**Name:**

| Noel Mrowiec |
| --- |

**Netid**:

| mrowiec3 |
| --- |

**CS 441 - HW 4: Trees and MLPs**

Complete the sections below. You do not need to fill out the checklist. **Do select all relevant pages in Gradescope.**

**Total Points Claimed [ ] / 170**

1. Model Complexity with Tree Regressors
   1. Depth vs. Error plot [ ] / 10
   2. Analysis [ ] / 20
2. MLPs with MNIST
   1. Loss Curves [ ] / 20
   2. Model Selection and Results [ ] / 20
3. Species Prediction
   1. Feature Analysis [ ] / 10
   2. Simple Rule [ ] / 10
   3. Model Design [ ] / 10
4. Stretch Goals
   1. Improve MNIST classification [ ] / 30
   2. A second simple rule [ ] / 10
   3. Positional encoding of RGB Image [ ] / 30
5. **Model Complexity with Tree Regressors**
6. **Include your plot below.**

|  |
| --- |

1. **Analyze your results:**
2. For a given max tree depth, which of regressor model (single tree, random forest, boosted tree) has the lowest bias (or most powerful)?

| Random Forest |
| --- |

1. For single regression trees, what tree depth achieves minimum validation error?

| 4 |
| --- |

1. A model “overfits” when increasing the complexity increases the validation error. Which model is least prone to overfitting? Why?

| Random forest is least prone. We can see that the error decreases or remains constant even as the tree depth increases. Random forests are less prone to overfitting because random forests average multiple deep decision trees which are trained on different parts of the same training set. Trees using different parts of data will overfit differently and thus reduce overfitting. |
| --- |

1. Do boosted trees seem to perform better with smaller or larger trees? Why?

| Boosted trees seem to perform better when the tree depth is less than 8, so it performs better on smaller trees. Boosted trees do better on smaller trees because there is overfitting as the max tree depth increases. |
| --- |

1. **MLPs with MNIST**
2. **Show the loss curves** for 3 learning rates (1E-2, 1E-1, 1E1) training for 100 epochs. An example of the loss curves is shown for LR=0.001.

| Learning Rate | Loss Curve Plot |
| --- | --- |
| 0.001 |  |
| 0.01 |  |
| 0.1 |  |
| 1 |  |

1. **Model selection and results**

**Select the best hyperparameters** (learning rate and number of epochs up to 100) based on minimizing the validation loss.

**Learning Rate**

|  |
| --- |

**Epochs**

|  |
| --- |

**Report the losses and errors for the model trained with these hyperparameters:**

Use scientific notation with one decimal place, e. 1.5E-3

| Training **Loss** | Validation **Loss** |  |
| --- | --- | --- |
|  |  |  |

Show two decimal places for percent

| Training **Error** (%) | Validation **Error** (%) | Test **Error** (%) |
| --- | --- | --- |
|  |  |  |

**3. Species Prediction**

1. **Visualization of Features**

**Include at least two scatterplots of pairs of features.**

| Visualization (labels should make clear which features are used) |
| --- |
|  |
|  |

You may extend the table if you have more results

**Of these three options, which two features (by themselves) are best able to classify the penguin species?**

1. Culmen Depth + Flipper Length
2. Flipper Length + Culmen Length
3. Flipper Length + Body Mass

|  |
| --- |

1. **Simple rule to identify Gentoo**

**Display your decision tree with labeled features and classes.**

|  |
| --- |

**Write down the simple two-part rule to identify Gentoo.** For example, the format should be “If Mass > 3000 and Culmen Depth < 17, then species is Gentoo”.

**If…**

|  |
| --- |

**and**

|  |
| --- |

**then species is Gentoo.**

**Rule precision**: fraction of penguins that satisfy this rule that are Gentoos (# gentoo predicted / # predicted)

| / |
| --- |

**Rule recall:** fraction of all Gentoo penguins that are identified as Gentoo using this rule (# gentoo predicted / # gentoo)

| / |
| --- |

1. **Model Design**

Describe the model that achieves best 5-fold cross-validation accuracy:

|  |
| --- |

5-fold Cross-Validation Accuracy: (xx.x%)

|  |
| --- |

**3. Stretch Goals**

1. **Improve MNIST Classification Performance using MLPs**

Report the classification val and test errors and details of your best method. Describe your approach and parameters. Feel free to change the MLP batch size, optimizer (e.g. try Adam), learning rate, number of epochs, hidden layer size, activation layer, or anything else.

**Description and key parameters**

| Optimizer =  Hidden layer(s) =  Learning rate =  Number of epochs =  Any other details: |
| --- |

| Validation **Error** (%) | Test **Error** (%) |
| --- | --- |
|  |  |

1. **Find a second simple rule to identify Gentoo**

Provide the second two-part rule here (that is substantially different from your first rule).

**If…**

|  |
| --- |

**and**

|  |
| --- |

**then species is Gentoo.**

**Rule precision**: fraction of penguins that satisfy this rule that are Gentoos (# gentoo predicted / # predicted)

|  |
| --- |

**Rule recall:** fraction of all Gentoo penguins that are identified as Gentoo using this rule (# gentoo predicted / # gentoo)

|  |
| --- |

1. **Positional encoding**

Show the RGB image obtained by predicting directly from (x,y) and the image obtained by predicting from the positional encoding.

**Input to network is (x,y)**

|  |
| --- |

**Input to network is pos\_enc(x, y)**

|  |
| --- |

**Acknowledgments / Attribution**

List any outside sources for code or improvement ideas or “None”.

<https://pytorch.org/tutorials/beginner/blitz/neural_networks_tutorial.html>

<https://pythonguides.com/pytorch-fully-connected-layer/>

https://pytorch.org/docs/stable/generated/torch.randn.html

<https://machinelearningmastery.com/building-multilayer-perceptron-models-in-pytorch/>

https://datascience.stackexchange.com/questions/31041/what-does-logits-in-machine-learning-mean