Lab - 6.

Random Forest.

16/04/2025

1] Start with the data.

Input

Training dataset D with n example and m features
Number of trees T to create

2] For each tree t from 1 to T

Step 1: Randomly select a bootstrap sample of size n from dataset D.
This subset is used to train tree t.

Step 2: For each node in the tree.

* Randomly select k features from total m features
* choose the best feature among k features to split the node based on a criterion
* Repeat until tree reaches the stop condition.

3] For each test sample x:

Pass x through each of the T decision tree T to get a prediction.

→ If it's a classification program each tree votes for a class

→ If its a regression problem, each tree provides a predicted value

④ Output:

For classification: The final prediction is the majority vote of all T trees.

For regression: the final prediction is the average of predictions for all T trees.

5] End.

16.04