

## SCHOOL OF INFORMATION SCIENCES SECOND SEMESTER MASTER OF ENIGINEERING - ME (Big Data and Data Analytics) **DEGREE EXAMINATION - NOVEMBER 2017** DATE : Wednesday, November 15, 2017 TIME : 10:00AM - 1:00PM Machine Learning [BDA 602]

Marks: 100

## Α

Answer all the questions.

1) (10)Write the definitions of Machine Learning given by Arthur Samuel and Tom Mitchell. Briefly describe about any two computer applications for which machine learning approaches seem appropriate.

(4+6 Marks)

Duration: 180 mins.

2) Implement Candidate-elimination algorithm to obtain most general and most <sup>(10)</sup> specific hypotheses for the training examples 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	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

3) (10) Write the futility of Bias-Free Learning and usefulness of Inductive learning.

(4+6 Marks)

(10) 4) a. Define entropy S, relative to the Boolean classification and information gain of an attribute A, relative to a collection of examples S. (5 Marks)

b. Explain K-Fold Cross-validation technique in training and testing with suitable example. (5 Marks)

(10)5) Suppose the hypothesis H contains conjunctions of constraints on up to nBoolean attributes, find the number of training examples required for the following:

ů l	
0.05 0.05	10
0.01 0.05	10
0.01 0.01	10
0.01 0.01	50

Write your comment on the result.

6) (5 (10) a. What is conditional probability? Give an example for the same. marks)

b. Define Bayes theorem. Write the meaning of each term in the theorem (5) marks)

(10) 7) Training data given in the table below is classifying days 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
D11	Sunny	Mild	Normal	Strong	Yes
D12	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)
<sup>8)</sup> Compute the following from a confusion matrix given in the following figure for a binary <sup>(10)</sup> classifier.

(2.5 \* 4 = 10)

- a. Accuracy
- b. Misclassification Rate
- c. Sensitivity
- d. False Positive Rate

n=165	Predicted: NO	Predicted: YES	
Actual: NO	TN = 50	FP = 10	60
Actual: YES	FN = 5	TP = 100	105
	55	110	

- <sup>9)</sup> Realize the logical OR and logical AND functions using McCulloch-Pitts neuron <sup>(10)</sup> model
- <sup>10)</sup> Explain about different types of trainings used in Artificial Neural Network <sup>(10)</sup> learning

-----End-----