Assignment 1 - CS747

Hoang Dung Bui
NetID: G01301478
Team name on Kaggle Leaderboard: Hoang-Dung Bui

September 16, 2021

1 Perceptron

There are hyperparameters for the softmax algorithms: epoch's number, learning rate and batch_size.

- There are numbers of epochs having tested: 5, 10, 20, 25, 30, 35, 40, 50, 80, 100, 120, 180, 200, and 300. The optimal epoch is 100.
- I also tested on fixed learning rate from 0.01 - 0.2, however, the result is lower than the varied learning rate in Eq. (1) with initial value at 0.3.

For learning rate, it varies with the numbers of epoch with formula:

$$lr = \frac{lr_{last}}{1 + i \cdot decay\_rate} \tag{1}$$

$$decay\_rate = lr_0 / n_{ep}$$

where i is the current epoch, n_{ep} is the epoch number for training, lr_{last} is the learning rate of the last epoch, lr_0 is the initial learning rate.

This varied learning rate will decrease as the epoch goes up, and it speed up the convergence of the algorithm.

The optimal hyperparameters and the corresponding accuracies for CIFAR DATASET:

Optimal epoch	25
Optimal learning rate	varied lr ($lr_0 = 0.3$)
Training Accuracy	0.3665
Validation Accuracy	0.305
Test Accuracy	0.2866

2 SVM

There are several hyperparameters for the svm algorithms: epoch's number, learning rate and batch_size.

- There are number of epochs having tested: 20, 30, 40, 50, 80, 120, 200, and 300. The optimal epoch is 200.
- I also tested on fixed learning rate from 0.006 - 0.3, however, the result is lower than the varied learning rate in Eq. (1) with initial value at 0.3.
- The batch_size were tested: 16, 32, 64, 128. The best one for this model is 32
- The regularization is tested from 0.02 - 0.5. The value 0.05 is the best for our model.

The optimal hyperparameters and accuracies for CIFAR DATASET:

Optimal epoch	50
Optimal learning rate	varied lr ($lr_0 = 0.08$)
Regularization	1.0
Optimal batch size	32
Training Accuracy	0.3651
Validation Accuracy	0.303
Test Accuracy	0.3012

3 Softmax

There are hyperparameters for the softmax algorithms: epoch's number, learning rate and batch_size.

- There are number of epochs having tested: 20, 50, 120, 200, and 300. The optimal epoch is 300.
- I also tested on fixed learning rate from 0.02 - 0.2, however, the result is lower than the varied learning rate in 1 with initial value at 0.4.
- The batch_size were tested: 16, 32, 64, 128. The best one for this model is 32

The optimal hyperparameters and accuracies for CIFAR DATASET:

Optimal epoch	100
Optimal learning rate	varied lr ($lr_0 = 0.4$)
Optimal batch size	32
Training Accuracy	0.3356
Validation Accuracy	0.252
Test Accuracy	0.2701

4 Bonus Points

4.1 Extra part 1

By running multiple times with randomly selecting any two of 10 classes, the highest accuracy of correctly classify every data point is around 0.8, the lowest one is around 0.3. Therefore, we can conclude that CIFAR-10 is highly non-linearly separable. The code is included in the submission with the name: *svm_extra_1.py*

4.2 Extra part 2

I use the function *make_gaussian_quantiles* from the packages *sklearn.datasets* to create a 2D dataset which is highly non-linearly separable. Using svm algorithm, the highest accuracy is 0.5.

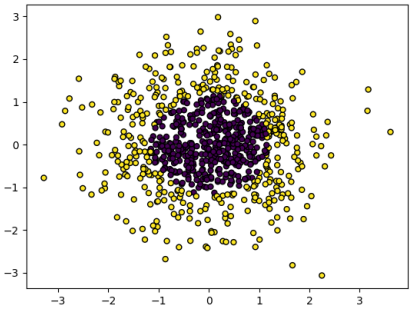

Figure 1: The Synthesis data

Because this dataset is highly non-linearly separable, I used Kernel SVM to classified the data. I used the polynomial kernel with degree = 8, and the model can classified 0.95 correctly the dataset. The code is submitted with the name: *svm_extra_2.py*