POSSESION OF MOBILES IN EXAM IS UFM PRACTICE

Name.....

Enrollment No. .....

10

Jaypee Institute of Information Technology, Noida

Test I Examination (ODD SEM 2023)

B. Tech, V Semester

Course Title: Fundamentals of Machine Learning

Course Code: 20B12CS331

Maximum Time: 1 Hour

Maximum Marks: 20

CO1	Understand the mathematical concepts of machine learning approaches.
CO2	Apply the fundamentals of linear algebra and probability theory to the machine learning problems.
CO3	Apply the concepts of regression analysis and vector calculus to the machine learning models.
CO4	Analyze the role of dimensionality reduction and density estimation for machine learning problems.
CO5	Evaluate and test the significance of machine learning results statistically.

Note: Attempt all questions in order.

Q1. You have to apply cosine similarity index on the following documents and identify which two documents are highest similar and less similar to each other.
Document 1: We went to the pizza place and you ate no pizza at all.
Document 2: I ate pizza with you yesterday at home.
Document 3: There's no place like home.

[CO1 (Understanding), 5 Marks]

Q2. Cluster the following eight points (with (x,y) representing locations) into three clusters: A1(2,10), A2(2,5), A3(8,4), A4(5,8), A5(7,5), A6(6,4), A7(1,2), A8(4,9). Initial cluster centers are: A1(2,10), A4(5,8) and A7(1,2). The distance function between two points a=(x1,y1) and b=(x2,y2) is defined as: $P(a,b)=|x2-x1| + |y2-y1|$. Use k-means algorithm to find the three cluster centers after the second iteration.

[CO1 (Understanding), 5 Marks]

Q3. Consider a linear transformation T: $R^4 \rightarrow R^3$ defined as follows:
$T(x,y,z,w) = (x+2y-z+3w, 2x-y+3z-2w, 3x+y-2z+w)$.
Find the kernel (null space) of the linear transformation T.

[CO2 (Applying), 5 Marks]

Q4. Find the SVD of a matrix A, where $A=\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$

[CO2 (Applying), 5 Marks]

POSSESSION OF MOBILES IN EXAM IS UFM PRACTICE

10

Name..... Enrollment No.....

Jaypee Institute of Information Technology, Noida
Test-2 Examination, 2023
B.Tech, V Semester

Course Title: Fundamentals of Machine Learning **Maximum Time: 1 Hour**
Course Code: 20B12CS331 **Maximum Marks: 20**

CO1	Understand the mathematical concepts of machine learning approaches.
CO2	Apply the fundamentals of linear algebra and probability theory to the machine learning problems.
CO3	Apply the concepts of regression analysis and vector calculus to the machine learning models.
CO4	Analyze the role of dimensionality reduction and density estimation for machine learning problems.
CO5	Evaluate and test the significance of machine learning results statistically.

Note: Attempt all questions in order.

Q1. You are working on a speech recognition system for a virtual assistant. The system distinguishes between two states: "Speech" (S) when the user is speaking and "Silence" (N) when the user is not speaking. You have the following parameters:

States: The states represent whether the user is speaking or in silence.
State 1: Speech (S)
State 2: Silence (N)
Observations (Emissions): The observations represent the sounds detected by the microphone.
Observation 1: Sound Detected (D)
Observation 2: No Sound Detected (X)
Initial Probabilities: The initial probabilities represent the probability of starting with each state.
P(Start with Speech)=0.4
P(Start with Silence)=0.6
Transition Probabilities:
P(S -> S) = 0.7
P(S -> N) = 0.3
P(N -> S) = 0.2
P(N -> N) = 0.8
Emission Probabilities:
P(D|S) = 0.9
P(X|S) = 0.1
P(D|N) = 0.2
P(X|N) = 0.8

a) Draw the transition diagram.
b) Given an observed sequence of microphone inputs: "D X D D X", use the Viterbi algorithm to find the most likely sequence of states.

[CO2 (Applying), 1+5=6 marks]

Q2. You are working on a classification project to predict whether a job applicant will be hired (1) or not hired (0) by a company based on their performance in a technical interview. You have a logistic regression model with the following equation:

Log(odds) = -0.5+0.8*Interview Score

where, Interview Score is the score the applicant received in the technical interview.

a) Predict the probability of a job applicant being hired if they scored 75 on the technical interview.
b) Determine the interview score a job applicant should achieve to have a 95% probability of being hired.

[CO3 (Applying), 2+2=4 marks]

POSSESSION OF MOBILES IN EXAM IS UFM PRACTICE

Q3. A random sample of eight drivers insured with a company and having similar auto insurance policies was selected. The following table lists their driving experiences (in years) and monthly auto insurance premiums. Calculate the two regression equations of X on Y and Y on X, taking deviations from actual means of X and Y.

Driving Experience (years)	Monthly Auto Insurance Premium ($)
5	64
2	87
12	50
9	71
15	44
6	56
25	42
16	60

[CO3 (Applying), 2.5+2.5=5 marks]

Q4. In a computer installation, 60% of programs are written in C++ and 40% in Java. 60% of the programs written in C++ compile on the first run and 80% of the Java programs compile on the first run.

a) What is the overall proportion of programs that compile on first run?

b) If a randomly selected program compiles on the first run, what is the probabilty that it was written in C++?

[CO2 (Applying), 2.5+2.5=5 marks]

(10)

POSSESION OF MOBILES IN EXAM IS UFM PRACTICE.

Name _____ _____ **Enrollment No.**____ _____

Jaypee Institute of Information Technology, Noida
End Term Examination, 2023 Odd Semester
B.Tech V Semester

Course Title: Fundamental of Machine Learning **Maximum Time :2 hrs**
Course Code: 20B12CS331 **Maximum Marks: 35**

C330-1.1	Understand the mathematical concept of machine learning approaches
C330-1.2	Apply the fundamentals of linear algebra and probability theory to the machine learning algorithm.
C330-1.3	Apply the concept of linear regression analysis and vector calculus to the machine learning models
C330-1.4	Analyze the role of dimensionality reduction and density estimation for machine learning problems.
C330-1.5	Evaluate and test the significance of machine learning results statistically.

Note: Attempt all questions in order.

Q1. Solve the equation 6x+15y+55z=76, 15x+55y+225z=295, 55x+225y+979z=1259 using Cholesky decomposition method. **[C330-1.1, 5 Marks]**

Q2. By using NLP, I can detect spam email in my inbox. Assume that the word 'offer' occurs in 80% of the spam message in my account. Also, let's assume 'offer' occurs in 10% of my desired email. If 30% of the received emails are considered as spam, and I will receive a new message which contains 'offer' then what is the probability that it is a spam? (Use Bayes theorem only) **[C330-1.2, 5 Marks]**

Q3. Consider the multivariate function $f(x,y) = e^x \log(1+y)$:

a) Expand f(x,y) by using taylor series around the point (0, 1) in terms of second degree.

b) Use the Taylor expression to estimate f(0.1,1.2) **[C330-1.3, 5 Marks]**

Q4. Find singular value decomposition for $A = \begin{matrix} 1 & 0 & 1 \\ -1 & 1 & 0 \end{matrix}$ **[C330-1.4, 5 Marks]**

P.T.O.

Q5. Consider the two-dimensional patter (2,1), (3,5), (4,3), (5, 6), (6,7), and (7,8).

a) Compute the principal components using PCA algorithm.

b) Compute the projected points on PCA space. **[C330-1.4, 5 Marks]**

Q6. Suppose a researcher is investigating the average IQ scores of a population. The historical mean I score for the general population is known to be 100 with a standard deviation of 15. The researcher collec a sample of 36 individuals and find that the sample mean IQ scores is 105. Conduct hypothesis test at a % significance level using the z -test.

a) State the null hypothesis and alternate hypothesis for z-test.

b) Calculate the test statistics for both z-test.

c) Determine the critical value for both z-test.

d) Make a decision about the null hypothesis.

e) Provide a conclusion in the context of the problem. **[C330-1.5, 5 Marks]**

Q7. A professor wants to know if her introductory statistics class has a good grasp of basic math. students are chosen at random from the class and given a math proficiency test. The Professor wants cl to be able to score above 70 in the test. The six students get scores of 62, 92, 75, 68, 83 and 95. Can professor have 90% confidence that the math score for the class on the test would be above 70?

[C330-1.5, 5 Marks]

****