

Due Date: Webcampus
How to submit: Webcampus

HW3-1: Rule-Based Classifier

Consider a training set that contains 70 positive examples and 200 negative examples. For each of the following candidate rules,

$R_1: A \rightarrow +$ (covers 9 positive and 0 negatives)

$R_2: B \rightarrow +$ (covers 65 positive and 40 negatives)

determine which is the best and the worst candidate rule according to:

- a) The Rule accuracy
- b) The Ratio statistic
- c) The Laplace measure
- d) The m-estimate measure

HW3-2. Naive Bayes Classifier

Consider Table 1.

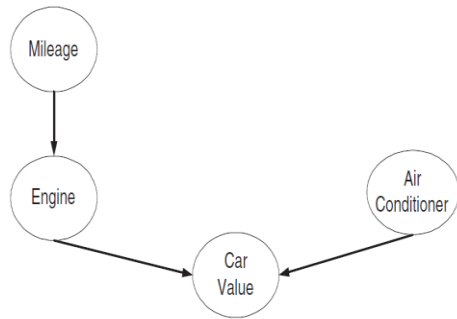
Instance	A	B	C	Class
1	0	0	0	+
2	1	0	1	+
3	1	1	0	+
4	1	1	0	+
5	1	1	0	+
6	0	1	0	+
7	1	0	0	+
8	0	0	1	-
9	1	1	1	-
10	0	0	1	-

Table 1

- a) Estimate the conditional probability for $P(A = 1|+)$; $P(B = 1|+)$; $P(C = 1|+)$; $P(A = 1|-)$; $P(B = 1|-)$ and $P(C = 1|-)$.
- b) Predict the class label for the test sample ($A = 0$, $B = 1$, $C = 1$).
- c) What is the time complexity for training Naive Bayes classifier?
- d) What is the testing time complexity of Naive Bayes classifier?

HW3-3. Bayesian Belief Network

Consider the following Bayesian belief network for the data set shown in Table 2,



Mileage	Engine	Air Conditioner	Number of Records with Car Value=Hi	Number of Records with Car Value=Lo
Hi	Good	Working	3	4
Hi	Good	Broken	1	2
Hi	Bad	Working	1	5
Hi	Bad	Broken	0	4
Lo	Good	Working	9	0
Lo	Good	Broken	5	1
Lo	Bad	Working	1	2
Lo	Bad	Broken	0	2

Figure 1: Bayesian belief network

Table 2

- Draw the probability table for each node in the network.
- Use the Bayesian network to compute $P(\text{Engine} = \text{Bad}, \text{Air Conditioner} = \text{Broken})$.