Q1 Information and Honor Code
0 Points

In this assignment, you will work on the Colab 7 notebook and obtain results from it. If your answers are float values, round the decimal number to the **nearest 0.001**. For example, 0.2435 would become `0.244`.

You can submit as many times as you want, and the last submission will be graded. Only the fully correct answer will receive 1 point. No late day is allowed for any Colab assignment.

Please verify that you have read the above instructions and the Stanford Honor Code and that you have not given or received unpermitted aid while completing this assignment.

If you have any questions about how the Honor Code applies to Colab assignments or other parts of the course, please contact the teaching staff for clarification.

⊙ I have read and understood the above information

Q2 Decision tree classification on the MNIST database
9 Points

Q2.1 Train size
1 Point

How large is the train dataset? (Integer)

Q2.2 Test size
1 Point

How large is the test dataset? (Integer)

Q2.3 Label for the fifth row in test set
1 Point

You then obtain the prediction on the first 10 digits. For the fifth data sample in test set, what is the label? (**Integer**)

Q2.4 Predicted label for the fifth row in test set
1 Point

For the fifth data sample in test set, what is the prediction label? (**Integer**)

Q2.5 Test accuracy
1 Point

Running the `MulticlassClassificationEvaluator` on the predictions obtained from the default Decision Tree you just trained, what is your evaluation accuracy on the test set? We will accept answer +/- 0.01 from the reference. (Float)

Q2.6 Depth of the tree
1 Point

What is the maximal depth of the tree? (Integer)

Q2.7 Number of nodes on depth 10

After you train 21 different DTs, varying the max depth from 0 to 20.

Approximately, how many nodes are on the decision tree of depth 10? Choose the best answer below.

○ 250

○ 500

○ 1000

○ 2000

○ 4000

Q2.8 Number of nodes on depth 20
1 Point

Approximately, how many nodes are on the decision tree of depth 20? Choose the best answer below.

○ 1000

○ 2000

○ 4000

○ 8000

○ 16000

Q2.9 Overfit
1 Point

In which range of the max depth hyperparameter do you start to observe evidence of model overfitting if any?

○ [0, 5]
○ [6, 10]
○ [11, 20]
○ There is no evidence of overfit