CS 6375- ASSIGNMENT 2 BONUS

Please read the instructions below before starting the assignment.

- You are given an opportunity to earn 10 bonus points if you complete everything mentioned in this assignment.

- You should use a cover sheet, which can be downloaded at:
  http://www.utdallas.edu/~axn112530/cs6375/CS6375_CoverPage.docx

- You are allowed to work in pairs i.e. a group of two students is allowed. Please write the names of the group members on the cover page.

- No free days are allowed for this assignment.

- You are free to use any of the following programming languages -
  Java, Python, C++, Ruby. You cannot use any machine learning packages or libraries. **The instructions for compiling and running your code must be specified in the README file.**

- **Your code will be checked for plagiarism. Do not copy code from online resources.**

ASSIGNMENT 2 BONUS

You are given an opportunity to earn 10 bonus points if you extend your assignment 2 as described below. *Note: You cannot use free days for this bonus part*.

Instead of using the ID3 algorithm to choose which attribute to select for splitting the data at each node, write a method that *randomly picks attributes* for each node. Every other requirement remains the same. <u>Also note that you cannot use the same attribute twice in a path</u>. There is no need to prune the tree.

Construct a new tree using *random* selection of attributes and compare the performance on the test dataset (in terms of accuracy) of the tree constructed using this approach to the one constructed earlier using ID3. You need to compare the trees *without pruning*.

Report your output as follows:

Tree parameters:

	Average Depth	Number of nodes
Tree constructed using ID3		
Tree constructed using random attribute selection		

Average depth will be computed as: $\dfrac{Sum\ of\ depth\ of\ the\ leaf\ nodes}{Total\ number\ of\ leaf\ nodes}$

Accuracy:
Run your random selection algorithm 5 times and report each run's accuracy.

Run #	Accuracy of tree constructed using random attribute selection
1	
2	
3	
4	
5	

Also, report the accuracy of the tree constructed using ID3.

- You are free to make any reasonable assumptions, provided you write them in the README file

**NOTE: To get the bonus points, you have to do everything specified above. Also, it is important that you have two different methods in your code – one for ID3 and the other for random attribute selection **

What to submit

- Complete code with the new method that randomly selects attributes.
- Report that contains comparison of the methods.
- README file indicating how to run your code and any assumption that you made