John Scebold

Homework-1

Question 1

You are given the following training data:

	WRITABLE	UPDATED	SIZE	CLASS
1	yes	no	small	text
2	yes	yes	large	text
3	no	yes	med	text
4	no	no	med	executable
5	yes	no	large	executable
6	no	no	large	executable

The target attribute is CLASS.

1

Compute the decision tree generated by first considering WRITABLE, then UPDATED, then SIZE.
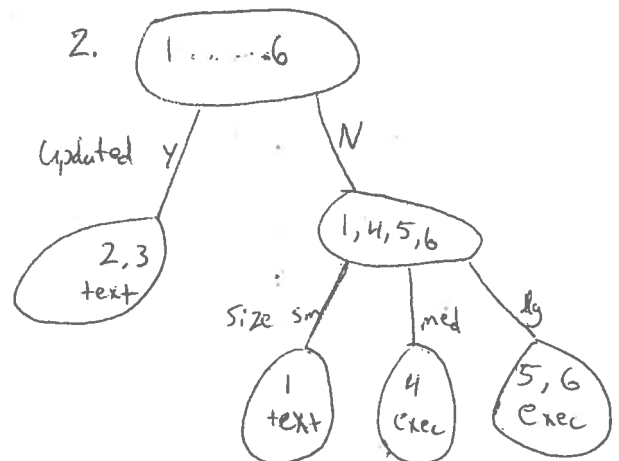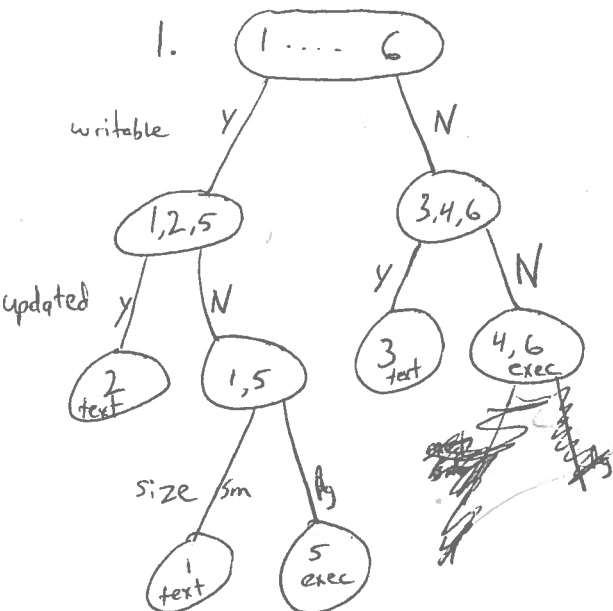
2

What tree is generated by using ID3?

3

How would these trees classify the following two instances:

	WRITABLE	UPDATED	SIZE	CLASS	Tree 1	Tree 2
1	yes	yes	large	?	text	text
2	no	no	small	?	exec	text

1.

2.

I wrote some python to calculate the gains

$gain(S_0 | Writable) = 0.0817$

$gain(S_0, updated) = 0.459148$ ✶

$gain(S_0, size) = 0.20752$

$gain(S_1, size) = 0.8112781$ ✶

$gain(S_1, writable) = 0.3112781$