Deep Learning Summary Report

Performance of the Deep Learning Model

- Data Preprocessing
  - The target variable for the model is the 'IS_SUCCESSFUL' feature in the csv file.
  - The features of the model are APPLICATION_TYPE, AFFILIATION, CLASSIFICATION, USE_CASE, ORGANIZATION, STATUS, INCOME, SPECIAL_CONSIDERATIONS, and ASK_AMT.
  - The EIN and NAME columns are not features or targets and have been removed.
- Compiling, Training, and Evaluation of the Model
  - Is the initial model, I selected 2 hidden layers, with 8 and 5 neurons, respectively. I chose these for a baseline assessment only.
  - With this initial model, I was only able to achieve a 72.9% accuracy, less than the desired 75%.
  - I created 5 additional neural models to attempt to achieve the 75% accuracy performance.
    - Attempt 2 – added a hidden layer with 5 neurons. 72.6% accuracy.
    - Attempt 3 - binned corporation and co-operative organization types. 72.6% accuracy
    - Attempt 4 – lowered epochs to 25 total. 73% accuracy
    - Attempt 5 – lowered 1st hidden layer to 5 neurons. 72.7 accuracy
    - Attempt 6 – added 2 additional hidden layers for 4 total. Changed neurons to 8 in first 3 hidden layers and 5 in 4th hidden layer. Changed to 50 epochs. 72.4% accuracy.
  - I attempted a different model using supervised learning and Random Forest Classification with Logistic Regression for attempt 7. It returned a 71.5 % testing score.
- In summary, the deep learning model attempted did not perform at the 75% accuracy requested. I believe an unsupervised model using PCA to reduce the features would be a likely candidate to improve the model accuracy as it may be a better classification tool for this dataset.