CS-5783 : Assignment 04

Question 1.

Calculation of conditional probability for each feature.

Methodology :

We will be using Gaussian distribution to get conditional probabilities.

(i) Calculate mean for each feature over each class.

(ii) Calculate standard deviation for each feature over each class.

(iii) Use normal distribution formulae to compute conditional probability, given by :-

$$P(X_j \mid C = c_i) = \frac{1}{\sqrt{2\pi\sigma_{ji}^2}} e^{\left(\frac{-(x_j - \mu_{ji})^2}{2\sigma_{ji}^2}\right)}$$

(μ_{ji} → mean of X_j's for class c_i; σ_{ji} → std of X_j's for class c_i)

where

$$\mu_{ji} = \frac{\sum X_j}{\text{count}(X_j)}$$

$$\sigma^2 = \frac{\sum (X_j - \mu_j)^2}{N} \quad ; N \text{ -> number of samples.}$$

$$\text{std} = \sigma = \sqrt{\frac{\sum (X_j - \mu_j)^2}{N}}$$

Based on these formulaes calculate the μ, σ^2, σ and conditional probability for all features for given class.

In detail calculations for two features is shown :-

Feature (1) : #Rooms

class → / stat ↓	Apartment	Condo	House
mean (μ)	$= \frac{7+8+5+6+9+6+7}{7}$ $= 6.85$	$= \frac{6+7+10+6+6+6}{6}$ $= 6.83$	$= \frac{6+6+6+5+6+7+7}{7}$ $= 6.14$
Variance (σ^2)	$= \frac{(7-6.85)^2+(8-6.85)^2+(5-6.85)^2+(6-6.85)^2+(9-6.85)^2+(6-6.85)^2+(7-6.85)^2}{6}$ $= 1.809$	$= \frac{(6-6.83)^2+(7-6.83)^2+(10-6.83)^2+(6-6.83)^2+(6-6.83)^2+(6-6.83)^2}{5}$ $= 2.56$	$= \frac{(6-6.14)^2+(6-6.14)^2+(6-6.14)^2+(5-6.14)^2+(6-6.14)^2+(7-6.14)^2+(7-6.14)^2}{6}$ $= 0.47$
Std (σ)	$= \sqrt{1.809} = 1.345$	$= \sqrt{2.56} = 1.60$	$= \sqrt{0.47} = 0.69$
Conditional Probabilities (for unique test case values of features)	$P(6|Apt) = \frac{1}{\sqrt{2\pi \times 1.809^2}} e^{\left(-\frac{(6-6.85)^2}{2\times 1.809^2}\right)}$ $= 0.24$ Similarly $P(7|Apt) = 0.29$ $P(8|Apt) = 0.20$	$P(6|Condo) = \frac{1}{\sqrt{2\pi(2.56)^2}} \times e^{\left(\frac{(6-6.83)^2}{2\times 2.56^2}\right)}$ $= 0.21$ $P(7|Condo) = 0.24$ $P(8|Condo) = 0.19$	$P(6|House) = \frac{1}{\sqrt{2\pi(0.47)^2}} \times e^{\left(\frac{(6-6.14)^2}{2\pi(0.47)^2}\right)}$ $= 0.56$ $P(7|House) = 0.26$ $P(8|House) = 0.015$

Feature 2 : # Bathrooms

class → / stats ↓	Apartment	Condo	House
Mean (μ)	$= \frac{1+1+1+2.5+1+1.5+1}{7}$ $= 1.28$	$= \frac{1+1+2.5+1.5+1+1}{6}$ $= 1.33$	$= \frac{1+1+1+1+1+1+1.5}{7}$ $= 1.07$
Variance (σ^2)	$= \frac{(1-1.28)^2+(1-1.28)^2+(1-1.28)^2+(2.5-1.28)^2+(1-1.28)^2+(1.5-1.28)^2+(1-1.28)^2}{6}$ $= 0.321$	$= \frac{(1-1.33)^2+(1-1.33)^2+(2.5-1.33)^2+(1.5-1.33)^2+(1-1.33)^2+(1-1.33)^2}{5}$ $= 0.366$	$= \frac{(1-1.07)^2+(1-1.07)^2+(1-1.07)^2+(1-1.07)^2+(1-1.07)^2+(1.5-1.07)^2}{6}$ $= 0.035$
Std (σ)	$= \sqrt{0.321} = 0.56$	$= \sqrt{0.366} = 0.60$	$= \sqrt{0.035} = 0.188$
Conditional Probabilities (for unique test set values of feature)	$P(1\|Apt) = \frac{1}{\sqrt{2\pi(0.56)^2}} \times e^{-\frac{(1-1.28)^2}{2(0.56)^2}}$ $= 0.62$ Similarly, $P(1.5\|Apt) = 0.65$	$P(1\|Condo) = \frac{1}{\sqrt{2\pi(0.6)^2}} \times e^{-\frac{(1-1.33)^2}{2(0.6)^2}}$ $= 0.57$ $P(1.5\|condo) = 0.63$	$P(1\|House) = \frac{1}{\sqrt{2\pi(0.188)^2}} \times e^{-\frac{(1-1.28)^2}{2(0.188)^2}}$ $= 1.97$ $P(1.5\|House) = 0.155$

For the remaining features the calculated values are tabulated in next page,

Features → / Class ↓		Local Price	Land Area	Living Area	# Garages	# Bedrooms	Age of home
Apartment :	μ	7.33	6.10	1.505	1.21	3.42	38.71
	σ^2	13.07	10.562	0.496	0.476	0.9409	215.502
	σ	3.61	3.25	0.704	0.69	0.97	14.68
Condo :	μ	7.41	6.024	1.553	1.33	3.33	39.66
	σ^2	21.26	6.452	0.857	0.2601	0.656	194.60
	σ	4.61	2.54	0.926	0.51	0.81	13.95
House :	μ	5.76	6.6309	1.391	1.07	3.00	34.28
	σ^2	0.32	5.0176	0.0449	0.689	0.3249	161.79
	σ	0.57	2.24	0.212	0.83	0.57	12.72

Conditional probabilities (CP) of different features.

Features ↓ / Class (CP) →	X_j values	$P(X_j \mid \text{Apartment})$	$P(X_j \mid \text{Condo})$	$P(X_j \mid \text{House})$
Local Price	6.0931	0.104	0.083	0.5
	8.3607	0.106	0.084	2.1×10^{-5}
	8.14	0.107	0.085	1×10^{-4}
	9.1416	0.097	0.080	1.59×10^{-8}
	12	0.047	0.52	6.62×10^{-27}
Land Area	6.7265	0.120	0.15	0.17
	9.15	0.02	0.07	0.09
	8.00	0.103	0.116	0.14
	7.3262	0.114	0.132	0.16
	5.00	0.115	0.144	0.13
Living Area	1.652	0.55	0.429	0.89
	1.177	0.52	0.419	0.35
	1.504	0.56	0.431	0.63
	1.831	0.50	0.413	0.21
	1.2	0.51	0.401	1.25
# Garages	1	0.552	0.63	0.47
	1.5	0.520	0.73	0.42
	2	0.300	0.33	0.25
# Bedrooms	3	0.37	0.45	0.69
	4	0.34	0.34	0.15
Age of home	44	0.0025	0.027	0.023
	48	0.0021	0.023	0.017
	3	0.001	0.0009	0.0015
	31	0.023	0.02	0.030
	30	0.022	0.022	0.029