Programming Assignment 1: Decision Trees

Part 1: Implementation [7 points]

Your job in this exercise is to predict whether you will have a good night-out in Jerusalem for the coming New Year's Eve. Assume that you have kept a record of your previous night-outs with the following attributes.

- How densely the place is usually **occupied** {High, Moderate, Low}
- How the **prices** are {Expensive, Normal, Cheap}
- Volume of the **music** {Loud, Quiet}
- The **location** {Talpiot, City-Center, Mahane-Yehuda, Ein-Karem, German-Colony}
- Whether you are a frequent customer (**VIP**) {Yes, No}
- Whether this place has your **favorite beer** {Yes, No}
- Whether you **enjoyed** {Yes, No}

We have provided a data file (dt_data.txt) that contains the relevant records.

(a) Write a program to construct a decision tree based on the idea of splitting by Information Gain.
(b) Run your program on the data file.
(c) Make a prediction for (occupied = Moderate; price = Cheap; music = Loud; location = City-Center; VIP = No; favorite beer = No).

You can write your program in any programming language. However, you will have to implement the decision tree learning algorithm yourself instead of using library functions. Provide a description of the data structures you use, any code-level optimizations you perform, any challenges you face, and of course, the requested prediction.

Your code should print the decision tree that it produces. Please describe the format in which you print the decision tree in your report.

Part 2: Software Familiarization [Optional - No Credit]

Do your own research and find out about a library function that offers a good implementation of the decision tree learning algorithm. Please learn how to use it. Compare it against your implementation and suggest some ideas for how you can improve your code. Describe all these in your report.

Part 3: Applications [Optional - No Credit]

Do your own research and describe some interesting applications of decision trees.

Submission Guidelines

In your submission report, please include the names of all group members and mention their individual contributions. The maximum number of the members in a team is 2. The report should be in PDF format.

Your submission should include the code as well as the report and is due before **09/13**, **11:59pm** in an archive in the zip, tar.gz or tar.xz format. Your source code should have a comment line that contains the names of all group members. Only one submission is required for each group by one of the group members. Please submit your homework on **BlackBoard** (do NOT email the homework to the instructor or the TA).

畫出樹 → https://treelib.readthedocs.io/en/latest/

決策樹ID3算法，手把手用Python實現→ https://www.cnblogs.com/techflow/p/12935130.html

https://blog.csdn.net/csqazwsxedc/article/details/65697652

https://blog.csdn.net/LY_ysys629/article/details/72639568

https://blog.csdn.net/keyue123/article/details/82253538?utm_medium=distribute.pc_relevant.none-task-blog-BlogCommendFromMachineLearnPai2-4.channel_param&depth_1-utm_source=distribute.pc_relevant.none-task-BlogCommendFromMachineLearnPai2-4.channel_param

https://blog.csdn.net/komed/article/details/106676517?utm_medium=distribute.pc_relevant.none-task-blog-BlogCommendFromMachineLearnPai2-10.channel_param&depth_1-utm_source=distribute.pc_relevant.none-task-BlogCommendFromMachineLearnPai2-10.channel_param

https://xiaoxubeii.github.io/articles/decision-tree-id3/

https://www.itread01.com/content/1548029014.html

https://blog.csdn.net/sun830910/article/details/89318085