UNIT 3
Bayesian Learning

MACHINE LEARNING

CONTENT

- File
- File Access Modes
- File Operations
- Text File Handling
- Binary File Handling
- CSV File Handling
- Excel File Handling

About Probabiliy

- Probability means possibility.
- It is a branch of mathematics that deals with the occurrence of a random event.
- Probability is a measure of the likelihood of an event to occur.

Example

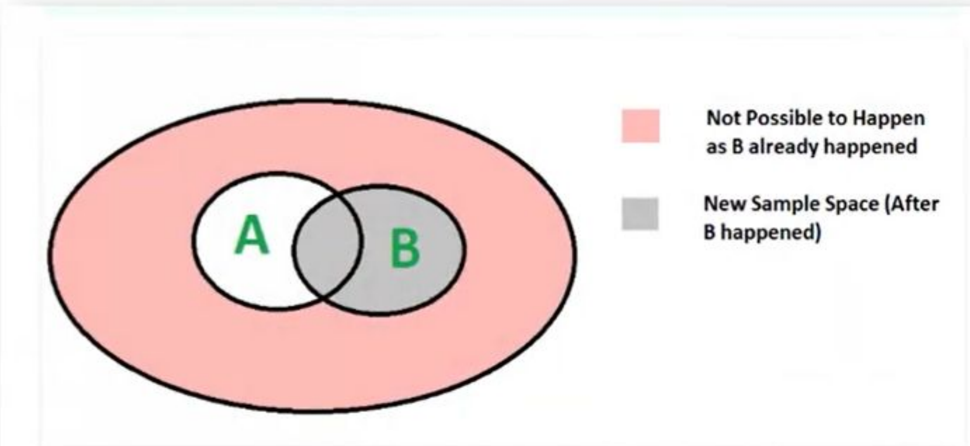
- When we toss a coin, either we get Head OR Tail, only two possible outcomes are possible (H, T).
- But if we toss two coins in the air, there could be three possibilities of events to occur, such as both the coins show heads or both show tails or one shows heads and one tail.
- i.e.(H, H), (H, T),(T, T).

Probability Formula:

$$P(A) = \frac{\text{Number of favorable outcomes to A}}{\text{Total number of outcomes}}$$

Conditional Probability

- The probability of an event occurring given that another event has already occurred is called a **conditional probability**.



Conditional Probability Formula

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

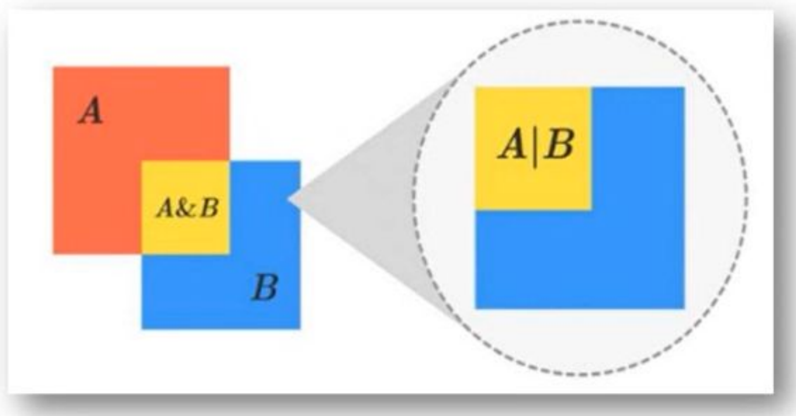
Probability of *A* given *B*

Probability of *A* and *B*

Probability of *B*

How To Find The Conditional Probability From A Word Problem?

- **Step 1:** Write out the Conditional Probability Formula in terms of the problem
- **Step 2:** Substitute in the values and solve.



Conditional Probability Example

- Neha took two tests.
- The probability of her passing both tests is 0.6.
- The probability of her passing the first test is 0.8.
- What is the probability of her passing the second test given that she has passed the first test?

Solution:

$$P(\text{second} \mid \text{first}) = \frac{P(\text{first and second})}{P(\text{first})} = \frac{0.6}{0.8} = 0.75$$

Independent Events

Independent Events

Events can be "Independent", meaning each event is **not affected** by any other events.

Example: Tossing a coin.

Each toss of a coin is a perfect isolated thing.

What it did in the past will not affect the current toss.

The chance is simply 1-in-2, or 50%, just like ANY toss of the coin.

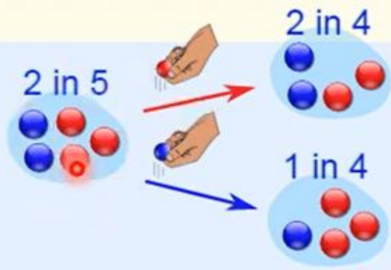
So each toss is an **Independent Event**.

Dependent Events

Dependent Events

But events can also be "dependent" ... which means they **can be affected by previous events** ...

Example: Marbles in a Bag



2 blue and 3 red marbles are in a bag.

What are the chances of getting a blue marble?

The chance is **2 in 5**

But after taking one out the chances change!

So the next time:

if we got a **red** marble before, then the chance of a blue marble next is **2 in 4**

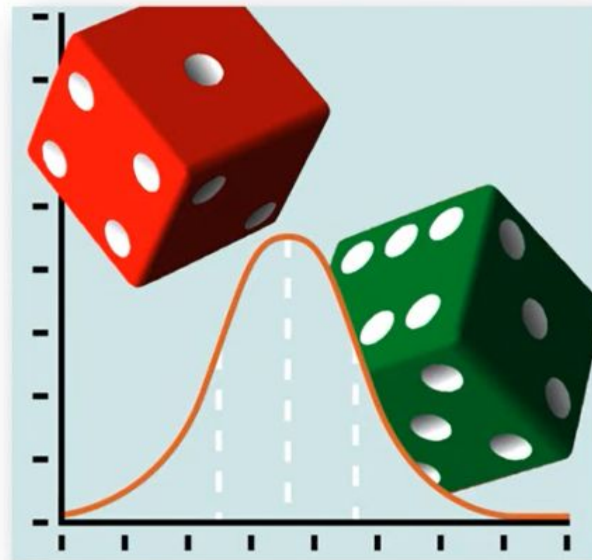
if we got a **blue** marble before, then the chance of a blue marble next is **1 in 4**

This is because we are **removing** marbles from the bag.

So the next event **depends on** what happened in the previous event, and is called **dependent**.

Contents

1. About Bayes Theorem
2. Bayes Theorem Formula
3. Example of Bayes Theorem
4. Applications of Bayes Theorem



About Bayes Theorem

- Bayes' theorem is also known as **Bayes' rule, Bayes' law**, or **Bayesian reasoning**, which determines the probability of an event with uncertain knowledge.
- Bayes' theorem was named after the British mathematician **Thomas Bayes**.
- It is a way to **calculate the value of P(B|A) with the knowledge of P(A|B).**
- Bayes' theorem allows updating the probability prediction of an event by observing new information of the real world.

Bayes Theorem Formula

Bayes' Theorem gives the conditional probability of an event A given another event B has occurred

Bayes Theorem

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

where:

$P(A|B)$ = Conditional Probability of A given B

$P(B|A)$ = Conditional Probability of B given A

$P(A)$ = Probability of event A

$P(B)$ = Probability of event A

Likelihood
Probability of collecting this data when our hypothesis is true

Prior
The probability of the hypothesis being true before collecting data

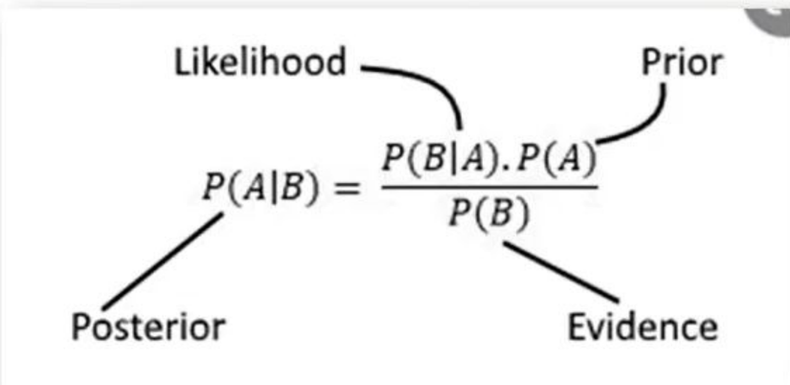
$$P(H|D) = \frac{P(D|H)\,P(H)}{P(D)}$$

Posterior
The probability of our hypothesis being true given the data collected

Marginal
What is the probability of collecting this data under all possible hypotheses?

About Bayes Theorem Formula

- **P(A|B)** is known as **Posterior**, which we need to calculate, and it will be read as Probability of hypothesis A when we have **occurred an evidence B.**
- **P(B|A)** is called the **Likelihood**, in which we consider that hypothesis is true, then we calculate the **probability of evidence**.
- **P(A)** is called the **Prior Probability**, probability of hypothesis **before considering the evidence**
- **P(B)** is called **Marginal Probability**, **pure probability of an evidence**.



$$\text{Posterior} = \frac{\text{prior x likelihood}}{\text{evidence}}$$

$$P(King|Face) = \frac{P(Face|King)P(King)}{P(Face)}$$

Example of Bayes Theorem

- From the deck of the cards, find the probability of the card being picked is king given that it is the face card.
- This can be represented as : **P(King|Face)**

- There are total 52 cards in a deck (n=52), from which 12 cards are face cards; king, queen and jack with club, diamond, heart and spade
- There is a set of face cards with 12 members.
- There is a subset of king cards with 4 members.
- So 4 face cards out of 12 face cards (4/12) are king cards.

Example of Bayes Theorem

- **Putting all in Bayes' formula:**

$$P(King|Face) = \frac{P(Face|King)P(King)}{P(Face)}$$

$$= \frac{1 * \frac{1}{13}}{\frac{3}{13}}$$

$$= \frac{1}{3}$$

➢ Bayes' Theorem proof:

$$P(King|Face) = P(King \cap Face)/P(Face)$$

So from the image above

$$P(King \cap Face) = 4$$

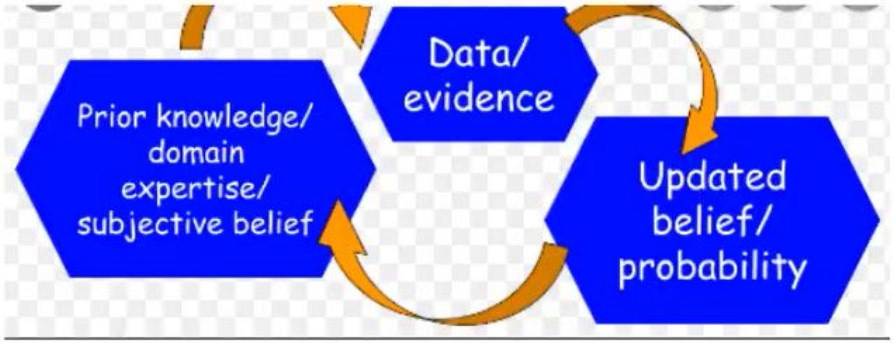
$$P(Face) = 12$$

$$P(King|Face) = 4/12 = 1/3$$

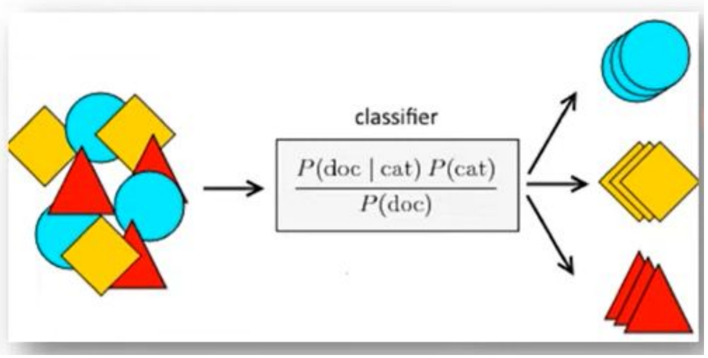
$$\therefore 4/12 = 1/3$$

Applications of Bayes Theorem

1. It is used to calculate the next step of the robot when the already executed step is given.
2. Bayes' theorem is helpful in weather forecasting.
3. Solving logical puzzles & games.

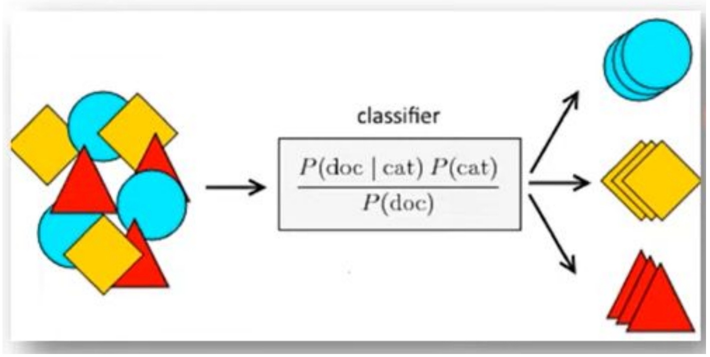


Naïve Bayes Classifier



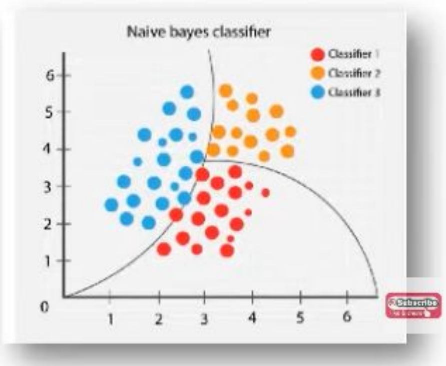
Contents

1. About Naïve Bayes Classifier
2. Why is it called Naïve Bayes?
3. Bayes theorem
4. Algorithm
5. Example
6. Advantages & Disadvantages
7. Applications



About Naïve Bayes Classifier

- Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems.
- It is mainly used in text classification that includes a high-dimensional training dataset.
- Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions.
- It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.
- Some popular examples of Naïve Bayes Algorithm are spam filtration, Sentimental analysis, and classifying articles.



Why is it called Naïve Bayes?

➢**Naïve**:

- It is called **Naïve** because it assumes that the occurrence of a certain feature is independent of the occurrence of other features.
- Such as if the fruit is identified on the bases of color, shape, and taste, then red, spherical, and sweet fruit is recognized as an apple.
- Hence each feature individually contributes to identify that it is an apple without depending on each other.

➢**Bayes**:

- It is called Bayes because it depends on the principle of Bayes' Theorem.

Bayes Theorem

Bayes' Theorem gives the conditional probability of an event A given another event B has occurred

Bayes Theorem

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

where:

P(A|B) = Conditional Probability of A given B
P(B|A) = Conditional Probability of B given A
P(A) = Probability of event A
P(B) = Probability of event A

Likelihood

Prior

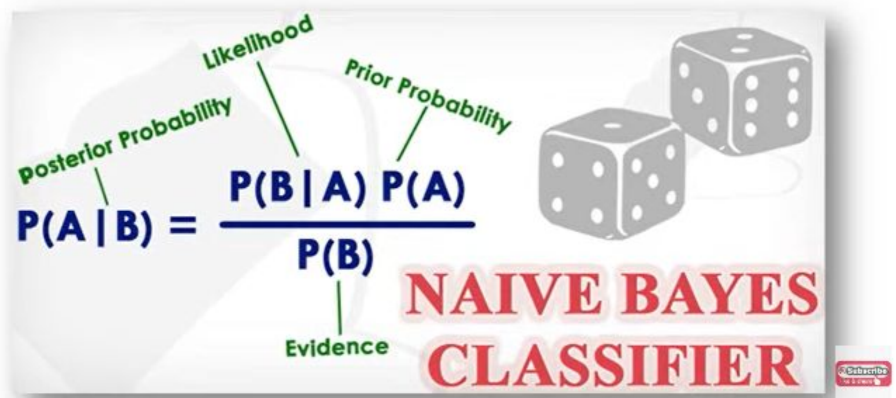
$$P(A|B) = \frac{P(B|A).P(A)}{P(B)}$$

Posterior

Evidence

Algorithm steps according to example

- **Step 1:** Compute the ‘Prior’ probabilities for each of the class of fruits.
- **Step 2:** Compute the probability of evidence that goes in the denominator.
- **Step 3:** Compute the probability of likelihood of evidences that goes in the numerator.
- **Step 4:** Substitute all the 3 equations into the Naive Bayes formula, to get the probability that it is a banana.



Example: Naïve Bayes Theorem

Here, **Training data set** of “Weather” and **Corresponding Target variable** “Play”.

- **Step 1:** Convert the data set into a frequency table.
- **Step 2:** Create Likelihood table by finding the probabilities.
- **Step 3:** Now, use Naive Bayesian equation to calculate the posterior probability for each class. The class with the highest posterior probability is the outcome of prediction.

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		

Problem: If the weather is sunny, then the Player should play or not?

Solution: To solve this, first consider the below dataset:

Applying Bayes'theorem:

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

P(Sunny|Yes)= 3/10= 0.3

P(Sunny)= 0.35

P(Yes)=0.71

So P(Yes|Sunny) = 0.3*0.71/0.35= **0.60**

P(No|Sunny)= P(Sunny|No)*P(No)/P(Sunny)

P(Sunny|NO)= 2/4=0.5

P(No)= 0.29

P(Sunny)= 0.35

So P(No|Sunny)= 0.5*0.29/0.35 = **0.41**

So as we can see from the above calculation that **P(Yes|Sunny)>P(No|Sunny)**

Hence on a Sunny day, Player can play the game.

Advantages & Disadvantages

➢Advantages of Naïve Bayes Classifier:

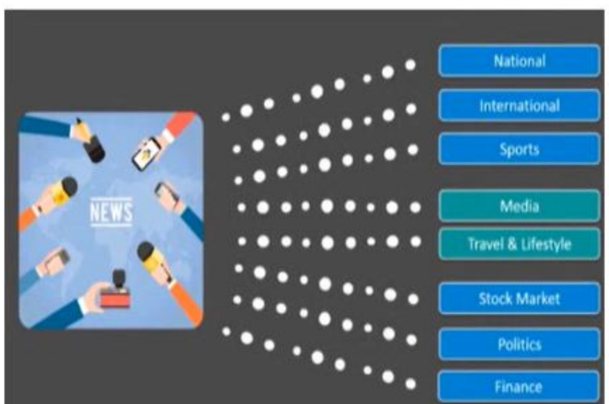
1. Naïve Bayes is one of the fast and easy ML algorithms to predict a class of datasets.
2. It can be used for Binary as well as Multi-class Classifications.
3. It performs well in Multi-class predictions as compared to the other Algorithms.
4. It is the most popular choice for text classification problems.

➢Disadvantages of Naïve Bayes Classifier:

1. Naive Bayes assumes that all features are independent or unrelated, so it cannot learn the relationship between features.

Applications of Naïve Bayes Classifier:

1. It is used for **Credit Scoring**.
2. It is used in **medical data classification**.
3. **Objects & Face** detection.
4. It is used in Text classification such as **Spam filtering** and **Sentiment analysis**.
5. **Recommendation System**



Bayesian Network

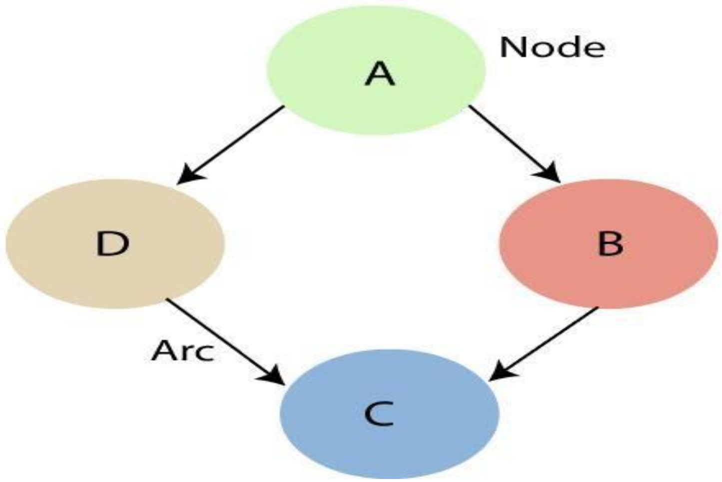
- "A Bayesian network is a probabilistic graphical model which represents a set of variables and their conditional dependencies using a directed acyclic graph."

- It is also called a **Bayes network, belief network, decision network**, or **Bayesian model**.

- Bayesian Network consists of two parts:

1. Directed Acyclic Graph
2. Table of conditional probabilities.

- A Bayesian network graph is made up of nodes and Arcs (directed links), where:



- Each node corresponds to the random variables, and a variable can be continuous or discrete.
- Arc or directedarrows represent the causal relationship or conditional probabilities between random variables

- In the above diagram, A, B, C, and D are random variables represented by the nodes of the network graph.

- If we are considering node B, which is connected with node A by a directed arrow, then node A is called the parent of Node B.

- Node C is independent of node A.

THANK YOU