

# Solving classification problems with naïve bayes

## How does naïve bayes algorithm work?

- Naive Bayes classifier algorithm is based on a famous theorem called “**Bayes theorem**”.
- It can help us find simple yet powerful solutions to many problems ranging from **text analysis** to **spam detection** and much more.

$$P(A|B) =$$

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

# Probability to describe how likely an event is to happen

A value between 0 and 1 represents the possibility of an event happening

0

**Less likely to  
happen**

1

**Most likely to  
happen**

# Bayes theorem is centered on conditional probability

## What is conditional probability?

**Conditional probability** is the probability of an event 'A' happening given that another event 'B' has already happened.

- ▶ The Bayes theorem is **an extension of conditional probability**. It allows us in a sense to use reverse reasoning.

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

CONDITIONAL PROBABILITY

# Understanding the Bayes theorem formula

- **Prior probability  $P(A)$**  –  
The probability of just 'A' occurring
- **Posterior probability  $P(A|B)$**  –  
The probability of event 'A' given  
that event 'B' occurs
- **$P(B|A)$**  - The probability of event B  
happening given that event A has  
occurred
- **$P(B)$**  - The probability of just B

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

BAYES' THEOREM

# What makes Naïve bayes algorithm naïve?

**When the model calculates the conditional probability of one feature given a class,**



...it doesn't take into account the effect of any other feature.



...it assumes that features are independent from each other.



...it gives us the flexibility to describe the probability of each feature.

# The algorithm's naivety has some advantages & limitations

## Advantages

- Quick & simple
- Produce good results with small amount of training data
- Used for benchmarking of a model
- Works well with continuous data by discretizing

## Disadvantages

- In most real-world situations some of the features are likely to be dependent on each other, which might cause wrong results.

# Three types of naïve bayes classifiers in sklearn

## Bernoulli

Used when data is binary like true or false, yes or no etc.

## Multinomial Naïve Bayes

Used when there are discrete values such as number of family members or pages in a book.

## Gaussian Naïve Bayes

Used when all features are continuous variables, like temperature or height.