Indian Institute of Technology Madras

ID5055 Foundations of Machine learning

Tutorial VI

Due date: 11:59 pm, October 8, 2023

Instruction

1. Assignment shall be submitted on the due date. Late submissions will not be entertained. If you cannot submit the assignment due to some reasons, please contact the instructor by email.
2. All the assignments must be the student's own work. The students are encouraged to collaborate or consult friends. In the case of collaborative work, please write every student's name on the submitted solution.
3. If you find the solution in the book or article or on the website, please indicate the reference in the solutions.

Problems

1. [Decision Tree Classifier]

   For the same dataset used in tutorial today :

   a. Set aside a test set.

   b. Experiment with different splitting criterion : Gini, Entropy and misclassification rate(log_loss).

   c. Also Hypertune only the max_depth parameter for each of 3 criterion. (look at tutorial 5)

   Hypertuneing methods may include k- fold validation or keeping a validation set aside method.

2. Your task inlcude:

   1. For each of the splitting criterion

      - report(plot) the value of max_depth parameter found after hypertuning
      - report(plot) the decision boundary of the decision tree
      - report the test error.
      - report which one works best and just give some intuition on why so.