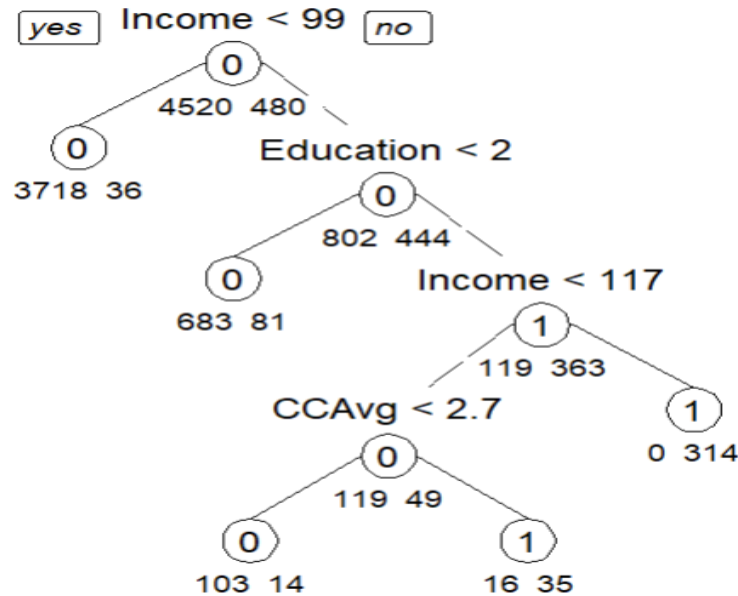# Homework 1

Due: 9/14

Problem 1 (10 points). Answer the following questions. Your answer must be clear, unambiguous, and accurate.

(1). Describe the difference between classification and clustering. Show one real-world example of classification and one real-world example of clustering.
(2). We illustrated association rule mining using a market-basket example. Show another real-world example of association rule mining.
(3). What are two basic principles of clustering (you must describe these two principles)?
(4). What is the most important challenge of mining from a very large amount of data?

Problem 2 (10 points). Consider the following decision tree, which we discussed in the class:

Classify the following three persons using the above decision tree:

Age	Income	CCAvg	Education	Mortgage	Personal Loan
62	142	1.5	1	0	
45	63	3.9	3	270	
33	105	3.1	2	450	

Problem 3 (10 points). Consider the following transactional database:

Transaction database

Customer	Items
C1	bread, chip, egg, milk
C2	beer, chip, egg, popcorn
C3	bread, chip, egg
C4	beer, bread, chip, egg, milk, popcorn
C5	beer, bread, milk
C6	beer, bread, egg
C7	bread, chip, milk
C8	bread, butter, chip, egg, milk
C9	butter, chip, egg, milk

(1). Determine the supports of the following itemsets:

{beer, egg}
{chip, egg, milk}

(2). Calculate the confidences of the following rules:

{butter} => {milk}
{bread, egg} => {milk}

Submission:

Submit the solutions in a single Word or PDF document and upload it to Blackboard. Use *LastName_FirstName_hw*1.*docx* or *LastName_FirstName_hw*1.*pdf* as the file name. If necessary, you may submit an additional file that shows how you obtained your answers.