

Full name: Tran Hai Nam
 ID: 11219279
 Class: DSEB 63

Problem 2:

No	Color	Legs	Height	Smelly	Species
1	White	3	Short	Yes	M
2	Green	2	Tall	No	M
3	Green	3	Short	Yes	M
4	White	3	Short	Yes	M
5	Green	2	Short	No	H
6	White	2	Tall	No	H
7	White	2	Tall	No	H
8	White	2	Short	Yes	H

- Estimate conditional probabilities of each attributes {color, legs, height, smelly} for the species classes: {M, H} using the data given in the table.
- Using these probabilities estimate the probability values for the new instance -

(Color=Green, legs=2, Height=Tall, and Smelly=No).

Sol:

$$P(M) = 4/8 = 0.5, \quad P(H) = 4/8 = 0.5$$

Color	M	H
White	2/4	3/4
Green	2/4	1/4

Legs	M	H
2	1/4	4/4
3	3/4	0/4

Height	M	H
Tall	3/4	2/4
Short	1/4	2/4

Smelly	M	H
Yes	3/4	1/4
No	1/4	3/4

New instance = (Color=Green, legs=2, Height=Tall, Smelly=No)

$$p(M | \text{New Instance}) = p(M) * p(\text{Color} = \text{Green} | M) * p(\text{Legs} = 2 | M) * p(\text{Height} = \text{tall} | M) * p(\text{Smelly} = \text{no} | M)$$

$$p(M | \text{New Instance}) = 0.5 * \frac{2}{4} * \frac{1}{4} * \frac{3}{4} * \frac{1}{4} = 0.0117$$

$$p(H | \text{New Instance}) = p(H) * p(\text{Color} = \text{Green} | H) * p(\text{Legs} = 2 | H) * p(\text{Height} = \text{tall} | H) * p(\text{Smelly} = \text{no} | H)$$

$$p(H | \text{New Instance}) = 0.5 * \frac{1}{4} * \frac{4}{4} * \frac{2}{4} * \frac{3}{4} = 0.047$$

Since: $p(H | \text{New Instance}) > p(M | \text{New Instance})$

⇒ **Hence the new instance belongs to species H**

Problem 3:

Person	Height (ft)	Weight (lbs)	Foot size (inches)
Male	6	180	12
Male	5.92	190	11
Male	5.58	170	12
Male	5.92	165	10
Female	5	100	6
Female	5.5	150	8
Female	5.42	130	7
Female	5.75	150	9

Based on the following data determine the gender of a person having height 6 ft., weight 130 lbs, and foot size 8 inch. (use Naive Bayes algorithm).

Sol:

$$P(\text{Male}) = 4/8 = 0.5$$

$$P(\text{Female}) = 4/8 = 0.5$$

Male:

$$\text{Mean (Height)} = \frac{(6+5.92+5.58+5.92)}{4} = 5.855$$

$$\text{Variance (Height)} = \frac{\sum_1^n (x_i - \text{mean})^2}{n-1} = 0.035055$$

Sex	Mean (height)	Variance (height)	Mean (weight)	Variance (weight)	Mean(foot size)	Variance (foot size)
Male	5.855	0.035033	176.25	122.92	11.25	0.91667
Female	5.4175	0.097225	132.5	558.33	7.5	1.6667

$$P(H|M) = 1.5789$$

$$P(W|M) = 5.9881e^{-6}$$

$$P(FS|M) = 1.3112e^{-3}$$

$$P(H|F) = 2.2346e^{-1}$$

$$P(WF) = 1.6789e^{-2}$$

$$P(FS|F) = 2.8669e^{-1}$$

$$\begin{aligned} \text{Posterior (Male)} &= \frac{P(M) * P(H|M) * P(W|M) * P(FS|M)}{\text{Evidence}} \\ &= 0.5 * 1.5789 * 5.9881e^{-6} * 1.3112e^{-3} \\ &= 6.1984e^{-9} \end{aligned}$$

$$\begin{aligned} \text{Posterior (Female)} &= \frac{P(F) * P(H|F) * P(W|F) * P(FS|F)}{\text{Evidence}} \\ &= 0.5 * 2.2346e^{-1} * 1.6789e^{-2} * 2.8669e^{-1} \\ &= 5.377e^{-4} \end{aligned}$$

Since : Posterior (Female) > Posterior (Male)

⇒ **New instance belong to Female group.**