

**Name(s): Ziyue Tang**

**NetID(s): ziyet3**

**Team name on Kaggle leaderboard: Ziyue Tang**

**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.*

Setting 1:

Learning rate = 0.5,      # of epochs = 10,      no decay

Setting 2:

Learning rate = 0.5,      # of epochs = 10,      decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Setting 3:

Learning rate = 0.5,      # of epochs = 12,      decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Setting 4:

Learning rate = 0.05,      # of epochs = 14,      decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Setting 5:

Learning rate = 0.05,      # of epochs = 10,      decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Setting 6:

Learning rate = 0.05,      # of epochs = 12,      decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Setting 7:

Learning rate = 0.05,      # of epochs = 14,      decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Note: Adding a learning rate decay can improve the accuracy.  
The decay function I use:  $lr = lr / (1 + epoch\_num * decay\_rate)$

#### RICE DATASET

Optimal hyperparameters:	Setting 7
Training accuracy:	99.807534
Validation accuracy:	99.835029
Test accuracy:	99.752543

#### Fashion-MNIST DATASET

Optimal hyperparameters:	Setting 3
Training accuracy:	84.510000
Validation accuracy:	82.640000
Test accuracy:	81.680000

### 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.*

#### Setting 1

Learning rate = 0.5, # of epochs = 10, decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )  
regularization constant = 0.05, batch size = 2

#### Setting 2

Learning rate = 0.5, # of epochs = 10, decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )  
regularization constant = 0.1, batch size = 2

#### Setting 3

Learning rate = 0.5, # of epochs = 10, decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )  
regularization constant = 0.15, batch size = 2

#### Setting 4

Learning rate = 0.5, # of epochs = 12, decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )  
regularization constant = 0.05, batch size = 2

#### Setting 5

Learning rate = 0.5, # of epochs = 12, decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )  
regularization constant = 0.1, batch size = 2

#### Setting 6

Learning rate = 0.5, # of epochs = 12, decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )  
regularization constant = 0.15, batch size = 2

#### Setting 7

Learning rate = 0.05, # of epochs = 10, decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )  
regularization constant = 0.05, batch size = 2

#### Setting 8

Learning rate = 0.05, # of epochs = 10, decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )  
regularization constant = 0.1, batch size = 2

#### Setting 9

Learning rate = 0.05, # of epochs = 10, decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )  
regularization constant = 0.15, batch size = 2

#### Setting 10

Learning rate = 0.05, # of epochs = 12, decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )  
regularization constant = 0.05, batch size = 2

#### Setting 11

Learning rate = 0.05, # of epochs = 12, decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )  
regularization constant = 0.1, batch size = 2

### Setting 12

Learning rate = 0.05, # of epochs = 12, decay\_rate = 8  
(decay function:  $lr = lr / (1 + \text{epoch\_num} * \text{decay\_rate})$ )  
regularization constant = 0.15, batch size = 2

### RICE DATASET

Optimal hyperparameters:	Setting 5
Training accuracy:	99.880854
Validation accuracy:	99.835029
Test accuracy:	99.862524

### Fashion-MNIST DATASET

Optimal hyperparameters:	Setting 10
Training accuracy:	83.910000
Validation accuracy:	82.300000
Test accuracy:	81.730000

### 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.*

### Setting 1

Learning rate = 0.5, # of epochs = 10, decay\_rate = 8  
(decay function:  $lr = lr / (1 + \text{epoch\_num} * \text{decay\_rate})$ )  
regularization constant = 0.05, batch size = 6

### Setting 2

Learning rate = 0.5, # of epochs = 10, decay\_rate = 8  
(decay function:  $lr = lr / (1 + \text{epoch\_num} * \text{decay\_rate})$ )

regularization constant = 0.1, batch size = 6

#### Setting 3

Learning rate = 0.5, # of epochs = 10, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

regularization constant = 0.15, batch size = 6

#### Setting 4

Learning rate = 0.5, # of epochs = 12, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

regularization constant = 0.05, batch size = 6

#### Setting 5

Learning rate = 0.5, # of epochs = 12, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

regularization constant = 0.1, batch size = 6

#### Setting 6

Learning rate = 0.5, # of epochs = 12, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

regularization constant = 0.15, batch size = 6

#### Setting 7

Learning rate = 0.05, # of epochs = 10, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

regularization constant = 0.05, batch size = 6

#### Setting 8

Learning rate = 0.05, # of epochs = 10, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

regularization constant = 0.1, batch size = 6

#### Setting 9

Learning rate = 0.05, # of epochs = 10, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

regularization constant = 0.15, batch size = 6

#### Setting 10

Learning rate = 0.05, # of epochs = 12, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

regularization constant = 0.05, batch size = 6

#### Setting 11

Learning rate = 0.05, # of epochs = 12, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )  
regularization constant = 0.1, batch size = 6

Setting 12

Learning rate = 0.05, # of epochs = 12, decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )  
regularization constant = 0.15, batch size = 6

#### RICE DATASET

Optimal hyperparameters:	Setting 4
Training accuracy:	99.844194
Validation accuracy:	99.780038
Test accuracy:	99.752543

#### Fashion-MNIST DATASET

Optimal hyperparameters:	Setting 10
Training accuracy:	85.016000
Validation accuracy:	83.180000
Test accuracy:	82.210000

### 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.*

Setting 1:

Learning rate = 0.5, # of epochs = 10, decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )  
Threshold = 0.5

Setting 2:

Learning rate = 0.5, # of epochs = 10, decay\_rate = 8  
(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Threshold = 0.6

Setting 3:

Learning rate = 0.5, # of epochs = 10, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Threshold = 0.7

Setting 4:

Learning rate = 0.5, # of epochs = 12, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Threshold = 0.5

Setting 5:

Learning rate = 0.5, # of epochs = 12, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Threshold = 0.6

Setting 6:

Learning rate = 0.5, # of epochs = 12, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Threshold = 0.7

Setting 7:

Learning rate = 0.05, # of epochs = 10, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Threshold = 0.5

Setting 8:

Learning rate = 0.05, # of epochs = 10, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Threshold = 0.6

Setting 9:

Learning rate = 0.05, # of epochs = 10, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Threshold = 0.7

Setting 10:

Learning rate = 0.05, # of epochs = 12, decay\_rate = 8

(decay function:  $lr = lr / (1 + epoch\_num * decay\_rate)$ )

Threshold = 0.5

Setting 11:

Learning rate = 0.05, # of epochs = 12, decay\_rate = 8  
(decay function:  $lr = lr / (1 + \text{epoch\_num} * \text{decay\_rate})$ )

Threshold = 0.6

Setting 12:

Learning rate = 0.05, # of epochs = 12, decay\_rate = 8  
(decay function:  $lr = lr / (1 + \text{epoch\_num} * \text{decay\_rate})$ )

Threshold = 0.7

#### RICE DATASET

Optimal hyperparameters:	Setting 6
Training accuracy:	99.871689
Validation accuracy:	99.835029
Test accuracy:	99.780038