Text

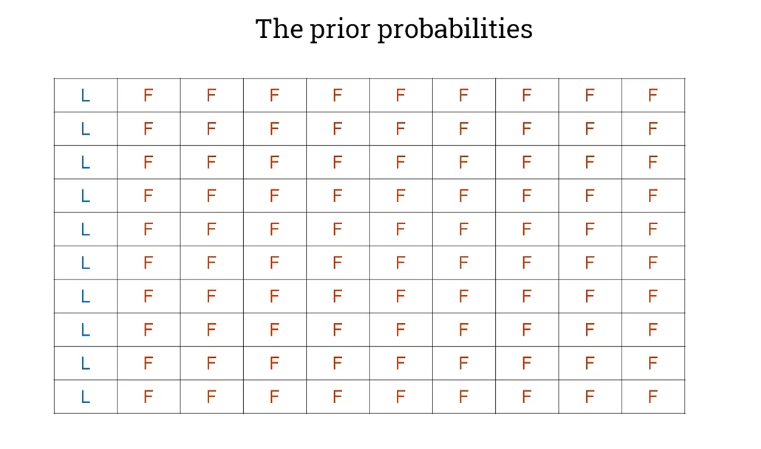
Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Text

Description automatically generated



So let's imagine we have 100 people sampled at random, and you can see, let's say that, taking those statistics that we had on the previous slide, where there's a ratio of nine to one in favor of farmers. So you can see there are 90 cells over here on the right that are farmers and there's ten cells over here that are librarians, those are what we call the prior probabilities.

Table

Description automatically generated

Table

Description automatically generated

So you can see there are 90 cells over here on the right that are farmers and there's ten cells over here that are librarians, those are what we call the prior probabilities. Now let's imagine that we overlay on this our assumption that we will say for the sake of argument that 40% of librarians fit this description, whereas only 10% of farmers do so. That gives us this coloring that we see here, where we're showing 40%. Four out of ten of the librarians fit this description, and one out of ten of the farmers fit this description. So now we can actually work out how likely it is given an individual that picked at random, that fits this description or the cells that are shown colored in here. What is the probability that that individual is a librarian? So we can now mathematically calculate this by looking at the proportion of these cells here that are marked with an L, divided by the proportion of all the cells that are marked in this highlight color. And you can see that it's simply this number here four divided by four, plus the nine that we have here, and that actually gives this value of brown 31 or 31%.

Table

Description automatically generated

We might say our confidence over here that the hypothesis is 50% for probably to the evidence given the hypothesis 50%, I possibly the probability that somebody fits this description is a library is 50% and we might argue that the other percentage over here is 20 rather than 10. But in either case, you can see that you can draw out in terms of areas, the relative proportions and therefore the relative probabilities.

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application, email, Teams

Description automatically generated

## Bayesian classification

Text

Description automatically generated

Table

Description automatically generated

We talked about the probability of a librarian. They were outnumbered nine to one in a general population. You could imagine that for every one librarian here, probability H, we have nine non librarians.

Table

Description automatically generated with medium confidence

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, Teams

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, Teams

Description automatically generated

## Multinomial naive bayes

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

## Simple linear regression

Chart, scatter chart

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Text

Description automatically generated with medium confidence

Chart

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

## Regularization

Chart, line chart

Description automatically generated

Chart

Description automatically generated

A screenshot of a computer

Description automatically generated with medium confidence

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

## Support Vector Machine

Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

A screenshot of a computer

Description automatically generated with medium confidence

Graphical user interface, text, application, email

Description automatically generated

Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated

Graphical user interface, chart

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

## Decision Trees

Diagram

Description automatically generated

Diagram

Description automatically generated

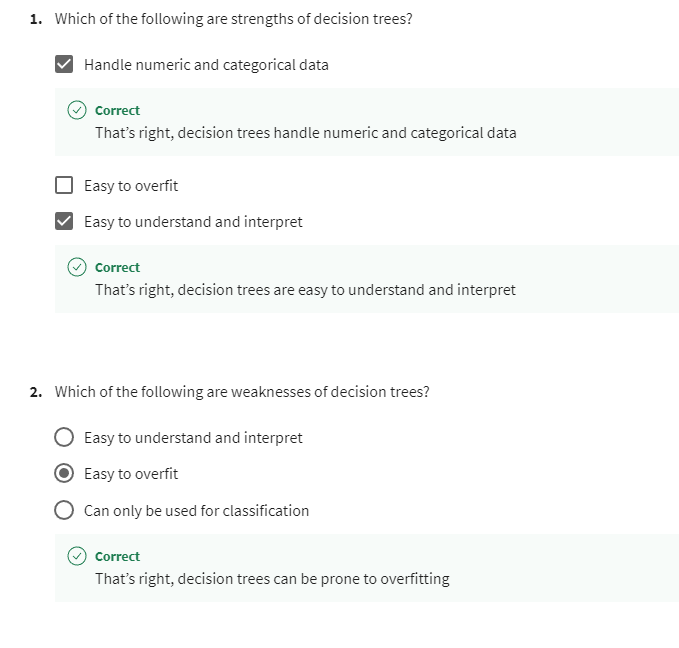
Chart, scatter chart

Description automatically generated

Qr code

Description automatically generated

Lowest entropy => lowest uncertainty



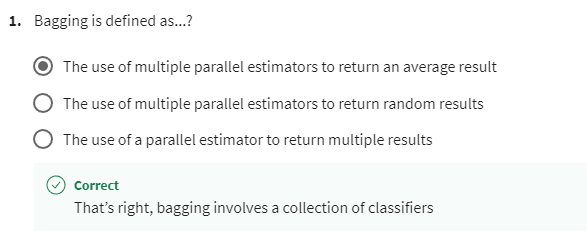
Graphical user interface, text, application, email, Teams

Description automatically generated

## Random forests

Graphical user interface, text, application, email

Description automatically generated



Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

## Introduction to clustering

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, application

Description automatically generated

## Expectation maximisation

Graphical user interface, text, application, email

Description automatically generated

A picture containing calendar

Description automatically generated

