

AI & ML Algorithm (315330)

MSBTE Diploma · K-Scheme · Model Question Paper — Set A

Time: 3 Hours
Max. Marks: 70
Instructions:

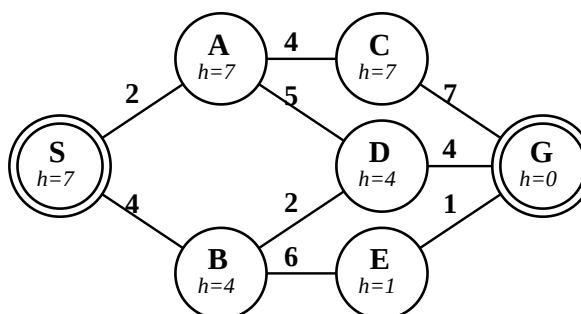
1. All questions are compulsory.
2. Answer each next main question on a new page.
3. Illustrate your answers with neat sketches wherever necessary.
4. Figures to the right indicate full marks.
5. Assume suitable data, if necessary.
6. Use of Non-programmable Electronic Pocket Calculator is permissible.
7. Mobile phone, pager and any other electronic communication devices are not permissible in the examination hall.

Q.1 Attempt any FIVE of the following:
10

- a) List any four applications of Artificial Intelligence.
- b) Define reasoning. Name any two types of reasoning.
- c) Differentiate between Data Analytics and Data Science (any two points).
- d) Draw a neat labelled diagram of the machine learning life cycle.
- e) Define unsupervised learning. Give one example.
- f) Define Precision and Recall.
- g) State the need of search algorithms in Artificial Intelligence.

Q.2 Attempt any THREE of the following:
12

- a) Explain A* algorithm. Apply it on the following graph to find the path from S to goal G. Edge costs: S–A = 2, S–B = 4, A–C = 4, A–D = 5, B–D = 2, B–E = 6, C–G = 7, D–G = 4, E–G = 1; heuristic values $h(n)$: S = 7, A = 7, B = 4, C = 7, D = 4, E = 1, G = 0.



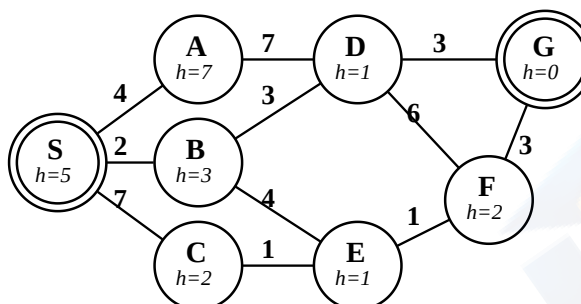
- b) State Bayes' theorem. In a city 2% of people have a disease. A test detects the disease correctly in 95% of sick people, but also gives a positive result for 8% of healthy people. If a person tests positive, find the probability that the person actually has the disease.

- c) Differentiate between training dataset and testing dataset (minimum four points).
- d) Explain k-fold cross-validation with a neat diagram. What are positive and negative classes?

Q.3 Attempt any THREE of the following:

12

- a) Explain Greedy Best-First search. Trace it on the following graph from S to G and comment on the path found. Edge costs: $S-A = 4$, $S-B = 2$, $S-C = 7$, $A-D = 7$, $B-D = 3$, $B-E = 4$, $C-E = 1$, $D-F = 6$, $E-F = 1$, $D-G = 3$, $F-G = 3$; $h(n)$: $S = 5$, $A = 7$, $B = 3$, $C = 2$, $D = 1$, $E = 1$, $F = 2$, $G = 0$.



- b) Describe the architecture of a knowledge-based agent with a neat diagram.
- c) Explain the different forms of data: data mining, data analytics and statistical data.
- d) A regression model gave the following predictions. Calculate MAE, MSE and RMSE.

Actual y	12	15	20	22	25	30
Predicted $\hat{y}$	10	16	18	23	27	28

Q.4 Attempt any THREE of the following:

12

- a) Explain agglomerative hierarchical clustering. Cluster the points A(1, 1), B(2, 1), C(4, 5), D(6, 5), E(7, 7) using single linkage (Euclidean distance) and draw the dendrogram.
- b) Compare forward chaining and backward chaining with an example.
- c) Explain overfitting and underfitting with a neat diagram.
- d) Apply K-means clustering ($k = 2$) on the points P1(2, 3), P2(3, 3), P3(6, 8), P4(7, 7), P5(2, 2), P6(8, 8), P7(3, 4). Take P1 and P3 as initial centroids. Show the iterations until the clusters do not change.
- e) Explain binary and multiclass classification with examples.

Q.5 Attempt any TWO of the following:

12

- a) Explain the intelligent agent in AI with respect to (i) structure of an agent and (ii) Turing test.

- b) A spam filter was tested on 100 e-mails. The results are: TP = 45, FN = 5, FP = 10, TN = 40. Draw the confusion matrix and calculate accuracy, precision, recall and F1-score. State the importance of the confusion matrix.
- c) Write a Python program using scikit-learn to implement multiple linear regression to predict the price of a house from its area and number of bedrooms. Display the coefficients, intercept, R^2 score and the predicted price for a new house.

Q.6 Attempt any TWO of the following:

12

- a) Differentiate between supervised and unsupervised learning (minimum six points).
- b) Using the ID3 algorithm (entropy and information gain), construct a decision tree for the following data. Target attribute: Play.

No.	Weather	Temperature	Wind	Play
1	Sunny	Hot	Weak	No
2	Sunny	Hot	Strong	No
3	Cloudy	Hot	Weak	Yes
4	Rainy	Mild	Weak	Yes
5	Rainy	Mild	Strong	No
6	Cloudy	Mild	Strong	Yes
7	Sunny	Mild	Weak	Yes
8	Rainy	Hot	Weak	Yes
9	Sunny	Mild	Strong	Yes
10	Cloudy	Hot	Strong	Yes

- c) Explain data cleaning in machine learning. Write a Python (pandas) program to handle missing values and detect outliers in a dataset.

Note: This is a **model question paper prepared by Ananda Technologies** only for practice and exam preparation, based on the pattern of previous MSBTE question papers and the K-scheme syllabus. It is **not an official MSBTE paper**, and it does **not** mean that these questions will come in the exam. Answers follow standard textbook points; marks shown are indicative. Free for students — anandatechnologies.in