

Lab Assignment 2 - Decision Tree Part 2 - CART

Build a **CART decision tree** using **Sklearn** for the following problem of finding whether a student is single or committed 😊

Part 1:

- Use the **Sklearn** package to implement the CART Decision tree for the following data. After training, finally visualize the tree, print the importance of features (Gini values), properties of the tree such as number of leaves, depth of the tree etc.

S.No	Branch	CGPA	Gamer	Movie Fanatic	Committed?
1	CSE	High	Yes	No	No
2	CSE	Low	Yes	No	No
3	CSE	High	Yes	Yes	No
4	CSE	High	No	No	Yes
5	CSE	Low	No	Yes	Yes
6	ECE	Low	Yes	No	No
7	ECE	High	Yes	Yes	Yes
8	ECE	Low	Yes	Yes	No
9	ECE	High	Yes	Yes	Yes
10	ECE	High	No	Yes	Yes
11	MECH	High	Yes	Yes	No
12	MECH	High	No	No	No
13	MECH	High	No	No	Yes
14	MECH	Low	No	No	Yes
15	MECH	Low	No	No	Yes

- Test the CART model with the below samples and calculate the accuracy. Print the decision path for each of the below samples.

1	CSE	High	No	Yes	Yes
2	ECE	Low	Yes	No	No
3	MECH	Low	No	Yes	No

Part 2:

- Download 1 classification dataset (<https://tinyurl.com/uciclass>) and 1 regression dataset (<https://tinyurl.com/ucireg>) of your choice. Each of you should have unique datasets with you.
- Load the data, pre-process the data. Split the dataset into training and testing sets using built-in sklearn functions. Build a CART model for classification as well as regression and do the procedure as in part 1.

Suggested Platform: Python: Azure Notebook/Google Colab Notebook, packages such as Numpy, Pandas

Submission: Submit your files in **Single ipython Notebook** in LMS before **Sunday 5th Aug, 11.59 pm.**

Marking: Marking is based on both **performance during the lab hours** as well as **complete submission in LMS.**