CISC 452 assignment 1 report

Zili Luo 20001744

**Self Based model**

I selected a 4-input unit fully connected to 3 output unit with sigmoid activate function as my structure of the ANN.

A

R

G

M

A

X

Inputs output

The initial values are all set to zeros, as i found randomized initial value bring very unstable performance. The final value of the weight tensor and bias tensor are listed as follows:

Weight : [[0.98 4.55372334 -9.00216866]

[1.54 -5.41183615 -9.5839529]

[-2.24 2.86118317 12.7408953]

[-1.17999983 -6.40912533 10.7291946]]

Bias: [[0.98 4.55372334 -9.00216866]

[1.54 -5.41183615 -9.5839529]

[-2.24 2.86118317 12.7408953]

[-1.17999983 -6.40912533 10.7291946]]

As I have used same update method for both tensor, so the final value are the same.

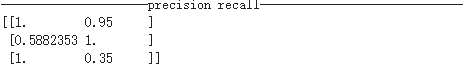
The training iterations is set to 200 epoch with 120 training data(full data) per epoch with initial learning rate 0.2 and multiply 0.99 per epoch, which brings a final learning rate at 0.0268

I also found that add a final argmax layer to filter the output helps the performance of the model

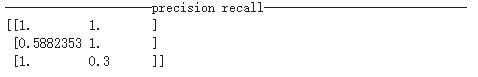
As the optimize algorithm is selected as pocket algorithm with Ratchet, so the training step have no early-stop until it runs through all the steps.

The precision-recall matrix are list as follows:

For training set:



For testing set:



The confusion matrix are list as follows:

For training set:

|  |  |  |  |
| --- | --- | --- | --- |
| act  prd | setosa | versicolor | virginica |
| setosa | 38 | 2 | 0 |
| versicolor | 0 | 40 | 0 |
| virginica | 0 | 26 | 14 |

For testing set.

|  |  |  |  |
| --- | --- | --- | --- |
| act  prd | setosa | versicolor | virginica |
| setosa | 10 | 0 | 0 |
| versicolor | 0 | 10 | 0 |
| virginica | 0 | 7 | 3 |

**The Tool Based Model**

For tool based model, i choose keras as the library. To build the same structure as my self crated model.But I choose different optimizer, loss function and activation function.

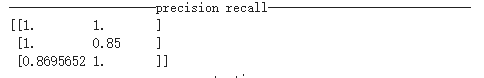
For the optimizer i choose SGD, and cross\_entropy as my loss function.

For the activation function a choose softmax which is similar as the final argmax function on my self crated model.

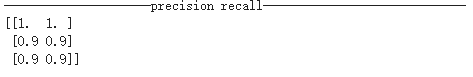
All the other parameter value are set to default.

The precision-recall matrix are list as follows:

For training set:



For testing set:



The confusion matrix are list as follows:

For training set:

|  |  |  |  |
| --- | --- | --- | --- |
| act  prd | setosa | versicolor | virginica |
| setosa | 40 | 0 | 0 |
| versicolor | 0 | 34 | 6 |
| virginica | 0 | 0 | 40 |

For testing set:

|  |  |  |  |
| --- | --- | --- | --- |
| act  prd | setosa | versicolor | virginica |
| setosa | 10 | 0 | 0 |
| versicolor | 0 | 10 | 0 |
| virginica | 0 | 7 | 3 |

PS. I provide both .ipynb and .py file. But since I coded on the colab. So .ipynb file is expected have a better performance.