Naive Bayes classification

Naive Bayes is a simple yet powerful classification algorithm. It is based on Bayes' theorem with the assumption of conditional independence

In simple terms, (class) conditional independence assumes that presence of a particular feature in a class is unrelated to the presence of any other feature.

Bayes' Theorem:

Given a hypothesis H and evidence E, the probability of the hypothesis before getting the evidence [P(H)] and the probability after obtaining the evidence is related as follows:

$$P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E)}$$

Bayes' Theorem is based on conditional probability.

Eg: In a deck of cards

$$P(King) = \frac{4}{52} = \frac{1}{13}$$

$$P(King|Face) = \frac{P(Face|King) \cdot P(King)}{P(Face)}$$

$$= \frac{1 \cdot \frac{1}{13}}{\frac{3}{13}} = \frac{1}{3}$$

Proof for Bayes' Theorem

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

$$\therefore P(A \cap B) = P(A|B) \cdot P(A) = P(B|A) \cdot P(B)$$

$$\Rightarrow P(A|B) = \frac{P(B|A) \cdot P(B)}{P(A)}$$

$$\overbrace{P(H|E)}^{\text{Posterior Probability}} = \frac{\overbrace{P(E|H)}^{\text{Likelihood}} \cdot \overbrace{P(H)}^{\text{Prior Probability}}}{\underbrace{P(E)}_{\text{Marginal Probability}}}$$

$P(H|E)$ – how probable the hypothesis is, given the observed evidence?

$P(E|H)$ – how probable the evidence is, given the hypothesis is true?

$P(H)$ – how probable was the hypothesis before observing the evidence?

$P(E)$ – how probable is the evidence under all possible hypotheses?

Naive Bayes: Worked Example

Day	Outlook	Humidity	Wind	Play
D1	Sunny	High	Weak	No
D2	Sunny	High	Strong	No
D3	Overcast	High	Weak	Yes
D4	Rain	High	Weak	Yes
D5	Rain	Normal	Weak	Yes
D6	Rain	Normal	Strong	No
D7	Overcast	Normal	Strong	Yes
D8	Sunny	High	Weak	No
D9	Sunny	Normal	Weak	Yes
D10	Rain	Normal	Weak	Yes
D11	Sunny	Normal	Strong	Yes
D12	Overcast	High	Strong	Yes
D13	Overcast	Normal	Weak	Yes
D14	Rain	High	Strong	No

Priors

Frequency Table	Yes	NO
play	9/14	5/14

Likelihood Tables

Frequency Table		Play Yes	Play NO	
Outlook	Sunny	2	3	5
	Overcast	4	0	~~8~~4
	Rainy	3	2	5
		9	5	14

$P(Sunny|yes) = 2/9$

$P(Sunny) = 5/14$

$P(yes) = \frac{9}{14}$

Frequency Table		Play Yes	Play NO	
Humidity	High	3	4	7
	Normal	6	1	7
		9	5	14

$P(No|High) = \frac{4}{5}$

$P(High) = \frac{7}{14}$

Frequency Table		Play Yes	Play NO	
Wind	strong	6	2	8
	Weak	3	3	6
		9	5	14

$P(Weak|No) = \frac{3}{5}$

$$(yes|sunny) = \frac{P(Sunny|yes)\, P(yes)}{P(sunny)} = \frac{\frac{2}{9} \times \frac{9}{14}}{\frac{5}{14}} = \frac{2}{5} = 0.4$$

$$P(no|Sunny) = \frac{P(Sunny|no)\, P(no)}{P(Sunny)} = \frac{\frac{3}{5} \times \frac{5}{14}}{\frac{5}{14}} = 0.6$$

classification

Outlook = Rain

Humidity = High

Wind = Weak

Play = ?

~~P(Outlook = Rain)~~

P(Play = Yes | Outlook = Rain, Humidity = High, Wind = Weak)

= P(Outlook = Rain | Yes) * P(Humidity = High | Yes) * P(Wind = Weak | Yes) * P(Yes)

$= \frac{3}{9} * \frac{3}{9} * \frac{3}{9} * \frac{9}{14} = \frac{1}{52} = \underline{0.0714}$ ✓

P(Play = No | Outlook = Rain, Humidity = High, Wind = Weak)

= P(Outlook = Rain | No) * P(Humidity = High | No) * P(Wind = Weak | No) * P(No)

$= \frac{3}{9} * \frac{3}{9} * \frac{3}{9} * \frac{5}{14} = 0.0132$

$\Rightarrow \boxed{\text{Play} = \text{Yes}}$

Naive Bayes classification

~~$P(H_i|E) = \underset{H_i}{\text{argmax}}$~~

$$P(H_i|E) = \underset{H_i}{\text{argmax}}\ P(E|H_i) \cdot P(H_i)$$

Applications of Naive Bayes classification

- Spam filtering
- News categorization
- Medical Diagnosis
- Weather Prediction