BITS PILANI, DUBAI CAMPUS
DUBAI INTERNATIONAL ACADEMIC CITY, DUBAI

FIRST SEMESTER 2023-24
COURSE: BITS F464 (Machine Learning)
COMPONENT: Practice Tutorial 2 **Date:** 10th October 2023

Q1: Consider a following confusion matrices, calculate precision, recall, F1-score, and accuracy by considering it.

55	5
10	30

Answer:

Precision = 0.8462

Recall = 0.91

F1-score = 0.88

Accuracy = 0.85

Q2: Calculate the best LDA (Linear Discriminant Analysis) projection vector for following dataset. (Show all steps of calculations)

X1 = (x1, x2) = {(1,4), (3,7), (2,4)}

X2 = (x1, x2) = {(11,10), (9,11), (8,9)}

Answer:

V1= -0.567, V2 = 0.8278

Q3: Consider the following dataset,

X1	**X2**	**Y**
10	20	1
5	5	0
15	25	1
13	22	1
8	10	0
11	7	0

1. Calculate the probability of Y=1 for X1 = 16, X2 = 23?

2. Calculate the probability of Y=0 for X1 = 5 and X2 = 8?

Answer:

Regression line: Y = -0.03027X1 + 0.0723X2 - 0.25967

1. 0.71

2. 0.54

Q4: Consider the dataset of Titanic survival (titanic_train.csv.csv) and perform following operations on it.

i. Read the csv file and bulid the dataframe **train**.

ii. Perform the exploratory data analytics using coutplot (i) Draw countplot with respect to "Survived" features (ii) Draw countplot with respect to "Survived" features by considering **hue=Sex** (iii) Draw countplot with respect to "Survived" features by considering **hue=Pclass**.

iii. Check for data carefully for missing values, nan, or nan values. Do data cleaning.

iv. Build logistic regression model for predicting/classifying either person is survived or not by considering PassengerID, Pclass, Age, SibSp, Parch, Fare, and Sex, as an independent variable.