

Name: Negar Nejatishahidin

NetID: 01207447

Name on Kaggle leaderboard: Negar Nejatishahidin

Perceptron

Briefly describe the hyperparameter settings you tried. In particular, you should list the different values for learning rate and number of epochs you tried. You should also mention whether adding a learning rate decay helped and how you implemented this decay. Report the optimal hyperparameter setting you found in the table below. Report your training, validation, and testing accuracy with your optimal hyperparameter setting.

For the perceptron, I used the update method which was mentioned in slides. I also add one element to X and also one to the weight to handle the bias. I found that when we update the ground truth class if we multiply it by the number of the wrong classes which got the better score, the accuracy increases. I tried several learning rates, first I started from 0.05 with 1 to 100 epochs. I also plot the accuracy. However, by reducing the alpha according to the number of epochs the accuracy increases. Since the initial weights are random, 2 different runs have two different accuracies. I also plot the accuracy and easily understand that some learning rates are too high and cause the learning to oscillate. Finally, with 10^{-9} and 100 epochs I got the accuracy

Optimal hyperparameters:	Alpha = 0.1 power (-6) with 1/epochs times decay at each epoch, Epoch = 100
Training accuracy:	29.21
Validation accuracy:	28.90
Test accuracy:	28.24

SVM

Describe the hyperparameter tuning you tried for learning rate, number of epochs, and regularization constant. Report the optimal hyperparameter setting you found in the table below. Also report your training, validation, and testing accuracy with your optimal hyperparameter setting.

For the SVM I used the update method which was described in the slides. I also multiply the change of the ground truth class by the number of classes that got better scores. I also reduced the amount of alpha with $(1/\text{epoches}) \cdot \alpha$. I have started from 0.05 and reduced alpha up to 0.1 power (-10). I plot the accuracy and loss to find if the alpha is big and cause oscillating. I also used the loss plot to find if it is reducing continuously. I changed the number of batches and 10 was good enough for the experiment.

Optimal hyperparameters:	Alpha = $0.1^{**(-10)}$, epochs = 100, regularization constant = 0.01
Training accuracy:	34.96
Validation accuracy:	36.40
Test accuracy:	35.14

Softmax

Once again, describe the hyperparameter tuning you tried for learning rate, number of epochs, and regularization constant. Report the optimal hyperparameter setting you found in the table below. Also report your training, validation, and testing accuracy with your optimal hyperparameter setting.

Same as 2 other algorithms, I used the update rule of slides. I plot the accuracy and the loss to make sure if they are working correctly. I did not use the idea of multiplying the update of the groundtruth by the number of classes which gets higher than ground truth score because it reduced the accuracy. I used the same range of epoches for softmax and SVM. I changed the number of batches and 10 was good enough for the experiment.

Optimal hyperparameters:	Learning rate = 0.1 Power (-7) epochs = 100, batch size = 1
Training accuracy:	44.58
Validation accuracy:	41.70
Test accuracy:	42.36