CSCI 580 Assignment #3

Submission

- Due: 10/15/2024, 11:59:59 PM
- PDF on Canvas
- Code Submission from github (see later section)

Problem

1. (50 pts) Use the provided data file (linear_regression_data.csv on Canvas) for linear regression. Write python code to find out the linear model and plot it along with the data points. Since this data is about one dependent variable and its relation with one independent variable, you can use the covariance approach to compute the model parameters. Implement that approach in python and submit a pointer of your github project (see more specifics from the Submission Requirements Section)

2. (50 pts) Decision Tree Development. Based on the data in Table 1, compute the Gini Index and weighted Gini for each criteria in Table 2. Decide the split criteria to pick for the first decision level. Then decide the split criteria to pick for the next level in the decision tree. Follow what we have done in class, clearly show how the Gini indices are computed. Put your solutions in a PDF file and submit it on Canvas.

Table 1: data collection

ID	Study Hours	Study Group	Homework Done	Pass
1	6	A	Yes	Yes
2	8	B	Yes	Yes
3	7	B	No	Yes
4	5	A	Yes	No
5	9	C	No	Yes
6	10	B	Yes	Yes
7	3	A	Yes	No
8	2	C	No	No
9	4	C	Yes	No
10	12	B	No	Yes

Table 2: Split criteria to consider for the first level of the decision tree

Feature	Threshold	Split Criteria
StudyHours	6	StudyHours ≤ 6, vs > 6
StudyHours	7	StudyHours ≤ 7, vs > 7
StudyHours	8	StudyHours ≤ 8, vs > 8
StudyGroup	A	StudyGroup from A, vs not A
StudyGroup	B	StudyGroup from B, vs not B
StudyGroup	C	StudyGroup from C, vs not C
HomeworkDone	Yes	HomeworkDone, vs not Done

Submission Requirements

- Create a Github project that includes the following:
  - Name of your repository: CS580_Fall24_<Firstname>_<Lastname>
  - Within your repository, create project: Assignment_3
  - Within your project, include the following:
    - Source code (either pure python code, or python code in jupyter notebook)
    - A readme file describes how to use your code
    - A screenshot of how your code is successfully executed
- Submission:
  - On Canvas, link to your Github repo/projects
  - Make sure you have invited me to your github repo/projects (see Additional Notes)

Additional Notes:

For submission of github links, you need to either make your github repo public, or add me as a collaborator. Here is an instruction on adding people as collaborators.

https://docs.github.com/en/account-and-profile/setting-up-and-managing-your-personal-account-on-github/managing-access-to-your-personal-repositories/inviting-collaborators-to-a-personal-repository

You can find me on github using my email: bshen2@csueastbay.edu or my github account name: boshen-csuchico.