HOMEWORK #1
Decision Trees

Group Members

Name	USC ID
Aayush Sinha	7268779355
Abhishek Jangalwa	8882321574
Radhika Agarwal	3865981998

STEPS TO RUN THE PROGRAMS:

1. To run the main program the following libraries need to be installed:
   - Pandas
2. To run the scikitlearn program the following libraries need to be installed:
   - Pandas
   - Scikit-Learn
   - Graphviz to view the dot file.
3. When the sklearn program is run, a dot file is created. This dot file is used to visualize the decision tree that is generated by Scikit-learn library. Graphviz needs to be used to open the dot file and view it.

4. The given dt-data.txt file was modified- the parenthesis from the header and the row numbers was removed. We run the program using modified.txt, which is attached.

5. From the terminal type: python Decision_Tree_HW1.py <training file> <test file> to run the decision tree

   $ python Decision_Tree_HW1.py modified.txt test.txt

6. From the terminal type: python DecisionTreeSklearn.py <trainingFile> <testFile>

   $ python DecisionTreeSklearn.py modified.txt test.txt

7. The output decision tree is shown below:

LEVEL 1: [Occupied]

LEVEL 2: [[High->Location] Moderate->Location] Low->Price]]

LEVEL 3: [[[Talpiot->Enjoy{No}] City-Center->Enjoy{Yes}] Mahane-Yehuda->Enjoy{Yes}] German-Colony->Enjoy{No}]] [City-Center->Enjoy{Yes}] German-Colony->VIP] Ein-Karem->Enjoy{Yes}] Mahane-Yehuda->Enjoy{Yes}] Talpiot->Price]] [Normal->Location] Cheap->Enjoy{No}] Expensive->Enjoy{No}]]]

LEVEL 4: [[[[] [] [] []] [[] [No->Enjoy{No}] Yes->Enjoy{Yes}]] [] [] [Cheap->Enjoy{No}] Normal->Enjoy{Yes}]]] [[Ein-Karem->Enjoy{No}] City-Center->Enjoy{No}]] [] []]]

8. Output:

Test case is:

	Occupied	Price	Music	Location	VIP	Favorite Beer
0	Moderate	Cheap	Loud	City-Center	No	No

The result is: Enjoy : Yes