Answers

1. What is the difference (if any) between the two decision trees?

J48 is not a binary tree whereas BF tree is a binary tree.

2. Why is there a difference?

J48 tree is based on C4.5 algorithm which splits on the attribute having the maximum information gain. Thus, the number of children at a node is equal to the number of different kinds of values that the node has.

BF tree is based on Gini index. It creates subsets and for each subset, a decision is made where the element belongs to the set or not, thus the tree becomes binary.

3. Take one example that you create on your own and explain how each decision tree will be used to predict the class for your example

Example – Let us try to classify the following tuple

No.	age	sex	region	income	married	children	car	save_act	current_act	mortgage
1	78	MALE	INNER_CITY	77546	NO	2	NO	YES	NO	NO

To predict the class, we inspect the nodes of the tree and traverse the tree by comparing the values in the tuple with the decision criterion labelled on each edge.

Using J48 –
   1) 'Children' is the root. We choose the edge children=2
   2) Now we inspect if 'income' > 30189.4
   3) We choose this path and reach the decision YES.

Using BF tree –
   1) 'Children' is the root. We choose the edge children=(3)|(0)|(2)
   2) We check for 'married' attribute. Since it isn't YES, we choose the path married!=(YES)
   3) Now we check for 'mortgage'. Since it isn't YES, we choose the path mortgage!=(YES)
   4) We again check for children and choose children=(3)|(2)|(1)
   5) Now we check for 'income' and since it's greater than 30340.85, we again check for children!=3 and hence the decision made is YES.