

DA assignment - 1

Q1. Calculating posterior and prior probabilities

Attribute	on time	Late	very late	Cancelled
Day:				
Weekday	9/14	1/2	3/3	0/1
Saturday	2/14	1/2	0/3	1/1
Sunday	1/14	0/2	0/3	0/1
Holiday	2/14	0/2	0/3	0/1
Fog:				
None	5/14	0/2	0/3	0/1
High	4/14	1/2	1/3	1/1
Normal	5/14	1/2	2/3	0/1
Rain:				
None	5/14	1/2	1/3	0/1
Slight	8/14	0/2	0/3	0/1
heavy	1/14	1/2	2/3	1/1
Season				
Spring	4/14	0/2	0/3	0/1
Summer	2/14	0/2	0/3	0/1
Autumn	2/14	0/2	1/3	0/1
Winter	2/14	2/2	2/3	0/1
Prior Probability	14/20	2/20	3/20	1/20

Instances Weekday, Winter, High, None ??

Case 1: class : Late

$$= 0.1 \times 0.5 \times 1 \times 0.5 \times 0.5$$

$$= 0.0125$$

Case 2: class = on time

$$= 0.7 \times 0.64 \times 0.14 \times 0.29 \times 0.36$$

$$= 6.547 \times 10^{-3} = 0.006547$$

Case 3: class = Very late

$$= 0.15 \times 1 \times 0.64 \times 0.33 \times 0.73$$

$$= 0.0109$$

Case 4: class = Cancelled

$$= 0.05 \times 0.0 \times 0.0 \times 1 \times 0$$

$$= 0$$

∴ Case 2 is strong
 ∴ The instance will be categorized under class Late

Q2. Using χ^2 (chi-square) calculation
degree of freedom = 1
By using formula

$$\chi^2 = \sum_{i=1}^m \sum_{j=1}^n \frac{(a_{ij} - e_{ij})^2}{e_{ij}}$$

$$\therefore \chi^2 = \frac{[250 - 90]^2}{90} + \frac{[50 - 210]^2}{210} + \frac{[200 - 360]^2}{360} + \frac{[1000 - 840]^2}{840}$$
$$= 507.93$$

Chi square value is large as compared to 2.706
(for 1 degree of freedom χ^2 value is rejected. the hypothesis)
Hence, the gender and preferred reading are strongly correlated.
We can reject the null hypothesis.