Question Paper

Exam Date & Time: 29-Nov-2023 (02:00 PM - 05:00 PM)

MANIPAL ACADEMY OF HIGHER EDUCATION

Manipal School of Information Sciences (MSIS), Manipal
First Semester Master of Engineering - ME (Big Data Analytics) Degree Examination - November / December 2023

Fundamentals of Machine Learning [BDA 5103]

Marks: 100 **Duration: 180 mins.**

Wednesday, November 29, 2023

Answer all the questions.

1) Interpret any four applications for which machine learning approaches seem to be appropriate. (10 Marks) (CO1-PO3-BL3) (10)

2) Apply the candidate-elimination algorithm to obtain most general and most specific hypotheses for the training examples given in the following table: (10 Marks) (CO1-PO3-BL3) (10)

Origin	Manufacturer	Color	Decade	Type	Example Type
Japan	Honda	Blue	1980	Economy	Positive
Japan	Toyota	Green	1970	Sports	Negative
Japan	Toyota	Blue	1990	Economy	Positive
USA	Chrysler	Red	1980	Economy	Negative
Japan	Honda	White	1980	Economy	Positive
Japan	Toyota	Green	1980	Economy	Positive
Japan	Honda	Red	1990	Economy	Negative

3) Analyze the reasons for overfitting in the decision tree. Examine any two methods to address the overfitting issues in the decision tree. (10 Marks) (CO2-PO4-BL4) (10)

4) Examine the methods to address the following issues in decision tree learning: (CO2-PO4-BL4) (10)

a). Choosing an appropriate attribute selection measure (3 Marks)
b). Handling training data with missing attribute values (3 Marks)

c). Handling attributes with differing costs, and improving computational efficiency (4 Marks)

5) Evaluate the true error of a hypothesis h with respect to the target concept with a suitable diagram. Describe the exhaustion of the version (10)

space with a suitable diagram. (10 Marks) (CO3-PO4-BL5)

6) Judge the trade-off between bias and variance with a suitable example. Describe possible methods to minimize bias and variance. (10 Marks) (CO3-PO4-BL5) (10)

7) The dataset given in the table that has two input variables (X1 and X2) and one output variable (Y). Apply logistic regression model to make predictions with maximum two epochs. use learning rate 1. (10 Marks) (CO4-PO3-BL3) (10)

Input X1	Input X2	Actual output (Y)	Predicted output (Y')
2.7810836	2.550537003	0	?

8) Interpret the functions of case-based reasoning approach. Explain the R4 model of the case-based reasoning (CBR) approach with a suitable example. (10 Marks) (CO4-PO3-BL3) (10)

9) Examine the Eigen values and Eigen vectors from the following data set: (10 Marks) (CO5-PO5-BL4) (10)

x	2.5	0.5	2.2	1.9	3.1	2.3	2	1	1.5	1.1
y	2.4	0.7	2.9	2.2	3	2.7	1.6	1.1	1.6	0.9

10) Interpret the boosting and bagging ensemble models with a suitable example (5 Marks) (CO5-PO5-BL4) (5)

a.)

b.) Interpret any two-distance metrics used in k-NN to find the distance between two points? Give an example. (5 Marks) (CO4-PO3-BL3) (5)

-----End-----