in the dataset.
- The following columns are features:
 - APPLICATION_TYPE
 - AFFILIATION
 - CLASSIFICATION
 - USE_CASE
 - ORGANIZATION
 - STATUS
 - INCOME_AMT
 - SPECIAL_CONSIDERATIONS
 - ASK_AMT
- The following columns were removed from the dataset, as they are not relevant to the outcome:
 - EIN
 - NAME

Compiling, Training, and Evaluating the Model

- The following image shows the summary of the model

Model: "sequential"

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 39)	1599
dense_1 (Dense)	(None, 37)	1480
dense_2 (Dense)	(None, 59)	2242
dense_3 (Dense)	(None, 81)	4860
dense_4 (Dense)	(None, 1)	82
Total params: 10,263		
Trainable params: 10,263		
Non-trainable params: 0		

-

- I was unable to reach target performance. I tried a couple of different methods for tuning the model, but I believe more data processing is required to achieve >75% accuracy.
- In order to reach the target performance, I tried increasing the number of epochs and the number of neurons. Then I used a function to test different combinations of neurons and activation functions. None of these combinations reached 75%.

Summary:

Overall, my model approached, but did not reach target performance. It is likely that additional data cleaning is required to reach the target performance. Perhaps making sure there is a normal distribution of certain features or removing outliers could improve model performance. Another model that could be useful in this situation is a Random Forest Classifier. It would be interesting to see the main drivers of the decisions and potentially put those features back into the neural network.