Deep Learning Homework Write up

1. Overview - the purpose of this analysis is to identify the charity organization with the highest chance of success based on a known data set of charity organizations. we will take this information and leverage machine learning and neural networks to predict the chances of success.

2.

Data Preprocessing

- What are the targets for the model? "is\_successful" is the targeted outcome

what variables are considered the features for your model? The application type and classification

- the variables that were neither targets or features included the columns of data that were removed. EIN, NAME, affiliation, use case, organization, income amount, and special considerations.

Compiling, Training, and Evaluating the Model

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

model 1:

neurones: 16

application cutoff - 500

classification cutoff - 500

hidden layers -1 (relu)

epochs - 50

Results:

loss: 0.5548

accuracy: 0.7252

Model 2

neurons: 16

application cutoff - 500

classification cutoff - 500

hidden layers -3 (added a second Relu and a Tanh layer)

epochs - 50

Results:

loss: 0.5535

accuracy: 0.7283

Model 3

neurones: 16

application cutoff - 80

classification cutoff - 500

hidden layers -3 (added a second Relu and a Tanh layer)

epochs - 75

Results:

loss: 0.5545

accuracy: 0.7301

the fit model history showed a more erratic trend with this setup compared to the prior variations. likely due to the accidental reversal of the application and classification cutoff scores.

Model 4

neurones: 16

application cutoff - 500

classification cutoff - 80

hidden layers -3 (added a second Relu and a Tanh layer)

epochs - 50

Results:

loss: 0.5500

accuracy: 0.7314

Were you able to achieve the target model performance?

- results of varied models landed in the range of 73%, unable to break through the 75% threshold.

What steps did you take to try and increase model performance?

- adjustments to the layers, cutoff scores to the applications and classification features were attempted, and adjustments to the epoch counts. the results showed to be roughly 73% accuracy.

Summary - the dataset had some variables that were removed as they didn’t have large enough counts, additionally some columns were dropped as they provided no value for the model, such as EIN and Name. Overall, predictions in the 73% accuracy range were obtained, while this is just under the line of 75% this model was close. Continued drill downs could obtain a higher rating with additional time and resources.