Lesson 3: Bias in, bias out

Experience AI

Raspberry Pi Foundation

What are the three types of machine learning?

1

2

3

What are the three types of machine learning?

Supervised learning

1

Unsupervised learning

2

Reinforcement learning

3

Lesson 3: Bias in, bias out

In this lesson, you will:

- Describe the impact of data on the accuracy of a machine learning (ML) model
- Explain the need for both training and test data
- Explain how bias can influence the predictions generated by an ML model

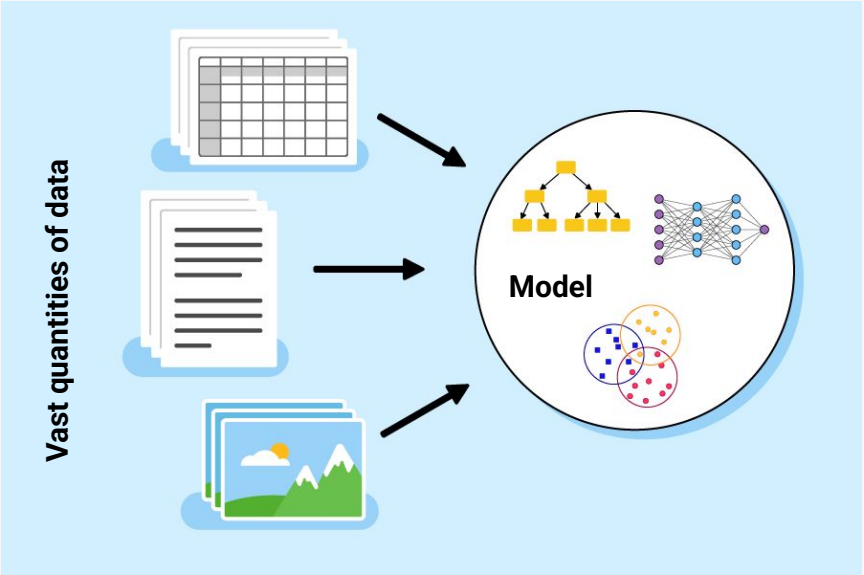
Supermarket AI

A national supermarket wants to save shoppers time at the checkout, and they believe that an AI application can help them achieve this.

They want to use the **cameras** around their stores to recognise the items customers have taken from the shelves and placed into their shopping baskets.

Machine learning models

The supermarket has asked you to create a machine learning **model** to test whether this idea could work.



Recap:

What is a model?

Supermarket model

A **model** is a representation of a real-world context.

The supermarket has asked you to create a model that will only recognise **apples** and **tomatoes**, to see if the idea will work.

What data do you think you will need?

Image data

You are about to train your first machine learning model using images of apples and tomatoes.

You are using **training data** to train your model, so which type of machine learning are you using?

Supervised learning	Unsupervised learning	Reinforcement learning
1	2	3

Image data

You are about to train your first machine learning model using images of apples and tomatoes.

You are using **training data** to train your model, so which type of machine learning are you using?

Supervised learning

1

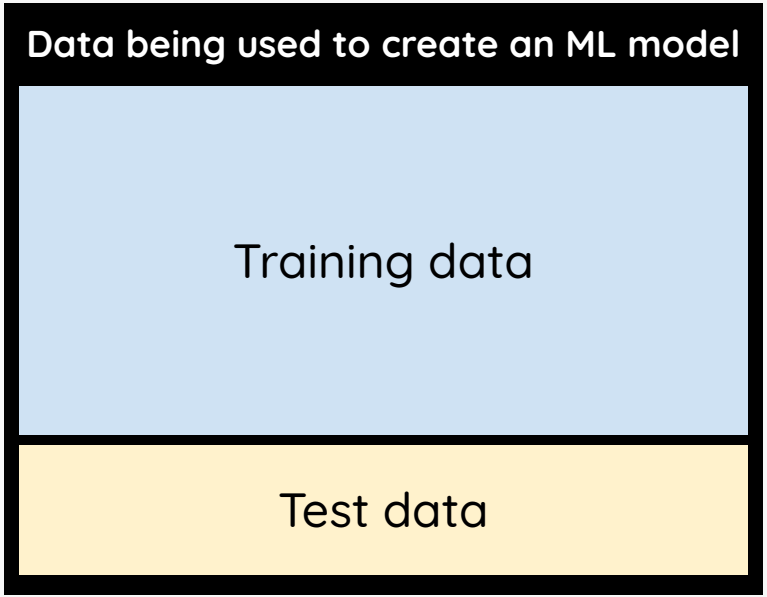
Data and machine learning models

A set of data is used to create an ML model.

The model is **trained** with some of the data,

and then **tested** with the remaining data.

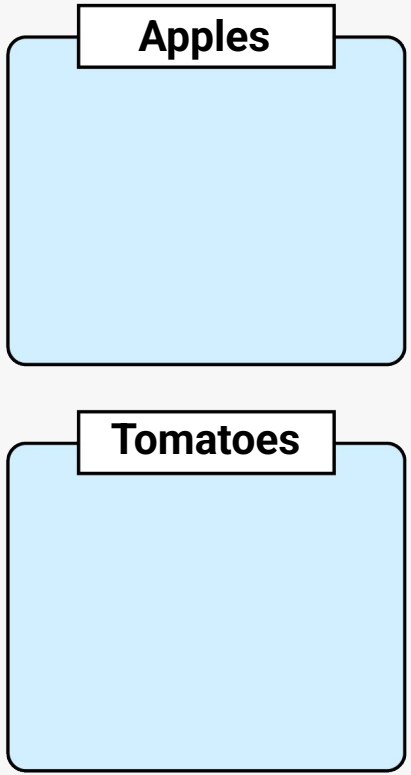
Once the ML model has been trained and tested, it is ready to use in the real world.



Create your classes

Follow the instructions on the worksheet to:

- Create a project that classifies images
- Create labels for **apples** and **tomatoes**
- Add training data
- Train your model

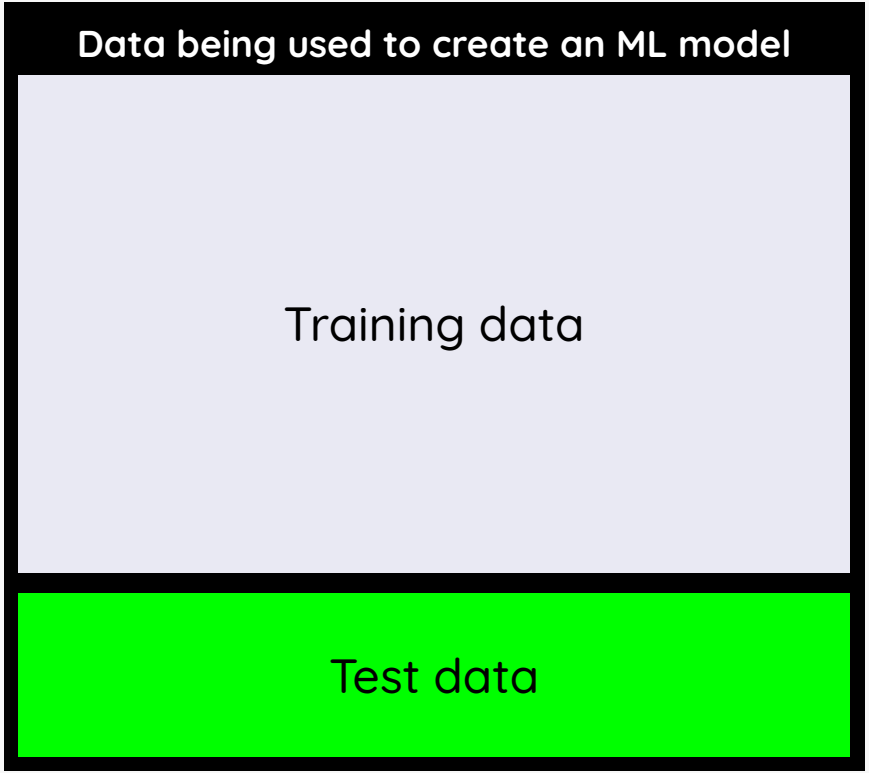


Testing your model

Now that you have trained your model, it is time to **test** it to see how accurate it is.

Some data has been kept aside to use as test data.

Follow the instructions on your worksheet to test your model, and answer the questions in the spaces provided.



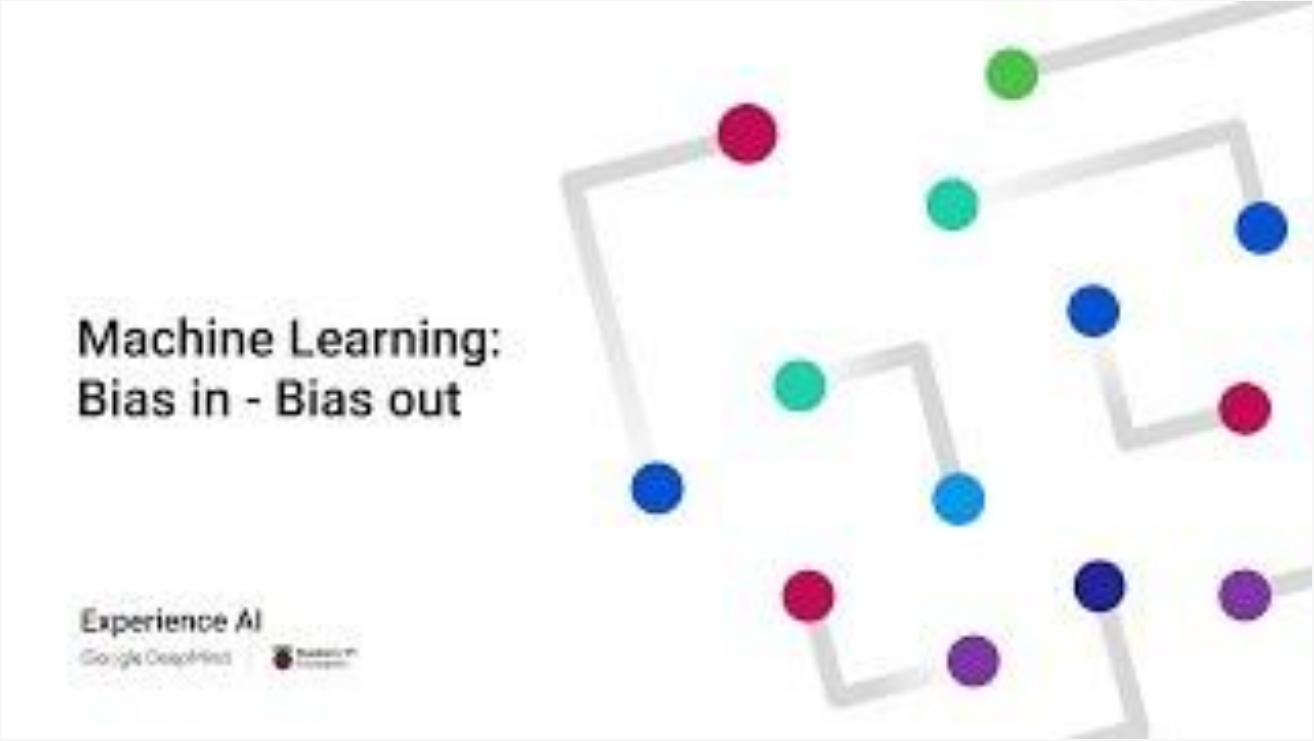
Accuracy

The supermarket has decided that at the moment, the model is not **accurate** enough to consider using.

- Why do you think the supermarket thought the model was not accurate enough?
- What could you do to improve the accuracy of your model?

Think, pair, share

Recognised as **Tomatoes** with 66% confidence



Watch the video on YouTube

Bias

Bias is when the output of a machine learning model favours some things and deprioritises or excludes others.

Examples of machine learning bias:

Computer programmer jobs only being shown to users that a social media platform has identified as male

A facial recognition system that is less accurate in recognising people with certain skin tones

Data bias

Occurs due to data being used that does not accurately represent what is being modelled.

How did data bias appear in your supermarket model?

Questions to consider about the data used to create the model

Was it enough data?

- Were enough image examples of apples and tomatoes used?

Was the ‘right’ data used?

- Only images of green apples were used to create the model.
- Were the images representative of what the fruit would look like in a supermarket?

Societal bias

This is when data being used to train a model reflects bias that exists in society.

Imagine there is an AI application that predicts the job a person might do.

The model has been trained using employment data from 1960.

What societal bias do you think could appear in the data?

Societal bias

In the data from 1960, which gender might have been associated with the following jobs?

- Professional football player
- Teacher
- Nurse
- Astronaut

If gender was in the data used to train the model, could this lead to the AI application making biased predictions?

Data or societal bias?

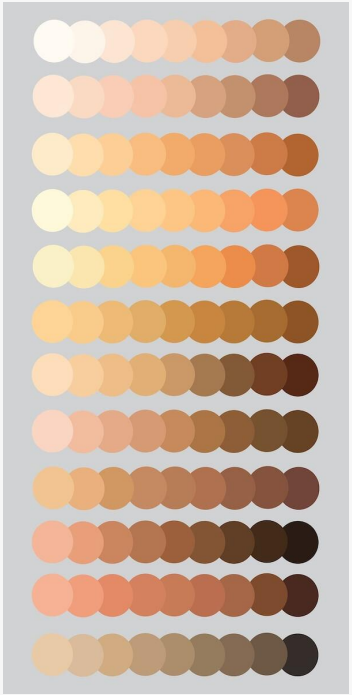
Computer programmer jobs only being shown to users that a social media platform has identified as male

Societal bias

Data or societal bias?

A facial recognition system that is less accurate in recognising people with certain skin tones

Data bias



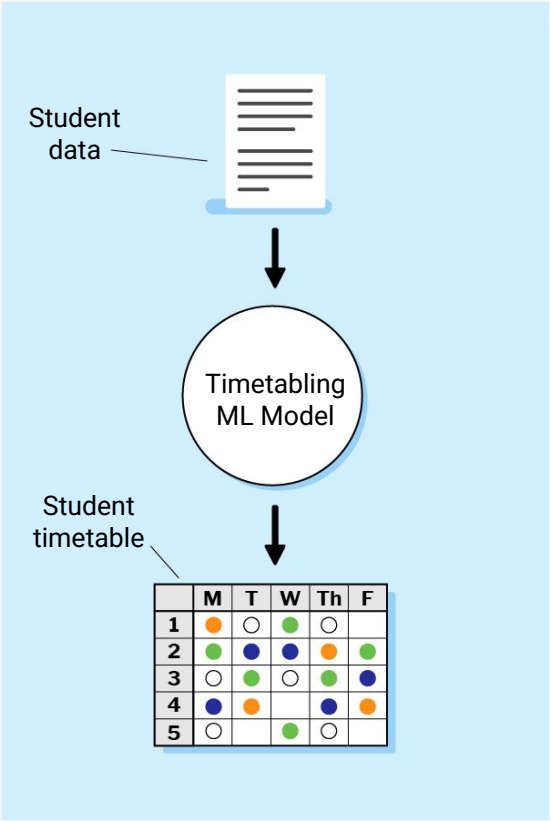
Societal bias can lead to data bias

The subjects you study

Your school is thinking of investing in timetabling software that uses an AI application.

The AI application is used to predict which subjects you will study.

This will save the school money as they will be able to better plan for the future.



Example choices

[Enter student name]

German | Music | Home economics

[Enter student name]

Physics | Biology | Computer science

[Enter student name]

Biology | Geography | Computer science

[Enter student name]

History | Drama | Computer science

[Enter student name]

Spanish | French | Home economics

[Enter student name]

Music | Media | Design and technology

What data was used for the model?

Sets of data

- Data from this school (last 10 years)
- Regional data (last 3 years)
- National data (1960–2000)
- International data

Worksheet questions:

- Which of these data sets would be suitable to use to train the model?
- Which data sets would not be suitable and why?

Meal choices

The school also wants to use an AI application to create a weekly menu for the school cafeteria.

What considerations should you make about collecting the data and training the model?

Reducing bias

It is impossible to avoid there being any bias in a data set. However, there are steps you must take to reduce bias:

- Use large and representative sets of data to train ML models
- Before training the model, separate a proportion of the data that will be used to create the model for testing
- During the process of designing the ML application, consider people with different backgrounds, experiences, and opinions

Next lesson

In this lesson, you...

Used a tool to classify data

Explored the difference between training and test data

Explored how bias can have an impact on the accuracy of an ML model

Next lesson, you will...

Learn how decision trees are used to build classification ML models