Naive Bayes Classifier

It is a classification technique based on Bayes' Theorem with an independence assumption among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

The Naïve Bayes classifier is a popular supervised machine learning algorithm used for classification tasks such as text classification. It belongs to the family of generative learning algorithms, which means that it models the distribution of inputs for a given class or category. This approach is based on the assumption that the features of the input data are conditionally independent given the class, allowing the algorithm to make predictions quickly and accurately.

In statistics, naive Bayes classifiers are considered as simple probabilistic classifiers that apply Bayes' theorem. This theorem is based on the probability of a hypothesis, given the data and some prior knowledge. The naive Bayes classifier assumes that all features in the input data are independent of each other, which is often not true in real-world scenarios. However, despite this simplifying assumption, the naive Bayes classifier is widely used because of its efficiency and good performance in many real-world applications.

Bayes Theorem

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

The above equation represents Bayes Theorem in which it describes the probability of an event occurring P(A) based on our prior knowledge of events that may be related to that event P(B).

Lets explore the parts of Bayes Theorem:

- P(A|B) - **Posterior Probability**
  - The conditional probability that event A occurs given that event B has occurred.
- P(A) - **Prior Probability**
  - The probability of event A.
- P(B) - **Evidence**
  - The probability of event B.
- P(B|A) - **Likelihood**
  - The conditional probability of B occurring given event A has occurred.

How Do Naive Bayes Algorithms Work?

Let's understand it using an example. Below I have a training data set of weather and corresponding target variable 'Play' (suggesting possibilities of playing). Now, we need to classify whether players will play or not based on weather condition. Let's follow the below steps to perform it.

1. **Convert the data set into a frequency table**

   In this first step data set is converted into a frequency table

2. **Create Likelihood table by finding the probabilities**

   Create Likelihood table by finding the probabilities like Overcast probability = 0.29 and probability of playing is 0.64.

Weather	Play
Sunny	No
Overcast	Yes
Rainy	Yes
Sunny	Yes
Sunny	Yes
Overcast	Yes
Rainy	No
Rainy	No
Sunny	Yes
Rainy	Yes
Sunny	No
Overcast	Yes
Overcast	Yes
Rainy	No

Frequency Table		
Weather	No	Yes
Overcast		4
Rainy	3	2
Sunny	2	3
Grand Total	5	9

Likelihood table				
Weather	No	Yes		
Overcast		4	=4/14	0.29
Rainy	3	2	=5/14	0.36
Sunny	2	3	=5/14	0.36
All	5	9		
	=5/14	=9/14		
	0.36	0.64		

3. **Use Naive Bayesian equation to calculate the posterior probability**

   Now, use Naive Bayesian equation to calculate the posterior probability for each class. The class with the highest posterior probability is the outcome of the prediction.

Problem: Players will play if the weather is sunny. Is this statement correct?

We can solve it using the above-discussed method of posterior probability.

P(Yes | Sunny) = P(Sunny | Yes) * P(Yes) / P (Sunny)

Here P(Sunny | Yes) * P(Yes) is in the numerator, and P (Sunny) is in the denominator.

Here we have P (Sunny |Yes) = 3/9 = 0.33, P(Sunny) = 5/14 = 0.36, P(Yes)= 9/14 = 0.64

Now, P (Yes | Sunny) = 0.33 * 0.64 / 0.36 = 0.60, which has higher probability.

The Naive Bayes uses a similar method to predict the probability of different class based on various attributes. This algorithm is mostly used in text classification (nlp) and with problems having multiple classes.

Naive Baye's → Part 2

Baye's Theorem

$$\boxed{P(B/A) = \frac{P(B) * P(A/B)}{P(A)}}$$ → BAYE'S THEOREM

⇓

Supervised ML

↓ CLASSIFICATION

X_1	X_2	X_3	X_4	Y O/P
—	—	—	—	No
—	—	—	—	Yes ✓

$$P(y/\overbrace{x_1,x_2,x_3,x_4 \ldots x_n}^{A}) = \frac{P(y) * P(x_1,x_2,x_3,x_4 \ldots x_n / y)}{P(x_1,x_2,x_3 \ldots x_n)}$$

$$= \frac{P(y) * P(x_1/y) * P(x_2/y) * P(x_3/y) \ldots P(x_n/y)}{\cancel{P(x_1)\, P(x_2)\, P(x_3) \ldots P(x_n)}}$$

Constant →

$$P(N/x_1,x_2,x_3 \ldots x_n) = \frac{P(N) * P(x_1/y) * P(x_2/y) * \cdots P(x_n/y)}{\cancel{P(x_1) * P(x_2) * P(x_3) * \ldots P(x_n)}}$$

Constant →

O/P ↓

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

Outlook P(Sunny/Yes)

	Yes	No	P(Y)	P(N)
Sunny	2	3	2/9	3/5 ✓
Overcast	4	0	4/9	0/5
Rain	3	2	3/9	2/5
	9	5 = 14		

Temperature

	Yes	No	P(Y)	P(N)
Hot	2	2	2/9	2/5
Mild	4	2	4/9	2/5
Cold	3	1	3/9	1/5
	9	5		

PLAY

		P(Y)	P(N)
Yes	9	9/14	5/14
No	5		
	14		

New Data → Test (Sunny , Hot) → ??

$$P(Yes \mid \overset{x_1}{Sunny}, \overset{x_2}{Hot}) = P(Yes) * P(Sunny/Yes) * P(Hot/Yes)$$

$$P(No \mid Sunny, Hot) = P(No) * P(Sunny/No) * P(Hot/No)$$

$$P(Yes/Sunny, Hot) = \frac{\cancel{9}}{\cancel{14}_{2}} * \frac{\cancel{2}^{1}}{9} * \frac{2}{\cancel{9}}$$

$$= \frac{2}{63} = 0.031$$

$$P(No/Sunny, Hot) = P(No) * P(Sunny/No) * P(Hot/No)$$

$$= \cancel{5}/\cancel{14}_{7} * 3/5 * \cancel{2}^{1}/\cancel{5}$$

$$= \frac{3}{35} = \underline{\underline{0.085}}$$

$$P(Yes/Sunny, Hot) = \frac{0.031}{0.031 + 0.085} = 27\%$$

$$P(No/Sunny, Hot) = \frac{0.085}{0.031 + 0.085} = 73\%$$

{ Naive Bayes Theorm } $\longrightarrow$ Sunny, Hot $\xrightarrow{O/p}$ $\underline{\underline{No}}$ $\underline{\underline{0.73}}$

$\underline{\text{Assignment}}$ $\rightarrow$ (Overcast, Mild) $\longrightarrow$ O/p ??