Independent University,Bangladesh.
CSE-417 Data mining and warehousing
Final Exam(Summer 2021)
Full Marks 30
Time- 3 H 00 M

1 Explain Bayesian theorem. How to avoid 0 probability problem?

The Bayes' theorem describes the probability of an event based on prior knowledge of the conditions that might be relevant to the event. The Bayes' theorem is expressed in the following formula:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Where:

- P(A|B) – the probability of event A occurring, given event B has occurred
- P(B|A) – the probability of event B occurring, given event A has occurred
- P(A) – the probability of event A
- P(B) – the probability of event B

One of the disadvantages of Naïve-Bayes is that if you have no occurrences of a class label and a certain attribute value together then the frequency-based probability estimate will be zero. And this will get a zero when all the probabilities are multiplied. An approach to overcome this 'zero-frequency problem' in a Bayesian environment is to add one to the count for every attribute value-class combination when an attribute value doesn't occur with every class value.