Naive Bayes Algorithm in Machine Learning

Simple, powerful, probabilistic classifier

Outline

1. Introduction To Naive Bayes Algorithm?
2. Types Of Naive Bayes Classifiers
3. Pros And Cons Of Naive Bayes Algorithm
4. When To Use And To Avoid Naive Bayes Algorithm
5. Comparison of Naive Bayes Algorithm
6. Applications Of Naive Bayes Algorithms

1. Introduction To Naive Bayes Algorithm

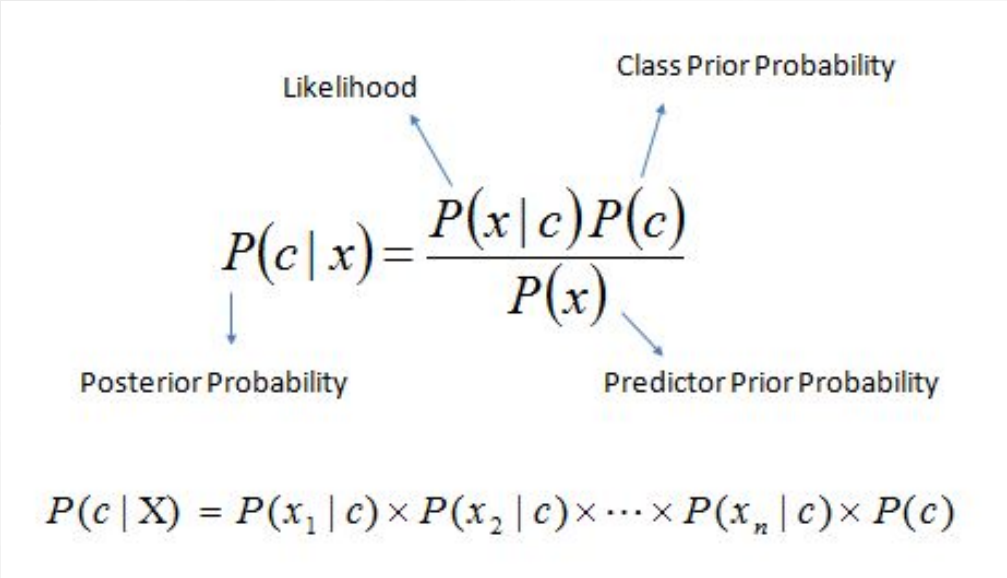
What is Naive Bayes Algorithm

- It is a probabilistic classification algorithm
- It is based on Naive Bayes Theorem
- It is widely used in text classification

What it means to be Naive?

- Naive Bayes assumes that each feature contributes independently to the final classification.
  For example, in spam detection for sentence “**Get free money now**”, it treats the words **"free"** and **"money"** as unrelated even if they often appear together. This means it multiplies their individual probabilities without considering their relationship.

What is the Naive Bayes Theorem



Example

Sentence = “I love this film”;
Positive words: {love, movie, film, amazing}
Negative words: {hate, movie, film, boring}
P(P) = 1/2; P(N) = 1/2;

P(love| Positive) = 1/4; P(film| Positive) = 1/4;
P(love| Negative) = 0; P(film| Negative) = 1/4;

P(Positive | sentence) = P(Positive) * P(love | Positive) * P(film | Positive);
P(Positive | sentence) = 1/2 * 1/4 * 1/4 = 0.03125;

P(Negative | sentence) = P(Negative) * P(love | Negative) * P(film | Negative);
P(Negative | sentence) = 1/2 * 0/4 * 1/4 = 0;

Final Prediction “I love this film” is classified as positive

What is Laplace Smoothing

- Solves zero probability problem
- Add a small constant (usually 1) to counts
- Prevents probabilities from being zero

Example

Sentence = “I love this film”;
Positive words: {love, movie, film, amazing}
Negative words: {hate, movie, film, boring}
P(P) = 1/2; P(N) = 1/2;

P(love| Positive) = 1 + 1/4 + 6 = 0.2; P(film| Positive) = 1 + 1/4 + 6 = 0.2;
P(love| Negative) = 0 + 1/ 4 + 6 = 0.1; P(film| Negative) = 1 + 1/4 + 6 = 0.2;

P(Positive | sentence) = 0.2/0.2 + 0.01 = 0.6667;

P(Negative | sentence) = 0.1/0.2 + 0.01 = 1/3 = 0.33

Final Prediction “I love this film” is classified as positive

2. Types Of Naive Bayes Algorithm

Types of Naive Bayes Algorithm

- Gaussian Naive Bayes
- Multinomial Naive Bayes
- Bemoulli Naive Bayes

Gaussian Naive Bayes

- It is used for continuous values (e.g., age, height)
- It assumes data follows a bell curve
- It’s formula uses mean & variance

Multinomial Naive Bayes

- It is designed for discrete/count data
- Good for text classification
- Works well with word counts in documents

Bernoulli Naive Bayes

- It is designed for binary or boolean features.
- Example: Word present or absent in email
- Similar to logistic regression but simpler

3. Pros And Cons Of Naive Bayes Algorithm

Pros of Naive Bayes Algorithm

- It is Simple & fast
- It performs surprisingly well with text data
- It requires little training data

Cons of Naive Bayes Algorithm

- Assumes independence of features
- Struggles with highly correlated features
- If probability is zero, model may fail (use smoothing)

4. When To Use And To Avoid Naive Bayes Algorithm

When To Use Naive Bayes Algorithm

- When features are mostly independent
- Text classification
- When you need a fast and simple classification model

When To Avoid Naive Bayes Algorithm

- Data with strong feature correlations
- Complex class boundaries
- Complex datasets where accuracy is critical

5. Comparison of Naive Bayes Algorithm

Naive Bayes vs Logistic Regression

- Naive Bayes: probabilistic, assumes independence
- Logistic usually performs better with correlated features
- Naive Bayes: works with small data

Naive Bayes vs Decision Trees

- Naive Bayes: simple, works well with text
- Decision Trees: handles feature interaction better
- Trees may overfit, Naive Bayes is more stable

Naive Bayes vs KNN

- Naive Bayes: probability-based, very fast
- KNN: distance-based, slower for big datasets
- KNN captures complex decision boundaries, Naive Bayes is simpler

5. Application Of Naive Bayes Algorithm

Application of Naive Bayes Algorithm

- Spam detection
- Sentiment analysis
- Document classification
- Recommendation systems

Thank You!