Deep Learning Homework- Charity Funding Predictor

**Objective:**

The objective of the Charity-funding-predictor project is to predict which applicants for funding will be successful. We processed the data by removing data which was not needed and converted other columns to numeric values using Get\_Dummies.

**Source Data Analysis:**

The target column for this dataframe is Is-Successful.

Feature columns include APPLICATION\_TYPE, AFFILIATION, CLASSIFICATION, USE\_CASE, ORGANIZATION, STATUS, INCOME\_AMT, SPECIAL\_CONSIDERATIONS, ASK\_AMT

Columns that are neither features or targets are EIN and NAME.

Below is a summary of some of the testing of the model that I performed.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Notebook | # of units | Number of hidden layers | Activation | # epochs | Final training accuracy | Testing Accuracy |
| Starter\_Code.ipynb | Layer 1=80  Layer 2= 30 | 2 | Relu | 15 | 0.9645 | 0.5939358472824097 |
| Starter\_Coderun\_2.ipynb | Layer 1=50  Layer 2=70 | 2 | Relu | 35 | 0.9678 | 0.6165597438812256 |
| Starter\_Coderun\_3.ipynb | Layer 1=100  Layer 2=100 | 2 | Relu | 75 | 0.9710 | 0.649562656879425 |
| Starter7 copy.ipynb | Layer 1=80  Layer 2=30  Layer 3=80 | 3 | Relu | 35 | 0.9676 | 0.7542856931686401  (Meets threshold) |
| Starter9.ipynb | Layer 1=80  Layer 2=80  Layer 3=80 | 3 | Relu | 100 | 0.9720 | 0.7336443066596985 |

The file which had a testing accuracy of .754 was starter7 copy.ipynb

**Conclusion:**

I ran an iterations changing parameters and got a testing accuracy of 0.7547, with is above the threshold rate of .75. To do this I added a third layer with 80 hidden nodes. This iteration only had 35 epochs. I chose these parameters after trying several iterations, and this 35 epochs seemed to be the right amount, along 3 hidden layers.

When I increased the second hidden layer 80 hidden nodes the testing accuracy decreased. For those runs with a total of 100 epochs the testing accuracy decreased slightly to .733, while the training accuracy improved slightly. Perhaps the model was overfit for the training data, and not as accurate for testing data.

Although the training accuracy was .968 the testing accuracy was only .754.

Recommendations to improve data accuracy further will be to add more hidden layers with the same number of nodes and about 35 epochs.

Recommendations to use a different model might include Logistic regression, since it uses a binary model to predict for a target. I this case the target will predict either be successful or not successful.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |