Homework 1 Decision Tree and Tree ensemble

Given a dataset as below, answer Question 1 to Question 3

Feature 1	Feature 2	Class
Sunny	Hot	No
Sunny	Mild	No
Overcast	Hot	Yes
Rainy	Cool	Yes
Rainy	Mild	Yes
Rainy	Hot	No

1. Compute Entropy of Feature 1 = '*Overcast*'

2. Compute Information Gain for Feature 2

3. Draw the decision tree for this dataset using ID3 algorithm

4. Given a dataset with 1,000 samples and 20 features, explain how to build random forest with 5 trees?

5. Describe Bagging (Random Forest) and Boosting (XGBoost).

6. Given a dataset with 10 samples and 2 features: X1 and X2, compute trees F1 and F2 using XGBoost algorithm

จุด	X_1	X_2	y (คลาส)
A	1.0	4.0	1
B	1.5	5.5	1
C	2.0	2.0	0
D	2.5	3.0	0
E	3.0	4.5	1
F	3.5	1.0	0
G	4.0	5.0	1
H	4.5	2.5	0
I	5.0	3.5	0
J	5.5	5.5	1

Given learning rate η = 0.5, F0 = 0, P0 = 0.5 and max depth = 1

Epoch	Criterion split
1	X1 < 3.0
2	X2 > 3.5

Homework 1 Decision Tree and Tree ensemble

Homework 1 Decision Tree and Tree ensemble