LAHORE UNIVERSITY OF MANAGEMENT SCIENCES
Department of Electrical Engineering

EE 514 (CS 535) Machine Learning
Quiz 8 Solutions

Name: _____

Campus ID: _____

Total Marks: 10
Time Duration: 15 minutes

Question 1 (4 marks)

1. What is the main assumption behind the Naïve Bayes classifier?
   a) Features are linearly dependent.
   b) Features are conditionally independent given the class.
   c) Features follow a Gaussian distribution.
   d) Features have equal variance.

   Solution: b) Features are conditionally independent given the class.

2. Which of the following is not a strength of Naïve Bayes?
   a) Handles missing values.
   b) Robust to outliers.
   c) Works well with redundant features.
   d) Easy to implement.

   Solution: c) Works well with redundant features.

3. In the context of Bayesian learning, what does the MAP estimation maximize?
   a) Only the prior probability of the parameters.
   b) Only the likelihood of the data.
   c) The posterior probability of the parameters given the data.
   d) The marginal likelihood of the data.

   Solution: c) The posterior probability of the parameters given the data.

4. You are working with a very small dataset that is likely not a good representative sample of the true data distribution. Considering this under-representation and assuming that a well-informed prior is used in MAP, which of the following best describes the likely effect of using MAP compared to ML?
   a) MAP will reduce variance without significantly increasing bias, leading to better generalization than ML.
   b) MAP will still have higher variance than ML due to the small dataset.
   c) ML will outperform MAP because it does not rely on any assumptions.
   d) Both MAP and ML will perform similarly since the dataset is small.

   Solution: (a) MAP will reduce variance without significantly increasing bias, leading to better generalization than ML.

Question 2 (6 marks)

You are given the following word frequency table extracted from a text classification task distinguishing between two categories: Positive and Negative reviews.

Word	Positive Count	Negative count
great	4	1
boring	0	3
acting	2	2
script	3	1
dull	0	2
fun	2	0

Assume equal class priors. You are given the test document: "great acting dull"

(a) [3 marks] Compute the class-conditional probabilities for each word in the document using Laplace smoothing.

(b) [2 marks] Calculate the log probability of the document belonging to each class.

(c) [1 mark] Based on the results, determine whether the document is more likely to be Positive or Negative.

Solution:

Vocabulary size: $V = 6$
Total words in Positive: 11
Total words in Negative: 9
With Laplace Smoothing:
Positive denominator: $11 + 6 = 17$
Negative denominator: $9 + 6 = 15$

(a) Class-Conditional Probabilities

Positive Class:

$$P(\text{great}|\text{Positive}) = \frac{4+1}{17} = \frac{5}{17}, \quad P(\text{acting}|\text{Positive}) = \frac{2+1}{17} = \frac{3}{17}, \quad P(\text{dull}|\text{Positive}) = \frac{0+1}{17} = \frac{1}{17}$$

Negative Class:

$$P(\text{great}|\text{Negative}) = \frac{1+1}{15} = \frac{2}{15}, \quad P(\text{acting}|\text{Negative}) = \frac{2+1}{15} = \frac{3}{15}, \quad P(\text{dull}|\text{Negative}) = \frac{2+1}{15} = \frac{3}{15}$$

(b) Log Probabilities of Document "great acting dull"

Positive Class:

$$\log P(\text{doc}|\text{Positive}) = \log\left(\frac{5}{17}\right) + \log\left(\frac{3}{17}\right) + \log\left(\frac{1}{17}\right) \approx -2.515$$

Negative Class:

$$\log P(\text{doc}|\text{Negative}) = \log\left(\frac{2}{15}\right) + \log\left(\frac{3}{15}\right) + \log\left(\frac{3}{15}\right) \approx -2.273$$

(c) Decision

Since $-2.273 > -2.515$, the document is more likely to belong to the **Negative** class.