Homework 2 (100 points)

Ensemble Learning

In this homework, you will solve the same problem from homework 1 using a different machine learning technique. Upload a .txt file with a link to your file as your submission on Submitty.

You need to perform the following tasks for this homework:

Task 1 (30 points): Implement a Decision Tree Classifier for your classification problem. You may use a built-in package to implement your classifier. Try modifying one or more of the input parameters and describe what changes you notice in your results. Clearly describe how these factors are affecting your output.

Task 2 (30 points): From the Bagging and Boosting ensemble methods pick any one algorithm from each category. Implement both the algorithms using the same data. Use k-fold cross validation to find the effectiveness of both the models. Comment on the difference/similarity of the results.

Task 3 (40 points): Compare the effectiveness of the three models implemented above. Clearly describe the metric you are using for comparison. Describe (with examples) Why is this metric(metrics) suited/appropriate for the problem at hand? How would a choice of a different metric impact your results? Can you demonstrate that?