

Name(s): Yixiao (Jimmy) Fang, Jingjie He

NetID(s): yixiaof2, jingjie8

Team name on Kaggle leaderboard: Artificial Intelligence

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.

Attempt 1:

- Learning Rate = 0.1
- Number of epochs = 10
- Learning rate decay = 1.0

The first attempt already has decent result, by using the hyperparameters when fitting the Fashion-MNIST dataset. So the first attempt is used as the final report.

RICE DATASET

Optimal hyperparameters:	lr = 0.1 n_epochs = 10 lr_decay= 1.0
Training accuracy:	99.21%
Validation accuracy:	99.23%
Test accuracy:	99.01%

Attempt 1:

- Learning Rate = 0.5
- Number of epochs = 10
- No learning rate decay
- Test accuracy: around 75%

Attempt 2:

- Learning Rate = 0.2
- Number of epochs = 10
- No learning rate decay
- Test accuracy: around 77%

Attempt 3:

- Learning Rate = 0.2
- Number of epochs = 10
- Learning rate decay = 1.0
- Test accuracy: around 80%

Fashion-MNIST DATASET

Optimal hyperparameters:	lr = 0.1 n_epochs = 10 lr_decay = 1.5
Training accuracy:	84.44%
Validation accuracy:	82.23%
Test accuracy:	81.54%

For the perceptron model, weight decay definitely affect the overall accuracy. We can compare attempt 2 and attempt 3 for Fashion-MNIST dataset. Adding a learning rate decay increases the test accuracy by almost 3 percent without changing other parameters. Theoretically, this makes sense because lower learning rate when about to converge can allow the model not to jump around near the global/local minimum of the loss function.

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.

Attempt 1:

- Learning Rate = 0.5
- Number of epochs = 10
- Regularization Constant = 0.05
- Test accuracy: around 79%

Attempt 2:

- Learning Rate = 0.1
- Number of epochs = 10
- Regularization Constant = 0.05
- Test accuracy: around 81%

RICE DATASET

Optimal hyperparameters:	reg_const = 0.05 lr = 0.01 n_epochs = 10
Training accuracy:	99.71%
Validation accuracy:	99.29%
Test accuracy:	99.09%

Attempt 1:

- Learning Rate = 0.5
- Number of epochs = 10
- Regularization Constant = 0.05
- Batch Size = 100
- Test accuracy: around 78%

Attempt 2:

- Learning Rate = 0.1
- Number of epochs = 10
- Regularization Constant = 0.01
- Batch Size = 128
- Test accuracy: around 75%

Attempt 3:

- Learning Rate = 0.1
- Number of epochs = 10
- Regularization Constant = 0.01
- Batch Size = 300
- Test accuracy: around 75%

Attempt 4:

- Learning Rate = 0.05
- Number of epochs = 10
- Regularization Constant = 0.05
- Batch Size = 200
- Test accuracy: around 80%

Fashion-MNIST DATASET

Optimal hyperparameters:	reg_const = 0.05 lr = 0.01 n_epochs = 10 batch_size = 200
Training accuracy:	84.22%
Validation accuracy:	82.92%
Test accuracy:	81.48%

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.

Attempt 1:

- Learning Rate = 0.05
- Number of epochs = 10
- Regularization Constant = 0.05
- Batch Size = 128
- Weight decay = 0.0
- Test accuracy: around 75%

Attempt 1:

- Learning Rate = 0.01
- Number of epochs = 10
- Regularization Constant = 0.05
- Batch Size = 64
- Weight decay = 1.0
- Test accuracy: around 98.5%

RICE DATASET

Optimal hyperparameters:	reg_const = 0.05 lr = 0.01 n_epochs = 10 batch_size = 16 weight_decay = 1.0
Training accuracy:	99.85%
Validation accuracy:	99.81%
Test accuracy:	99.78%

Attempt 1:

- Learning Rate = 0.1
- Number of epochs = 10
- Regularization Constant = 0.05
- Batch Size = 100
- Weight decay = 1.0
- Test accuracy: around 77%

Attempt 2:

- Learning Rate = 0.01
- Number of epochs = 10
- Regularization Constant = 0.05
- Batch Size = 100
- Test accuracy: around 76%

Attempt 3:

- Learning Rate = 0.01
- Number of epochs = 10
- Regularization Constant = 0.01
- Batch Size = 100
- Test accuracy: around 75%

Attempt 4:

- Learning Rate = 0.01
- Number of epochs = 10
- Regularization Constant = 0.05
- Batch Size = 32
- Test accuracy: around 80%

Fashion-MNIST DATASET

Optimal hyperparameters:	reg_const = 0.05 lr = 0.01 n_epochs = 10 batch_size = 64 weight_decay = 1.0
Training accuracy:	84.59%
Validation accuracy:	83.21%
Test accuracy:	82.12%

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.

Attempt 1:

- Learning Rate = 0.5
- Number of epochs = 10
- Threshold = 0.5
- Test accuracy: around 84%

Attempt 2:

- Learning Rate = 0.1
- Number of epochs = 10
- Threshold = 0.3
- Test accuracy: around 95%

RICE DATASET

Optimal hyperparameters:	lr = 0.1 n_epochs = 10 threshold = 0.5
Training accuracy:	99.63%
Validation accuracy:	99.62%
Test accuracy:	99.67%