| U.S.N. | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|

# B.M.S. College of Engineering, Bengaluru-560019

**Autonomous Institute Affiliated to VTU**

## October 2024 Supplementary Examinations

**Programme: B.E.** **Semester: IV**
**Branch: AI & DS / Computer Science & Engineering(Data Science)** **Duration: 3 hrs.**
**Course Code: 23DS4PCMLG** **Max Marks: 100**
**Course: Machine Learning**

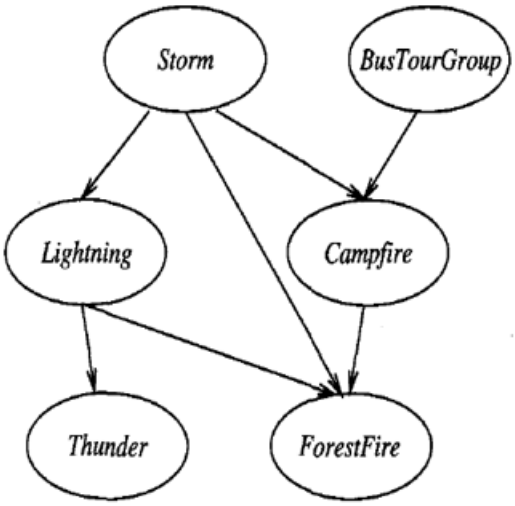
**Instructions**: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

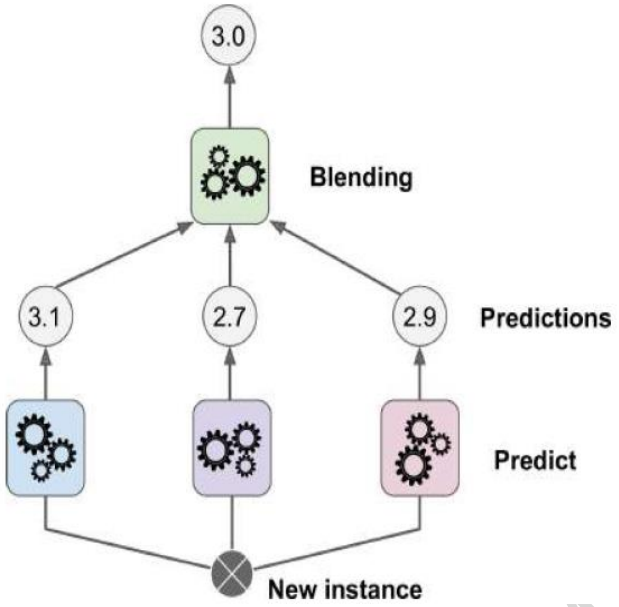
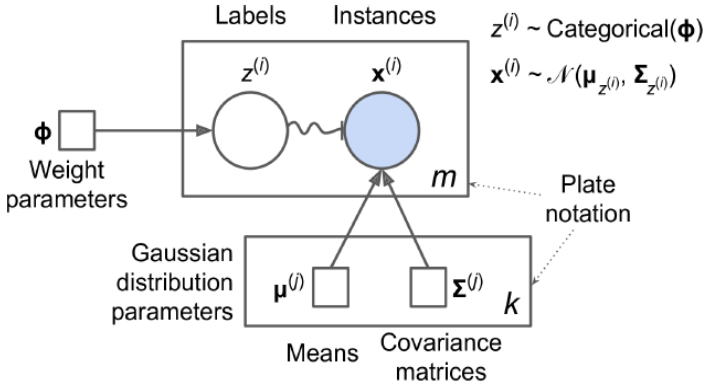




| | | UNIT – I | CO | PO | Marks |
|---|---|---|---|---|---|
| 1 | a) | Imagine you are the head of a cybersecurity team for an email service provider. Your team has been receiving numerous complaints from users about an increasing number of spam emails that are making it into their inboxes. The traditional rule-based filters your company has been using seem to be less effective against these sophisticated spam campaigns.<br><br>Given this scenario, how would implementing machine learning improve the accuracy and efficiency of spam detection compared to the current rule-based system? | CO2 | PO2 | 10 |
| | b) | Construct a decision tree using the ID3 algorithm for the following data.<br>Instance \| Classification \| a1 \| a2<br>1 \| + \| T \| T<br>2 \| + \| T \| T<br>3 \| - \| T \| F<br>4 \| + \| F \| F<br>5 \| - \| F \| T<br>6 \| - \| F \| T | CO3 | PO3 | 10 |
| | | **UNIT – II** | | | |
| 2 | a) | Enumerate on Linear Support Vector Machines and distinguish between soft margin and hard margin. | CO2 | PO2 | 06 |
| | b) | Elucidate on KNN technique using instance-based Learning. | CO1 | PO1 | 04 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | c) | Given the dataset, predict the outcome of an instance.<br>The features are BMI and Age, and the target variable is Diabetic.<br>NOTE: Diabetic = 1 indicates the patient is diabetic, and 0 indicates non-diabetic.<br><br>Design a k-NN algorithm-based system. Assume k=3.<br><br>**Sl. No. / BMI / Age / Diabetic**<br>1 / 33.6 / 50 / 1<br>2 / 26.6 / 30 / 0<br>3 / 23.4 / 40 / 0<br>4 / 43.1 / 67 / 0<br>5 / 35.3 / 23 / 1<br>6 / 35.9 / 67 / 1<br>7 / 36.7 / 45 / 1<br>8 / 25.7 / 46 / 0<br>9 / 23.3 / 29 / 0<br>10 / 31 / 56 / 1<br>11 / 43.6 / 40 / ? | ***CO3*** | ***PO3*** | **10** |
| | | | **UNIT - III** | | | |
| 3 | a) | | Summarize on Bayesian Learning for hypothesis "This pneumonia patient has 93% chance of complete recovery". | ***CO1*** | ***PO1*** | **5** |
| | b) | | Enumerate on<br>i. Bayes Optimal Classifier<br>ii. GIBBS Algorithm<br><br>Illustrate with an example on when to choose (i) and (ii). | ***CO2*** | ***PO2*** | **5** |
| | c) | | Analyze the working of the Naïve Bayes classifier. Find the probability to play tennis or not on the 15th day, if the attributes are as follows: OUTLOOK =Sunny, Temperature= Cool, Humidity= High, Wind= Strong. Justify<br><br>**Day / Outlook / Temperature / Humidity / Wind / PlayTennis**<br>D1 / Sunny / Hot / High / Weak / No<br>D2 / Sunny / Hot / High / Strong / No<br>D3 / Overcast / Hot / High / Weak / Yes<br>D4 / Rain / Mild / High / Weak / Yes<br>D5 / Rain / Cool / Normal / Weak / Yes<br>D6 / Rain / Cool / Normal / Strong / No<br>D7 / Overcast / Cool / Normal / Strong / Yes<br>D8 / Sunny / Mild / High / Weak / No<br>D9 / Sunny / Cool / Normal / Weak / Yes<br>D10 / Rain / Mild / Normal / Weak / Yes<br>D11 / Sunny / Mild / Normal / Strong / Yes<br>D12 / Overcast / Mild / High / Strong / Yes<br>D13 / Overcast / Hot / Normal / Weak / Yes<br>D14 / Rain / Mild / High / Strong / No | ***CO2*** | ***PO2*** | **10** |
| | | | **OR** | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 4 | a) | Given the dataset, predict the outcome of an instance.<br>NOTE: The target variable is Species.<br><br>Solve using the Naïve Bayes Classifier.<br><br>**Sl. No. \| Color \| Legs \| Height \| Smelly \| Species**<br>1 \| White \| 3 \| Short \| Yes \| M<br>2 \| Green \| 2 \| Tall \| No \| M<br>3 \| Green \| 3 \| Short \| Yes \| M<br>4 \| White \| 3 \| Short \| Yes \| M<br>5 \| Green \| 2 \| Short \| No \| H<br>6 \| White \| 2 \| Tall \| No \| H<br>7 \| White \| 2 \| Tall \| No \| H<br>8 \| White \| 2 \| Short \| Yes \| H<br>9 \| Green \| 2 \| Tall \| No \| ? | ***CO1*** | ***PO1*** | **10** |
| | | b) | Analyze the Bayesian Belief Network given below.<br>Consider the node: Campfire and expound on conditional dependence/ independence.<br><br>  <br> | ***CO2*** | ***PO2*** | **04** |
| | | c) | Outline the advantages, disadvantages, and applications of Expectation Maximization Algorithm. | ***CO1*** | ***PO1*** | **06** |
| | | | **UNIT – IV** | | | |
| | 5 | a) | i. Illustrate Hard voting and Soft voting classifiers with example.<br>ii. Write a Python program (use Python libraries) to demonstrate the working of Voting classifiers. | ***CO1*** | ***PO1*** | **12** |
| | | b) | Identify the difference between bagging and boosting. | ***CO2*** | ***PO2*** | **08** |
| | | | **OR** | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 6 | a) | Examine the idea of Stacking (stacked generalization) in Machine Learning using the flow diagram below.<br>  <br> | *CO2* | *PO2* | **10** |
| | | b) | Demonstrate the working of Gradient Boosting with a Python program. | *CO3* | *PO3* | **10** |
| | | | **UNIT – V** | | | |
| | 7 | a) | The diagram below illustrates the Gaussian Mixture Model.<br>  <br><br>Deduce the meaning of the above graphical representation. | *CO2* | *PO2* | **08** |
| | | b) | Enumerate how the Reinforcement learning problem is different from other function approximation tasks. | *CO2* | *PO2* | **06** |
| | | c) | Write the Q learning algorithm. | *CO1* | *PO1* | **06** |

**********