Neural Network Modeling for a funding decision

Overview: The purpose of this analysis is to help select applicants for funding to Alphabet Soup Charity who are most likely to be successful. By training a model on past data, the model attempts to predict the success or failure of a new applicant based on that new candidate’s information.

Results:

Data Preprocessing

* + - The variable IS\_SUCCESSFUL is the target for the model.
    - The variables APPLICATION\_TYPE, AFFILIATION, CLASSIFICATION, USE\_CASE, ORGANIZATION, STATUS, INCOME\_AMT, SPECIAL\_CONSIDERATIONS, AND ASK\_AMT are features of the model. APPLICATION\_TYPE was removed for one optimization attempt.
    - EIN and NAME were removed in preprocessing because they are neither targets nor features.

Compiling, Training and Evaluation:

* + - I initially selected two hidden layers and increased this to four in one attempt at optimization. I chose relu and sigmoid functions initially, and gelu in one attempt at optimization because it is a significantly different function which weights inputs by value according to documentation: https://keras.io/api/layers/activations/
    - I was not able to achieve the target model performance of 75% or greater.
    - Steps I took included increasing numbers of hidden layers, using gelu instead of relu activation function and removing application type from consideration.

Summary: The overall results were not effective. It is possible the data is not conducive to accuracy with machine learning. I can recommend trying an ensemble method which involves a decision tree. Toshi Takeuchi presents an example using what in some ways is a similar data set to ours (multiple categorical columns in the data set, and one numeric wanting to predict a binary outcome) https://blogs.mathworks.com/loren/2015/06/18/getting-started-with-kaggle-data-science-competitions/#539630b4-502f-404d-bf05-423ebb2248b6