

Name(s): Li-Kai Chuang

NetID(s): likaikc2

Team name on Kaggle leaderboard: kaiwenjon

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.

I primarily are tuning batch size and learning rate and fix the decay rate to 0.99. $lr = lr * (1/(1+epoch*decayRate))$. Also I shuffle the dataset before training.

RICE DATASET

Optimal hyperparameters:	batchSize: 50, lr: 0.01, epoch: 500
Training accuracy:	99.91%
Validation accuracy:	99.81%
Test accuracy:	99.78%

Fashion-MNIST DATASET

Optimal hyperparameters:	batchSize: 300, lr: 0.01, epoch: 50
Training accuracy:	84.64%
Validation accuracy:	83.17%
Test accuracy:	82.36%

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.

I primarily are tuning batch size and learning rate and fix the decay rate to 0.99. $lr = lr * (1/(1+epoch*decayRate))$. Also I shuffle the dataset before training.

RICE DATASET

Optimal hyperparameters:	batchSize: 200, lr: 0.001, epoch: 50, reg: 0.05
Training accuracy:	99.89%
Validation accuracy:	99.89%
Test accuracy:	99.84%

Fashion-MNIST DATASET

Optimal hyperparameters:	batchSize: 200, lr: 0.1, epoch: 30, reg: 0.001
Training accuracy:	84.79%
Validation accuracy:	82.94%
Test accuracy:	82.07%

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.

I've tried adam, but it seems that normal SGD along with learning rate decay gives the best performance. I also early stop training to prevent overfitting instead of tuning regularization constant. For the formula of learning rate decay: $lr = lr * (1/(1+epoch*decayRate))$. Also I shuffle the dataset before training.

RICE DATASET

Optimal hyperparameters:	Batch_size: 50, lr:0.01, epochs: 100, decay rate: 0.99
Training accuracy:	99.93%
Validation accuracy:	99.86%
Test accuracy:	99.89%

Fashion-MNIST DATASET

Optimal hyperparameters:	Batch_size: 200, lr:0.01, epochs: 50, decay rate: 0.99
Training accuracy:	85.13%
Validation accuracy:	84.21%
Test accuracy:	82.97%

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.

I didn't use decay learning rate in this implementation. Just tuning learning rate and batch size will suffice. Also I shuffle the dataset before training.

RICE DATASET

Optimal hyperparameters:	Batch_size: 30, lr: 0.1, epoch: 300
Training accuracy:	99.91%
Validation accuracy:	99.89%
Test accuracy:	99.95%