

JK Lakshmipat University Jaipur
INSTITUTE OF ENGINEERING AND TECHNOLOGY

End Term Examination, December 2022
Bachelor of Computer Application, Semester V, 2022-23

Roll No.....

CS1128: Applied Artificial Intelligence and Machine Learning

Time: 2 hours

Max. Marks: 20

Instructions to students:

1. Do not write anything other than your roll number on question paper.
2. Mention all the assumptions for your answers clearly.

- Q. 1 Consider following simple linear regression model to predict salary of graduating MBA students versus their percentage marks in grade 10.

4
(CO2,
3)

	Coefficients	Standard Error	t-stat	p-value
Intercept	61555.35534	66701.901	0.9228	0.3607
Percentage in grade 10	3076.177438	1031.5258	2.9821	0.0044

Write interpretation of simple linear regression coefficients and explain other results.

- Q. 2 In a study, 165 patients were being tested for the presence of a disease. There are two possible predicted classes: "yes" and "no". If we were predicting the presence of a disease, for example, "yes" would mean they have the disease, and "no" would mean they don't have the disease. Out of those 165 cases, the classifier predicted "yes" 110 times, and "no" 55 times. In reality, 105 patients in the sample have the disease, and 60 patients do not.

10
(CO2,
4)

n=165	Predicted: NO	Predicted: YES
Actual: NO	50	10
Actual: YES	5	100

Calculate following rates from given confusion matrix (describe the terms also):

- 2.1) True positive rate
- 2.2) True negative rate
- 2.3) False positive rate (Type I error)
- 2.4) False negative rate (Type II error)

Based on the calculation of 2.1-2.4, computed following from a confusion matrix for a binary classifier (describe the terms also):

- 2.5) Accuracy
- 2.6) Miss-classification rate
- 2.7) Sensitivity (recall)
- 2.8) Specificity
- 2.9) Precision
- 2.10) Prevalence

- Q. 3 Table shows 5 customers and their movie downloads from a portal. The data consists of genre of the movies, maximum rating given by the customer, and the marital status (code 1 implies married and 0 otherwise). For example, customer 1 downloaded 23 action, 5 romance, 15 comedy, and 0 Sci-fi movies and his maximum rating was 4. 3
(CO5)

Customer	Number of Movies Downloaded Under Each Genre				Maximum Rating (k = 5)	Marital Status (k = 6)
	Action	Romance	Comedy	Sci-fi		
	(k = 1)	(k = 2)	(k = 3)	(k = 4)		
1	23	5	15	0	4	0
2	5	18	16	2	5	1
3	25	0	0	15	5	0
4	2	30	15	0	4	1
5	45	0	0	10	5	0

Calculate Gower's Similarity Coefficient.

- Q.4 Consider movie DVD purchases made by two customers as given by the following sets 3
(CO5)

Customer 1 = {Jungle Book (JB), Iron Man (IM), Kung Fu Panda (KFP), Before Sunrise (BS), Bridge of spies (BoS), Forest Gump (FG)}

Customer 2 = {Casablanca (C), Jungle Book (JB), Forrest Gump, Iron Man (IM), Kung Fu Panda (KFP), Schindler's List (SL), The God Father (TGF)}

In this case, each movie is an attribute. The purchases made by the two customers are shown in Table

Movie Title	BS	BoS	C	FG	IM	JB	KFP	SL	TGF
Customer 1	1	1	0	1	1	1	1	0	0
Customer 2	0	0	1	1	1	1	1	1	1

Calculate Jaccard Similarity Coefficient.