**Name(s):** Haoyang Liu,

**NetID(s):** hl57, yifeil6

**Team name on Kaggle leaderboard:** Tuner

**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.*

Learning rate decay W.R.T. the epoch iteration. Lr = lr/(1+it^0.2)

MUSHROOM DATASET

|  |  |
| --- | --- |
| Optimal hyperparameters: | Lr=1e-4, epochs=40 |
| Training accuracy: | 92.162495 |
| Validation accuracy: | 91.384615 |
| Test accuracy: | 91.938462 |

CIFAR DATASET

|  |  |
| --- | --- |
| Optimal hyperparameters: |  |
| Training accuracy: |  |
| Validation accuracy: |  |
| Test accuracy: |  |

**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.*

Used grid search with learning rate=1e-6 ~ 10e-3; regularization constant=1e-2 ~ 1e6 following the power of 10. Trained with 20 epochs and 40 epochs. Chose the best hypermapameter setting from them, and trained for 80 epochs, with 1/t learning rate decay that was launched after the 20th epoch. The factor was 1/(1+0.1\*(i-20)). Tried the decay rates 0.2 and 0.1.

MUSHROOM DATASET

|  |  |
| --- | --- |
| Optimal hyperparameters: | lr=0.001, reg\_const=0.01, n\_epochs=80 |
| Training accuracy: | 91.034 |
| Validation accuracy: | 90.031 |
| Test accuracy: | 90.338 |

CIFAR DATASET

|  |  |
| --- | --- |
| Optimal hyperparameters: | lr=1e-6, reg\_const=1000.0, n\_epochs=80 |
| Training accuracy: | 41.203 |
| Validation accuracy: | 37.200 |
| Test accuracy: | 37.730 |

**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.*

Used grid search with learning rate=1e-6 ~ 10e-3; regularization constant=1e-2 ~ 1e4 following the power of 10. Trained with 20 epochs and 40 epochs. Chose the best hypermapameter setting from them, and trained for 80 epochs, with 1/t learning rate decay that was launched after the 20th epoch. The factor was 1/(1+0.1\*(i-20)). Tried the decay rates 0.2 and 0.1.

MUSHROOM DATASET

|  |  |
| --- | --- |
| Optimal hyperparameters: | lr=0.001, reg\_const=0.01, n\_epochs=80 |
| Training accuracy: | 87.731 |
| Validation accuracy: | 86.954 |
| Test accuracy: | 86.215 |

CIFAR DATASET

|  |  |
| --- | --- |
| Optimal hyperparameters: | lr=1e-6, reg\_const=100.0, n\_epochs=80 |
| Training accuracy: | 44.253 |
| Validation accuracy: | 39.270 |
| Test accuracy: | 39.510 |

**Logistic**

*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.*

MUSHROOM DATASET

|  |  |
| --- | --- |
| Optimal hyperparameters: | First 10 epochs: 1e-3, next 5 epochs:1e-5 |
| Training accuracy: | 93.249897 |
| Validation accuracy: | 91.876923 |
| Test accuracy: | 93.723077 |