JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATION- 2024

B.Tech-VIII Semester (CSE/IT)

COURSE CODE (CREDITS): 19B1WCI832 (3) MAX. MARKS: 35

COURSE NAME: Probabilistic Graphical Models

COURSE INSTRUCTORS: Dr. Nancy Singla MAX. TIME: 2 Hours

Note: *(a) All questions are compulsory.*
(b)Marks are indicated against each question in square brackets.
(c)The candidate is allowed to make Suitable numeric assumptions wherever required for solving problems

---

Q1. Explain the concept of structured prediction in machine learning. Provide an example of a real-world problem where structured prediction is applicable, and discuss why traditional machine learning approaches might not be sufficient for solving such a problem effectively. [5] [CO1]

Q2. (a) Describe the key components of a Hidden Markov Model (HMM) and explain how they relate to the Virtebri algorithm. [3+7] [CO4]

(b) Lea is our imaginary friend and during the day she does either of these four things: Painting; Cleaning the house; Biking and Shopping for groceries. Now, let's suppose that in four days Lea does the following: painting, cleaning, shopping, biking. From this observation sequence, use the Viterbi algorithm to know whether the day has been sunny or rainy.
Consider the following properties of the Hidden Markov Model for the above scenario:

$$S = \{ S_{sunny}, S_{rainy} \} \quad (Hidden\ States)$$

$$O = \{ O_{clean}, O_{bike}, O_{shop}, O_{paint} \} \quad (Observables)$$

$$\pi = | 0.6 \quad 0.4 | \quad (Initial\ Probabilities)$$

$$A = \begin{vmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{vmatrix} \quad (Transition\ Probabilities)$$

$$B = \begin{vmatrix} 0.4 & 0.1 & 0.2 & 0.3 \\ 0.3 & 0.45 & 0.2 & 0.05 \end{vmatrix} \quad (Emission\ Probabilities)$$

Q3. (a) What is the difference between exact and approximate probabilistic inference methods in the context of Bayesian networks? [3+2+5] [CO3]
(b) Why Monte Carlo methods often struggle to perform well in high-dimensional spaces, particularly in the context of probabilistic inference.

(c) Provide examples of techniques used to efficiently sample high-dimensional distributions and explain how they address the issues associated with high dimensionality.

Q4. Suppose you are a data scientist working for a pharmaceutical company. Your team is tasked with estimating the parameters of a new drug's efficacy model based on clinical trial data. The drug is designed to lower blood pressure in patients with hypertension. The efficacy of the drug is modeled using a normal distribution with unknown parameters μ (mean) and σ (standard deviation). Explain how you would use Maximum Likelihood Estimation (MLE) to estimate the parameters μ and σ of the drug's efficacy model based on the observed clinical trial data. [5] [CO2]

Q5. Describe the two main steps involved in each iteration of the Expectation Maximization (EM) algorithm and discuss its convergence properties. [5] [CO3]