

CSE352

HOMEWORK 2 (Decision Tree Learning 1) – 15pts

TRAINING DATA SET FOR THE HOMEWORK: Class Attribute: **Buys Computer**

Age	Income	Student	Credit Rating	Buys Computer
<=30	high	No	Fair	No
<=30	high	No	Excellent	No
31...40	high	No	Fair	Yes
>40	medium	No	Fair	Yes
>40	Low	Yes	Fair	Yes
>40	low	Yes	Excellent	No
31...40	low	Yes	Excellent	Yes
<=30	medium	No	Fair	No
<=30	low	Yes	Fair	Yes
>40	medium	Yes	Fair	Yes
<=30	medium	Yes	Excellent	Yes
31...40	medium	No	Excellent	Yes
31...40	high	Yes	Fair	Yes
>40	medium	No	Excellent	No

Problem 1

Use the Training Data to create two decision trees:

1. one with **general majority** voting , as defined in lecture notes, i.e. majority voting at any node of your choice.

Use CREDIT RATING as the root attribute, and nodes attributes of your choice;

2. one without general majority voting; i.e use ID3 algorithm (without Information Gain). Use INCOME as root attribute, and nodes attributes of your choice;

EVALUATE Predictive accuracy for each of your trees (sets of rules) – use the TEST Dataset below.

TEST DATA SET

Obj	Age	Income	Student	Credit_Rating	Class
1	<=30	High	Yes	Fair	Yes
2	31...40	Low	No	Fair	Yes
3	31...40	High	Yes	Excellent	No
4	>40	Low	Yes	Fair	Yes
5	>40	Low	Yes	Excellent	No
6	<=30	Low	No	Fair	No

Problem 2

Create test data sets for your sets rules corresponding to trees 1 and 2 that guarantees 100% predictive accuracy.

Problem 3

Compute the predictive accuracy of the set of discriminant rules in the lecture notes L8 with respect of the TEST Dataset from Problem 1.