Bahria University, Islamabad Campus

Course Title: [CSC 411] Artificial Intelligence
Instructor: Dr.Kashif Sultan
Date: 19th May, 2024
Due Date: 25th May 2024

Assignment # 03

(CLO3)

Instructions:

- Each student is expected to submit their own original work. Plagiarism or copying from other sources, including fellow students' assignments, is strictly prohibited.
- You are encouraged to explore various sources of information, including internet resources, books, and scholarly articles, to gather relevant information for your answers.
- Your answers should be written in your own words, demonstrating your understanding of the concepts presented in the scenarios. Avoid direct copying of text from external sources.
- Ensure that your answers are concise and to the point. Provide relevant examples or samples where necessary to support your explanations.
- If you refer to specific sources in your answers, be sure to provide appropriate citations or references.
- The viva of assignments will be conducted at the end of the course.

Question # 01:

Interpret the difference between supervised and unsupervised learning with a real-world example related to customer segmentation

Question # 02:

Use a real-world example to explain how a classification algorithm can be applied to diagnose medical conditions based on patient data.

Question # 03:

A biased coin (with probability of obtaining a Head equal to $p > 0$) is tossed repeatedly and independently until the first head is observed. Compute the probability that the first head appears at an even numbered toss.

Question # 04:

A medical test for a rare disease has 98% sensitivity and 97% specificity. The disease occurs in only 0.5% of the population. If a person tests positive, what is the probability they actually have the disease**?** Use Bayes' Theorem to solve it and explain the steps clearly.

Question # 05:

You're developing a robotic navigation system using Reinforcement Learning.

- Describe what would be the states**,** actions, and rewards in this context.
- How would the robot learn the best path in an unknown grid environment?

Question # 06:

In a city, 3% of the population uses illegal software. A digital monitoring system is used to detect such users. The system is 95% accurate in identifying actual illegal software users (true positive rate) and falsely flags 4% of legal users (false positive rate). If a person is flagged by the system, what is the probability that they are actually using illegal software? Use Bayes' Theorem and show all steps.

Question # 07:

A diagnostic test for a virus has a 90% sensitivity (true positive rate) and 95% specificity (true negative rate). If the virus occurs in 1% of the population**,** and a randomly selected person tests positive**,** what is the probability that the person actually has the virus? Apply Bayesian reasoning to compute the result.

Question # 08:

1% of a population have a certain disease and the remaining 99% are free from this disease. A test is used to detect this disease. This test is positive in 95% of the people with the disease and is also (falsely) positive in 2% of the people free from the disease. If a person, selected at random from this population, has tested positive, what is the probability that she/he has the disease?

Question # 09:

You are building an email spam detection system using the Naïve Bayes classifier.
The following dataset is available:

Word	P(word|Spam)	P(word|Not Spam)
"win"	0,08	0.01
"money"	0.05	0.02
"offer"	0.07	0.01

Assume:

- P(Spam) = 0.4
- P(Not Spam) = 0.6
- A new email contains the words "win", "money", and "offer"

Using Naïve Bayes, compute the probability that this email is spam**.** Show all calculations.

Question # 10:

Use the Naïve Bayes classifier to predict if a given fruit is a 'Banana' or 'Orange' or 'Other' when only the 3 features (long, sweet and yellow) are known. The dataset is given below.

Type	Long	Not Long	Sweet	Not Sweet	Yellow	Not Yellow	**Total**
Banana	400	100	350	150	450	50	500
Orange	0	300	150	150	300	0	300
Other	100	100	150	50	50	150	200
Total	500	500	650	350	800	200	1000

**********END**********