

# Naive Bayes Classifier

## **Group Members**

Zein Samira 610724

Farouk Minaour Kiyini 610725

Samuel Bwambale 610726

# What we have learned

- Bayes' Theorem provides a way that we can calculate the probability of a piece of data belonging to a given class, given our prior knowledge.

$$P(\text{class} | \text{data}) = (P(\text{data} | \text{class}) * P(\text{class})) / P(\text{data})$$

Where  $P(\text{class} | \text{data})$  is the probability of class given the provided data.

# 5 Steps

- **Separate the data by class.** In our case we have classes: Dog, Horse and Elephant with their height and length.
- **Summarize the dataset** by calculating the mean and standard deviation for each column.
- **Summarize the data by class.** We achieved this by calculating the statistics (mean and standard deviation) for each class.
- **Gaussian Probability Function.** Calculating the probabilities is possible if we assume that the data follows the bell curve/Gaussian distribution which can be summarized using the mean and standard deviation.

## 5 Steps cont.

- **Class probabilities.** We use the statistics obtained from the training data to calculate the probability of a piece of data belonging to a given class.

Our dataset contains 10 entries for each class. Each entry contains the height and length of an animal. From this data we were able to correctly predict from the new data if an animal is a **Dog, Horse or Elephant**

Supplied data: [0.58,1.29]

Prediction:

```
[Horse] => 0
```

```
[Dog] => 1
```

```
[Elephant] => 2
```

```
Data=[0.58, 1.29], Predicted: 1
```