Naive Bayes Classifier

NB classifier applies to learning tasks where each instance x is described by a conjunction of attribute values & where target function f(x) can only take any value from finite set V.

$$v_{MAP} = \text{argmax}\ P(v_j \mid a_1, a_2, \ldots a_n)$$

v_j → class label; $a_1, \ldots a_n$ → attributes (data)

↓ rewrite as

$$v_{MAP} = \underset{v_j \in V}{\text{argmax}} \frac{P(a_1, a_2, a_3 \ldots a_n \mid v_j) P(v_j)}{P(a_1, a_2, \ldots a_n)}$$

$P(a_1, a_2, \ldots a_n)$ → this is same across classes

$$= \underset{V}{\text{argmax}}\ P(a_1, a_2, \ldots, a_n \mid v_j) P(v_j)$$

→ prob of attribute values given class × Prob of class

- $P(v_j)$: class proportions in training data
- Assuming all attributes are independent of each other:

$$P(a_1, a_2, \ldots a_n \mid v_j) = \prod_i P(a_i \mid v_i)$$

↓ plugging these in

$$v_{NB} = \underset{v_j}{\text{argmax}}\ P(v_j) P(a_i \mid v_j)$$ } Naive Bayes Classifier

→ target value output by NB classifier

ex// Instance to predict:

Outlook	Temp	Humid	Wind
Sunny	Cool	High	Strong

Target = Play Tennis (yes / no)

1. Estimate $P(v_j)$ from class proportions

   $P(\text{PlayTennis} = \text{yes}) = 0.64$

   $P(\text{PlayTennis} = \text{No}) = 0.36$

2. Estimate conditional probabilities

   $P(\text{wind} = \text{Strong} \mid \text{Play Tennis} = \text{yes}) = 0.33$

   $P(w = \text{strong} \mid PT = \text{no}) = 0.6$

$P(\text{yes} \mid \text{inst}) = P(\text{yes}) \times P(\text{sunny} \mid \text{yes}) \times P(\text{cool} \mid \text{yes}) P(\text{high} \mid \text{yes}) P(\text{str} \mid \text{yes})$

$= 0.0206$

Calculate same for other labels:

$$P(no \mid inst) = \underbrace{P(no)}_{P(v_j)} \underbrace{P(sunny \mid no)\, P(cool \mid no)\, P(high \mid no)\, P(Strong \mid no)}_{\prod_i P(a_i \mid v_j)}$$

$= 0.0206 \rightarrow$ Our prediction (larger than "yes")

Normalizing these: $\frac{0.0206}{0.0206 + 0.0053} = 0.795$