School of Computer Science Engineering and Technology

Lab No. - 10

Course-B. Tech.	**Type-** Core
Course Code- CSET301	**Course Name-** Artificial Intelligence and Machine Learning
Year- 2025	**Semester-** Odd
Date-	**Batch-** 2023-2027

CO-Mapping

	CO1	CO2	CO3	CO4	CO5	CO6
Q1		✓		✓		

Lab – Random Forest classifier **Total Marks: 1**

Objective:

To Implement Random Forest classifier in Python Programming Language on Wine Quality Dataset.

Problem Statement:

About Dataset:

Data Set Characteristics:	Multivariate	**Number of Instances:**	4898	**Area:**	Business
Attribute Characteristics:	Real	**Number of Attributes:**	12	**Date Donated**	2009-10-07
Associated Tasks:	Classification, Regression	**Missing Values?**	N/A	**Number of Web Hits:**	1942447

The list of attributes with description is given below:
Input variables (based on physicochemical tests):

1 - fixed acidity
2 - volatile acidity
3 - citric acid
4 - residual sugar
5 - chlorides
6 - free sulfur dioxide
7 - total sulfur dioxide
8 - density
9 - pH
10 - sulphates
11 - alcohol

Output variable (based on sensory data):
12 - quality (score between 0 and 10)

Questions

Dataset: Download the dataset from the link.
https://archive.ics.uci.edu/ml/datasets/wine+quality

1. Read the dataset.
2. Extract the Independent and Dependent Variable.
3. Convert the Output column quality (score between 0 and 10) into the three categories i.e., best, average, and poor.
4. Split the dataset into training and testing using 75-25 division.
5. Perform normalization on numerical features.
6. Build a Random Forest classification model using Sklearn with default parameters. Predict the target values in the testing set.
7. Check the performance of model using confusion matrix.
8. Calculate the accuracy, Precision, Recall and F1-score.
9. Compare the performance of the Random Forest model with Decision Tree Model.