

MANIPAL UNIVERSITY

FIRST SEMESTER MASTER OF ENGINEERING- ME (BIG DATA AND DATA ANALYTICS) DEGREE EXAMINATION- NOVEMBER 2017 DATE : Monday, November 20, 2017 TIME : 10:00AM - 1:00PM

Fundamentals of Machine Learning [BDA 611]

Duration: 180 mins.

Marks: 100

A

Answer all the questions.

- Describe the method of choosing a Representation for the ⁽⁶⁾ Target Function in designing a machine learning system (6 MARKS)
 - ^{B)} Describe the method of choosing a Function Approximation ⁽⁴⁾ Algorithm for a machine learning system.

(4 MARKS)

²⁾ Obtain the most general and most specific hypotheses using ⁽¹⁰⁾ Candidate-Elimination algorithm for the training data contains errors as given in the following table.

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

3)

4)

- Write the futility of Bias-Free Learning and usefulness of ⁽¹⁰⁾ Inductive learning. (4+6 Marks)
- Obtain the decision tree for the following training examples ⁽¹⁰⁾ using ID3 algorithm.

Instance	Classification	a_1	<i>a</i> ₂
1	+	т	т
2	+	т	Т
3	-	Т	F
4	+ .	F	F
5	-	F	т
6	-	F	Т

5)	Solve for number of training examples " m " sufficient for successful learning of any target concept in the <i>finite hypotheses</i> space?	(10)
6)	Derive the <i>Sample Complexity</i> for <i>Infinite Hypothesis</i> Spaces.	(10)
7)	Design a Brute-Force MAP learning algorithm for concept	(10)

- ⁷⁾ Design a Brute-Force MAP learning algorithm for concept ⁽¹⁰⁾ learning to output the *maximum a posteriori hypothesis* (h_{MAP}) , based on Bayes theorem.
- ⁸⁾ Training data given in the table below is classifying days ⁽¹⁰⁾ according to whether someone will play tennis.

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
011	Sunny	Mild	Normal	Strong	Yes
012	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Use Naive Bayes classifier and the training data from this table to predict the target value (**yes** or **no**) of the target concept **PlayTennis** for the following new instance:

(Outlook = sunny, Temperature = cool, Humidity = high, Wind = strong) What is confusion matrix? Define the following

terms $(2.5 \times 4 = 10 \text{ Marks})$

(10)

a. True Positive

9)

10)

- b. True Negative
- c. False Positive
- d. False Negative
- Obtain the covariance matrix and principle components of ⁽¹⁰⁾ the following data set

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

-----End-----