Machine Learning

Naive Bayes

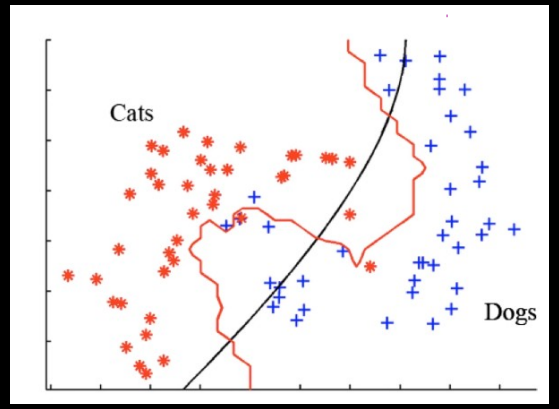
1. Naive Bayes - Classification

- The Naive Bayes algorithm gets its name because it is based on the Bayes Probability Theorem.

- The algorithm aims to calculate the probability that an unknown sample belongs to each of the possible classes. That is, predict the most likely class.

- This type of prediction is called a statistical classification because it is completely based on probability.

- Considering that the effect of the value of an attribute on a given class is independent of the values of the other attributes, which simplifies the calculations involved.



Some examples:

- SPAM detection.
- Automatic email organization.
- Identification of pages with adult content.
- Detection of expressions and feelings.

Classification

Classification consists of the process of finding, through machine learning, a model or function that describes different classes of data.

It is a simple and easy to build model, very useful for large data sets. Superior to other highly sophisticated classification methods, such as artificial neural networks.

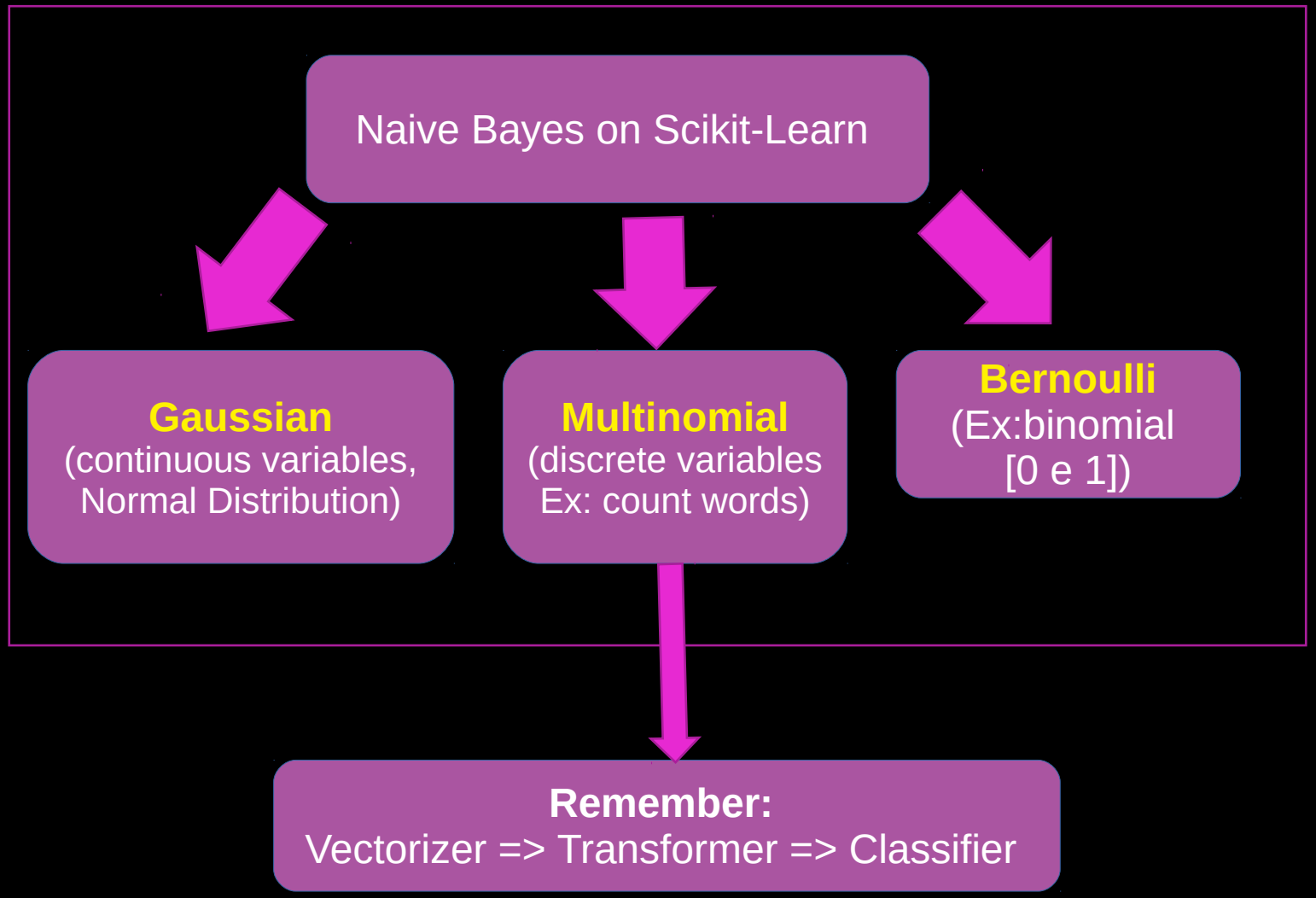
2. Application of the Naive Bayes Algorithm

- Multi-class predictions.

- Text classification / spam filtering / sentiment analysis.

- Real-time forecasts.

Recommendation systems (patterns in the data).



3. The Probability Theory

1 event!

Probability is the study of experiments that, even under very similar conditions, present results that are not possible to predict.

We study **Probability** with the intention of predicting the possibility of the occurrence of a certain situation or fact.

Associated concepts:

Random Experiment
Example: launch a coin that has different sides.

Event
In probability the occurrence of a fact or situation is called an event.

Sample space
Possible results.

Probability ratio
Is given by the possibility of an event taking into account its sample space.

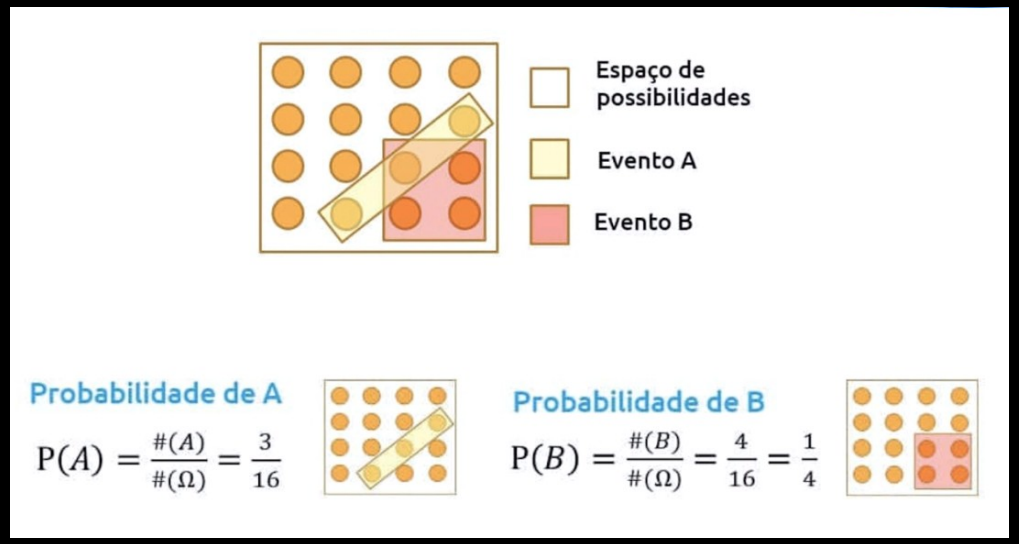
$$\frac{\text{number of elements of the event}}{\text{number of elements of the sample space}}$$

$$P(E) = \frac{n(E)}{n(S)} \qquad 0 \leq P(E) \leq 1$$

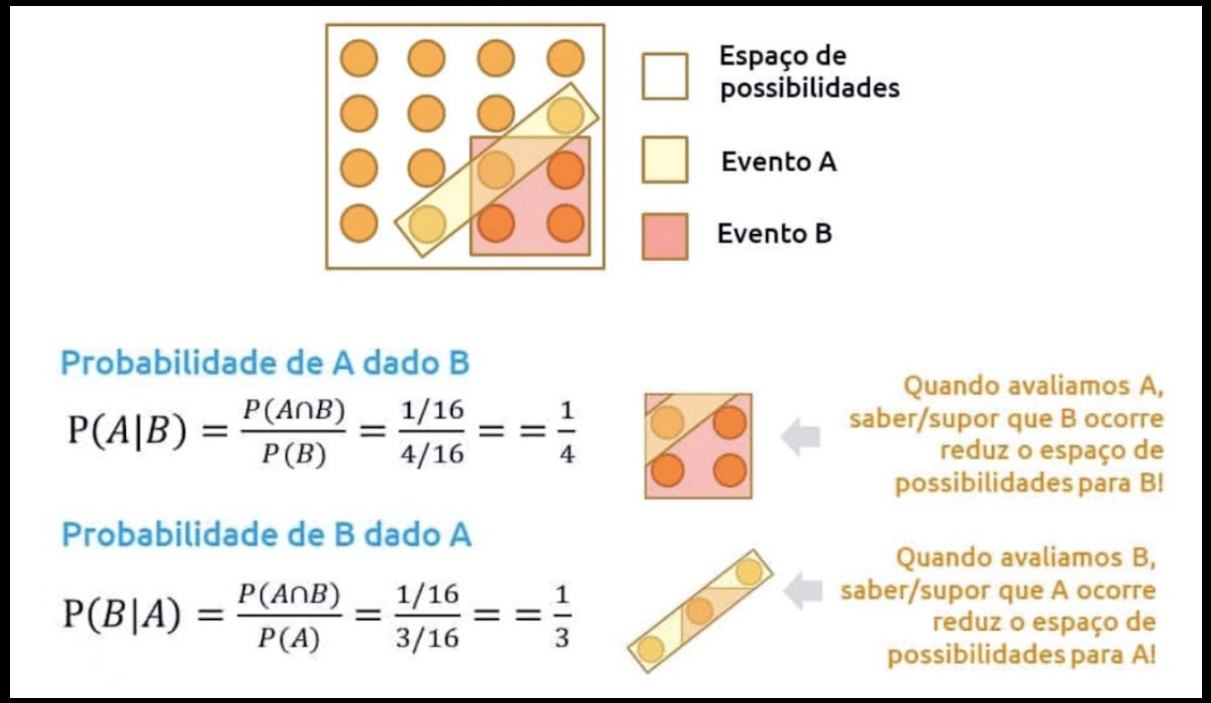
Com $n(S) \neq 0$

4. Conditional Probability

multiple events!



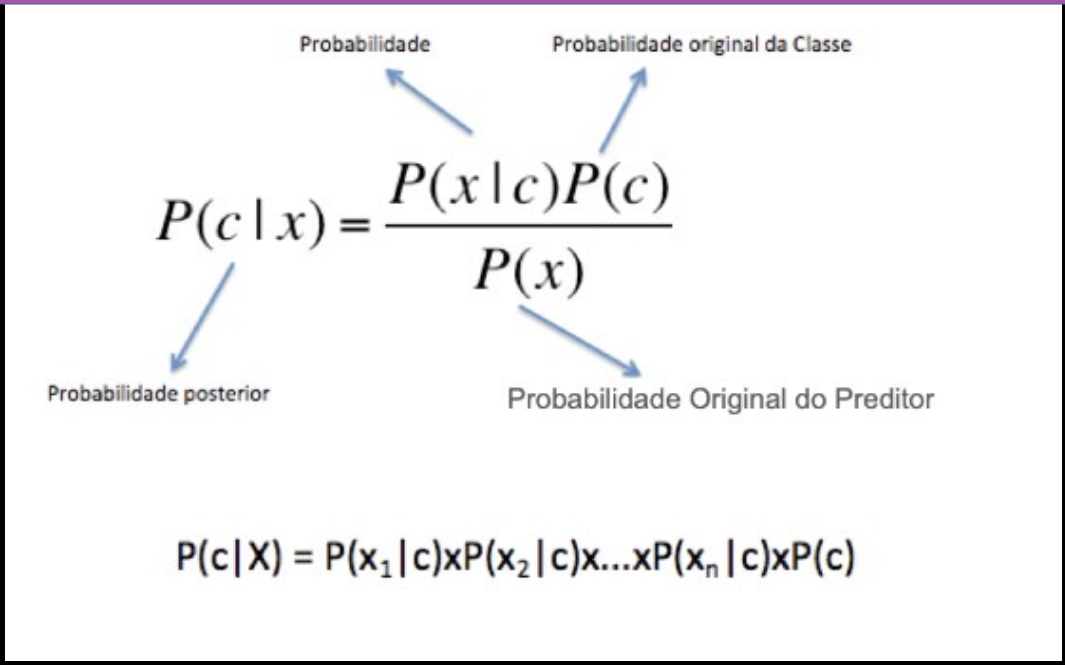
To understand Bayes' theorem!



5. Bayes' theorem

A theorem describing how the conditional probability of each of a set of possible causes for a given observed outcome can be computed from knowledge of the probability of each cause and the conditional probability of the outcome of each cause.

Bayes' rule shows how to change a-priori probabilities taking into account new evidence in order to obtain a pos-teriori probabilities.



6. The Naive Bayes Algorithm

- The **Naive Bayes algorithm** does not have any configuration parameters, so a much **larger treatment** must be done on the data!
- This algorithm is very useful for text classification!

- The naive bayes is a probabilistic model, it needs to receive the probabilities, it calculates the probabilities from the frequency table.
And **tf-idf** allows us to convert the mass of data within a text (frequency of occurrence) into a frequency table.

Naive Bayes on Scikit-Learn

GaussianNB

When dealing with continuous data, a typical assumption is that the continuous values associated with each class are distributed according to a Gaussian distribution (normal distribution).

MultinomialNB

The Naive Bayes Multinomial classifier is suitable for classification with discrete variables (**for example, word counts for text classification)**. Multinomial distribution typically requires counts of entire entities. However, in practice, fractional counts like tf-idf can also work.

BernoulliNB (binomial [0 e 1])

Like MultinomialNB, the BernoulliNB classifier is suitable for discrete data. The difference is that while MultinomialNB works with occurrence counts, BernoulliNB is designed for binary / Boolean resources.