

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST 2 EXAMINATION- APRIL 2019

B.Tech (VIII) Semester

COURSE CODE: 17B1WCI814

MAX. MARKS: 25

COURSE NAME: Design and Analysis of Real World Algorithms

COURSE CREDITS: 3

MAX. TIME: 1 Hr 30 Min

Note: All questions are compulsory.

1. [3 + 2 Marks]

a. Explain the following forms of learning:

- i. Reinforcement learning
- ii. Evolutionary learning

b. Give one real-life application example for the following learning tasks:

- i. Regression
- ii. Clustering

2. [5 Marks]

Table A shows a confusion matrix for medical data where the data values are yes or no for a class label attribute. The numbers within the four cells indicate number of patients. Based on the given information, answer the following:

- (a) Determine the misclassification rate. Can we use misclassification rate as a performance measure for the given data? Justify your answer.
- (b) For the given confusion matrix, determine sensitivity and specificity.
- (c) Which measure has better ability to detect the important class members correctly: misclassification rate; sensitivity; specificity?

Table A

	Predicted class = "yes"	Predicted class = "no"
Actual class = "yes"	90	210
Actual class = "no"	140	9560

3. [5 Marks]

Consider the dataset D given in Table B. Use the naïve Bayes classifier and the training data from this table to classify the following instance: {F, 1.75}

Table B

	Gender x1	Height x2	class	y
s ⁽¹⁾	F	1.6 m	Short	Y1
s ⁽²⁾	M	2 m	Tall	Y3
s ⁽³⁾	F	1.9 m	Medium	Y2
s ⁽⁴⁾	F	1.88 m	Medium	Y2
s ⁽⁵⁾	F	1.7 m	Short	Y1
s ⁽⁶⁾	M	1.85 m	Medium	Y2
s ⁽⁷⁾	F	1.6 m	Short	Y1
s ⁽⁸⁾	M	1.7 m	Short	Y1
s ⁽⁹⁾	M	2.2 m	Tall	Y3
s ⁽¹⁰⁾	M	2.1 m	Tall	Y3
s ⁽¹¹⁾	F	1.8 m	Tall	Y3
s ⁽¹²⁾	M	1.95 m	Medium	Y2
s ⁽¹³⁾	F	1.9 m	Medium	Y2
s ⁽¹⁴⁾	F	1.8 m	Medium	Y2
s ⁽¹⁵⁾	F	1.78 m	Medium	Y2

4. [5 Marks]

Describe the steps involved in the design process for the following clustering methods:

- K-means
- Fuzzy K-means

5. [2+ 2 + 1 Marks]

- Parameter k in k-NN algorithm could be a very large value or a very small value. Give the drawbacks, in any, of each choice.
- Show that logistic regression is a non linear regression problem.
- Write short note on : Sample error and True error