

Wesleyan University, Fall 2021, COMP 411
Homework 3: Decision Trees Again
Due by 11:59pm on October 8, 2021

1. RETHINKING YOUR DECISION TREE IMPLEMENTATION

Use cross-validation to evaluate your decision tree implementation. Check whether you are overfitting to your data. If you are, try the pre-pruning and/or the post-pruning methods described in lecture 6, and discuss if anything changes.

2. SCIKIT DECISION TREE

Use scikit's decision tree implementation to train a model to perform classification using your dataset. Scikit allows you to choose between using the Gini index and entropy as measures of impurity. Try both. How do your results compare with your own decision tree implementation?

<https://scikit-learn.org/stable/modules/tree.html>

<https://scikit-learn.org/stable/modules/generated/sklearn.tree.DecisionTreeClassifier.html>

3. FEEDBACK ON SLIDES

Let me know what you found easy and hard to understand on the slides. Any comments or criticisms are appreciated.

4. THOUGHTS ON OTHER TOPICS TO COVER LATER IN THE COURSE

If there are specific topics related to neural networks that you would like to see covered, let me know. If there are other topics you might be interested in seeing covered, also tell me, though it may be less likely that I can cover them.