

**Data Mining and Web Algorithms**  
**15B22CI621**

**Tutorial 6**

**Question 1** Consider the following data set having the attribute acid durability and strength to classify whether a special paper tissue (  $X_1 = 3$  and  $X_2 = 7$ ) using euclidian as well as manhattan distance is good or not .( Use  $K=3$ )

$X_1$ = Acid Durability (seconds)	$X_2$ = Strength (kg/square meter)	Y = Classification
7	7	Bad
7	4	Bad
3	4	Good
1	4	Good

**Question 2** Consider the following data set

Example No.	Color	Type	Origin	Stolen?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

Classify the following example using naïve bayes classifier:

- (i) Example 11: ( Red, SUV, Domestic)
- (ii) Example 12: (Yellow, Sports, Imported)

**Question3.** Consider the following data set and classify following test samples using Naïve bayes classifier:

Tid	Refund	Marital Status	Taxable Income	Evade
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

- (i) Refund = yes, Status="Single", Income=125K)
- (ii) (Refund = No, Status="Married", Income=120K)