CS7641 ML Practice Quiz

Module SL 5: Ensemble B & B

Created by Kyle Nakamura with the help of ChatGPT

Question 1

Which of the following statements accurately reflect the concepts discussed in Ensemble Learning and Boosting?
A. Ensemble learning combines multiple complex algorithms like decision trees and neural networks to make more accurate decisions.
B. Simple rules in ensemble learning, such as email characteristics, can individually determine if a message is spam with high accuracy.
C. Ensemble learning involves learning over subsets of data and then combining the learned rules into a single, more complex rule.
D. In ensemble learning, analyzing subsets of data never reveals simple rules that may be missed when considering the entire dataset.
E. Boosting focuses on examples that are difficult to classify and improves performance by focusing on these challenging cases.

Question 2

Regarding the Ensemble Learning Algorithm and its application, which of the following are true?
A. Combining rules in ensemble learning is typically done by averaging them, especially in regression tasks.
B. The effectiveness of ensemble learning is diminished by the complexity of combining algorithms, such as averaging.
C. Ensemble learning algorithms uniformly and randomly choose subsets of data for analysis in every scenario.
D. In ensemble learning, the distribution of examples is considered when defining the error, especially in boosting.
E. Boosting as a form of ensemble learning randomly selects subsets of data without any focus on example difficulty.

Question 3

Concerning the concepts of boosting in ensemble learning, which statements are correct?
A. Boosting involves giving equal importance to all examples, regardless of their difficulty.
B. In boosting, a weighted mean is used to combine rules, addressing the limitations of simple averaging.
C. Boosting does not require consideration of previously mastered examples when focusing on difficult ones.
D. Error in boosting is defined only as the number of mismatches and does not consider the distribution of examples.
E. Boosting aims to improve performance by iteratively focusing on examples that the current level of learning has not yet mastered.

Question 4

What are the key aspects and implications of the weak learning concept in boosting?
A. A weak learner is an algorithm that performs poorly on the training set, regardless of the data distribution.
B. The expected error of a weak learner is always less than half, making it better than random guessing.
C. Weak learners are irrelevant in the boosting process as they do not contribute significantly to the final hypothesis.
D. In boosting, a weak learner must perform well on a variety of data, not just on the last set of difficult examples.
E. Boosting relies on the concept that there cannot be many examples that are continuously misclassified.

Question 5

With regard to the final hypothesis in ensemble learning, which of the following statements are accurate?
A. The final hypothesis is a simple average of all the weak classifiers without considering their individual weights.
B. The weights for each weak classifier are determined by alpha sub T, which reflects the classifier's relative performance.
C. The final hypothesis in boosting does not involve any thresholding function and directly uses the weighted sum.
D. In ensemble learning, the final hypothesis combines weak classifiers without considering their error rates or weights.
E. Boosting involves constructing a distribution and combining weak classifiers to create the final hypothesis.

Answer Key

1. C, E
2. A, D
3. B, E
4. B, D, E
5. B, E