Notes by :- jpwebdevelopers.in

Naive Bayes

Naive Bayes is a supervised learning algorithm based on Bayes' Theorem used for classification tasks.

Naive :- It is called "naive" because it assumes that the features are independent of each other which is rarely the case in real-world data.

Bayes :- It is called "Bayes" because it uses Bayes Theorem to calculate the Probability of a class based on given features.

Bayes' Theorem

- Bayes' theorem is also known as Bayes' Rule or Bayes' law, helps us find the chance of something happening by using what we know.

- It works by using Conditional Probability :-

which means the chance of one thing happening, given that something else has already happened.

Formula

$$P(\text{class} \mid \text{Data}) = \frac{P(\text{Data} \mid \text{class}) \cdot P(\text{class})}{P(\text{Data})}$$

#JPNotes

$P(\text{class} \mid \text{Data})$ → Probability of a class given the data (Posterior)

$P(\text{Data} \mid \text{class})$ → Probability of data given the class (Likelihood)

$P(\text{class})$ → Probability of class (Prior)

$P(\text{Data})$ → Probability of data (evidance)

Working

① Calculate prior Probabilities for each class.

(2) Calculate likelihood of each feature give class.

(3) Use Bayes' Theorem to compute posterior probabilities.

(4) Select the class with the highest posterior probability.

#JPNotes

Types of Naive Bayes

Gaussian

→ Continuous data (assumes normal distribution)

Example

→ Temperature

Multinomial

→ Discrete data (e.g. word counts)

Example

Text classification

Bernouli

→ Binary data (0 or 1 values)

Example

→ Spam or not spam

Notes By:-
JPWebdevelopers.in

Naive Bayes Example:- Should the Player Play on a sunny Day?

Problem: We have a dataset showing different weather conditions and whether the player played or not.

We want to predict if the Player will play when the weather is sunny using Naive Bayes.

#JPNotes

@JPwebdevelopers

Step 1 Dataset Summary

Outlook	Play
Rainy	Yes
Sunny	Yes
Overcast	Yes
Overcast	Yes
Sunny	No
Rainy	Yes
Sunny	Yes
Overcast	Yes
Rainy	No
Sunny	No

Outlook	Play
Sunny	Yes
Rainy	No
Overcast	Yes
Overcast	Yes

✓The Good Paper

Step2 Frequency Table

Weather	Yes	No
Overcast	5	0
Rainy	2	2
Sunny	3	2
Total	10	4

#JPNotes

Step3 Probabilities

P(Yes) = 10/14 = 0.71
P(No) = 4/14 = 0.29
P(Sunny) = 5/14 = 0.35
P(Sunny | Yes) = 3/10 = 0.3
P(Sunny/No) = 2/4 = 0.5

Step4 Apply Bayes' Theorem

#JPNotes

P(Yes | Sunny):

$$= \frac{P(Sunny \mid Yes) \cdot P(Yes)}{P(Sunny)} = \frac{0.3 \cdot 0.71}{0.35} = 0.60$$

P(No | Sunny):

$$= \frac{P(Sunny/No) \cdot P(No)}{P(Sunny)} = \frac{0.5 \cdot 0.29}{0.35} = 0.41$$

Final Decision

Since P(Yes | Sunny) = 0.60 > 0.41
The player should play on a sunny day.

Applications

→ Spam detection
→ Document classification
→ Student performance prediction
→ Medical diagnosis.

Advantages

→ Simple & fast.
→ Works well with high-dimensional data
→ Good for text-classification.

Disadvantages

→ Assumes feature independence.
→ Not great with correlated features.