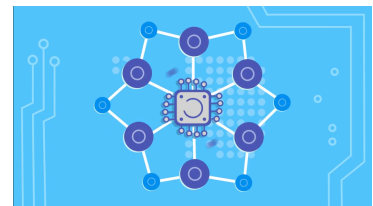# Student Glossary

Accuracy

How often machine learning predictions are correct, usually given as a percentage.

Artificial intelligence (AI)

The design and study of systems that appear to be intelligent.

Artificial intelligence applications

Applications that allow a user to use an AI system to solve a problem.

For example, a smartphone might have a facial recognition application that unlocks the phone when the owner is identified. This application will use a model that allows the application to identify the particular patterns that make up the owner's face.

Artificial intelligence project lifecycle

The AI project lifecycle is the set of steps people follow to design, build, and improve a machine learning model.

Bias

Bias is when you favour one thing over another. Machine learning bias is when the output of a model is more or less accurate for one group of data, which could represent people, objects, or outcomes.

Class

The groups that human developers create when developing a classification model (see below). Training data is put into these classes by developers who apply labels to the data. The model is trained on these classes to identify patterns in the data.
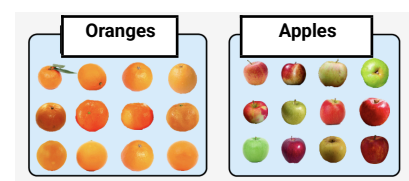
Classification model

A type of supervised learning model (see below) that is created by being trained on a large set of data that has been labelled by humans. Human developers group labelled data into classes. A classification model is created by being trained on the labelled data to identify patterns in each class. This allows the model to

predict what class new data belongs to. For example, developers might label a large number of reviews as positive or negative. A model would then be created by being trained on this data. The model could then produce a prediction about whether a new review is positive or negative.

Computer vision

An area of study of AI that concentrates on identifying data patterns in digital pictures and videos.

A person can use computer vision to identify objects in a picture.

Examples include facial recognition and object recognition systems.

Confidence score

A confidence score shows how likely it is that a predicted label is correct.

The more closely the new data matches data patterns identified in training data, the higher the score.

Prediction:
Tomatoes
Confidence: 68%

Confidence threshold

A percentage that helps us decide which predictions to accept. Predictions with a confidence score below the threshold are ignored.
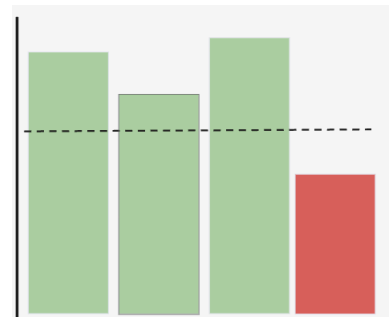
Data

Values that can be stored and used in a computer program. There are many types of data, such as text, pictures, and sound.

Data bias

This appears when a model is trained on data that does not accurately represent what is being modelled.

For example, imagine that an application is created to identify what illness a person has. It uses a model that has mostly been trained on data from adults. Because of this, the application may not be accurate when used to identify a child's illness.

Data cleaning

Before creating a machine learning model, you must identify and correct missing, incorrect, or repeated data in the training set.

Data-driven model

A representation of something in the real world that can be used to solve a problem.

Data-driven models are created by identifying patterns in large amounts of training data.

For example, an AI chatbot application uses a data-driven model, which has been trained using many example conversations, to generate replies to users.
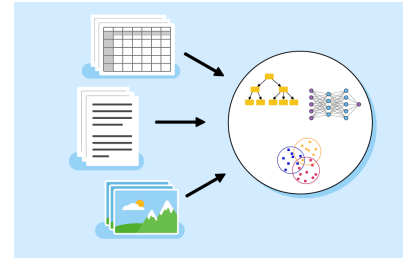
Generative artificial intelligence

An AI system that is designed to generate content, such as text, pictures, or sound. They are trained on large amounts of data and existing content.

Label

When developing a classification model, developers give a label to a single piece of training data to identify which class the data belongs to. Once a model is trained, new data can be provided to the model. The model will produce a label for this data, which is a prediction of which class the data belongs to, and a confidence score.
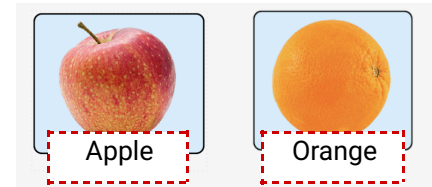
Machine learning

An area of AI where data-driven models produce predictions based on patterns in large amounts of data.

Model card

A way AI developers can help users understand the intended uses and limitations of their AI model.

Prediction

The result produced by an AI system based on patterns in data. AI systems use the patterns to identify what output is most likely.

Predictive artificial intelligence

A type of AI system that is trained on large amounts of past data to predict future outcomes and trends, to classify data, or to make recommendations.

Rule-based system

A system created using a set of instructions or rules to make a decision or solve a problem. A rule-based system does not "learn" (identify patterns) from lots of examples like a data-driven model; instead, it uses rules that human developers have created.

Move 1:
Place an X in a corner

Move 2:
IF the other player did not place an O in the opposite corner
THEN place an X in the opposite corner to move 1
ELSE place an X in a free corner

Societal bias

When the data that is used to train a model contains a bias that exists in society, such as racial or gender bias.

Supervised learning

A way to train machine learning models. Developers use large amounts of data that has been labelled by people to train a machine learning model to produce accurate predictions.

Test data

Data used to calculate the accuracy of a machine learning model. Usually this is data that is removed from the training data before the model is created and then used to test it afterwards.

Data being used to create an ML model

Training data

Test data
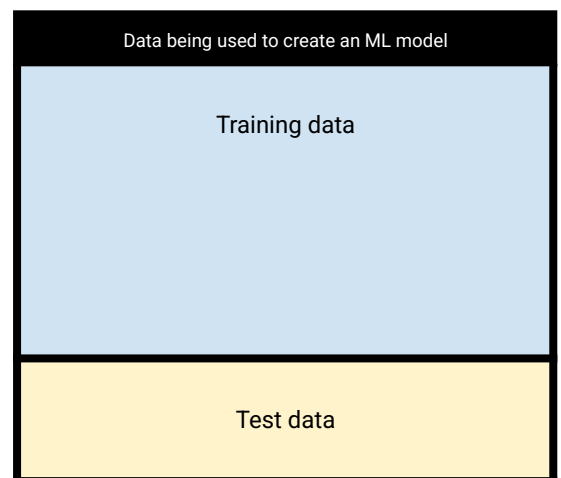
Training

During training, a machine learning method (such as supervised learning) identifies patterns in the data, which the model then uses to make predictions about new data.

User-focused

Making the user's problems, wants, and needs the most important areas when designing a system.