

# Machine Learning



# Introduction to Machine Learning



# Learning Objectives

By the end of this lesson, you will be able to:

- 👁 Explain the basics of machine learning
- 👁 Examine the different types of machine learning
- 👁 Learn about the Python packages used in machine learning
- 👁 Research the Python packages used in machine learning



## Business Scenario

ABC Inc., an e-commerce company, is struggling with a surge in fraudulent transactions on its website. The manual review process for transactions has caused delays in order processing and a negative customer experience.

To tackle this, ABC Inc. will use machine learning algorithms to detect fraudulent transactions in real time, which will be integrated into ABC Inc.'s existing transaction processing system to flag suspicious transactions and prevent fraud.

Additionally, the company will use machine learning algorithms to predict customer behavior based on past purchase history, improving its recommendation engine's performance. With machine learning, ABC Inc. can provide a more personalized customer experience, streamline order processing and increase sales.





# **What Is Machine Learning?**

# Discussion: Machine Learning

Duration: 10 minutes



- What is machine learning?
- In which types of problems machine learning is used?
- What is the difference between machine learning, artificial intelligence, and deep learning?

# Machine Learning

According to IBM: "Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy."



The term machine learning (ML) was coined by Arthur Samuel in 1959.

It allows programs to learn automatically, making computers more intelligent without any human intervention.

The adoption of machine learning has accelerated in recent times.

# Machine Learning

Traditional program approach	Machine learning model approach
Explicitly programmed by human developers using predefined rules and algorithms	Learns from data to make predictions or take actions without being explicitly programmed
Relies on explicitly defined logic and rules to perform tasks	Uses statistical techniques and optimization algorithms to learn patterns and make decisions
Requires manual feature engineering, where relevant features must be identified and extracted by human experts	Automatically learns features from raw data, reducing the need for manual feature engineering
Limited ability to handle complex and unstructured data without significant preprocessing	Can handle complex and unstructured data, such as images, text, and audio, without extensive preprocessing.
Performance depends on the accuracy and completeness of the predefined rules and algorithms	Performance improves with more data and learning iterations, resulting in enhanced accuracy and generalization.

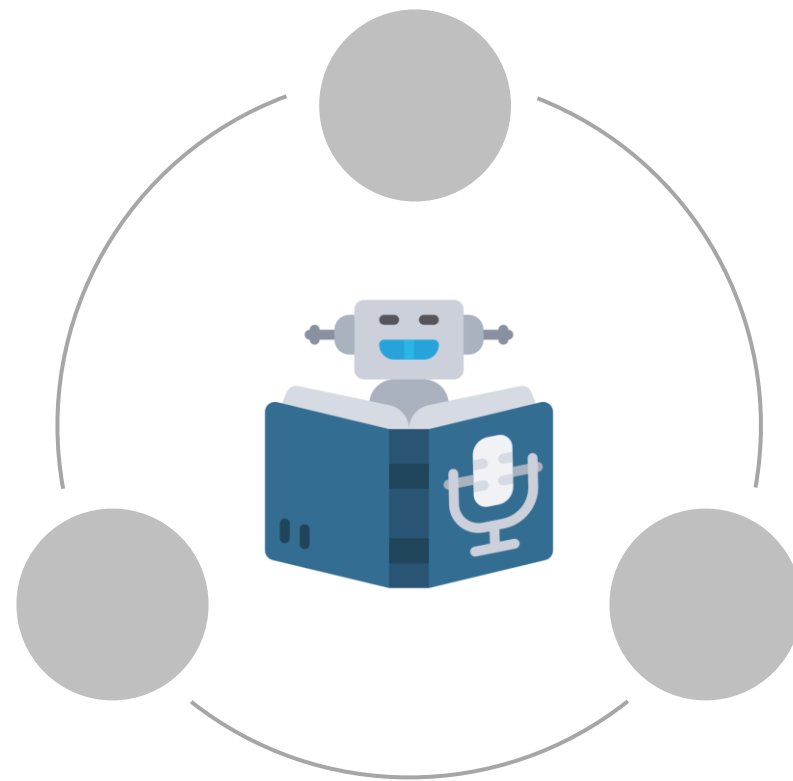


# Machine Learning

The various factors that contributed to its adoption include:

Increased data capture through smart devices, phones and the Internet of Things

Better internet connectivity and bandwidth



Increase of computation power with edge devices

# Comparison Between ML,DL and AI

Some key differences between the three are:

## Machine Learning

- It refers to algorithms that learn and perform based on the data exposed to them.
- **Example:** Amazon Alexa

## Deep Learning

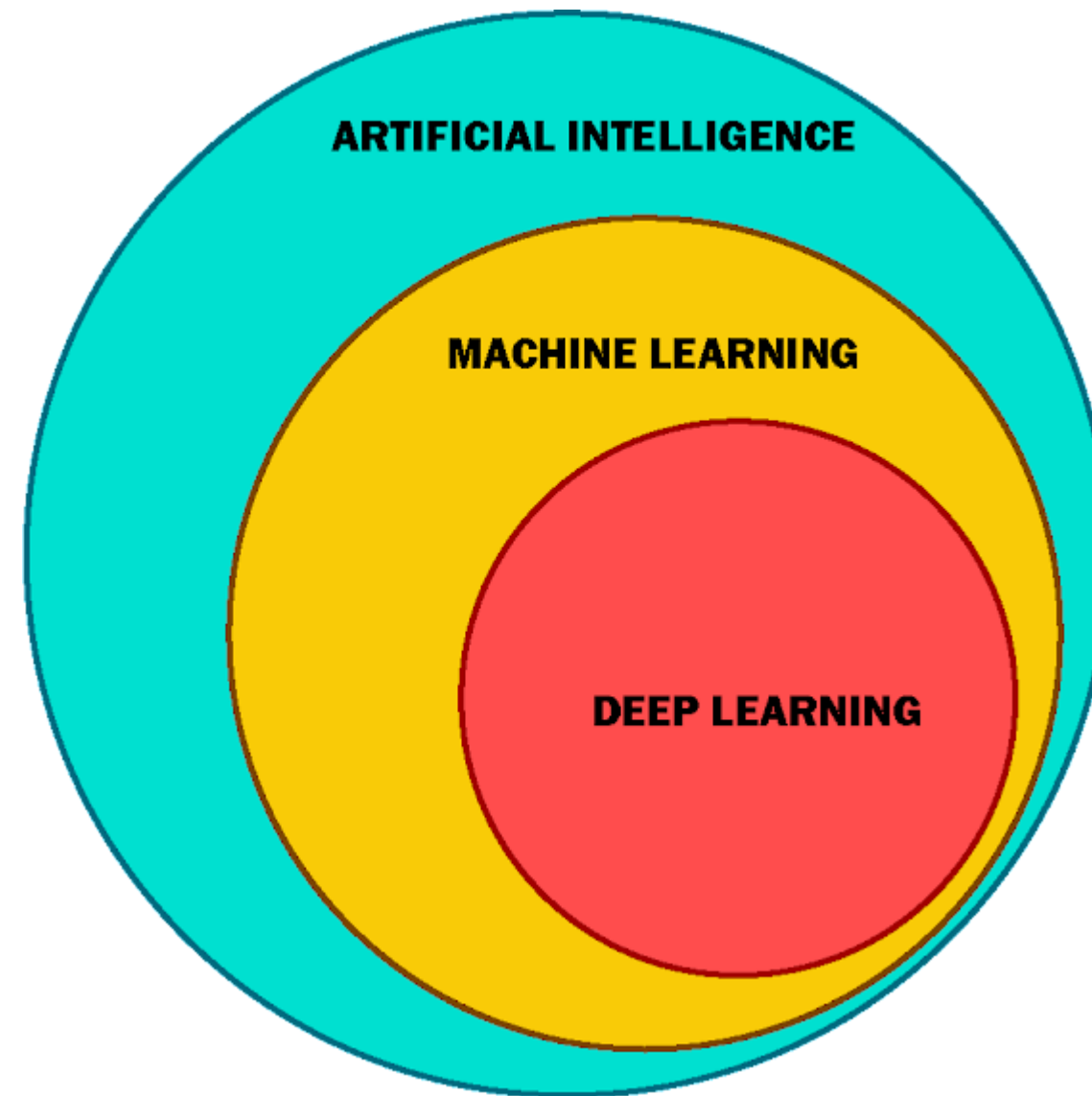
- It refers to layers of neural networks built with machine learning algorithms to solve complex problems.
- **Example:** Spam detection

## Artificial Intelligence

- It leverages different techniques, including ML and DL, to solve problems.
- **Example:** Number plate detection

# Comparison Between ML, DL, and AI

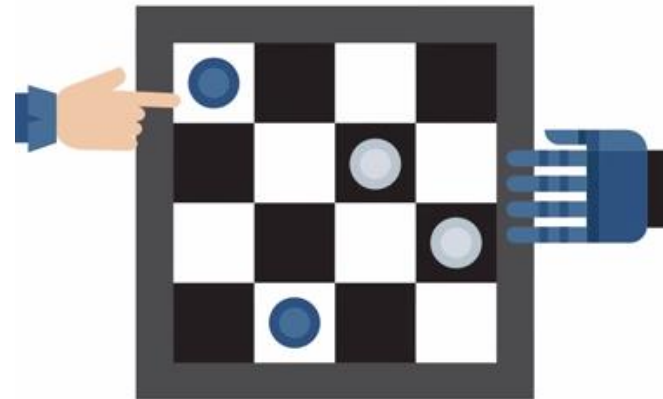
Machine learning (ML), deep learning (DL), and artificial intelligence (AI) are used interchangeably.



# Use Case of Machine Learning

**Example:** When a bot and a human are playing chess, the software program powering the bot requires AI to classify data, predict moves and make decisions.

AI makes the decisions by using a multilayered neural network that uses deep learning.



It uses machine learning to classify the profile of the bot's competitor as beginner, intermediate, or advanced.

The classification is made based on the data parameters used to train this model.

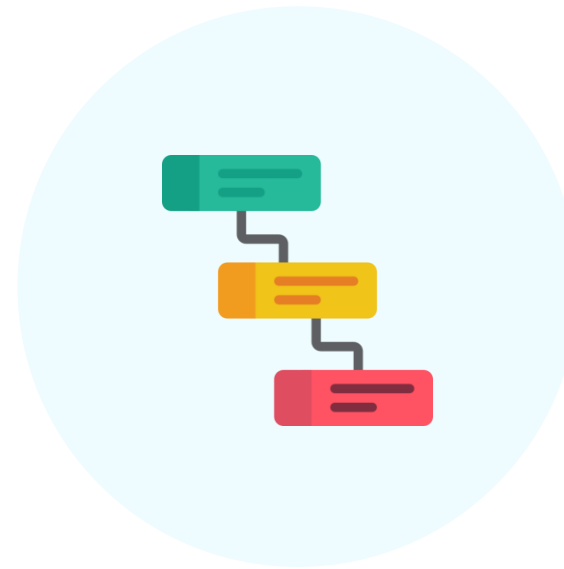
# Machine Learning Algorithms

Algorithms are a set of instructions used to solve problems.

Machine learning algorithms help a software application to:



Predict the  
outcome



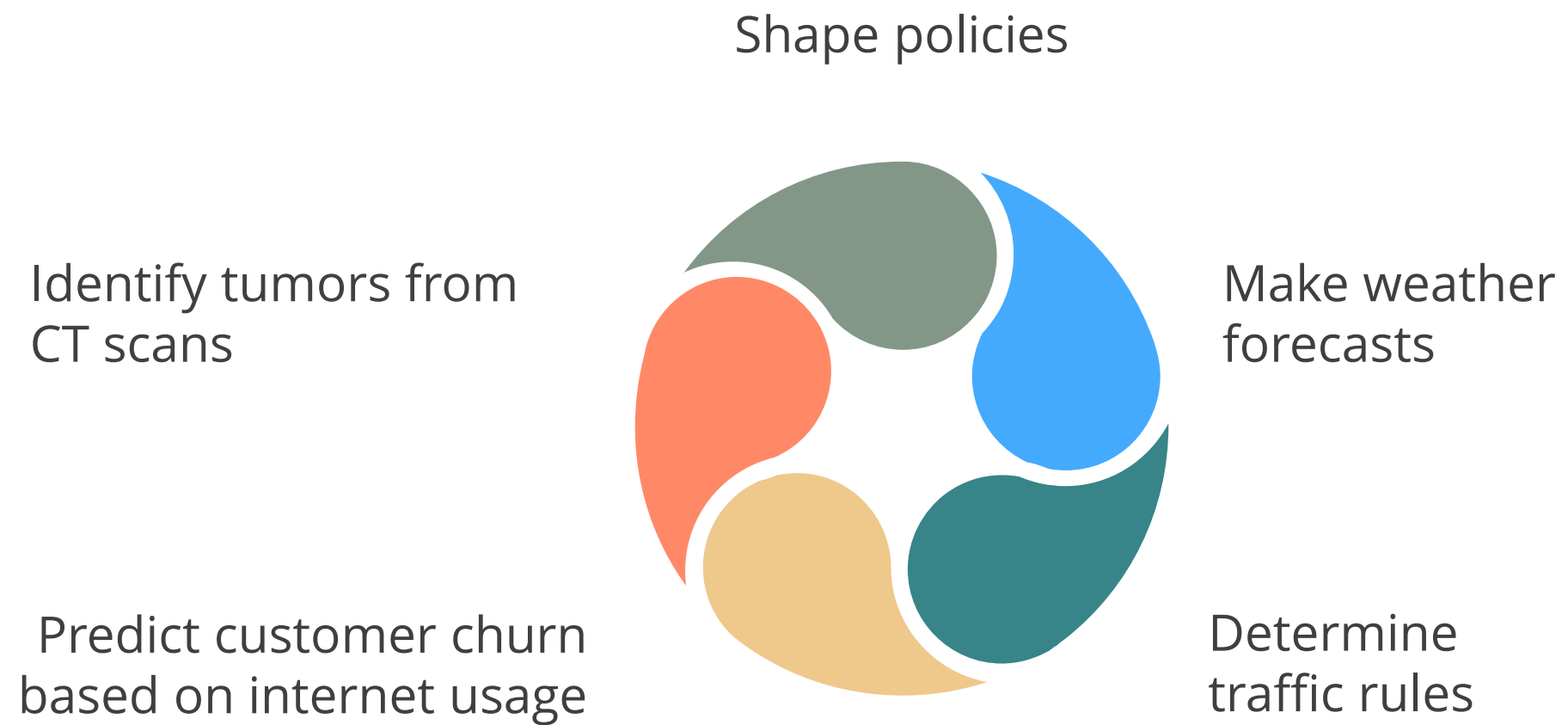
Classify the  
target feature



Improve the  
performance

# Uses of Machine Learning

Some examples of machine learning applications include:



# Data and Machine Learning

The performance of the algorithm is evaluated based on the quality of the data provided.



The chart shows positive correlation between data quality and algorithm performance.

Machine learning is dependent on data.

# Discussion: Machine Learning

Duration: 10 minutes

- What is machine learning?

It is a branch of artificial intelligence in which we deal with data and algorithms to find patterns in data and for the prediction of future data.

- In which types of problems machine learning is used?

Answer: Machine learning is used to:

- Classify different categories of data
  - Predict future sales
  - Forecast the weather
- What is the difference between machine learning, artificial intelligence, and deep learning?
    - In machine learning, the focus lies in studying algorithms and addressing data prediction.
    - Deep learning involves tackling complex machine learning problems with the help of neural networks.
    - Artificial intelligence is about making automated decisions after applying machine learning and deep learning concepts.







# **Types of Machine Learning**

# Discussion: Types of Machine Learning

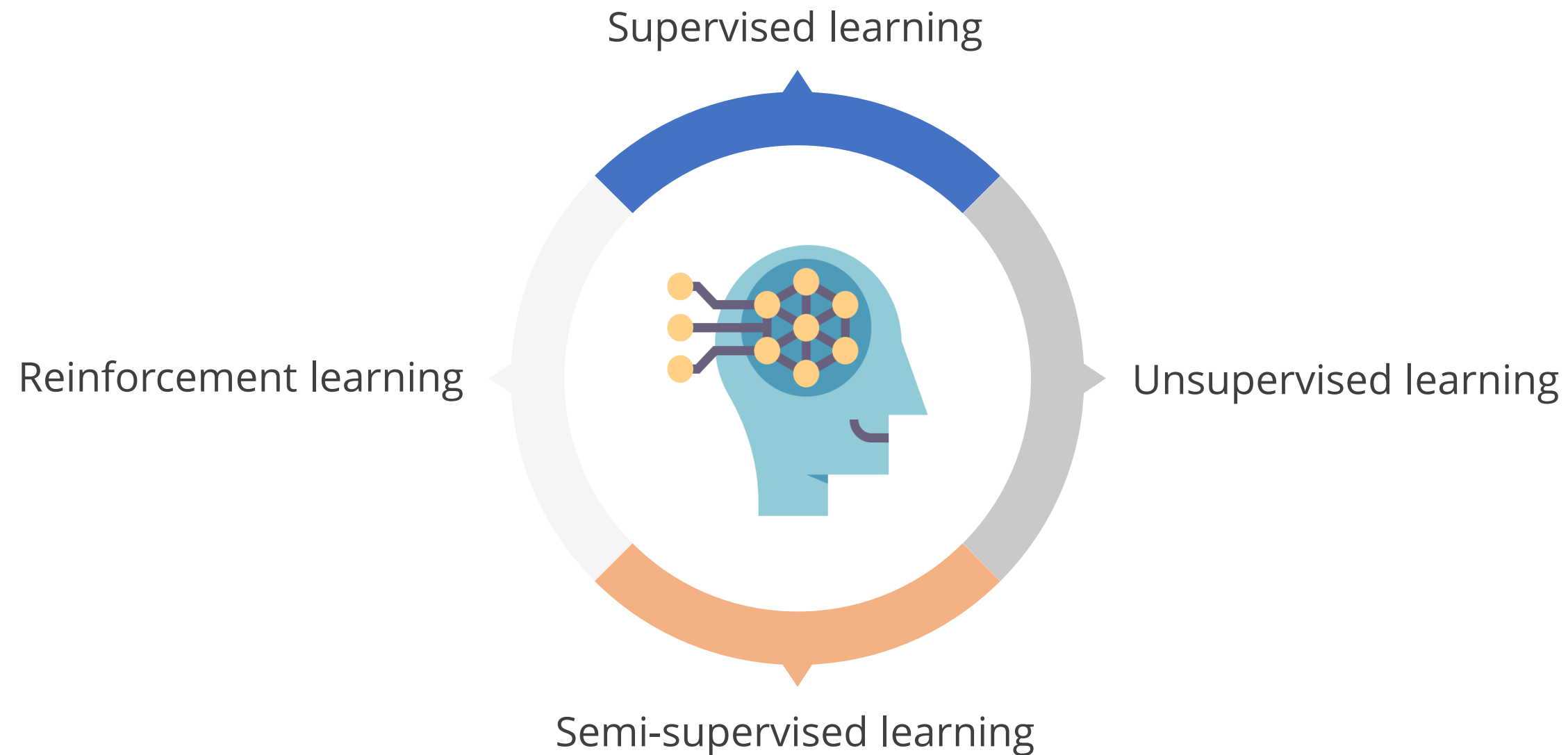
Duration: 10 minutes



- What are the different types of problems that can be solved with machine learning?
- What are the steps involved in applying a machine learning model to a problem?
- How supervised and unsupervised machine learning are different from each other?

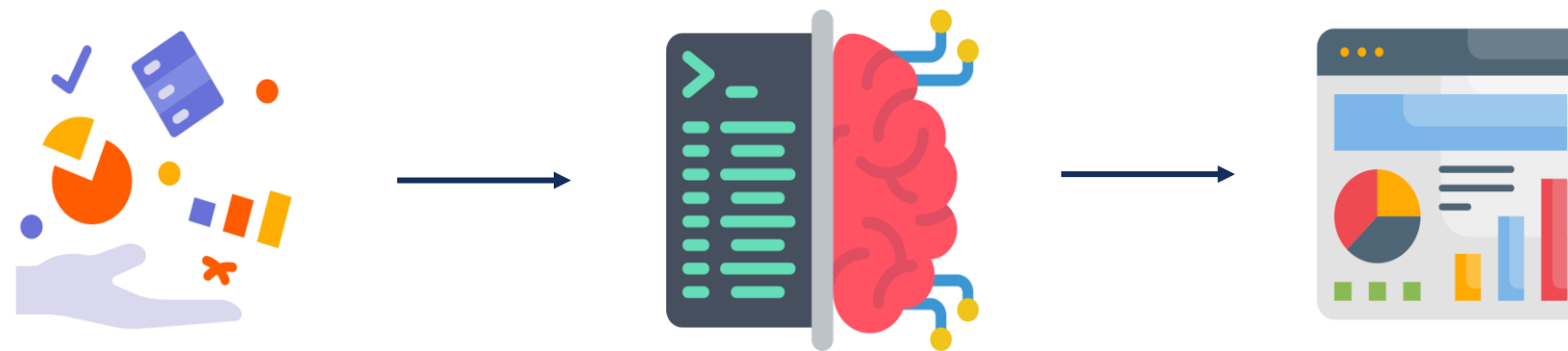
# Types of Machine Learning

There are four main categories of algorithms based on their ability to self-train and predict a condition or identify patterns to derive outcomes. They are:



# Supervised Learning

It is the most common type of machine learning. The computer learns to make predictions based on a set of labeled data.



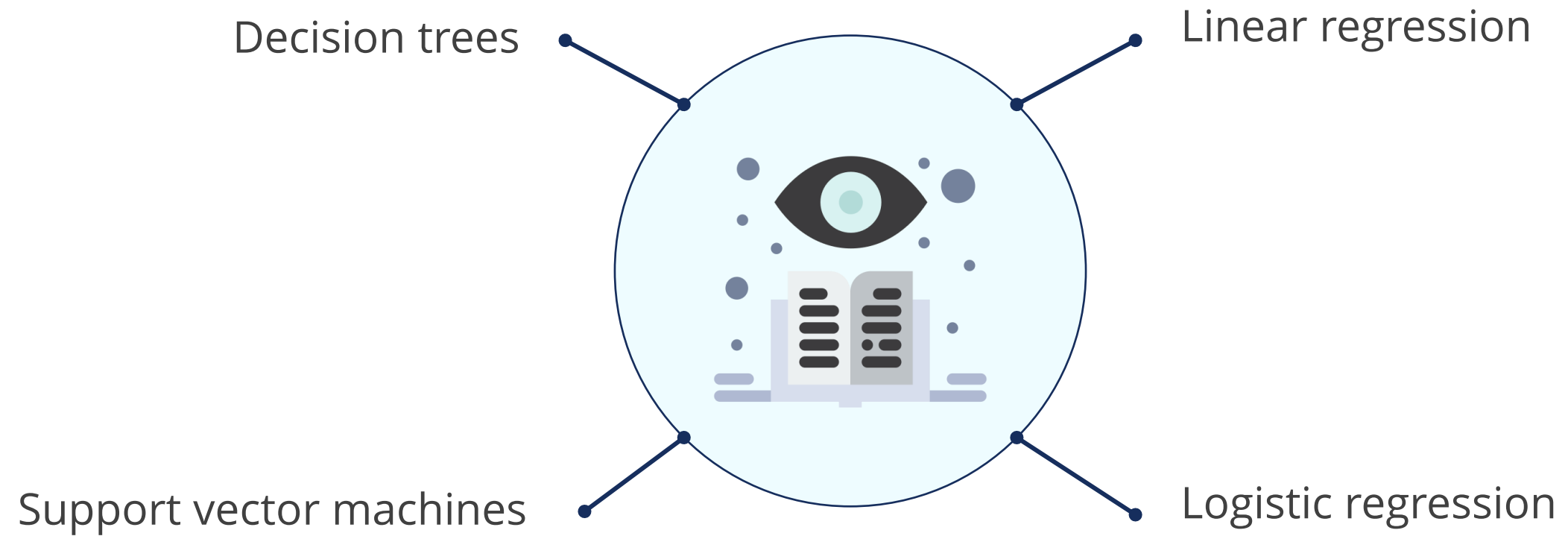
Here, the inputs and outputs are known.

## Example

A supervised learning algorithm could be used to train a machine to classify images of apples and oranges.

# Supervised Learning Methods

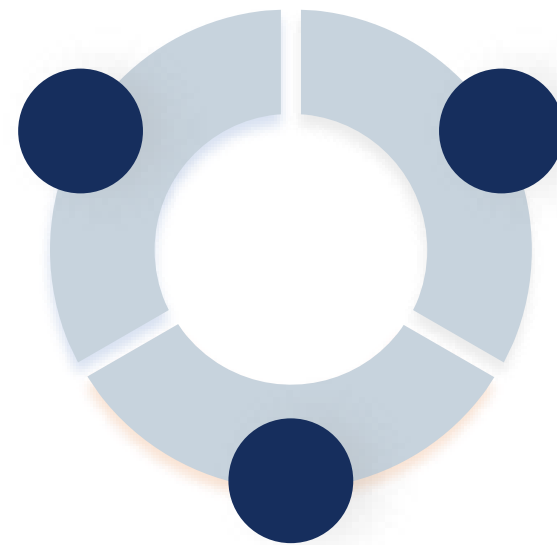
Some commonly known supervised learning methods are:



# Supervised Learning: Examples

Some examples of supervised learning are:

Prediction of temperature increase when year-wise temperature increase is known



Prediction of crop yield when season-wise crop quality deterioration is known

Classification of waste when the waste items (input) and the type of waste are known

Spam filtering: The algorithms are trained on a dataset of labeled emails (spam and non-spam) and utilize this data to predict whether or not a new email is spam.

# Unsupervised Learning

It is used when the machine is not given any labeled data, and the output is usually unknown.

Some examples are:



Identification and grouping of crop rot

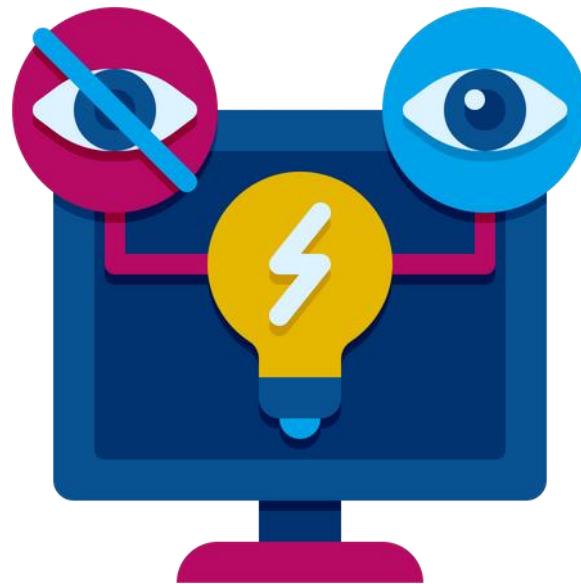
Identification of user groups based on commonalities

Identification of anomalies over geographical landscapes based on the data patterns

The unlabeled dataset is provided to an unsupervised learning algorithm that needs to find hidden patterns and recognize their relation.

# Semi-Supervised Learning

It blends both labeled and unlabeled data to train the model.



Like supervised learning, it aims to learn a function that can accurately predict the output variable from the input variables.

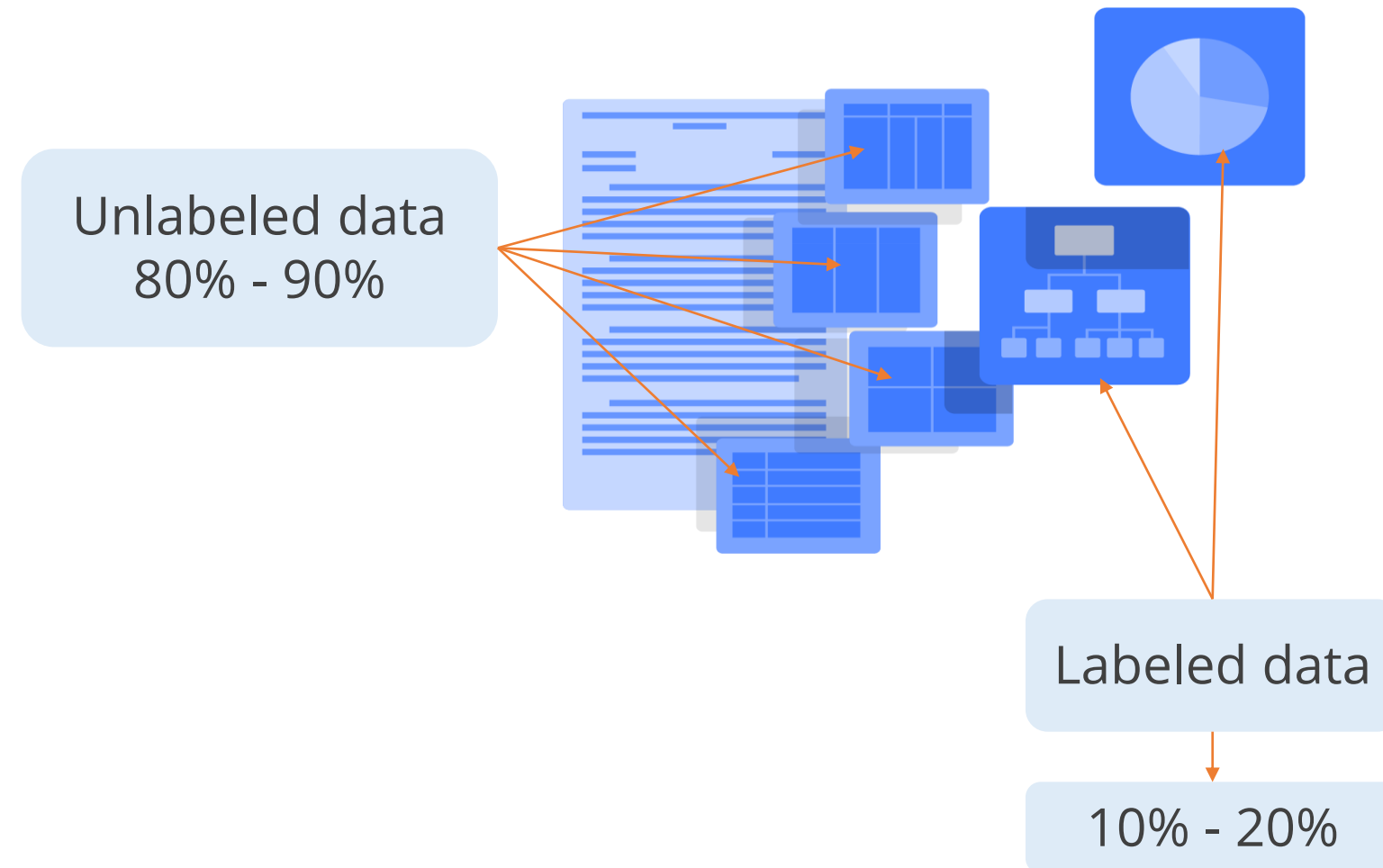
It uses unlabeled input to assist the learning process by either collecting more information or improving model generalization.

Hence, it is a type of machine learning that falls between supervised and unsupervised learning.



# Semi-Supervised Learning

Consider the data set given below:



This data set contains both labeled and unlabeled data. Hence, semi-supervised learning is applied.

# Semi-Supervised Learning: Example

Google Photos is a popular example of semi-supervised learning.

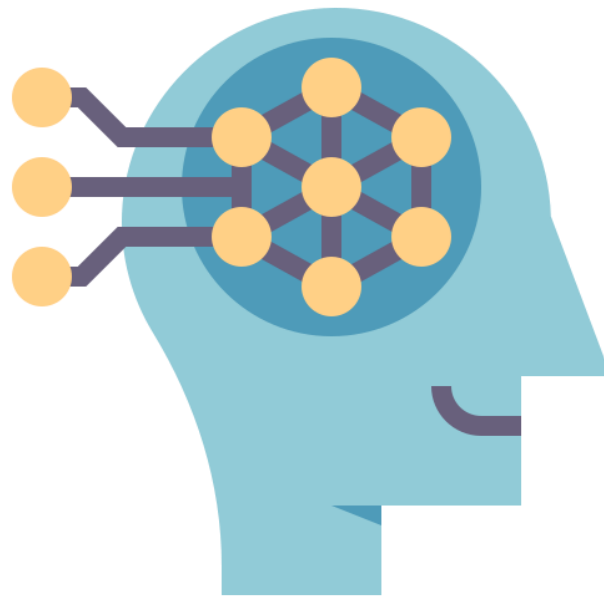
Whenever a picture is clicked, it gets stored in the Google Cloud platform or a database.



In many cases, the uploader tries to tag the image with a name. Although Google does not know the names of the images, it has an algorithm that can recognize them.

# Reinforcement Learning

Here, the program learns from its previous errors.



If the program finds the correct solution, the interpreter rewards the algorithm.

If the outcome is not up to the mark, the algorithm must reiterate until it finds a better result.

## Reinforcement Learning: Scenario

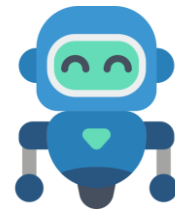
This type of learning is best seen in YouTube recommendations where a user searches for a particular song and the program shows the list of songs available.



When the user selects a specific song, the system trains itself to remember and deliver a similar result for future searches.

# Reinforcement Learning: Examples

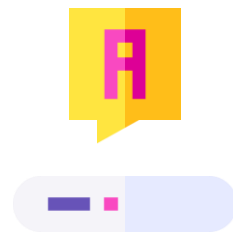
Some examples of reinforcement learning are:



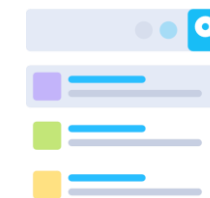
Games where players can play with bots



Self-driving cars



Autocorrect tools



Search recommendation engines

# Discussion: Types of Machine Learning

Duration: 10 minutes



- What are the different types of problems that can be solved with machine learning?

Machine learning can be used to forecast weather, predict future sales, and predict customer churn.

- What are the steps involved in applying a machine learning model to a problem?
  - Feed some data to machine learning to train it on that data
  - Do the prediction with new data
  - Check the accuracy of the model and enhance it through hyperparameter tuning
- How supervised and unsupervised machine learning are different from each other?
  - In supervised learning, data is fed to the model for training, and the prediction is done.
  - In unsupervised learning, the model identifies patterns in the data and performs classification accordingly.



# **Machine Learning Pipeline, MLOps and AutoML**

# Discussion: Machine Learning Pipeline

Duration: 10 minutes



- How to implement machine learning in real-life projects?
- What are the steps included in MLOps?



# Machine Learning Pipeline

A machine learning pipeline is a series of sequential steps used to codify and automate machine learning workflows to produce machine learning models.

It is an end-to-end construct that includes:



Data extraction

Raw data input

Preprocessing

Outputs

Features

Model parameter

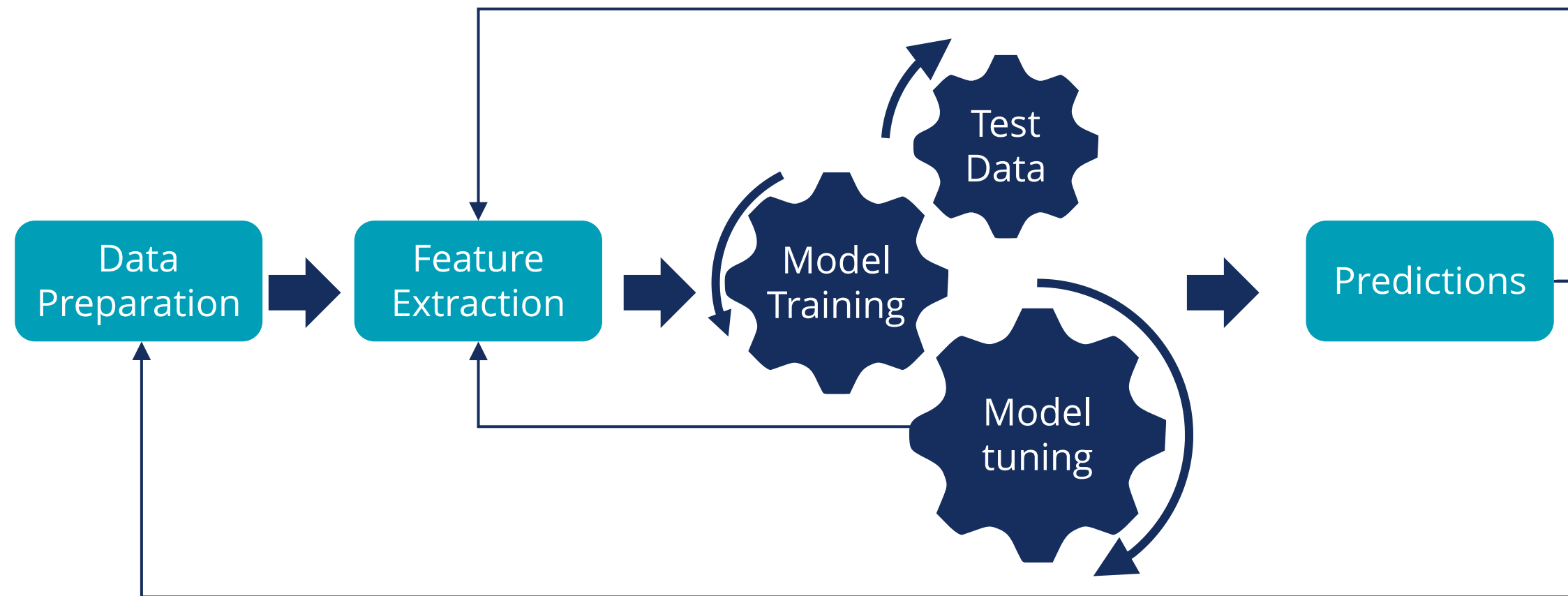
Model training

Deployment

Prediction outputs

# Machine Learning Pipeline

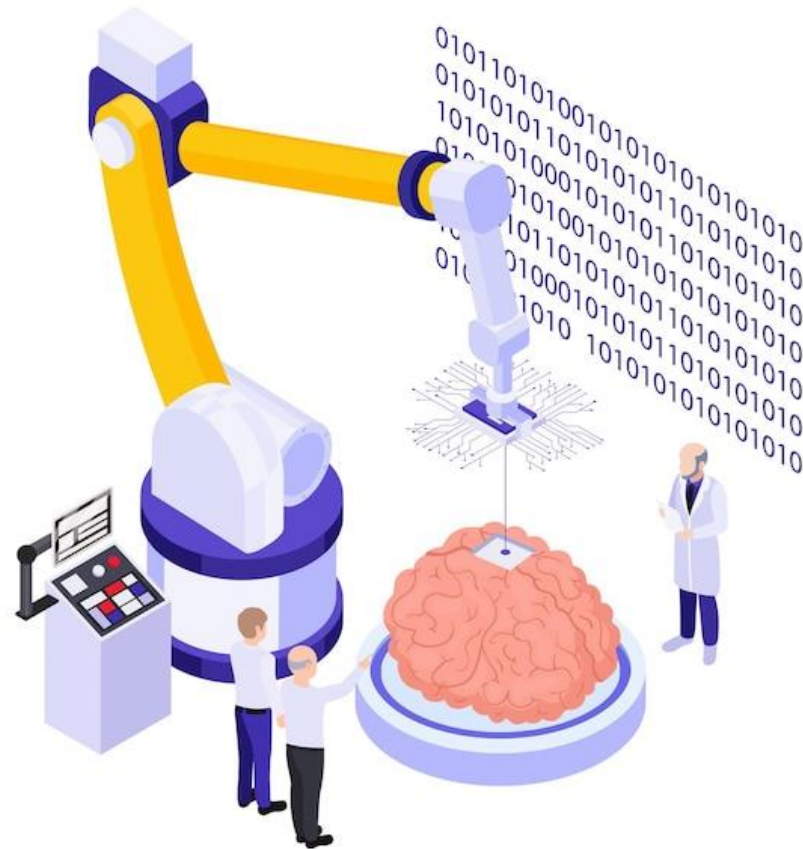
Although the term **pipeline** sometimes implies a one-way data flow, machine learning pipelines are cyclical and iterative.



Every step in the sequence is repeated several times to continuously improve the accuracy of the model. It is done until a successful algorithm is achieved on specific outcomes validated by the supervisor.

# Machine Learning and Operations (MLOps)

MLOps is a set of practices that combines machine learning, DevOps, and data engineering.



It ensures reliable and efficient deployment and maintenance of machine learning models in production systems.

It aims at improving communication and collaboration between machine learning and operations professionals.

It helps in shortening and managing the complete development life cycle.

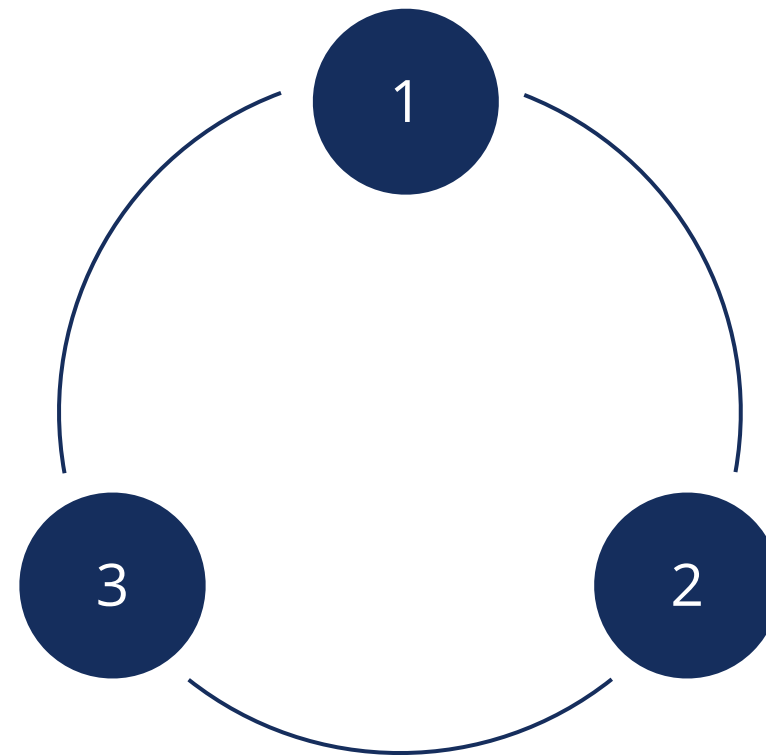
It provides continuous delivery of high-quality predictive services.

# Phases of MLOps

MLOps consists of three phases. They are:

**Design:** understanding the business and data to design the ML-powered software

**Operations:** delivering the developed ML model in production



**Model Development:** verifying the applicability of ML by implementing POC (proof of concept) for the ML model

All three MLOps phases are interconnected and influence one another.

# Tools of MLOps

Some of the tools used for MLOps are:

Kubeflow

MLFlow

Pachyderm

Data Version Control  
(DVC)

Metaflow

Seldon Core

Kedro

Amazon SageMaker

# CI/CD Pipeline Automation

Continuous integration and continuous delivery (CI/CD) system are used to test and deploy new pipeline implementations automatically.



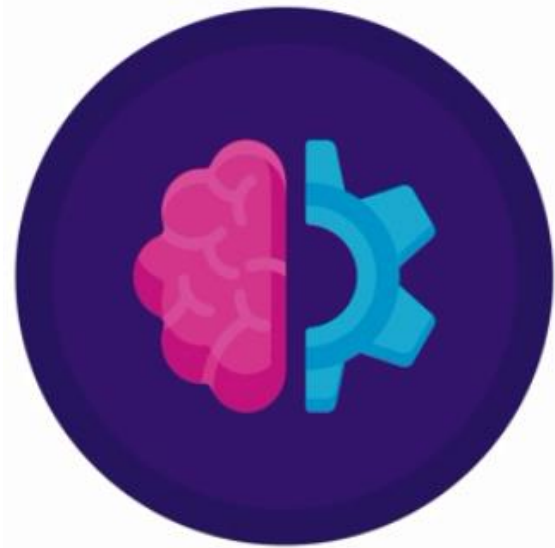
It provides a rapid and reliable update of the pipelines in production.

It assists in dealing with dynamic changes in the data, business environments, and model adjustments.

# Automated Machine Learning (AutoML)

AutoML is the process of applying machine learning models to real-world problems using automation. It also enhances the functionality of machine learning models.

It enables organizations to build and deploy ML models using:



Predefined templates

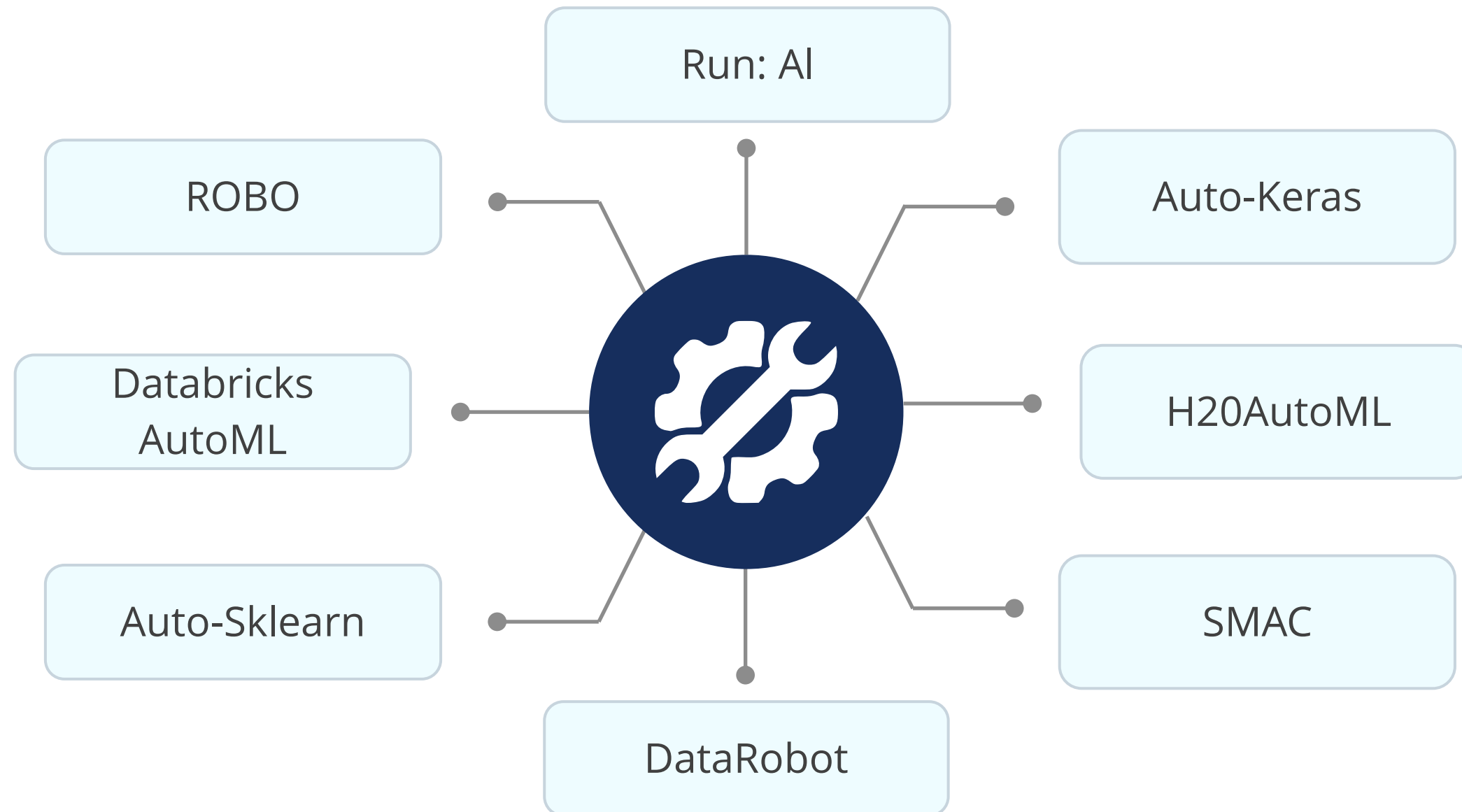
Frameworks

Processes to speed up time to completion

It combines the best practices in automation and machine learning to make machine learning accessible to a wide range of users.

# Tools of AutoML

Some AutoML tools used to automate machine learning processes are:





# Discussion: Machine Learning Pipeline

Duration: 10 minutes



- How to implement machine learning in real-life projects?  
The steps involved in the implementation of machine learning in real-life projects are:
  - Designing the model
  - Development of the model
  - Deploying the model in production
- What are the steps included in MLOps?  
The steps involved in MLOps are data extraction, raw data input, preprocessing, model building on training data, model parameters, prediction with model, fine tuning model, model deployment



# **Introduction to Python Packages for Machine Learning**

# Discussion: Python for Machine Learning

Duration: 10 minutes



- How is Python used in the machine learning process?
- What are Python libraries, and how are they related to machine learning?
- How to import and use different Python libraries?

# Python Libraries Used in Machine Learning

Some Python libraries used in machine learning are:



NumPy and Pandas manage preparation, loading and manipulation of data.



TensorFlow and Aesara perform fast numerical computing using data flow graphs.



Matplotlib performs data visualization and graphical plotting.

# Python Libraries Used in Machine Learning

Some Python libraries used in machine learning are:



SciPy solves mathematical equations and algorithms.



Scikit-learn provides efficient versions of common algorithms that help develop machine learning models.



Keras makes the implementation of neural networks easy.



PyTorch specializes in deep learning applications and accelerates the path from prototyping to deployment.

# Google Colab

It helps set up a coding environment on local machines or cloud-based solutions.

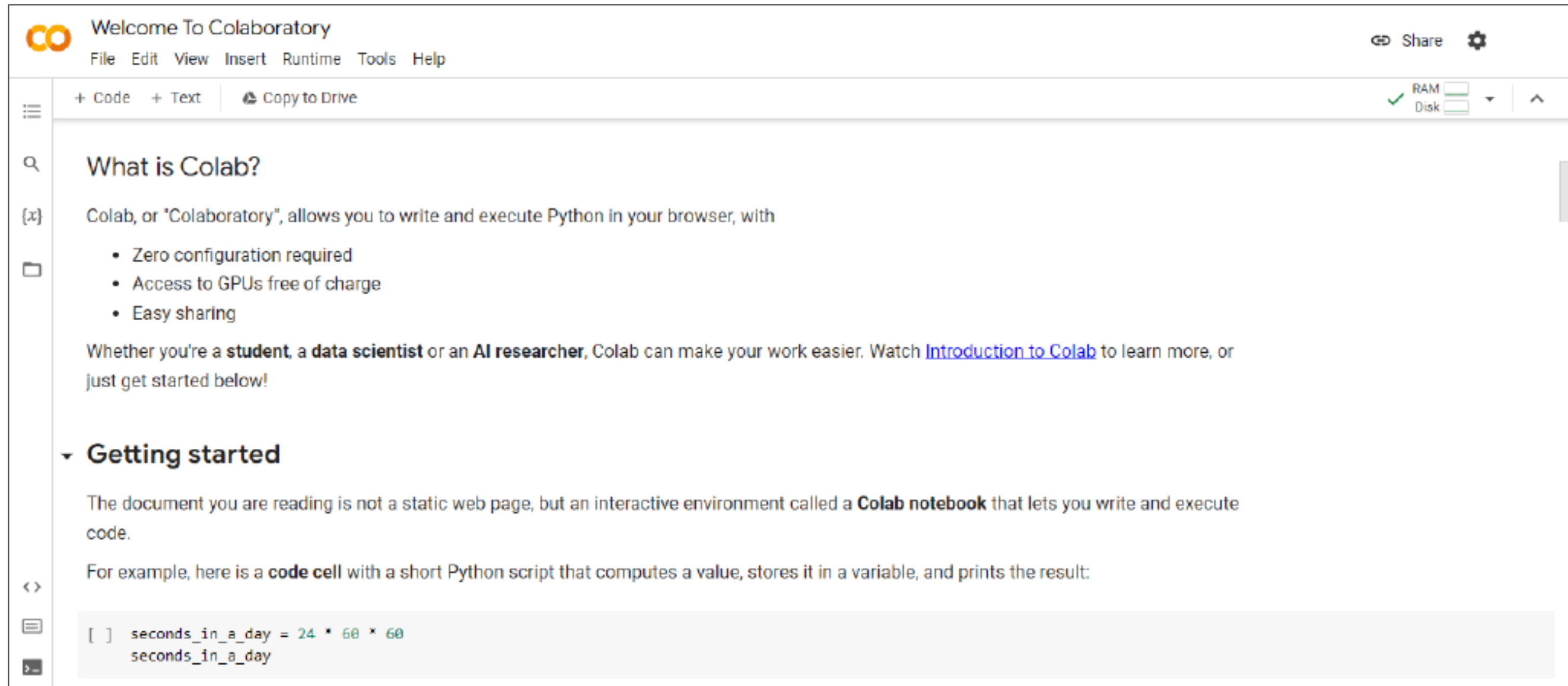


Google Colab can be used only for research and learning and not commercial purposes.

It enables running Python notebooks on the browser using the Google Cloud Platform.

# Google Colab

Google Colab will be used for the tutorial in the course.



The screenshot displays the Google Colaboratory web interface. At the top, the 'Welcome To Colaboratory' header is visible, along with a 'Share' button and a settings gear icon. Below the header is a menu bar with 'File', 'Edit', 'View', 'Insert', 'Runtime', 'Tools', and 'Help'. The main content area is titled 'What is Colab?' and describes Colab as a platform for writing and executing Python code in a browser. It lists three key features: zero configuration required, free access to GPUs, and easy sharing. A link to 'Introduction to Colab' is provided. Below this, a section titled 'Getting started' explains that the document is an interactive 'Colab notebook' and provides an example of a code cell. The code cell contains a Python script that calculates the number of seconds in a day and prints the result.

CO Welcome To Colaboratory

File Edit View Insert Runtime Tools Help

+ Code + Text Copy to Drive

RAM Disk

## What is Colab?

Colab, or "Colaboratory", allows you to write and execute Python in your browser, with

- Zero configuration required
- Access to GPUs free of charge
- Easy sharing

Whether you're a **student**, a **data scientist** or an **AI researcher**, Colab can make your work easier. Watch [Introduction to Colab](#) to learn more, or just get started below!

### Getting started

The document you are reading is not a static web page, but an interactive environment called a **Colab notebook** that lets you write and execute code.

For example, here is a **code cell** with a short Python script that computes a value, stores it in a variable, and prints the result:

```
[ ] seconds_in_a_day = 24 * 60 * 60
    seconds_in_a_day
```

# Google Colab

The following image shows how to install packages in Google Colab to run ML-relevant Python packages.

The screenshot displays the Google Colab web interface. At the top, there's a 'Welcome To Colaboratory' header with a menu bar (File, Edit, View, Insert, Runtime, Tools, Help) and a 'Share' button. Below the header, a toolbar shows '+ Code', '+ Text', and 'Copy to Drive'. On the right, RAM and Disk usage are indicated. The main area shows a code cell with the command `! pip install tensorflow`. The output of this command is a list of requirements already satisfied, including tensorflow (2.12.0), libclang (16.0.0), grpcio (1.53.0), astunparse (1.6.3), termcolor (2.2.0), tensorboard (2.12.1), gast (0.4.0), opt-einsum (3.3.0), jax (0.4.8), h5py (3.8.0), numpy (1.22.4), absl-py (1.4.0), tensorflow-io-gcs-filesystem (0.32.0), setuptools (67.6.1), wrapt (1.14.1), keras (2.12.0), typing-extensions (4.5.0), google-pasta (0.2.0), flatbuffers (23.3.3), protobuf (4.21.0), packaging (23.0), tensorflow-estimator (2.12.0), six (1.16.0), wheel (0.40.0), and ml-dtypes (0.0.4). The cell execution is marked as 'completed at 3:20 PM'.

```
Welcome To Colaboratory
File Edit View Insert Runtime Tools Help

+ Code + Text Copy to Drive

! pip install tensorflow

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: tensorflow in /usr/local/lib/python3.9/dist-packages (2.12.0)
Requirement already satisfied: libclang>=13.0.0 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (16.0.0)
Requirement already satisfied: grpcio<2.0,>=1.24.3 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (1.53.0)
Requirement already satisfied: astunparse>=1.6.0 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (1.6.3)
Requirement already satisfied: termcolor>=1.1.0 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (2.2.0)
Requirement already satisfied: tensorboard<2.13,>=2.12 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (2.12.1)
Requirement already satisfied: gast<=0.4.0,>=0.2.1 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (0.4.0)
Requirement already satisfied: opt-einsum>=2.3.2 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (3.3.0)
Requirement already satisfied: jax>=0.3.15 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (0.4.8)
Requirement already satisfied: h5py>=2.9.0 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (3.8.0)
Requirement already satisfied: numpy<1.24,>=1.22 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (1.22.4)
Requirement already satisfied: absl-py>=1.0.0 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (1.4.0)
Requirement already satisfied: tensorflow-io-gcs-filesystem>=0.23.1 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (0.32.0)
Requirement already satisfied: setuptools in /usr/local/lib/python3.9/dist-packages (from tensorflow) (67.6.1)
Requirement already satisfied: wrapt<1.15,>=1.11.0 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (1.14.1)
Requirement already satisfied: keras<2.13,>=2.12.0 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (2.12.0)
Requirement already satisfied: typing-extensions>=3.6.6 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (4.5.0)
Requirement already satisfied: google-pasta>=0.1.1 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (0.2.0)
Requirement already satisfied: flatbuffers>=2.0 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (23.3.3)
Requirement already satisfied: protobuf!=4.21.0,!4.21.1,!4.21.2,!4.21.3,!4.21.4,!4.21.5,<5.0.0dev,>=3.20.3 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (4.21.0)
Requirement already satisfied: packaging in /usr/local/lib/python3.9/dist-packages (from tensorflow) (23.0)
Requirement already satisfied: tensorflow-estimator<2.13,>=2.12.0 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (2.12.0)
Requirement already satisfied: six>=1.12.0 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (1.16.0)
Requirement already satisfied: wheel<1.0,>=0.23.0 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (0.40.0)
Requirement already satisfied: ml-dtypes>=0.0.3 in /usr/local/lib/python3.9/dist-packages (from tensorflow) (0.0.4)

4s completed at 3:20 PM
```



# Discussion: Python for Machine Learning

Duration: 10 minutes



- How is Python used in the machine learning process?  
There are very powerful libraries available in Python that make the machine learning process very easy.
- What are Python libraries, and how are they related to machine learning?  
NumPy, Pandas, Matplotlib, Seaborn, Scikit learn, and TensorFlow are the Python libraries. These libraries help in applying machine learning concepts to real-time data.
- How to import and use different Python libraries?  
Python libraries can be imported into any Python Integrated Development Environment (IDE) simply by utilizing the "Import" keyword, after which they become available for use.

# Key Takeaways

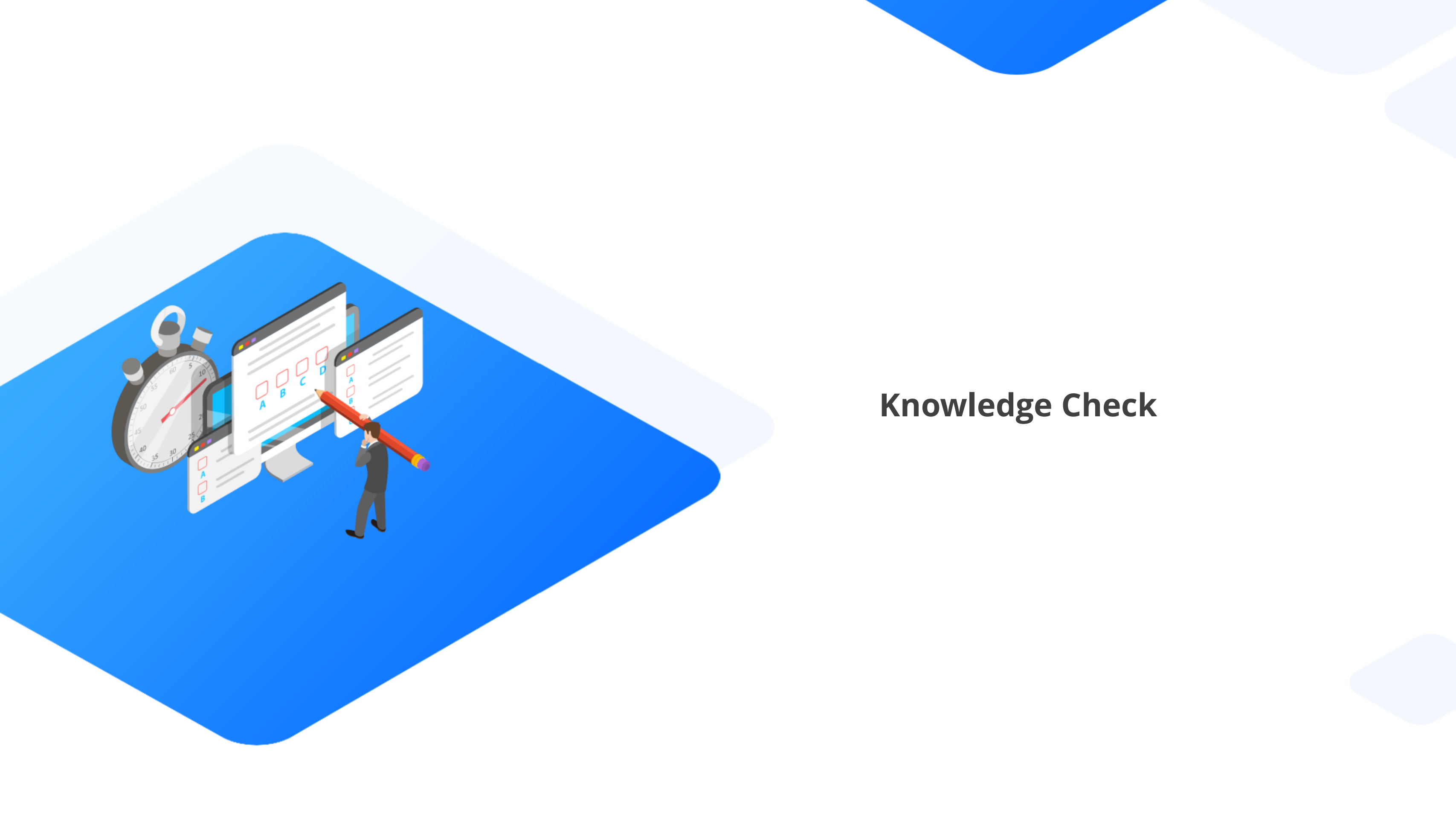
- 👁 The term machine learning refers to the ability of a machine to learn from and replicate human behavior.
- 👁 There are three main types of machine learning: supervised learning, unsupervised learning and reinforcement learning.
- 👁 The machine learning pipeline is a series of sequential steps to codify and automate machine learning workflows to produce machine learning models.
- 👁 MLOps is a set of practices that combines machine learning, DevOps and data engineering.



# Key Takeaways

- AutoML combines the best practices in automation and machine learning to make machine learning accessible to a wide range of users.
- Python packages are folders and modules that form the building blocks in Python-based programming.





## Knowledge Check

## Knowledge Check

1

### What is machine learning?

- A. The ability of a machine to replicate human behavior
- B. The ability of a machine to learn from and replicate human behavior
- C. The ability of a machine to predict human behavior
- D. The ability of a machine to replace human behavior



## Knowledge Check

1

### What is machine learning?

- A. The ability of a machine to replicate human behavior
- B. The ability of a machine to learn from and replicate human behavior
- C. The ability of a machine to predict human behavior
- D. The ability of a machine to replace human behavior



---

The correct answer is **B**

---

**Machine learning is the ability of a machine to learn from and replicate human behavior.**

## Knowledge Check

2

**What are the main categories of machine learning?**

- A. Supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning
- B. Deep learning, artificial intelligence and supervised learning
- C. Regression learning, classification learning and unsupervised learning
- D. Neural networks, algorithms and decision trees



## Knowledge Check

2

What are the main categories of machine learning?

- A. Supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning
- B. Deep learning, artificial intelligence and supervised learning
- C. Regression learning, classification learning and unsupervised learning
- D. Neural networks, algorithms and decision trees



---

The correct answer is **A**

---

**Machine learning is categorized into four main categories: supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning.**



## Knowledge Check

3

Which Python package is used to plot data?

- A. NumPy
- B. SciPy
- C. Matplotlib
- D. Pandas



## Knowledge Check

3

Which Python package is used to plot data?

- A. NumPy
- B. SciPy
- C. Matplotlib
- D. Pandas

---

The correct answer is **C**

---

**Matplotlib is a Python package that is used to plot data.**





**Thank You!**