CS 1656 - Introduction to Data Science - Fall 2016

Prof. Alexandros Labrinidis - Department of Computer Science - University of Pittsburgh

13 - Classification - Oct 12, 2016

Assume the following table with information about different people and whether they were given a loan or not from a bank.

Age	Credit Rating	Sex	LOAN_OK?
25	High	Male	Yes
25	Medium	Female	No
39	Low	Male	No
24	High	Female	Yes
35	Medium	Male	No
24	Low	Female	No
47	High	Male	Yes
29	Medium	Female	No
59	Low	Male	No
51	High	Female	Yes

- **(Q1)** Given the above table, which **ONE** attribute can be used to accurately predict the LOAN_OK attribute?
 - Age
 - Credit Rating
 - Sex
 - None of the above

Assume that additional people were given loans, as follows:

Age	Credit Rating	Sex	LOAN_OK?
42	Medium	Female	Yes
45	Medium	Female	Yes

- (i.e., we now need to consider both tables).
- **(Q2)** Given the updated table, can only **ONE** attribute still be used to accurately predict the LOAN_OK attribute?
 - Yes
 - No
- (Q3) If you answered no to the previous question, given the updated table, which additional attribute needs to be used to accurately predict the LOAN_OK attribute?
 - Age
 - Credit Rating
 - Sex