Name: Anowarul Kabir
NetID: akabir4
Name on Kaggle leaderboard: Anowarul Kabir

Perceptron

There are two hyperparameters in the basic perceptron implementation namely learning rate (alpha) and epochs. I started as alpha=.1 and epochs=30. Then I used the step decay learning rate. The equation is used as follows:

$$lr = lr0 * drop^{floor(t / epochs\_drop)}$$

lr0 = initial learning rate
lr = new learning rate
t = epoch
drop = learning rate dropping rate
epochs_drop = after how many epochs learning rate will be dropped

1. lr0=.10, drop=.3, epochs_drop=2.0, alpha=50
2. lr0=1.0, drop=.3, epochs_drop=5.0, alpha=80
3. lr0=1.0, drop=.5, epochs_drop=5.0, alpha=50
4. lr0=3.0, drop=.5, epochs_drop=5.0, alpha=50
5. lr0=3.0, drop=.7, epochs_drop=3.0, alpha=50
6. lr0=3.0, drop=.7, epochs_drop=5.0, alpha=50

Optimal hyperparameters:	lr0=3.0, drop=.5, epochs_drop=5.0, alpha=50
Training accuracy:	30.37
Validation accuracy:	29.80
Test accuracy:	29.90

SVM

I have used exponential learning rate from here. The equation used is as follows:

$lr = lr0 * e^{(-kt)}$

where
lr0 = initial learning rate
k = hyperparameter
t = epoch
lr = new learning rate

Another hyperparameter is regularization constant (reg_const). = 1e-4 #

1. lr0=1.0, k=.5, reg_const=0, epochs=20
2. lr0=1.0, k=.5, reg_const=n, epochs=30 {here, n=number of train examples=49000}
3. lr0=1.0, k=.5, reg_const=.1, epochs=30
4. lr0=0.1, k=.5, reg_const=.1, epochs=30

5. lr0=.01, k=.5, reg_const=.1, epochs=50
6. lr0=.01, k=.5, reg_const=1e-4, epochs=50
7. lr0=.1, k=.5, reg_const=1e-4, epochs=50
8. lr0=.1, k=.3, reg_const=1e-4, epochs=50
9. lr0=.1, k=.3, reg_const=1e-4, epochs=40

Optimal hyperparameters:	lr0=.1, k=.3, reg_const=1e-4, epochs=40
Training accuracy:	34.59
Validation accuracy:	34.4
Test accuracy:	35.08

Softmax
I have also used exponential decay of learning rate here and same settings of hyperparameters from the SVM classifier.

1. lr0=.1, k=.3, reg_const=1e-4, epochs=40
2. lr0=.1, k=.5, reg_const=1e-4, epochs=40
3. lr0=.1, k=.5, reg_const=1e-3, epochs=40
4. lr0=1.0, k=.5, reg_const=1e-4, epochs=40

Optimal hyperparameters:	lr0=.1, k=.3, reg_const=1e-4, epochs=40
Training accuracy:	33.4
Validation accuracy:	33.8
Test accuracy:	33.7