CS-596 Machine Learning Homework Assignment 4

Hsuan Yu, Liu Red ID:823327369

**Part I: FNN Model,**

**Question 1:**

1. First experiment🡺

Hyperparameter:

Number of hidden layers : 2

Number of neural units for each layer: 120

Learning rate : 0.001

Activation function : Softmax, ReLU

Output function : Linear

Other parameters:

Epoch : 20

Batch size : 100

Accuracy: 0.8761

Confusion matrix:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Pred: 0 | Pred: 1 | Pred: 2 | Pred: 3 | Pred: 4 | Pred: 5 | Pred: 6 | Pred: 7 | Pred: 8 | Pred: 9 |
| GT:0 | 951 | 0 | 8 | 6 | 3 | 6 | 10 | 1 | 12 | 0 |
| GT:1 | 0 | 1107 | 9 | 2 | 2 | 8 | 1 | 3 | 12 | 0 |
| GT:2 | 15 | 12 | 804 | 15 | 20 | 6 | 35 | 15 | 25 | 6 |
| GT:3 | 7 | 8 | 34 | 857 | 1 | 55 | 4 | 17 | 38 | 10 |
| GT:4 | 2 | 7 | 5 | 2 | 835 | 3 | 18 | 1 | 6 | 49 |
| GT:5 | 15 | 11 | 11 | 44 | 14 | 703 | 35 | 6 | 37 | 9 |
| GT:6 | 9 | 6 | 30 | 0 | 12 | 18 | 909 | 0 | 9 | 0 |
| GT:7 | 10 | 17 | 13 | 3 | 8 | 2 | 2 | 942 | 11 | 38 |
| GT:8 | 5 | 30 | 26 | 52 | 15 | 31 | 13 | 11 | 825 | 27 |
| GT:9 | 5 | 3 | 6 | 16 | 52 | 7 | 0 | 59 | 12 | 828 |

Recall:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 0.95386 | 0.96766 | 0.84365 | 0.83123 | 0.89978 | 0.79435 | 0.91541 | 0.90057 | 0.7971 | 0.83806 |

Precision:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 0.93327 | 0.92173 | 0.84989 | 0.85958 | 0.86798 | 0.8379 | 0.8851 | 0.89289 | 0.83587 | 0.85626 |

2. Second experiment🡺

Hyperparameter:

Number of hidden layers : 2

Number of neural units for each layer: 120

Learning rate : **0.1 (Edited)**

Activation function : Softmax, ReLU

Output function : Linear

Other parameters:

Epoch : 20

Batch size : 100

Accuracy: 0.9699

Confusion matrix:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Pred: 0 | Pred: 1 | Pred: 2 | Pred: 3 | Pred: 4 | Pred: 5 | Pred: 6 | Pred: 7 | Pred: 8 | Pred: 9 |
| GT:0 | 987 | 0 | 0 | 1 | 1 | 2 | 3 | 0 | 2 | 1 |
| GT:1 | 0 | 1119 | 3 | 2 | 0 | 3 | 0 | 4 | 12 | 1 |
| GT:2 | 9 | 5 | 913 | 2 | 5 | 1 | 4 | 5 | 7 | 2 |
| GT:3 | 0 | 0 | 8 | 968 | 0 | 27 | 0 | 1 | 23 | 4 |
| GT:4 | 1 | 1 | 0 | 0 | 882 | 1 | 5 | 1 | 4 | 33 |
| GT:5 | 2 | 0 | 2 | 2 | 1 | 865 | 4 | 0 | 5 | 4 |
| GT:6 | 6 | 3 | 1 | 0 | 3 | 5 | 967 | 1 | 7 | 0 |
| GT:7 | 3 | 1 | 3 | 0 | 0 | 2 | 0 | 1008 | 2 | 27 |
| GT:8 | 4 | 1 | 2 | 2 | 1 | 4 | 1 | 1 | 1011 | 8 |
| GT:9 | 0 | 0 | 0 | 1 | 2 | 2 | 0 | 2 | 2 | 979 |

Recall:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 0.98997 | 0.97815 | 0.95803 | 0.93889 | 0.95043 | 0.9774 | 0.97382 | 0.96367 | 0.97681 | 0.99089 |

Precision:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 0.9753 | 0.99027 | 0.97961 | 0.98978 | 0.98547 | 0.94846 | 0.98272 | 0.98534 | 0.94047 | 0.92446 |

3. Third experiment🡺

Hyperparameter:

Number of hidden layers : 2

Number of neural units for each layer: **300(Edited)**

Learning rate : 0.001

Activation function : Softmax, ReLU

Output function : Linear

Other parameters:

Epoch : 20

Batch size : 100

Accuracy: 0.9048

Confusion matrix:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Pred: 0 | Pred: 1 | Pred: 2 | Pred: 3 | Pred: 4 | Pred: 5 | Pred: 6 | Pred: 7 | Pred: 8 | Pred: 9 |
| GT:0 | 957 | 0 | 6 | 3 | 2 | 9 | 11 | 3 | 6 | 0 |
| GT:1 | 0 | 1115 | 7 | 1 | 1 | 8 | 1 | 4 | 7 | 0 |
| GT:2 | 9 | 8 | 834 | 8 | 19 | 5 | 21 | 14 | 31 | 4 |
| GT:3 | 4 | 4 | 32 | 894 | 0 | 45 | 3 | 8 | 29 | 12 |
| GT:4 | 1 | 6 | 7 | 1 | 852 | 4 | 8 | 3 | 7 | 39 |
| GT:5 | 10 | 5 | 3 | 33 | 23 | 744 | 22 | 3 | 26 | 16 |
| GT:6 | 8 | 6 | 9 | 1 | 6 | 17 | 932 | 2 | 12 | 0 |
| GT:7 | 9 | 15 | 17 | 4 | 10 | 4 | 0 | 947 | 2 | 38 |
| GT:8 | 7 | 20 | 9 | 25 | 14 | 29 | 13 | 2 | 901 | 15 |
| GT:9 | 4 | 2 | 1 | 16 | 44 | 6 | 0 | 35 | 8 | 872 |

Recall:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 0.95988 | 0.97465 | 0.87513 | 0.86712 | 0.9181 | 0.84068 | 0.93857 | 0.90535 | 0.87053 | 0.88259 |

Precision:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 0.94846 | 0.94412 | 0.90162 | 0.90669 | 0.87745 | 0.85419 | 0.92186 | 0.92752 | 0.87561 | 0.8755 |

**Question 2:**

Top ranked model🡪

Hyperparameter:

Number of hidden layers : 2

Number of neural units for each layer: 120

Learning rate : **0.1 (Edited)**

Activation function : Softmax, ReLU

Output function : Linear

Other parameters:

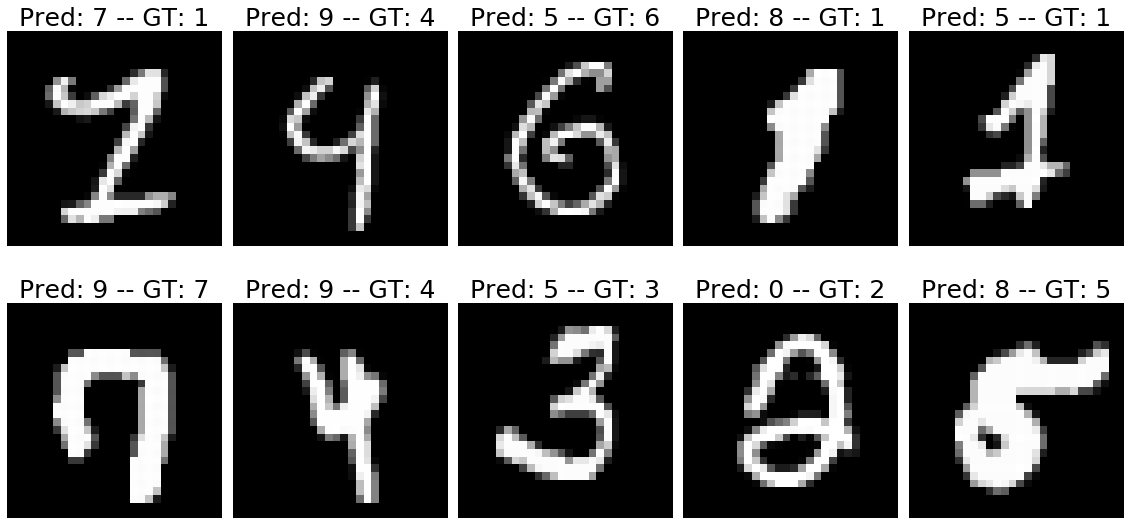
Epoch : 20

Batch size : 100

Accuracy: 0.9699

Confusion matrix, Per-class Recall , and Per-class precision are in the Question 1.

The images which the model made wrong predications.

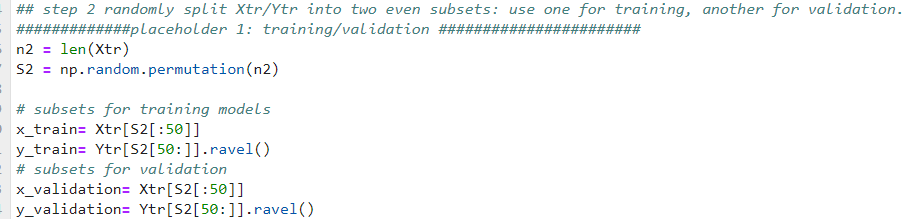


The numbers of above images are scratchy. Even though, using artificial inspection is still hard to recognize a number. Take [1,2] for example, it really likes 9.

**Part II,**

Step 2: Randomly split Xtr/Ytr into two even subsets.

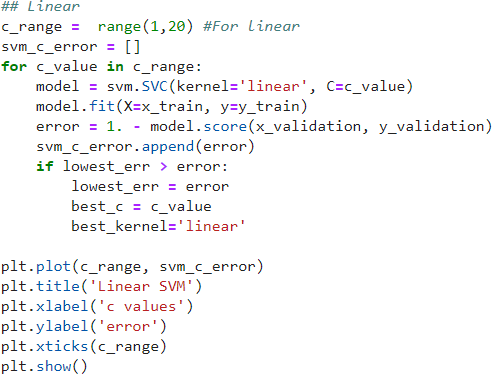
Setting the train and validation data.

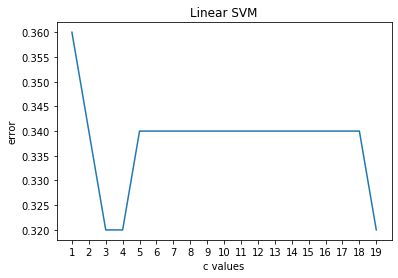
****

Step 3: Model selection over validation set

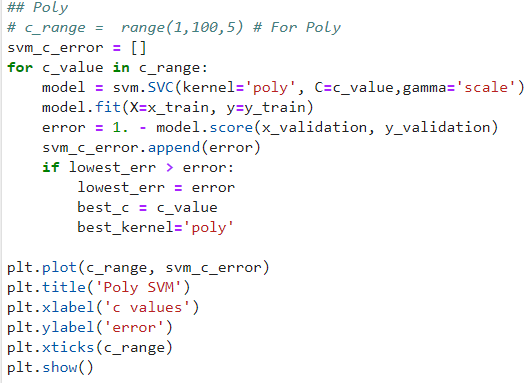
I respectively ran three kernels to find the best c value and kernel.

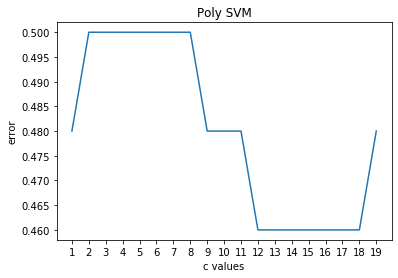
Linear:



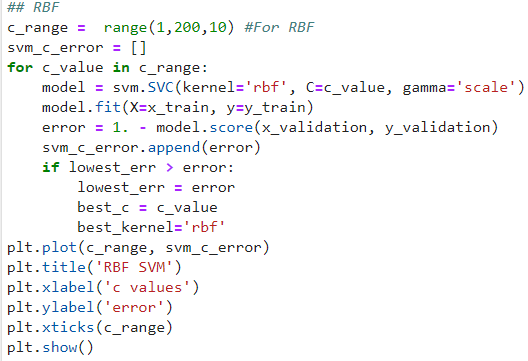


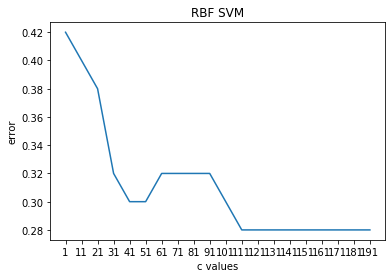
Poly:





RBF:





Confusion Matrix:

[[15 35]

[33 17]]

Average Accuracy: 0.32

Per-Class Precision: [0.3125 0.32692308]

Per-Class Recall: [0.3 0.34]

Success samples:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | 3 | 4 | 5 | GT | Pred |
| 0.0 | 15.0 | 12.3 | 30.1 | 33.3 | 14.0 | 1 | 1 |
| 1.0 | 15.7 | 13.9 | 33.6 | 38.5 | 14.1 | 1 | 1 |
| 1.0 | 8.8 | 7.7 | 18.1 | 20.8 | 7.4 | -1 | -1 |
| 0.0 | 14.0 | 12.8 | 28.8 | 32.4 | 12.7 | 1 | 1 |
| 0.0 | 10.7 | 9.7 | 21.4 | 24.0 | 9.8 | 1 | 1 |

Failure samples:

0 1 2 3 4 5 GT Pred

0 1.0 15.9 12.7 34.0 38.9 14.2 -1.0 1.0

1 0.0 20.1 13.7 40.6 44.5 18.0 -1.0 1.0

2 0.0 19.7 16.7 39.9 43.6 18.2 1.0 -1.0

3 1.0 16.9 13.2 37.3 42.7 15.6 -1.0 1.0

4 0.0 16.3 11.6 31.6 34.2 14.5 -1.0 1.0