Assignment # 2

Due by June 11, 2018 (By 11:59 PM)

Decision Tree Induction and Naïve Bayes Classifier

For the following problems (P1 and P2) you are required to use the Mushroom dataset and a programming language of your choice (Java, C++, Python, Matlab, etc). I have included the dataset (agaricus-lepiota.data) within the assignment folder. I have also included a PDF file containing useful description of the dataset downloaded from the following link: https://archive.ics.uci.edu/ml/datasets/mushroom

For the following problems, you are required to randomly split the dataset into training (80%) and testing (20%) sets. For simplicity in training/testing your models, you may ignore the instances (records) containing missing values.

P1. (25 Points)

Induce a decision tree using ID3 algorithm from the Mushroom dataset. Report the accuracy on the test set.

P2. (25 Points)

Create a Naïve Bayes classifier using the dataset. Report the accuracy on the test set.

Submission Notes:

Email your assignment in a zipped folder containing the following at **cs360ai@gmail.com**

1. A softcopy(PDF) of your report including your code, screenshots, etc.
2. A folder containing your code files
3. Subject line of your email **MUST** be your ID followed by assignment number, e.g., BCSF15M500-Assignment2
4. The name of your attachment (zipped folder) **MUST** be a combination of your ID and assignment number, e.g., BCSF15M500-Assignment2