Naives Bayes Algorithm

Naive Bayes is a classification algorithm based on Bayes' Theorem, assuming independence between predictors.

Bayes Theorem- The probability of a hypothesis based on new evidence. It relates the conditional and marginal probabilities of random events.

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

A, B = events

$P(A|B)$ = probability of A given B is true

$P(B|A)$ = probability of B given A is true

$P(A), P(B)$ = the independent probabilities of A and B

Assumption of Independence

- The "naive" aspect refers to the strong (naive) assumption that features are mutually independent given the class label.
- Despite this strong assumption, Naive Bayes performs well in many scenarios, even if the assumption is not entirely true

Types of Naive Bayes Classifiers:

- Gaussian Naive Bayes: Assumes that features follow a normal distribution.
- Multinomial Naive Bayes: Used for discrete counts such as word counts in text classification.
- Bernoulli Naive Bayes: Used for binary/Boolean features.

Training Process

- Calculate the prior probability for each class.
- Calculate the likelihood of each feature given each class.

- Compute the posterior probability for each class using Bayes' theorem.

Prediction

- For a given feature vector, calculate the posterior probability for each class.
- Assign the class with the highest posterior probability to the feature vector.

Categorical Features in Naive Bayes

- **Definition**
  - Features that take on discrete values (e.g., colour: red, green, blue; type: spam, not spam).
  - Naive Bayes can handle categorical features by calculating the frequency of each category given the class label.
- **Handling Categorical Features**
  - **Multinomial Naive Bayes**: Often used for text data where features are word counts or term frequencies.
  - **Bernoulli Naive Bayes**: Used when features are binary (presence or absence of a feature).
- **Example**
  - In a spam detection system, words in emails are categorical features.
  - The classifier calculates the likelihood of each word given that the email is spam or not spam.

Continuous Features in Naive Bayes

- **Definition**
  - Features that take on continuous values (e.g., height, weight, temperature).
  - Requires different handling compared to categorical features.
- **Handling Continuous Features**
  - **Gaussian Naive Bayes**: Assumes that the continuous features follow a normal distribution.
  - For each feature, the mean and standard deviation are calculated for each class.
  - The likelihood of a feature value given a class is then calculated using the Gaussian (normal) distribution formula.
- **Example**
  - In a medical diagnosis system, features like age and blood pressure are continuous.
  - The classifier uses the Gaussian distribution to estimate the likelihood of these features given each possible diagnosis.

Example of Categorical Data and Application of Naive Bayes'

Training Data

Fruit	Colour	Shape	Class
1	Red	Round	Apple
2	Yellow	Long	Banana
3	Green	Round	Apple
4	Yellow	Round	Apple
5	Yellow	Long	Banana

- **Applying Naive Bayes**

1. **Calculate Priors:**

   - $P(Apple) = \frac{3}{5} = 0.6$
   - $P(Banana) = \frac{2}{5} = 0.4$

2. **Calculate Likelihoods:**

   - $P(Color = Red|Apple) = \frac{1}{3} = 0.33$
   - $P(Color = Red|Banana) = \frac{0}{2} = 0.0$
   - $P(Color = Yellow|Apple) = \frac{1}{3} = 0.33$
   - $P(Color = Yellow|Banana) = \frac{2}{2} = 1.0$
   - $P(Color = Green|Apple) = \frac{1}{3} = 0.33$
   - $P(Color = Green|Banana) = \frac{0}{2} = 0.0$
   - $P(Shape = Round|Apple) = \frac{2}{3} = 0.67$
   - $P(Shape = Round|Banana) = \frac{0}{2} = 0.0$
   - $P(Shape = Long|Apple) = \frac{1}{3} = 0.33$
   - $P(Shape = Long|Banana) = \frac{2}{2} = 1.0$

3. **Predict New Fruit**:

- Given a new fruit: {Color: Yellow, Shape: Round}

4. **Posterior Calculation**:

- $P(Apple|Color = Yellow, Shape = Round) \propto P(Color = Yellow|Apple) \times P(Shape = Round|Apple) \times P(Apple)$
  - $\propto 0.33 \times 0.67 \times 0.6$
  - $\propto 0.132$
- $P(Banana|Color = Yellow, Shape = Round) \propto P(Color = Yellow|Banana) \times P(Shape = Round|Banana) \times P(Banana)$
  - $\propto 1.0 \times 0.0 \times 0.4$
  - $\propto 0.0$
- Since $P(Apple|Color = Yellow, Shape = Round) > P(Banana|Color = Yellow, Shape = Round)$, classify the new fruit as 'Apple'.

Example of Continuous Data and Application of Naive Bayes

Example: Flower Classification

- **Continuous Data**
  - Features: Petal Length, Petal Width
  - Classes: 'Iris Setosa' and 'Iris Versicolor'

Flower	Petal length	Petal width	Class
1	1.4	0.2	Iris Setosa
2	4.7	1.4	Iris Versicolor
3	1.3	0.2	Iris Setosa
4	4.5	1.5	Iris Versicolor
5	1.5	0.2	Iris Setosa

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$f(x)$ = probability density function
σ = standard deviation
μ = mean

- **Applying Gaussian Naive Bayes**

1. **Calculate Priors**:

   - $P(IrisSetosa) = \frac{3}{5} = 0.6$
   - $P(IrisVersicolor) = \frac{2}{5} = 0.4$

2. **Calculate Likelihoods** (using Gaussian distribution):

   - For 'Iris Setosa':
     - Mean and standard deviation for Petal Length: $\mu = 1.4$, $\sigma = 0.1$
     - Mean and standard deviation for Petal Width: $\mu = 0.2$, $\sigma = 0.0$ (treated as a very small number for calculation)
   - For 'Iris Versicolor':
     - Mean and standard deviation for Petal Length: $\mu = 4.6$, $\sigma = 0.1$
     - Mean and standard deviation for Petal Width: $\mu = 1.45$, $\sigma = 0.05$

3. **Predict New Flower**:
   - Given a new flower: {Petal Length: 1.4, Petal Width: 0.2}
4. **Posterior Calculation**:
   - $P(PetalLength = 1.4|IrisSetosa) = \frac{1}{\sqrt{2\pi(0.1)^2}} \exp\left(-\frac{(1.4-1.4)^2}{2(0.1)^2}\right) = \frac{1}{\sqrt{2\pi(0.1)^2}} \approx 3.989$
   - $P(PetalWidth = 0.2|IrisSetosa) = \frac{1}{\sqrt{2\pi(0.01)^2}} \exp\left(-\frac{(0.2-0.2)^2}{2(0.01)^2}\right) \approx 39.89$
   - $P(IrisSetosa|PetalLength = 1.4, PetalWidth = 0.2) \propto 3.989 \times 39.89 \times 0.6 \approx 95.77$
   - $P(PetalLength = 1.4|IrisVersicolor) = \frac{1}{\sqrt{2\pi(0.1)^2}} \exp\left(-\frac{(1.4-4.6)^2}{2(0.1)^2}\right) \approx 0.0$
   - $P(PetalWidth = 0.2|IrisVersicolor) = \frac{1}{\sqrt{2\pi(0.05)^2}} \exp\left(-\frac{(0.2-1.45)^2}{2(0.05)^2}\right) \approx 0.0$
   - $P(IrisVersicolor|PetalLength = 1.4, PetalWidth = 0.2) \propto 0.0 \times 0.0 \times 0.4 = 0.0$
   - Since $P(IrisSetosa|PetalLength = 1.4, PetalWidth = 0.2) > P(IrisVersicolor|PetalLength = 1.4, PetalWidth = 0.2)$, classify the new flower as 'Iris Setosa'.