Name(s): Ziheng Chen, Yutong Zhang
NetID(s): zihengc2, yutongz7
Team name on Kaggle leaderboard: zihengc2 & yutongz7

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.

At first, we noticed that the training accuracy could not even beat the benchmark, so we introduced learning rate decay into our model. The learning rate is multiplied by 0.85 each epoch. We chose exponentially decrease over other formulas since it was easy to implement. We then noticed the high training accuracy with lower validation accuracy and even lower testing accuracy. We suspected that our model was overfitting, so we made our model more robust by introducing a regularization constant. We slowly increased it from 0.1 to 8, and the gap between training and validation grew smaller. Finally, to further avoid overfitting, we further limited the number of epochs. We found that the training error started to converge fairly quickly, so we kept the number of epoch to 10.

RICE DATASET

Optimal hyperparameters:	lr = 0.3, n_epochs = 20
Training accuracy:	99.899184
Validation accuracy:	99.835029
Test accuracy:	99.835029

Fashion-MNIST DATASET

Optimal hyperparameters:	lr = 0.5, n_epochs = 10
Training accuracy:	84.082000

Validation accuracy:	82.020000
Test accuracy:	81.880000

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.

After getting a desired training error, we again faced the issue of overfitting. We increased the batch size to 1000 to make our model better against outliers, We kept the learning rate decay with the learning rate multiplied by 0.85 each epoch.

RICE DATASET

Optimal hyperparameters:	lr = 1, n_epochs = 50, reg_const = 0.05
Training accuracy:	79.048666
Validation accuracy:	78.581248
Test accuracy:	79.323618

Fashion-MNIST DATASET

Optimal hyperparameters:	lr = 1, n_epochs = 30, reg_const = 0.05
Training accuracy:	84.308000
Validation accuracy:	82.840000
Test accuracy:	82.140000

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.

We start by trying the default learning rate of 0.5, which results in high training accuracy but low validating and testing accuracy. To avoid overfitting, we decreased the learning rate and further tried 0.05, 0.005, and 0.0005, in which 0.005 finally had the better result. Then, we test different epochs number: 20, 50, 80, 100, 150, and 200, where we find out that it tends to converge after 100 epochs, where RICE dataset has the best performance at 150 epochs, and Fashion-MNIST dataset has the best performance at 200 epochs. Besides, we decreased the learning rate at each epoch to avoid overfitting,

RICE DATASET

Optimal hyperparameters:	lr = 0.005, n_epochs = 150, reg_const = 1
Training accuracy:	99.917514
Validation accuracy:	99.807534
Test accuracy:	99.862524

Fashion-MNIST DATASET

Optimal hyperparameters:	lr = 0.005, n_epochs = 200 , reg_const = 1
Training accuracy:	88.38800
Validation accuracy:	84.51000
Test accuracy:	83.50000

Logistic

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

Firstly, we find out the accuracy is always around 50%, then we find out that the rice labels are 0 and 1 rather than -1 and 1, so we change the label. Then the accuracy reaches around 98% with the default setting. Since its training accuracy continues growing and looks like not converged yet, so we increased the epochs to 20 and 50, and which RICE dataset has the best performance at 50 epochs.

RICE DATASET

Optimal hyperparameters:	lr = 0.5, n_epochs = 50, threshold = 0.5
Training accuracy:	99.560077
Validation accuracy:	99.560077
Test accuracy:	99.505087