





Naive Bayes Classifier: A Quick Overview

The Naive Bayes classifier is one of the most simple yet powerful probabilistic models used for classification tasks!  

 It's based on Bayes' Theorem which gives us the probability of a class given the input features. It assumes that the features are independent (hence “naive”) – a simplifying assumption that works surprisingly well in many scenarios. 

Key Insights:

Formula:

$$P(C|X) = P(C) * P(X|C) / P(X)$$

Where:


$P(C|X)$ is the probability of class C given the features X.


$P(X|C)$ is the likelihood of observing features X given class C.


$P(C)$ is the prior probability of class C.

$P(X)$ is the probability of the features X.


Types of Naive Bayes:


Gaussian Naive Bayes (for continuous data) 

Multinomial Naive Bayes (for discrete data like word counts in NLP) 

Bernoulli Naive Bayes (for binary features, like spam classification) 

Pros:

Fast & simple 

Works well with a large number of features 

Great for high-dimensional data like text classification 


Cons:


The “naive” assumption of feature independence doesn't always hold true 

Doesn't perform well if features are strongly correlated 

Common Applications:

Spam email classification 

Sentiment analysis 

Medical diagnosis 

Document categorization 

Key Takeaway:

Naive Bayes is a go-to algorithm for classification when speed and simplicity are priorities.

Despite its simplistic assumptions, it often works surprisingly well in practice! 

GitHub Code: <https://github.com/NafisAnsari786/Machine-Learning-Algorithms/tree/main/13%20Naive%20Bayes>