Decision trees

1) Given dataset Golf with 4 attributes Outlook, Temp, Humidity, Windy and an attribute Play (class).

Outlook	Temperature	Humidity	Windy	Class
sunny	85	85	false	Don't Play
sunny	80	90	true	Don't Play
overcast	83	78	false	Play
rain	70	96	false	Play
rain	68	80	false	Play
rain	65	70	true	Don't Play
overcast	64	65	true	Play
sunny	72	95	false	Don't Play
sunny	69	70	false	Play
rain	75	80	false	Play
sunny	75	70	true	Play
overcast	72	90	true	Play
overcast	81	75	false	Play
rain	71	80	true	Don't Play

- How to build the decision tree model fro classifying the dataset
- How many inductive rules are there in the decision tree model
- Use the decision tree model to classify 3 examples as follows:

Outlook	Temperature	Humidity	Windy	Class
overcast	63	70	false	?
rain	73	90	true	?
sunny	70	73	true	?

2) Implement the program using **Decision Tree Classifier** in **scikit-learn** library. The program requires 2 parameters:

- file name of trainset
- file name of testset

The program reports the classification results (accuracy, confusion matrix) for 5 datasets:

- Iris (.trn: trainset, .tst: testset)
- Optics (.trn: trainset, .tst: testset)
- Letter (.trn: trainset, .tst: testset)
- Leukemia (.trn: trainset, .tst: testset)
- Fp (.trn: trainset, .tst: testset)

3) Implement the program using **AdaBoost Classifier, Bagging Classifier, Random Forest Classifier** in **scikit-learn** library.

4) Comparison of classification results between decision trees and ensemble-based methods.

5) Why ensemble-based models improve the classification correctness of any single tree model?