UNIVERSITY OF TEXAS AT AUSTIN

Homework Assignment 9

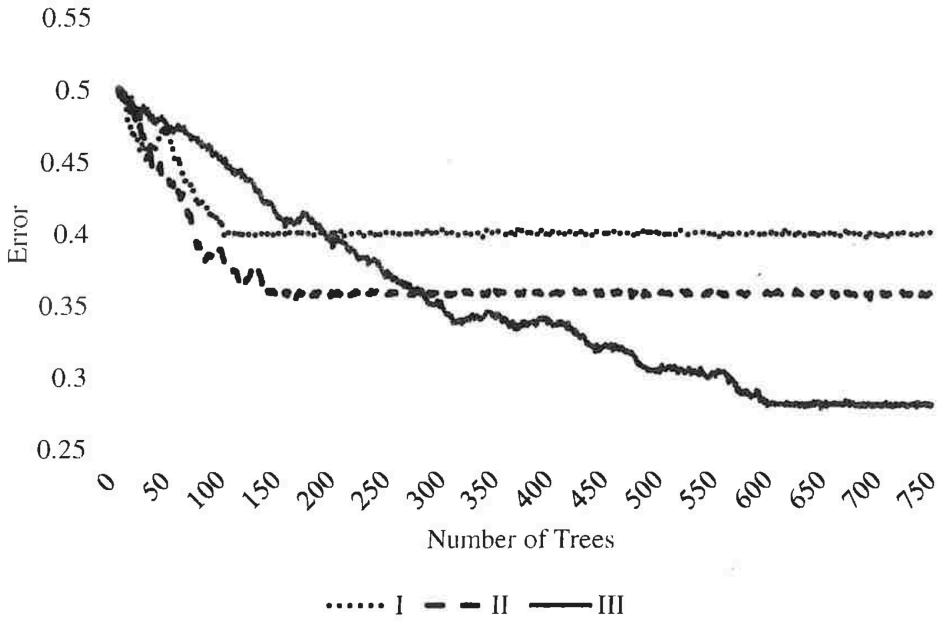
More on trees.

Please, provide your **complete solutions** to the following problems. Final answers only, even if correct will earn zero points for those problems.

Problem 9.1. (10 points) Solve Problem **8.4.2** (pp.361) from the textbook.

Problem 9.2. (5 points) Solve Problem **8.4.5** (pp.362) from the textbook.

Problem 9.3. (5 points) *Source: MAS-II, Fall 2018.*
An actuary creates three tree-based models using bagging, boosting, and random forests. The error on the test data set, as a function of the number of trees in each model, is plotted on the graph below:

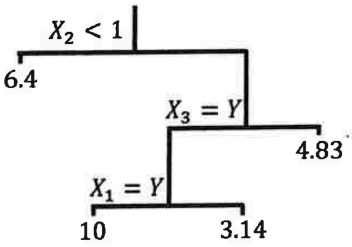


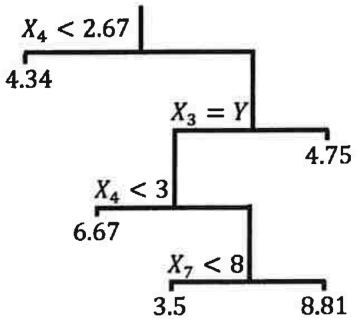
Determine the type of model most likely to have created each of the lines on the graph.

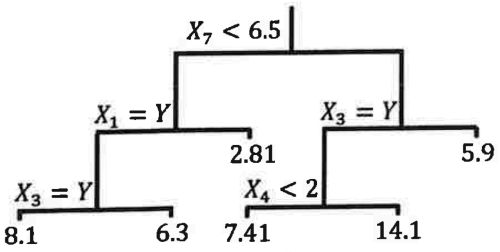
(a) I: Boosting, II: Bagging, III: Random forest
(b) I: Bagging, II: Boosting, III: Random forest
(c) I: Bagging, II: Random forest, III: Boosting
(d) I: Random forest, II: Bagging, III: Boosting
(e) None of the above.

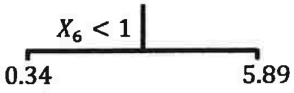
Problem 9.4. (15 points) *Source: MAS-II, Spring 2019.*
A boosted tree model is defined by:

- $\lambda = 0.2$
- The following four trees:









You are given the following record:

X_1	X_2	X_3	X_4	X_5	X_6	X_7
N	6	Y	4	0.5	0.25	6

Calculate the prediction of the boosted tree model for this record.

Problem 9.5. (10 points) A classification tree was constructed in order to predict the value of a categorical random variable with levels I, II, and III. A split of a specific node in the tree yielded these two regions:

Region	Count of I	Count of II	Count of III
R_1	40	10	10
R_2	5	25	10

Calculate the Gini index, the cross-entropy, and the classification error for this split.

Problem 9.6. (10 points) Consider a node in a classification tree. You are considering making a split at that node. Right now, there are 30 lapses and 10 non-lapses at that node. What would be the total reduction in the Gini index if this node were split into two regions so that one of them has 20 lapses and 4 non-lapses? Was this a meaningful split in and of itself?