Name(s): Won Woo Lyu, Albert Yoo
NetID(s): wlyu2, alberty3
Team name on Kaggle leaderboard: WWL

For each of the sections below, your reported test accuracy should approximately match the accuracy reported on Kaggle.

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.

Learning Rate	Epochs	Train Acc (CIFAR)
0.5	0	9.965
0.5	10	32.157
0.5	20	32.14
0.5	30	32.2125
0.5	40	32.235
0.5	50	32.2475
0.1	0	9.965
0.1	10	32.1575
0.1	20	32.14
0.1	30	32.2125
0.1	40	32.235
0.1	50	32.2475
0.01	0	9.965

0.01	10	32.1575
0.01	20	32.14
0.01	30	32.2125
0.01	40	32.235
0.01	50	32.2475

Above table shows how training accuracy for perceptrons differ based on the changing of learning rate. We can tell that an increase of number of epochs gives increase of accuracy but reducing learning rate did not affect the accuracy.

Learning Rate	Epochs	Train Acc (CIFAR)
0.01	0	9.965
0.01	10	32.1575
0.01	20	32.14
0.01	30	32.2125
0.01	40	32.235
0.01	50	32.2475
0.01	60	32.2375
0.01	70	32.25
0.01	80	32.2525
0.01	90	32.265
0.01	100	32.275
0.01	110	32.2675
0.01	120	32.2775
0.01	130	32.255
0.01	140	32.245
0.01	150	32.2475

0.01	160	32.2625
0.01	170	32.255
0.01	180	32.2575
0.01	190	32.255
0.01	200	32.26

Next step was to find out which number of epochs have the highest number of train accuracy. Within 200 numbers of epochs, it shows that 120 has the highest.

Finding out that change of learning rate does not affect accuracy and number of 120 epochs produce highest number of accuracy, next step is to find out whether adding learning decay rate increases the accuracy. To test this, we decayed the learning rate by half every 10 epochs.

Learning Rate	Epochs	Train Acc (CIFAR)
0.01	0	9.965
0.005	10	32.1575
0.0025	20	32.1825
0.00125	30	32.2125
0.000625	40	32.235
0.0003125	50	32.2425
0.00015625	60	32.235
7.8125e-5	70	32.2425
3.90625e-05	80	32.2324

Comparing two tables above, we can tell that the decaying learning rate by half for every 10 epochs, increased accuracy when the number of epochs was 20, but most of the accuracy was the same or decreased. This result tells that decaying learning rate really does not affect the accuracy.

Above data was based on CIFAR. But best accuracy in CIFAR does not produce the best result in Mushroom. With a learning rate of 0.01 and number of epochs 120, Mushroom dataset has accuracy of 85.1456. However, when the learning rate is 0.5 with a number of epochs is 120, the accuracy is 93.926959. By this we can tell that the mushroom dataset works better with a higher learning rate than CIFAR does.

MUSHROOM DATASET

Optimal hyperparameters:	Learning rate = 0.5 number of epochs = 120
Training accuracy:	93.926959%
Validation accuracy:	94.153846%
Test accuracy:	94.338462%

CIFAR DATASET

Optimal hyperparameters:	Learning rate = 0.01, number of epochs= 120
Training accuracy:	32.2775%
Validation accuracy:	31.65%
Test accuracy:	32.45%

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.

Learning Rate	Epochs	Regularization	Train Acc (CIFAR)
0.5	0	0.01	7.42749
0.5	10	0.01	26.0
0.5	20	0.01	26.36
0.5	30	0.01	25.94

0.5	40	0.01	27.38
0.5	50	0.01	26.2125
0.1	0	0.01	10.9925
0.1	10	0.01	30.975
0.1	20	0.01	31.7575
0.1	30	0.01	31.3199
0.1	40	0.01	32.129
0.1	50	0.01	31.0275
0.01	0	0.01	8.9175
0.01	10	0.01	36.8025
0.01	20	0.01	37.6775
0.01	30	0.01	37.37
0.01	40	0.01	36.8725
0.01	50	0.01	36.99
0.001	0	0.01	8.2325
0.001	10	0.01	40.4325
0.001	20	0.01	41.5
0.001	30	0.01	41.835
0.001	40	0.01	41.75
0.001	50	0.01	41.835

Above table shows that as the learning rate gets smaller, better accuracy is produced for CIFAR.
Next step is to check how changes in the number of epochs affect the accuracy.

Learning Rate	Epochs	Regularization	Train Acc (CIFAR)
0.001	0	0.01	7.79
0.001	10	0.01	40.47

0.001	20	0.01	41.2825
0.001	30	0.01	41.4925
0.001	40	0.01	41.5899
0.001	50	0.01	41.8875
0.001	60	0.01	41.5650
0.001	70	0.01	41.79
0.001	80	0.01	41.5625
0.001	90	0.01	41.635
0.001	100	0.01	41.85749
0.001	110	0.01	42.04
0.001	120	0.01	41.8975
0.001	130	0.01	41.7074
0.001	140	0.01	41.8925
0.001	150	0.01	41.5375
0.001	160	0.01	41.8925
0.001	170	0.01	41.7125
0.001	180	0.01	41.91
0.001	190	0.01	42.04
0.001	200	0.01	41.79

Above table shows that when the number of epochs is 110 or 190, it produces highest accuracy. Since the average of accuracy gets higher as the number of epochs is higher, the optimal parameter can be 190.

Next step is to find out the change of accuracy over the change of regularization.

Learning Rate	Epochs	Regularization	Train Acc (CIFAR)
0.001	190	0.001	44.1225

0.001	190	0.05	39.1275
0.001	190	0.01	41.8975
0.001	190	0.1	37.2425

Above table shows that the lowest number of regularizations leads to higher accuracy of prediction in CIFAR.

Like CIFAR, Mushroom also gives highest accuracy with the same hyperparameters.

With a change of Learning rate to 0.1, it produces 79.9958% of training accuracy, 78.4% validation accuracy and 79.4461 test accuracy. This tells 0.001 learning rate is better than 0.1 one.

With a change of regularization of 0.1, it produces 80.0369%, 78.4615% and 79.5077%. Furthermore, regularization of 0.01 produces 80.1189%, 78.52307% and 79.5692%. Regularization of 0.001 is similar to that of 0.01, but better on testing accuracy since it was 79.569231%

MUSHROOM DATASET

Optimal hyperparameters:	Learning rate = 0.001 Number of epochs = 190 Regularization = 0.001
Training accuracy:	80.118999%
Validation accuracy:	78.584615%
Test accuracy:	79.569231%

CIFAR DATASET

Optimal hyperparameters:	Learning rate = 0.001 Number of epochs = 190 Regularization = 0.001
Training accuracy:	44.1225%
Validation accuracy:	37.41%
Test accuracy:	36.94%

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.

Learning Rate	Epochs	Train Acc (CIFAR)
1	50	34.375
0.75	50	33.34
0.5	50	34.99
0.1	50	34.8025
0.01	50	32.767
0.001	50	34.192

The learning rate of 0.5 gives the highest accuracy on the CIFAR dataset when softmax is run for 50 epochs.

Learning Rate	Epochs	Train Acc (CIFAR)
0.5	50	34.99
0.5	100	35.707
0.5	150	35.0475

Running softmax for 100 epochs gives the highest accuracy on the CIFAR dataset. Running for 50 epochs and 150 epochs gives a lower accuracy.

MUSHROOM DATASET

Optimal hyperparameters:	Learning Rate = 0.5 Number of Epochs = 100 Regularization Constant = 0.001
Training accuracy:	81.904%
Validation accuracy:	79.815%
Test accuracy:	82.523%

CIFAR DATASET

Optimal hyperparameters:	Learning Rate = 0.5 Number of Epochs = 100 Regularization Constant = 0.001
Training accuracy:	35.707%
Validation accuracy:	28.82%
Test accuracy:	28.44%

Logistic

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.

Learning Rate	Epochs	Train Acc (Mushroom)
0.5	50	84.283956
0.1	50	89.7825
0.01	50	89.7825
0.001	50	93.2704
0.0001	50	90.43906

Looking at the above table, logistic produces best accuracy when learning rate is 0.001.

Learning Rate	Epochs	Train Acc (Mushroom)
0.001	50	93.2704
0.001	100	93.8654
0.001	150	94.1731
0.001	200	94.5219
0.001	250	94.6655

0.001	300	94.6245

From the above table, we can find out that when the number of epochs is more than 150, the training accuracy is at 94% and slightly increases when number epochs increase. But when it reaches 300, it slightly decreases. This tells that 250 is the optimal for the number of epochs.

MUSHROOM DATASET

Optimal hyperparameters:	Learning Rate = 0.001 Number of Epochs = 250
Training accuracy:	94.6655%
Validation accuracy:	93.8461%
Test accuracy:	94.2769%