Name(s): Jon Vincent Medenilla, Kulbir Singh Ahluwalia
NetID(s): jdm13, ksa5
Team name on Kaggle leaderboard: Terra_Pack

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 the 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.

Ans: Description for the hyperparameter tuning: For the perceptron model, we found that the learning rate of 0.55 and epoch number of 10 gave us the best result. We tried various values for the learning rate within the range of 0.10 to 0.7, and for the number of epochs from 8 to 20. We did not add any decaying function for our hyperparameters.

MUSHROOM DATASET

Optimal hyperparameters:	`learning_rate = 0.55` `Number_of_epochs = 10`
Training accuracy:	`94.521953`
Validation accuracy:	`94.030769`
Test accuracy:	`94.215385`

Fashion-MNIST DATASET

Optimal hyperparameters:	`learning_rate = 0.55` `Number_of_epochs = 10`
Training accuracy:	`82.242000`
Validation accuracy:	`81.630000`
Test accuracy:	`80.790000`

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.

Ans: Description for the hyperparameter tuning:
We decrease the learning rate as the epochs increase using `learning_rate_exponent_decay = 0.2`. And increase the regularization constant using "self.reg_const /= 0.9" with increasing epochs which helps in suppressing the weight vector so that the gradient does not explode.
At a higher number of epochs, the weights stop changing much because the learning rate becomes small and the regularization constant becomes larger. So we train only till 10 epochs.

MUSHROOM DATASET

Optimal hyperparameters:	`learning_rate = 0.001`
	`Number_of_epochs = 10`
	`regularization_constant = 0.6`

Training accuracy:	`90.254411`
Validation accuracy:	`89.046154`

Test accuracy:	88.861538

Fashion-MNIST DATASET

Optimal hyperparameters:	learning_rate = 0.005 Number_of_epochs = 10 regularization_constant = 0.3 Learning Rate Schedule for exponential decay: learning_rate_exponent_decay = 0.2 reg_const /= 0.9
Training accuracy:	84.008000
Validation accuracy:	82.790000
Test accuracy:	81.670000

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.

Ans: Description for the hyperparameter tuning: For the learning rate, we have tried 0.001 which gave us very low accuracy in combination with number of epochs from 10 to 20, where 20 gave us the least performance and 14 the best, and regularization constant of 0.55. We have experimented with the regularization constant from 0.005 (which gave us the worst performance) to 0.70 (which gave us slightly below than the best one, 0.55). We also added a decaying function for the learning rate by doing self.lr -= iter*self.lr/5 and an increasing function for reg_const of self.reg_const /= 0.9.

MUSHROOM DATASET

Optimal hyperparameters:	`learning_rate = 0.5`
	`Number_of_epochs = 10`
	`regularization_constant = 0.05`

Training accuracy:	`95.219532`
Validation accuracy:	`94.523077`
Test accuracy:	`95.323077`

Fashion-MNIST DATASET

Optimal hyperparameters:	`learning_rate = 0.01`
	`Number_of_epochs = 14`
	`regularization_constant = 0.55`

Training accuracy:	`85.404000`
Validation accuracy:	`81.820000`
Test accuracy:	`81.140000`

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.

Ans: Description for the hyperparameter tuning:
With a lower number of epochs, we get lower accuracy in the range of 70-85%.
This is because the weight vector does not get the chance to be updated.

With a lower learning rate than 0.6, the accuracy dropped as well because the delta weight updates were not large enough to change the weight vector.

Changing the threshold to values like 0.2,0.3,0.5 or even 0.9 had little effect on the accuracy.

MUSHROOM DATASET

Optimal hyperparameters:	`learning_rate = 0.6`
	`Number of epochs = 20`
	`threshold = 0.5`

Training accuracy:	`94.870743`
Validation accuracy:	`94.153846`
Test accuracy:	`94.461538`