MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY, WEST BENGAL

Paper Code : PEC-CS701E   Machine learning

Time Allotted : 3 Hours

Full Marks :70

The Figures in the margin indicate full marks.
Candidate are required to give their answers in their own words as far as practicable

Group-A (Very Short Answer Type Question)

1. Answer any ten of the following :

[1 x 10 = 10]

(I) _____ is a classification algorithm used to assign observations to a discrete set of classes.

(II) The number of nodes in the input layer is 10 and the hidden layer is 5. The maximum number of connections from the input layer to the hidden layer are

(III) True or False: Hierarchical clustering is slower than non-hierarchical clustering?

(IV) True or False: Ensemble learning can only be applied to supervised learning methods.

(V) A collection of individual models that learn to predict a target by combining their strengths and avoiding the weaknesses of each is called _____

(VI) Semi-supervised learning algorithm deals with which types of data

(VII) In an election, N candidates are competing against each other and people are voting for either of the candidates. Voters don't communicate with each other while casting their votes. Which of the following ensemble method works similar to above-discussed election procedure?

(VIII) A feature F1 can take certain value: A, B, C, D, E, & F and represents grade of students from a college. Feature F1 is an example of _dependent_ variable.

(IX) Imagine a Newly-Born starts to learn walking. It will try to find a suitable policy to learn walking after repeated falling and getting up. Specify what type of machine learning is best suited?

(X) The selling price of a house depends on many factors. For example, it depends on the number of bedrooms, number of kitchen, number of bathrooms, the year the house was built, and the square footage of the lot. Given these factors, predicting the selling price of the house is an example of which type of linear regression. https://www.makaut.com

(XI) Targeted marketing, Recommended Systems, and Customer Segmentation are applications in which algorithm?

(XII) The _____ is the difference between a sample statistic used to estimate a population parameter and the actual but unknown value of the parameter.

Group-B (Short Answer Type Question)
Answer any three of the following

[5 x 3 = 15]

2. Explain Matrix Factorization and where it is uesd. [5]

3. Why ensemble learning is used? What is the general principle of an ensemble method and what is bagging and boosting in ensemble method? [5]

4. Explain the Difference Between Classification and Regression? [5]

5. Compare K-means and KNN Algorithms. [5]

6. How do we decide the value of "K" in KNN algorithm? Why is the odd value of "K" preferable in KNN algorithm? [5]

Group-C (Long Answer Type Question)
Answer any three of the following

[15 x 3 = 45]

7. (a) Discuss the different types of Machine Learning? [6]

   (b) What are parametric and non-parametric model? [5]

   (c) How is machine learning related to AI? [4]

8. (a) Explain Generative Mixture model [4]

   (b) With a proper diagram explain the steps of a genarative mixture model [5]

   (c) Write down the steps of PCA (Principal Component Analysis) [6]

9. (a) Explain the Confusion Matrix with Respect to Machine Learning Algorithms with an suitable example [5]

(b) Calculate the accuracy percentage for the given Confusion Matrix.

[5]

Actual

	Yes	No
Yes	12	3
No	1	9

Predicted

Confusion Matrix

(c) Explain True Positive, True Negative, False Positive, and False Negative in Confusion Matrix with an example. [5]

10. (a) Explain the three techniques under supervised feature Selection [6]

(b) Explain the benefits of using feature selection in machine learning [4]

(c) Explain the curse of dimensionality [5]

11. (a) What is Artificial Intelligence and why do we need it? [5]

(b) What is Deep Learning, and give some of its example that is used in real-world? [4]

(c) Differentiate between Artificial intelligence, Machine Learning, and Deep Learning [6]

*** END OF PAPER ***

Printed pages: 2   **Sub Code: MTCS031**

Paper Id | 210201 | **Roll No.** | | | | | | | | | |

M. TECH.

(SEM-II) THEORY EXAMINATION 2017-18
MACHINE LEARNING

Time: 3 Hours   *Total Marks: 70*

Note: Attempt all Sections. If require any missing data; then choose suitably.

SECTION A

1. **Attempt *all* questions in brief.**   **2 x 7 = 14**

    a. What are the applications of machine learning?
    b. What is Machine learning? When to use it?

    c. List the differences between supervised and unsupervised learning?
    d. Give the example of classification problem.
    e. What is linear regression?
    f. What are the Matrix multiplication properties?
    g. What is meant by decision boundary? What is non linear decision boundary?

SECTION B

2. **Attempt any *three* of the following:**   **7 x 3 = 21**

    a. What are neural network? Explain its working with the help of an example.
    b. Explain dimensionality reduction as an unsupervised learning. Explain PCA.
    c. Write short note on K-means algorithm.
    d. Explain the types of learning along with their characteristics.
    e. What is data compression? Explain data compression techniques

SECTION C

3. **Attempt any *one* part of the following:**   **7 x 1 = 7**

    (a) Explain PCA algorithm. What are the applications of PCA
    (b) How a high accuracy learning system can be designed? Explain.

4. **Attempt any *one* part of the following:**   **7 x 1 = 7**

    (a) Explain how error analysis in Machine learning takes place.
    (b) Write a short note on anomaly detection in machine learning? What are its application areas?

5. **Attempt any *one* part of the following:**   **7 x 1 = 7**

    (a) Explain collaborative filtering algorithm in detail.
    (b) Explain how to deal with big datasets? Also state high bias problem.

6. **Attempt any *one* part of the following:**   **7 x 1 = 7**

    (a) What is online learning? Explain how machine learning is related to online

learning? What are its applications?

(b)   Write short note on Artificial data synthesis.

7.   Attempt any *one* part of the following: **7 x 1 = 7**

(a)   Explain how a ROC curve works. Define precision and recall

(b)   Why is "Naive" Bayes naive? Explain the difference between L1 and L2 regularization.

MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY, WEST BENGAL

Paper Code : MCAN-E305G Machine Learning

UPID : 003894

McA

Time Allotted : 3 Hours   *Full Marks :70*

The Figures in the margin indicate full marks.
Candidate are required to give their answers in their own words as far as practicable

Group-A (Very Short Answer Type Question)

1. Answer *any ten* of the following :   [1 x 10 = 10]

   (I) How do you handle missing or corrupted data in a dataset?

   (II) "Gradient descent algorithm is always better fit than a least square method". State if the statement is true or false.

   (III) Type of matrix decomposition model is_____

   (IV) Who was the inventor of the first neurocomputer?

   (V) "Regression, classification, subgroup discovery all are the types of supervised learning." State if the statement is true or false.

   (VI) A feature F1 can take certain value: A, B, C, D, E, & F and represents grade of students from a college. Here feature type is_____

   (VII) The problem of finding hidden structure in unlabeled data is called_____ learning.

   (VIII) "A Support Vector Machine (SVM) is a discriminative classifier defined by a separating hyperplane." State if the statement is true or false.

   (IX) How many types of Artificial Neural Networks?

   (X) "Euclidean distance is powerful distance metrics used by Geometric model." State if the statement is true or false.

   (XI) How can you prevent a clustering algorithm from getting stuck in bad local optima?

   (XII) "The computational complexity challenge related to learning half-space in high dimensional feature spaces can be solved using the method of kernels." . State if the statement is true or not.

Group-B (Short Answer Type Question)

Answer *any three* of the following :   [5 x 3 = 15]

2. Discuss how are Bayes nets and Markov nets used for representing dependencies?   [5]

3. Explain the concept of Bagging with its uses.   [5]

4. What is adaboost algorithm?   [5]

5. Discuss how random forest algorithm works to achieve its goal.   [5]

6. Explain in details how Expectation Maximization (EM) is achieved for soft clustering.   [5]

Group-C (Long Answer Type Question)

Answer *any three* of the following :   [15 x 3 = 45]

7. Explain about dimension reduction and its techniques.   [15]

8. What is the difference between supervised and unsupervised machine learning?   [15]

9. Elaborate on the concept of boosting in ensemble, with neat relevant example.   [15]

10. Explain SVM classifier with suitable example.   [15]

11. Explain the linear regression algorithm in details with relevant example.   [15]

*** END OF PAPER ***

Printed Pages : 1   **Roll No.** ☐☐☐☐☐☐☐☐☐☐   **MTCS-031**

M. TECH.

THEORY EXAMINATION (SEM–II) 2016-17

MACHINE LEARNING

Time : 3 Hours   *Max. Marks : 70*

Note : *Be precise in your answer. In case of numerical problem assume data wherever not provided.*

SECTION- A

1.   **Attempt all parts of this section:**   7×2=14
   (a)   Explain the various types of issues in machine earning.
   (b)   Describe the Artificial Neural Networks (ANN).
   (c)   Define the learning classifiers.
   (d)   Differentiated between Bayesian Learning and Instance based Learning.
   (e)   Discuss the complexity of in finite hypothesis spaces.
   (f)   Write the at least five applications of machine learning.
   (g)   Discuss the regression model.

SECTION- B

2.   **Attempt any three parts of the following:**   3×7=21
   (a)   What is a core point in DBSCAN? What role do core points play in forming clusters?
   (b)   What objective function do regression trees minimize?
   (c)   What is the task of the E-step of the EM-algorithm? Give a verbal description (and not (just) formulas) how EM accomplishes the task of the E-step.
   (d)   EM uses a mixture of *k* Gaussian for clustering; what purpose does the k Gaussian serve?
   (e)   Describe brain-in-a box model. Compare with it a recurrent network.

SECTION- C

3.   **Attempt all questions in this section:**   5×7=35
   (a)   What is learning? Write any four learning techniques and in each case give the expression for weight- updating.
                                         OR
         Discuss various Artificial Neural Network Architectures.
   (b)   Explain back propagation algorithm and derive expressions for weight update relations.
                                         OR
         With help of a suitable diagram discuss functioning of a simple artificial neuron. Explain how the functionality affected if two such neuron are connected in series.
   (c)   Define gradient. Using steepest descent rule to the following function
$$f(x) = x_1^2 + 5x_1x_2 + 10x_2^2$$
         Determine first three points of trajectory starting from
$$x_0 = [0.5, 0.5]$$
                                         OR
          Write an algorithm to implement simulated annealing.
   (d)   Explain Perceptron training algorithm for linear classification. And explain its equation using homogeneous coordinates.
                                         OR
         Explain geometric models and probabilistic models of machine learning with suitable examples.
   (e)   What are neighbors? Why is it necessary to use nearest neighbor while classifying justify the answer with suitable example.
                                         OR
         Explain how genetic algorithms are influenced by knowledge based techniques. Also discuss the how Genetic Algorithm is different from traditional algorithms?

Questions Bank

Subject Name: Machine Learning
Subject Code: 15CS73
Sem: VII

Module -1 Questions.

1. De4fine the following terms:
   a. Learning
   b. LMS weight update rule
   c. Version Space
   d. Consistent Hypothesis
   e. General Boundary
   f. Specific Boundary
   g. Concept

2. What are the important objectives of machine learning?

3. Explain find –S algorithm with given example. Give its application.

Table 1

Example	Sky	AirTemp	Humidity	Wind	Water	Forecast	EnjoySport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

4. What do you mean by a well –posed learning problem? Explain the important features that are required to well –define a learning problem.

5. Explain the inductive biased hypothesis space and unbiased learner

6. What are the basic design issues and approaches to machine learning?

7. How is Candidate Elimination algorithm different from Find-S Algorithm

8. How do you design a checkers learning problem

9. Explain the various stages involved in designing a learning system

10. Trace the Candidate Elimination Algorithm for the hypothesis space H' given the sequence of training examples from Table 1.

H'= < ?, Cold, High, ?,?,?>v<Sunny, ?, High, ?,?,Same>

11. Differentiate between Training data and Testing Data

12. Differentiate between Supervised, Unsupervised and Reinforcement Learning

13. What are the issues in Machine Learning

14. Explain the List Then Eliminate Algorithm with an example

15. What is the difference between Find-S and Candidate Elimination Algorithm

16. Explain the concept of Inductive Bias

17. With a neat diagram, explain how you can model inductive systems by equivalent deductive systems

18. What do you mean by Concept Learning?

Module -2 Questions.

1. Give decision trees to represent the following boolean functions:
   (a) $A \wedge \sim B$
   (b) $A \vee [B \wedge C]$
   (c) A XOR B
   (d) $[A \wedge B] \vee [C \wedge D]$

2. Consider the following set of training examples:

Instance	Classification	a1	a2
1	+	T	T
2	+	T	T
3	-	T	F
4	+	F	F
5	-	F	T
6	-	F	T

   (a) What is the entropy of this collection of training examples with respect to the target function classification?

   (b) What is the information gain of a2 relative to these training examples?

3. NASA wants to be able to discriminate between Martians (M) and Humans (H) based on the following characteristics: Green $\in\{N, Y\}$, Legs $\in\{2,3\}$, Height $\in\{S, T\}$, Smelly $\in\{N, Y\}$

   Our available training data is as follows:

	Species	Green	Legs	Height	Smelly
1	M	N	3	S	Y
2	M	Y	2	T	N
3	M	Y	3	T	N
4	M	N	2	S	Y
5	M	Y	3	T	N
6	H	N	2	T	Y
7	H	N	2	S	N
8	H	N	2	T	N
9	H	Y	2	S	N
10	H	N	2	T	Y

a) Greedily learn a decision tree using the ID3 algorithm and draw the tree.

b) (i) Write the learned concept for Martian as a set of conjunctive rules (e.g., if

(green=Y and legs=2 and height=T and smelly=N), then Martian; else if ... then Martian;...; else Human).

(ii) The solution of part b)i) above uses up to 4 attributes in each conjunction. Find a set of conjunctive rules using only 2 attributes per conjunction that still results in zero error in the training set. Can this simpler hypothesis be represented by a decision tree of depth 2? Justify.

4. Discuss Entropy in ID3 algorithm with an example

5. Compare Entropy and Information Gain in ID3 with an example.

6. Describe hypothesis Space search in ID3 and contrast it with Candidate-Elimination algorithm.

7. Relate Inductive bias with respect to Decision tree learning.

8. Illustrate **Occam's razor and** relate the importance of **Occam's razor** with respect to ID3 algorithm.

9. List the issues in Decision Tree Learning. Interpret the algorithm with respect to **Overfitting the data.**

10. Discuss the effect of **reduced Error pruning** in decision tree algorithm.

11. What type of problems are best suited for decision tree learning

12. Write the steps of ID3Algorithm

13. What are the capabilities and limitations of ID3

14. Define (a) Preference Bias   (b) Restriction Bias

15. Explain the various issues in Decision tree Learning

16. Describe Reduced Error Pruning

17. What are the alternative measures for selecting attributes

18. What is Rule Post Pruning

Module -3 Questions.

1) What is Artificial Neural Network?

2) What are the type of problems in which Artificial Neural Network can be applied.

3) Explain the concept of a Perceptron with a neat diagram.

4) Discuss the Perceptron training rule.

5) Under what conditions the perceptron rule fails and it becomes necessary to apply the delta rule

6) What do you mean by Gradient Descent?

7) Derive the Gradient Descent Rule.

8) What are the conditions in which Gradient Descent is applied.

9) What are the difficulties in applying Gradient Descent.

10) Differentiate between Gradient Descent and Stochastic Gradient Descent

11) Define Delta Rule.

12) Derive the Backpropagation rule considering the training rule for Output Unit weights and Training Rule for Hidden Unit weights

13) Write the algorithm for Back propagation.

14) Explain how to learn Multilayer Networks using Gradient Descent Algorithm.

15) What is Squashing Function?

Module -4 Questions.

1) Explain the concept of Bayes theorem with an example.

2) Explain Bayesian belief network and conditional independence with example.

3) What are Bayesian Belief nets? Where are they used?

4) Explain Brute force MAP hypothesis learner? What is minimum description length principle

5) Explain the k-Means Algorithm with an example.

6) How do you classify text using Bayes Theorem

7) Define (i) Prior Probability   (ii) Conditional Probability (iii) Posterior Probability

8) Explain Brute force Bayes Concept Learning

9) Explain the concept of EM Algorithm.

10) What is conditional Independence?

11) Explain Naïve Bayes Classifier with an Example.

12) Describe the concept of MDL.

13) Who are Consistent Learners.

14) Discuss Maximum Likelihood and Least Square Error Hypothesis.

15) Describe Maximum Likelihood Hypothesis for predicting probabilities.

16) Explain the Gradient Search to Maximize Likelihood in a neural Net.

Module -5 Questions.

1. What is Reinforcement Learning?

2. Explain the Q function and Q Learning Algorithm.

3. Describe K-nearest Neighbour learning Algorithm for continues valued target function.

4. Discuss the major drawbacks of K-nearest Neighbour learning Algorithm and how it can be corrected

5. Define the following terms with respect to K - Nearest Neighbour Learning :

   i) Regression   ii) Residual   iii) Kernel Function.

6. Explain Q learning algorithm assuming deterministic rewards andactions?

7. Explain the K – nearest neighbour algorithm for approximating a discrete – valued

functionf : Hn→ V with pseudo code

8. Explain Locally Weighted Linear Regression.

9. Explain CADET System using Case based reasoning.

10. Explain the two key difficulties that arise while estimating the Accuracy of Hypothesis.

11. Define the following terms

 a. Sample error   b. True error   c. Random Variable

 d. Expected value   e. Variance   f. standard Deviation

12. Explain Binomial Distribution with an example.

13. Explain Normal or Gaussian distribution with an example.

14.Explain the Central Limit Theorem with an example.

15. Write the Procedure for estimating the difference in error between two learning methods.

   Approximate confidence intervals for this estimate

MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY, WEST BENGAL

Paper Code : MCAN-E305G Machine Learning
UPID : 003894

Time Allotted : 3 Hours   *Full Marks :70*

The Figures in the margin indicate full marks.
Candidate are required to give their answers in their own words as far as practicable

Group-A (Very Short Answer Type Question)

1. Answer *any ten* of the following :   [1 x 10 = 10]

   (I) ANN is made up of ___ neurons.

   (II) SVM is an example of_____

   (III) Three companies A, B and C supply 25%, 35%,and 40% of the notebooks to a school. Past experience shows that 5%, 4% and 2% of the notebooks produced by these companies are defective. If a notebook is found to be defective, what is the probability that the notebook is supplied by A?

   (IV) K-means clustering algorithm is an example of which type of clustering method?

   (V) A computer program is said to learn from _____ E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with E.

   (VI) The probabilistic model used in machine learning is closely related to which field?

   (VII) The slope which always moves upward on a graph from left to right is called_____

   (VIII) _____ is a measure of the randomness in the information being processed in the Decision Tree.

   (IX) Bootstrap aggregation is known for_____.

   (X) How can learning process be stopped in backpropagation rule?

   (XI) A generalization of the concept of a plane which is used in SVM implementation for object separation is known as_____

   (XII) Suppose box A contains 4 red and 5 blue coins and box B contains 6 red and 3 blue coins. A coin is chosen at random from the box A and placed in box B. Finally, a coin is chosen at random from among those now in box B. What is the probability a blue coin was transferred from box A to box B given that the coin chosen from box B is red?

Group-B (Short Answer Type Question)

Answer *any three* of the following :   [5 x 3 = 15]

2. What does the conditional probability mean? What is its formula ?   [5]

3. What is classification? Explain the key differences between classification and regression.   [5]

4. What are the strengths and weaknesses of the decision tree method?   [5]

5. What are the similarities and differences among Bagging, Boosting, and Stacking?   [5]

6. Define deep learning. How are deep learning, AI and machine learning inter-related?   [5]

Group-C (Long Answer Type Question)

Answer *any three* of the following :   [15 x 3 = 45]

7. a) What is machine learning?   [3+6+6]
    b) Explain the different types of machine learning techniques with examples.
    c) Write some of the applications of machine learning algorithms.

8. There are two random variables X and Y with joint PMF given in Table below: [3+4+4+4]
    a) Find $P(X \leq 2, Y \leq 4)$.
    b) Find the marginal PMFs of X and Y.
    c) Find $P(Y = 2 | X = 1)$.
    d) Are X and Y independent?

L	Y=2	Y=4	Y=5
X=1	1/12	1/24	1/24
X=2	1/6	1/12	1/8
X=3	1/4	1/8	1/12

9. The weight and systolic blood pressure of 6 randomly selected males in the age group 45-50 are shown   [8+7]
below. Assume that weight and blood pressure are jointly normally distributed.
a) Find a regression line relating systolic blood pressure to weight.
b) Estimate the correlation coefficient.

Weight	165	167	180	155	212	175
Blood Pressure	130	133	150	128	151	146

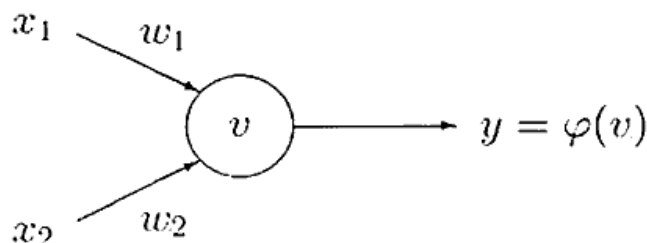10. a) Explain the different steps used in Gradient Descent Algorithm.   [4+3+4+4]
b) Explain the different types of Gradient Descent in detail.
c) Logical operators (i.e. NOT, AND, OR, XOR, etc) are the building blocks of any computational device. Logical functions return only two possible values, true or false, based on the truth or false values of their arguments. For example, operator AND returns true only when all its arguments are true, otherwise (if any of the arguments is false) it returns false. If we denote truth by 1 and false by 0, then logical function AND can be represented by the following table:

$$\begin{array}{c|cccc} x_1 : & 0 & 1 & 0 & 1 \\ x_2 : & 0 & 0 & 1 & 1 \\ \hline x_1 \text{ AND } x_2 : & 0 & 0 & 0 & 1 \end{array}$$

This function can be implemented by a single–unit with two inputs:

if the weights are w1 = 1 and w2 = 1 and the activation function is:

$$\varphi(v) = \begin{cases} 1 & \text{if } v \geq 2 \\ 0 & \text{otherwise} \end{cases}$$

Note that the threshold level is 2 ($v \geq 2$).
Test how the neural AND function works.
d)   Suggest how to change either the weights or the threshold level of this single–unit in order to implement the logical OR function (true when at least one of the arguments is true):

$$\begin{array}{c|cccc} x_1 : & 0 & 1 & 0 & 1 \\ x_2 : & 0 & 0 & 1 & 1 \\ \hline x_1 \text{ OR } x_2 : & 0 & 1 & 1 & 1 \end{array}$$

11. a) Explain Naïve Bayes classifier with an example and its use in practical life.   [6+3+6]
b) Write down the Naïve Bayes classifier algorithm .
c) Suppose box A contains 4 red and 5 blue coins and box B contains 6 red and 3 blue coins. A coin is chosen at random from the box A and placed in box B. Finally, a coin is chosen at random from among those now in box B. What is the probability a blue coin was transferred from box A to box B given that the coin chosen from box B is red?

*** END OF PAPER ***

MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY, WEST BENGAL

Paper Code : MCAN-E305G Machine Learning
UPID : 003894

Time Allotted : 3 Hours   *Full Marks :70*

The Figures in the margin indicate full marks.
Candidate are required to give their answers in their own words as far as practicable

Group-A (Very Short Answer Type Question)

1. Answer *any ten* of the following :   [1 x 10 = 10]

   (I) ANN is made up of ___ neurons.

   (II) SVM is an example of_____

   (III) Three companies A, B and C supply 25%, 35%,and 40% of the notebooks to a school. Past experience shows that 5%, 4% and 2% of the notebooks produced by these companies are defective. If a notebook is found to be defective, what is the probability that the notebook is supplied by A?

   (IV) K-means clustering algorithm is an example of which type of clustering method?

   (V) A computer program is said to learn from _____ E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with E.

   (VI) The probabilistic model used in machine learning is closely related to which field?

   (VII) The slope which always moves upward on a graph from left to right is called_____

   (VIII) _____ is a measure of the randomness in the information being processed in the Decision Tree.

   (IX) Bootstrap aggregation is known for_____.

   (X) How can learning process be stopped in backpropagation rule?

   (XI) A generalization of the concept of a plane which is used in SVM implementation for object separation is known as_____

   (XII) Suppose box A contains 4 red and 5 blue coins and box B contains 6 red and 3 blue coins. A coin is chosen at random from the box A and placed in box B. Finally, a coin is chosen at random from among those now in box B. What is the probability a blue coin was transferred from box A to box B given that the coin chosen from box B is red?

Group-B (Short Answer Type Question)

Answer *any three* of the following :   [5 x 3 = 15]

2. What does the conditional probability mean? What is its formula ?   [5]

3. What is classification? Explain the key differences between classification and regression.   [5]

4. What are the strengths and weaknesses of the decision tree method?   [5]

5. What are the similarities and differences among Bagging, Boosting, and Stacking?   [5]

6. Define deep learning. How are deep learning, AI and machine learning inter-related?   [5]

Group-C (Long Answer Type Question)

Answer *any three* of the following :   [15 x 3 = 45]

7. a) What is machine learning?   [3+6+6]
   b) Explain the different types of machine learning techniques with examples.
   c) Write some of the applications of machine learning algorithms.

8. There are two random variables X and Y with joint PMF given in Table below: [3+4+4+4]
   a) Find $P(X \leq 2, Y \leq 4)$.
   b) Find the marginal PMFs of X and Y.
   c) Find $P(Y = 2 | X = 1)$.
   d) Are X and Y independent?

L	Y=2	Y=4	Y=5
X=1	1/12	1/24	1/24
X=2	1/6	1/12	1/8
X=3	1/4	1/8	1/12

9. The weight and systolic blood pressure of 6 randomly selected males in the age group 45-50 are shown below. Assume that weight and blood pressure are jointly normally distributed. [8+7]
   a) Find a regression line relating systolic blood pressure to weight.
   b) Estimate the correlation coefficient.

Weight	165	167	180	155	212	175
Blood Pressure	130	133	150	128	151	146

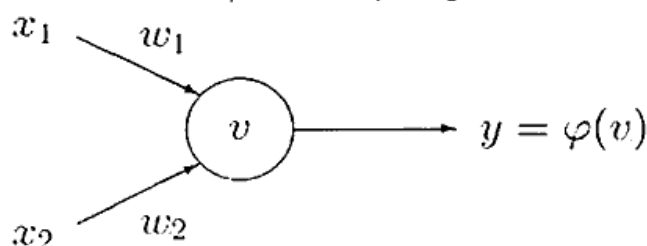10. a) Explain the different steps used in Gradient Descent Algorithm. [4+3+4+4]
    b) Explain the different types of Gradient Descent in detail.
    c) Logical operators (i.e. NOT, AND, OR, XOR, etc) are the building blocks of any computational device. Logical functions return only two possible values, true or false, based on the truth or false values of their arguments. For example, operator AND returns true only when all its arguments are true, otherwise (if any of the arguments is false) it returns false. If we denote truth by 1 and false by 0, then logical function AND can be represented by the following table:

$$
\begin{array}{c|cccc}
x_1: & 0 & 1 & 0 & 1 \\
x_2: & 0 & 0 & 1 & 1 \\
\hline
x_1 \text{ AND } x_2: & 0 & 0 & 0 & 1
\end{array}
$$

This function can be implemented by a single–unit with two inputs:

if the weights are w1 = 1 and w2 = 1 and the activation function is:

$$
\varphi(v) = \begin{cases} 1 & \text{if } v \geq 2 \\ 0 & \text{otherwise} \end{cases}
$$

Note that the threshold level is 2 ($v \geq 2$).
Test how the neural AND function works.

    d)   Suggest how to change either the weights or the threshold level of this single–unit in order to implement the logical OR function (true when at least one of the arguments is true):

$$
\begin{array}{c|cccc}
x_1: & 0 & 1 & 0 & 1 \\
x_2: & 0 & 0 & 1 & 1 \\
\hline
x_1 \text{ OR } x_2: & 0 & 1 & 1 & 1
\end{array}
$$

11. a) Explain Naïve Bayes classifier with an example and its use in practical life. [6+3+6]
    b) Write down the Naïve Bayes classifier algorithm .
    c) Suppose box A contains 4 red and 5 blue coins and box B contains 6 red and 3 blue coins. A coin is chosen at random from the box A and placed in box B. Finally, a coin is chosen at random from among those now in box B. What is the probability a blue coin was transferred from box A to box B given that the coin chosen from box B is red?

*** END OF PAPER ***