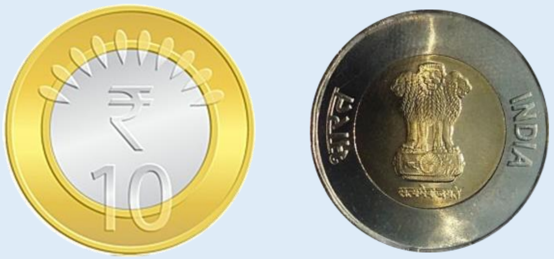
Naïve Bayes

Classification Algorithm

Probability Basics



There are two 10 Rs coins, when they are flipped together, the possible outcome could be

HH HT TT TH

This will be called as a sample space: $\{HH, HT, TT, TH\}$

Probabilities of occurrence of various event can be given as:

P(Getting two heads) = 1/4
P(Getting two tails) = 1/4
P(Getting at least one tail) = 3/4
P(Getting at least one head) = 3/4

P(Second coin being head given first is tail) = 1/2

P(Getting two head given first coin is head) = 1/2

Conditional Probability

Bayes Theorem

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

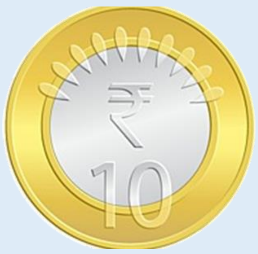
$P(A|B) = Conditional\ Probability\ of\ A\ given\ B$

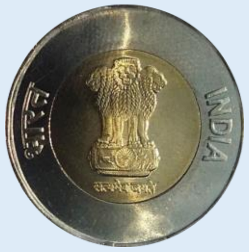
$P(B|A) = Conditional\ Probability\ of\ B\ given\ A$

$P(A) = Probability\ of\ event\ A$

$P(B) = Probability\ of\ event\ B$

Conditional Probability





The sample space: $\{HH, HT, TT, TH\}$

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

Bayes Theorem

$$P(Second\ coin\ being\ head\ given\ first\ is\ tail) = P(A|B)$$

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

$$= \frac{P(First\ Coin\ being\ Tail\ given\ second\ is\ head) \times P(Second\ Coin\ being\ head)}{P(First\ Coin\ Being\ Tail)}$$

$$= \frac{\frac{1}{2} \times \frac{2}{4}}{\frac{2}{4}}$$

$$= \frac{1}{2} = 0.5$$

Bayes Theorem

Likely hood

Posterior Probability $$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$ Class Prior Probability

Predictor Prior Probability

$P(A|B) = Conditional\ Probability\ of\ A\ given\ B$

$P(A|B) = Conditional\ Probability\ of\ B\ given\ A$

$P(A) = Probability\ of\ event\ A$

$P(B) = Probability\ of\ event\ B$

	OUTLOOK	TEMP	HUMIDITY	WINDY	PLAY
1	Sunny	Hot	High	False	No
2	Sunny	Hot	High	True	No
3	Overcast	Hot	High	False	Yes
4	Rainy	Mild	High	False	Yes
5	Rainy	Cool	Normal	False	Yes
6	Rainy	Cool	Normal	True	No
7	Overcast	Cool	Normal	True	Yes
8	Sunny	Mild	High	False	No
9	Sunny	Cool	Normal	False	Yes
10	Rainy	Mild	Normal	False	Yes
11	Sunny	Mild	Normal	True	Yes
12	Overcast	Mild	High	True	Yes
13	Overcast	Hot	Normal	False	Yes
14	Rainy	Mild	High	True	No

Likely hood

Posterior Probability $$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$ Class Prior Probability

Predictor Prior Probability

Here, we are trying to predict whether there will be a play, on the basis of weather condition, hence weather condition is a predictor here, whereas Yes and No is the class we are trying to predict

Class Prior Probability:

$$P(A) = P(Yes) = \frac{9}{14}$$

$$P(A) = P(No) = \frac{5}{14}$$

Predictor Prior Probability:

$$P(B) = P(Synny) = \frac{5}{14} \qquad P(B) = P(Hot) = \frac{4}{14} \qquad P(B) = P(High) = \frac{7}{14} \qquad P(B) = P(False) = \frac{8}{14}$$

$$P(B) = P(Overcast) = \frac{4}{14} \qquad P(B) = P(Mild) = \frac{6}{14} \qquad P(B) = P(Normal) = \frac{7}{14} \qquad P(B) = P(True) = \frac{6}{14}$$

$$P(B) = P(Rainy) = \frac{5}{14} \qquad P(B) = P(Cool) = \frac{4}{14}$$

	OUTLOOK	TEMP	HUMIDITY	WINDY	PLAY
1	Sunny	Hot	High	False	No
2	Sunny	Hot	High	True	No
3	Overcast	Hot	High	False	Yes
4	Rainy	Mild	High	False	Yes
5	Rainy	Cool	Normal	False	Yes
6	Rainy	Cool	Normal	True	No
7	Overcast	Cool	Normal	True	Yes
8	Sunny	Mild	High	False	No
9	Sunny	Cool	Normal	False	Yes
10	Rainy	Mild	Normal	False	Yes
11	Sunny	Mild	Normal	True	Yes
12	Overcast	Mild	High	True	Yes
13	Overcast	Hot	Normal	False	Yes
14	Rainy	Mild	High	True	No

Likely hood

Posterior Probability $P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$ Class Prior Probability

Predictor Prior Probability

Frequency Table		**Play**	
		Yes	No
Outlook	Sunny	2	3
	Overcast	4	0
	Rainy	3	2

Frequency Table		**Play**	
		Yes	No
Temp	Hot	2	2
	Mild	4	2
	Cool	3	1

Frequency Table		**Play**	
		Yes	No
Humidity	High	3	4
	Normal	6	1

Frequency Table		**Play**	
		Yes	No
Windy	True	3	3
	False	6	2

Frequency Table		Play	
		Yes	No
Outlook	Sunny	2	3
	Overcast	4	0
	Rainy	3	2

Likelihood Table		Play	
		Yes	No
Outlook	Sunny	2/9	3/5
	Overcast	4/9	0/5
	Rainy	3/9	2/5

Frequency Table		Play	
		Yes	No
Temp	Hot	2	2
	Mild	4	2
	Cool	3	1

Likelihood Table		Play	
		Yes	No
Temp	Hot	2/9	2/5
	Mild	4/9	2/5
	Cool	3/9	1/5

Frequency Table		Play	
		Yes	No
Humidity	High	3	4
	Normal	6	1

Likelihood Table		Play	
		Yes	No
Humidity	High	3/9	4/5
	Normal	6/9	1/5

Frequency Table		Play	
		Yes	No
Windy	True	3	3
	False	6	2

Likelihood Table		Play	
		Yes	No
Windy	True	3/9	3/5
	False	6/9	2/5

Likely hood

Posterior Probability

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

Class Prior Probability

Predictor Prior Probability

Likelihood Table		Play	
		Yes	No
Outlook	Sunny	2/9	3/5
	Overcast	4/9	0/5
	Rainy	3/9	2/5

Likelihood Table		Play	
		Yes	No
Temp	Hot	2/9	2/5
	Mild	4/9	2/5
	Cool	3/9	1/5

Likelihood Table		Play	
		Yes	No
Humidity	High	3/9	4/5
	Normal	6/9	1/5

Likelihood Table		Play	
		Yes	No
Windy	True	3/9	3/5
	False	6/9	2/5

Likely hood

Posterior Probability

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

Class Prior Probability

Predictor Prior Probability

$$P(Yes|Sunny) = \frac{P(Sunny|Yes) \times P(Yes)}{P(Sunny)}$$

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

Let’s predict probability of Play = Yes on the a certain day when the weather conditions are like

Outlook = Sunny, Temp = Cool, Humidity = High, Windy = Yes

$$P(A|B) = P(Yes|Sunny, cool, Humidity = High, Windy = Yes)$$

Let’s predict probability of Play = Yes on the a certain day when the weather conditions are like
Outlook = Sunny, Temp = Cool, Humidity = High, Windy = Yes

Likely hood

Posterior Probability $P(A|B) = \dfrac{P(B|A) \times P(A)}{P(B)}$ Class Prior Probability

Predictor Prior Probability

$$P(A|B) = P(Yes|Sunny, cool, Humidity = High, Windy = Yes)$$

$$P(Yes|Sunny, cool, Humidity = High, Windy = Yes) = \frac{P(Sunny|Yes) \cdot P(Cool|Yes) \cdot P(Humidity = High|Yes) \cdot P(Windy = Yes|Yes) \cdot P(Yes)}{P(Sunny)P(Cool)P(Humidity = High)P(Windy = Yes)}$$

$$P(Yes|Sunny, cool, Humidity = High, Windy = Yes) = \frac{\frac{2}{9} \cdot \frac{3}{9} \cdot \frac{3}{9} \cdot \frac{3}{9} \cdot \frac{9}{14}}{\frac{5}{14}\frac{4}{14}\frac{7}{14}\frac{6}{14}} = 0.2419$$

Where the Naïve Bayes is used?

- News Classification
- Document Classification
- Email Spam Classifier
- Face Recognition
- Sentiment Analysis
- Medical Diagnosis
- Weather Predication

Types of Naïve Bayes

- Gaussian Naïve Bayes
- Multinomial Naïve Bayes
- Bernoulli Naïve Bayes

Gaussian Naïve Bayes

- Used when the features are continuous
- The data assumed to have gaussian distribution / normal distribution

Multinomial Naïve Bayes

- The data is discreate
- We have text classification problem (NLP)
- When the TF IDF Model is used, we use Multinomial Naïve Bayes
- TF >> Term Frequency , IDF >> Inverse Document Frequency

Bernoulli Naïve Bayes

- Used when feature vectors are binary
- Bag of words model where the features are ones and zeros

Discrete Data	Continuous Data
The type of data that has clear spaces between values is discrete data.	Continuous information is information that falls into a continuous series.
Discrete data is countable.	Continuous data is measurable
There are distinct or different values in discrete data.	Every value within a range is included in continuous data.
The bar graph is used to graphically represent discrete data.	A histogram is used to graphically represent continuous data.
Ungrouped frequency distribution of discrete data is performed against a single value	Grouped distribution of continuous data tabulation frequencies is performed against a value group.
Points in a graph of the discrete function remain unconnected.	The points are associated with an unbroken line.