

Lab 2: 28-01-2026

1. To implement the bootstrap resampling technique to estimate the performance of a machine learning classifier.

Problem Statement:

Using the Iris dataset, apply the bootstrapping method to evaluate the accuracy of a classification model i.e., decision tree model. Generate multiple bootstrap samples, train the model on each sample, and compute the mean accuracy and standard deviation. Compare the result with a simple train–test split.

2. Consider a Students marksheet dataset containing students' marks. Calculate their percentage, percentile, grade (M.Tech (AI)). Rank the students based on each of these measures and observe how the ranking differs for each method.

Use these calculated values and ranks to train a machine learning model and analyze whether different performance metrics and ranking orders affect the model's behavior or results.

Range	Percentage Labels	Percentile Labels
91-100	C1	P1
81-90	C2	P2
71-80	C3	P3
61-70	C4	P4
50-60	C5	P5
<50	F	F

Note: <https://www.kaggle.com/datasets/rohithmahadevan/students-marksheet-dataset>