Sunday, May 5, 2024 9:42 AM

## Bayes Theorem

$$P(A/B) = \frac{P(B/A) \times P(A)}{P(B)}$$

- $P(A/B)$ → Posterior
- $P(B/A)$ → likelihood
- $P(A)$ → Prior
- $P(B)$ → evidence

***Probability & likelihood ? → difference

height → dataset

① $P(h=170\,cm / \mu=150\,cm, \sigma=10\,cm) \Rightarrow$ Prob.

② $P(\mu=150\,cm, \sigma=10\,cm / h=170\,cm) \Rightarrow$ likelihood

"Teenagers" ← Naive Bayes

innocent,
ignorant,
lack of wisdom

→ All features of independent of each other (input)

$$P(C_x/X_i) = \frac{P(X_i/C_x)\,P(C_x)}{P(X_i)}$$

$C_x \rightarrow$ class label

$X_i \Rightarrow$ input features

$$P(Yes/outlook) = \frac{P(outlook/Yes)\,P(Yes)}{P(outlook)}$$

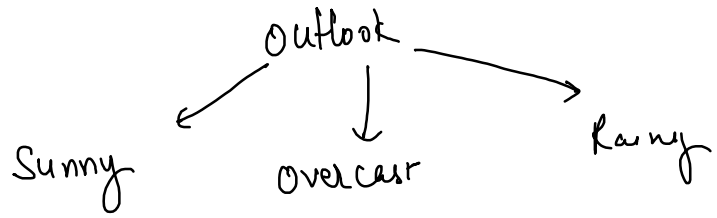
$$P(No/outlook) = \frac{P(outlook/No)\,P(No)}{P(outlook)}$$

| Outlook | Temperature | Humidity | Windy | PlayTennis |
|---|---|---|---|---|
| Sunny | Hot | High | False | No |
| Sunny | Hot | High | True | No |
| Overcast | Hot | High | False | Yes |
| Rainy | Mild | High | False | Yes |
| Rainy | Cool | Normal | False | Yes |
| Rainy | Cool | Normal | True | No |
| Overcast | Cool | Normal | True | Yes |
| Sunny | Mild | High | False | No |
| Sunny | Cool | Normal | False | Yes |
| Rainy | Mild | Normal | False | Yes |
| Sunny | Mild | Normal | True | Yes |
| Overcast | Mild | High | True | Yes |
| Overcast | Hot | Normal | False | Yes |
| Rainy | Mild | High | True | No |

$$P(Yes) = \frac{9}{14} \qquad P(No) = \frac{5}{14}$$

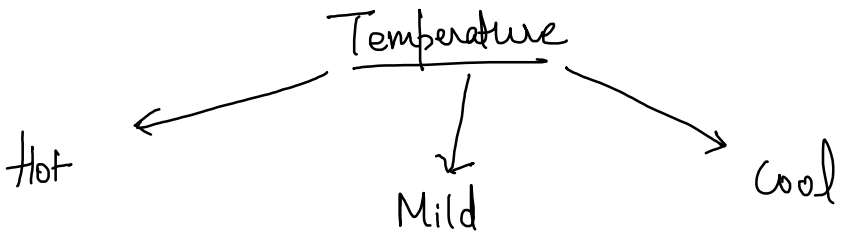
Working

NB fit (X_train, Y_train)




$$P(Sunny/Yes) = \frac{2}{9} \qquad P(Overcast/Yes) = \frac{4}{9} \qquad P(Rainy/Yes) = \frac{3}{9}$$

$$P(Sunny/No) = \frac{3}{5} \qquad P(Overcast/No) = \frac{0}{5} = 0 \qquad P(Rainy/No) = \frac{2}{5}$$



Mild

$P(Hot/Yes) = \frac{2}{9}$ $P(Mild/Yes) = \frac{4}{9}$ $P(cool/Yes) = \frac{3}{9}$

$P(Hot/No) = \frac{2}{5}$ $P(Mild/No) = \frac{2}{5}$ $P(cool/No) = \frac{1}{5}$

Humidity

High Normal

$P(High/Yes) = \frac{3}{9}$ $P(Normal/Yes) = \frac{6}{9}$

$P(High/No) = \frac{4}{5}$ $P(Normal/No) = \frac{1}{5}$

Windy

True False

$P(True/Yes) = \frac{3}{9}$ $P(false/Yes) = 6/9$

$P(True/No) = \frac{3}{5}$ $P(false/No) = 2/5$

Q outlook → Sunny , temp → cool , humidity → high, wind → true

Will 9 play?

Sol. $P(Yes/sunny, cool, high, true) = P(Yes) \times P(Sunny/Yes) \times P(cool/Yes) \times P(high/Yes) \times P(true/Yes)$

$$= \frac{9}{14} \times \frac{2}{9} \times \frac{3}{9} \times \frac{3}{9} \times \frac{3}{9}$$

$$= 0.00529$$

$P(No/sunny, cool, high, true) = P(No) \times P(sunny/No) \times P(cool/No) \times P(high/No) \times P(true/No)$

$$= \frac{5}{14} \times \frac{3}{5} \times \frac{1}{5} \times \frac{4}{5} \times \frac{3}{5} = 0.02$$

| $P(Yes/X_i)$ | | $P(No/X_i)$ |
|---|---|---|
| 0.0052 | < | 0.02 |

No, 9 will not play

Q outlook → rainy, temp → cool, humidity → high, windy → false
Will 9 play?

Sol. $P(Yes/rainy, cool, high, false) = 0.015$

$P(No/rainy, cool, high, false) = 0.22$

No, 9 will not play

Problem of zero probability:

Q $P(\text{Yes} \mid \text{cloudy}, \text{cool}, \text{high}, \text{true}) = 0$

$P(\text{No} \mid \text{cloudy}, \text{cool}, \text{high}, \text{true}) = 0$

(I) Ignore cloudy ✗

$$\begin{bmatrix} P(\text{cloudy} \mid \text{Yes}) = 1 \\ P(\text{cloudy} \mid \text{No}) = 1 \end{bmatrix} \times$$

$P(\text{Yes} \mid \text{cloudy}, \text{cool}, \text{high}, \text{true}) = P(\text{Yes}) \times P(\text{cool} \mid \text{Yes}) \times P(\text{high} \mid \text{Yes}) \times P(\text{true} \mid \text{Yes}) \times 1$

$P(\text{No} \mid \text{cloudy}, \text{cool}, \text{high}, \text{true}) = P(\text{No}) \times P(\text{cool} \mid \text{No}) \times P(\text{high} \mid \text{No}) \times P(\text{true} \mid \text{No}) \times 1$

(II) Laplace Smoothing: (var_smoothing)

$$P(\text{Cloudy} \mid \text{Yes}) = \frac{0 + \alpha}{n + \alpha K}$$

$\alpha \rightarrow$ smoothing parameter

$n \rightarrow$ # datapoints ($y = \text{yes}$)

$K \rightarrow$ # distinct values your column can take

effect of $\alpha$:

10000 → words ⇒ $y = 1$ 2 rare words

(I) $\alpha = 0$

(I) $\alpha = 0$

$$P(RW/y=1) = \frac{2}{1000} \approx 0$$

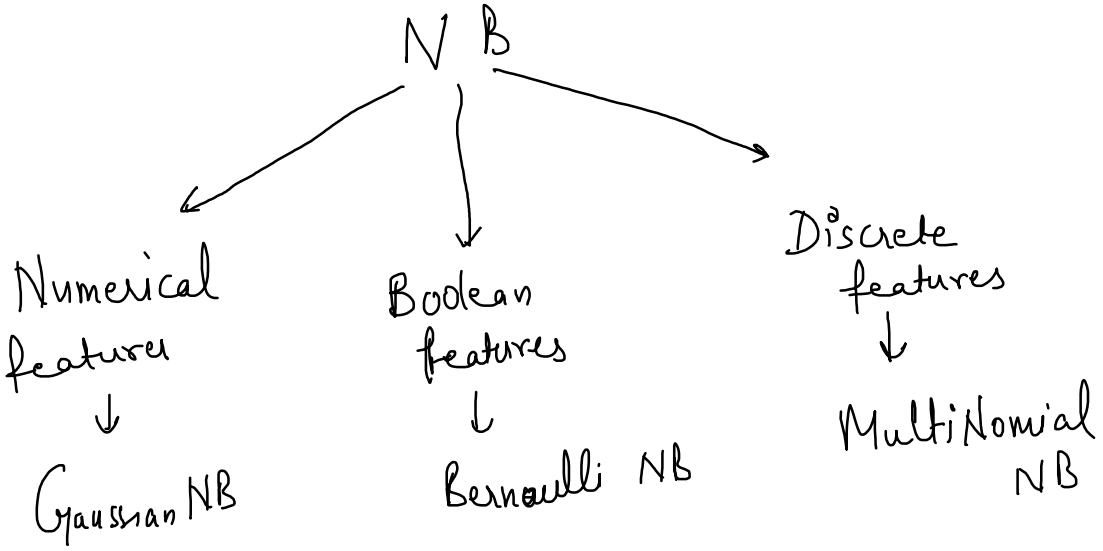
overfitting

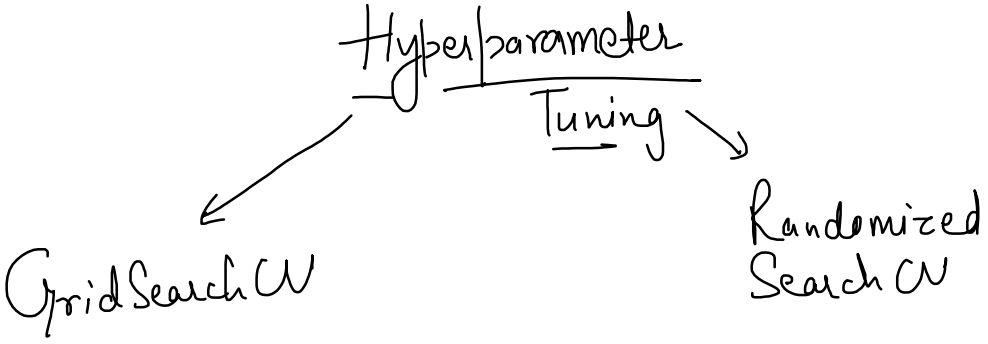
(II) $\alpha = 10000$

underfitting

$$P(RW|y=1) = \frac{2+10000}{1000+2\times10000} = \frac{10002}{21000} \approx 0.5$$

Thumb Rule: $\alpha = 1$




N_i : [100, 200, 300, 400] => 4

$\alpha_1$ : [100, 200, 300, 400] ⇒ 4

$\alpha_2$ : [1, 2, 3, 4] ⇒ 4

$\alpha_3$ : [0.01, 0.02, 0.03, 0.04, 0.05] ⇒ 5

$\alpha_4$ : [20, 30, 40] = 3

Total No. of models ⇒ 4 X 4 X 5 X 3 = 240

→ gives exact best parameters

= same as LHS

→ Out of 240 models any 10 models will be selected randomly

→ approx. best parameters