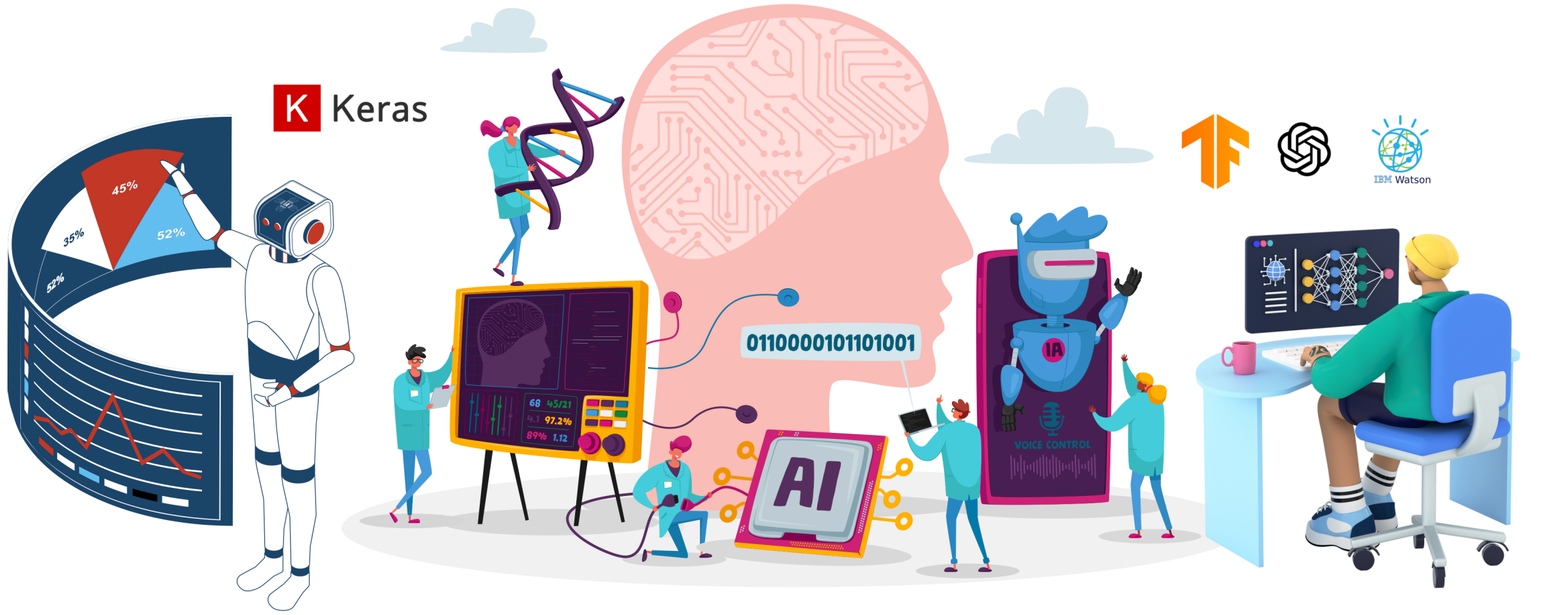
MACHINE LEARNING

EXPLAINED



Keras

LightGBM

NVIDIA
CUDA

XGBoost

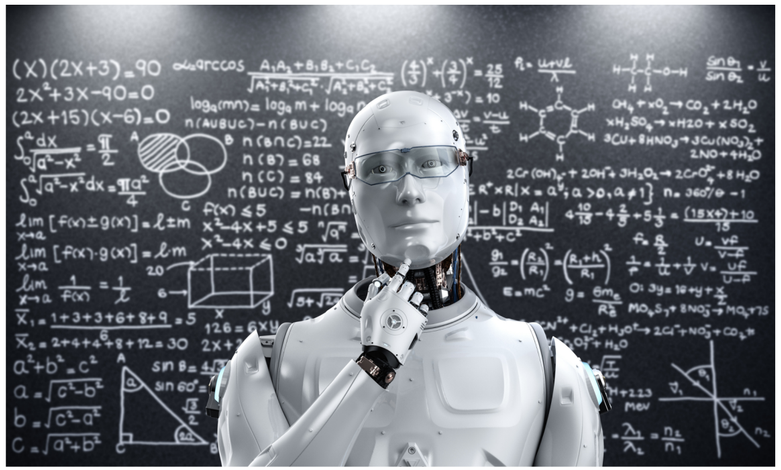
IBM Watson

PyTorch

MATLAB

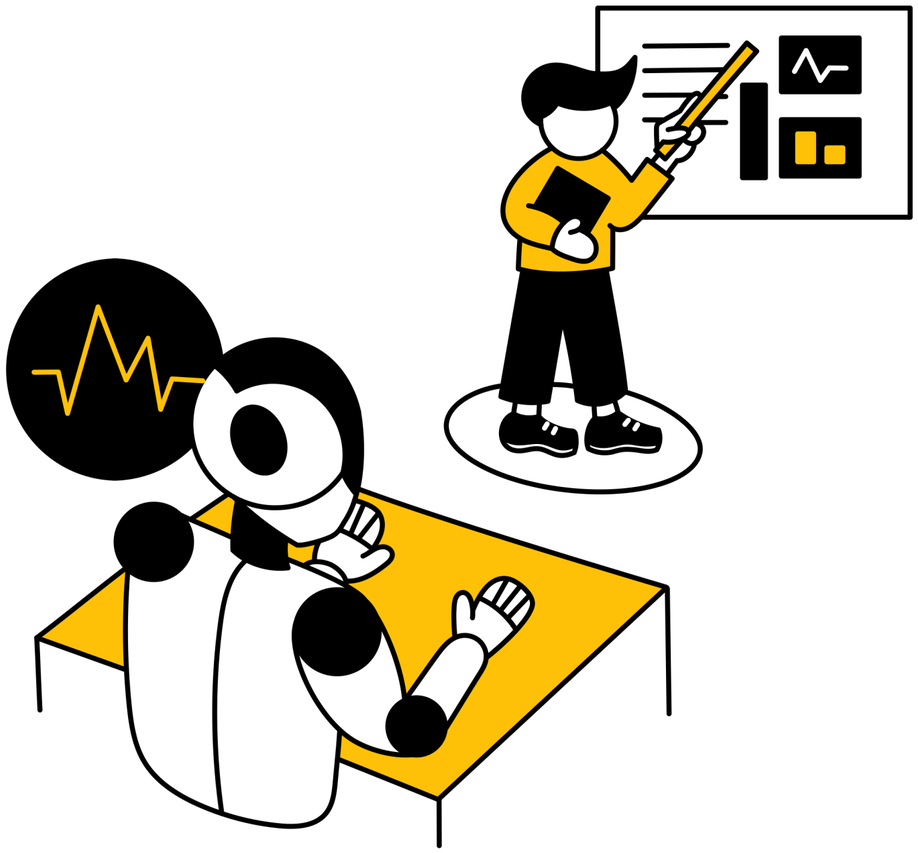
A machine learning model hardly learns anything about a problem or its human context. Instead, the machine learning model typically focuses on the problem's data and its interpretation as questions and answers.

A model applies some mathematical method (statistical, probablistic etc.) to learn to find answers for the questions of the data.



It trains itself by 'working on' the given data and learning how the given answers were obtained for the questions.

SUPERVISED MACHINE LEARNING

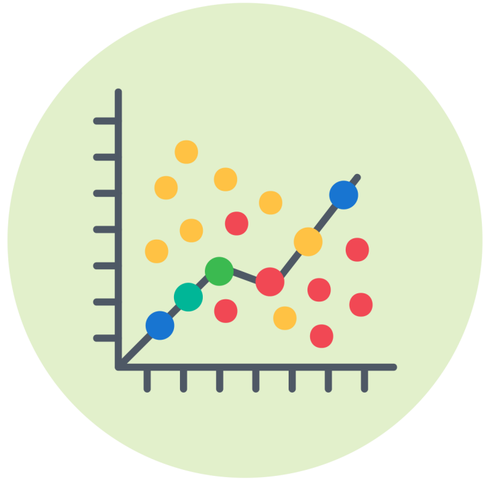


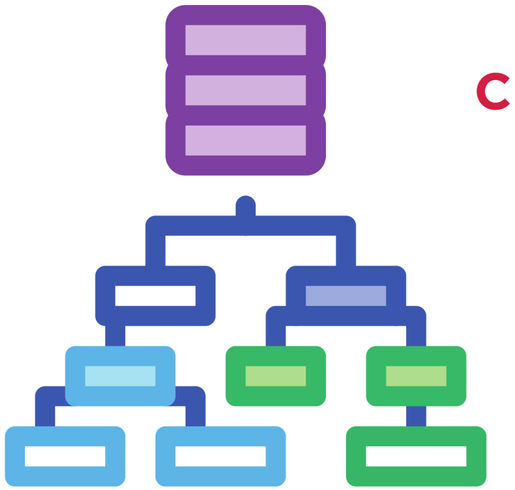
Supervised machine learning involves using **well labelled data for training** the model to find mapping or relationships between input and output features.

- Can involve tasks like :

Regression : Predicting continuous *numerical* values based on input features

Example : Predicting house prices based on features like area, no. of bedrooms, and location





Classification : Assigning category labels / classifications to data

Example : Classify emails as ***spam*** or ***not spam***

Anomaly Detection:
Identifying outliers or unusual patterns in data that deviate from normal behavior.

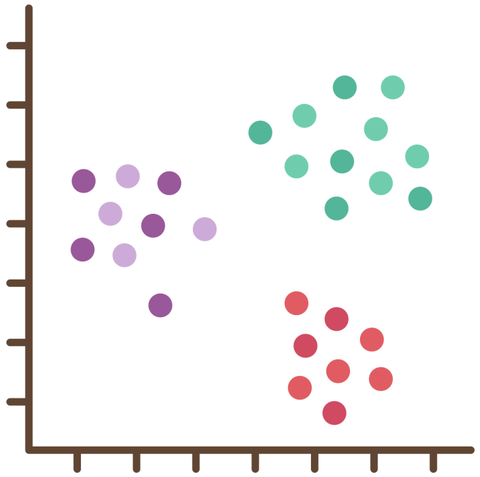
Example : Credit card fraud detection

UNSUPERVISED MACHINE LEARNING

Unsupervised machine learning involves the model learns patterns and structures from unlabeled data without explicit supervision to uncover hidden relationships, group similar data points, or reduce the dimensionality of the data.

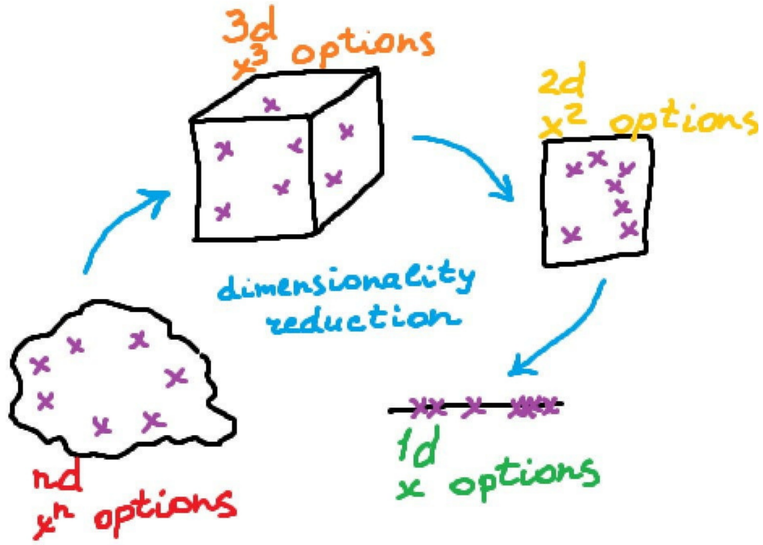
CLUSTERING

Group similar data points together into clusters or segments based on their inherent patterns or characteristics



DIMENSIONALITY REDUCTION

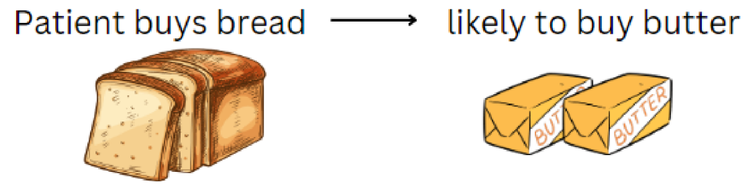
Reduce the number of features in a dataset while preserving its essential structure and characteristics



https://miro.medium.com/v2/resize:fit:959/1*kK4aMPHQ89ssFEus6RT4Yw.jpeg

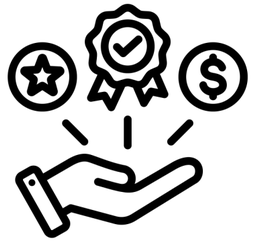
ASSOCIATION RULE LEARNING

Discovering interesting relationships or associations between variables



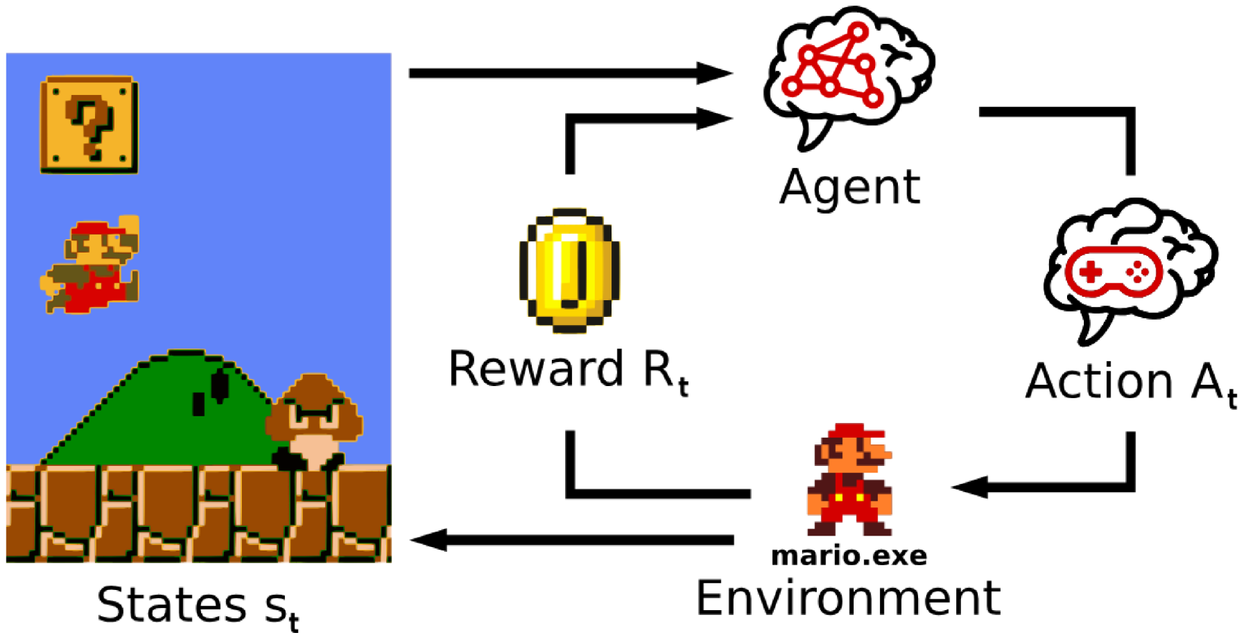
REINFORCEMENT LEARNING

Reinforcement learning involves an agent which learns to make sequential decisions by interacting with an environment through trial and error.



It receives feedback in the form of rewards or penalties based on its actions.

Reward: is a signal provided by the environment to indicate the desirability of the agent's actions. The goal of the agent is to maximize the cumulative reward over time.



The algorithm learns to optimize the agent's policy through experience, typically using iterative trial-and-error processes.

It can choose actions which will give larger long term rewards compared to short term ones.

COMMON TASKS OF MACHINE LEARNING

PREDICTION

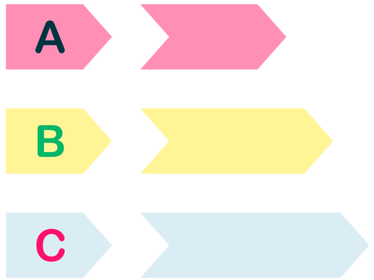
Derive a prediction for a future or unknown occurrence by identifying patterns and trends within current and past data.

FORECASTING

Forecasting involves estimating future events or trends based on historical and statistical data.

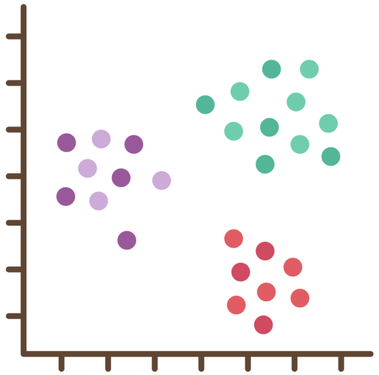
Forecasting predicts outcomes over a longer time frame, often over months, years, or even decades, often with time-series data.
Predictions are short-term and immediate, often in the near future, up to a year.

CLASSIFICATION



Classify given one or multiple data entities (text, numbers, image etc.) to one of given classes, for example, classifying an image as a picture of truck.

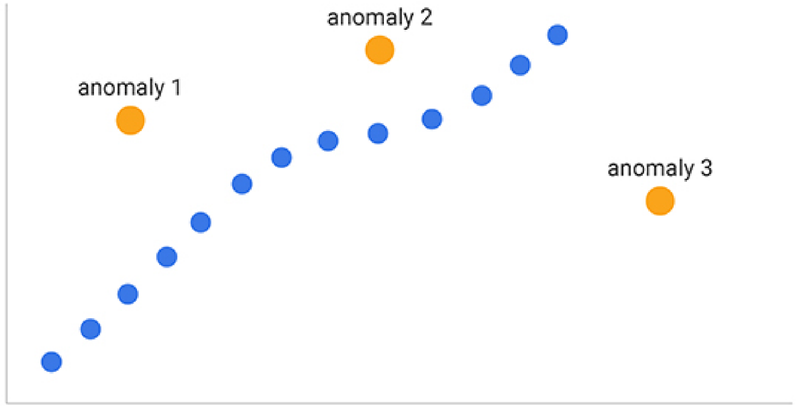
CLUSTERING



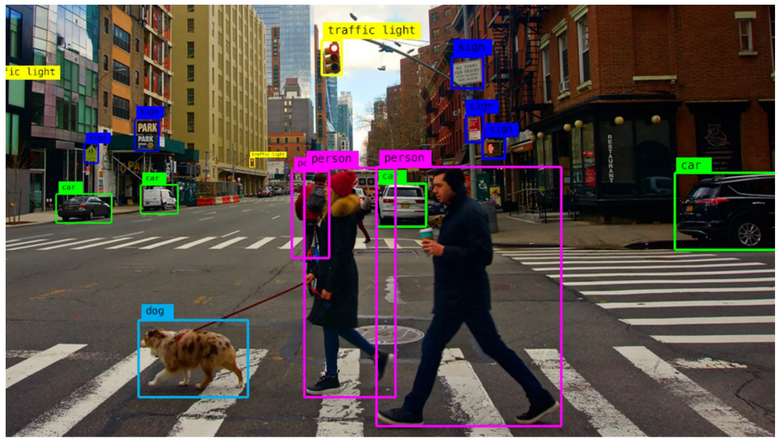
Given data entities, group data entities into clusters based on related features and differences, for example, given an random group of students, group them on the basis of ages.

ANOMALY DETECTION

Within the given data, identify anomalies, outliers, and rare occurences within the data, which differ from usual behavior of the data.



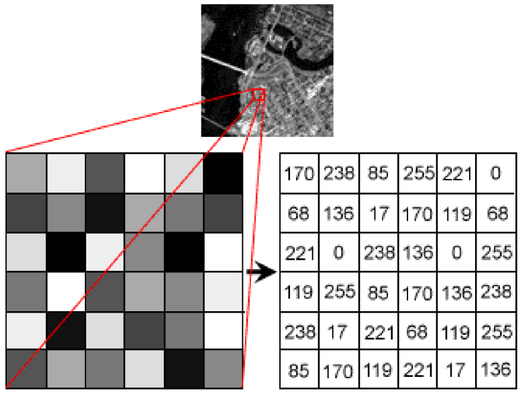
ELEMENT DETECTION



Within the given data, detect the presence of a specific data entity, often corresponding to some sort of classification.

For example, find an object in an image or video, detect cracks in the image of a processor, detecting a face in a camera and classifying the mood of the subject from the facial expressions.

ENCODING



Convert and represent data obtained in one format to different formats, often identifying key features which help in compressing and reducing the dimensionality of the data.

For example, converting images to arrays of pixels, large documents of text into vectors.

Covers topics like natural language processing, image processing etc. and corresponds with DECODING - which generates original representations from encoded representations of data.

GENERATION

Given input prompt, generate data or (text, images etc.) or prompt for action, in accordance to prompt and parameters.

OpenAI Gemini Hugging Face BLACKBOX AI

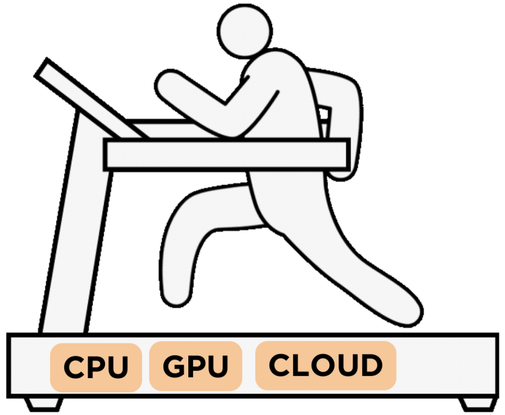
SIMPLE ML PIPELINE

STEP 1 : WHAT AND HOW

1. Define the problem statement.
2. Identify the task of machine learning model.
3. Identify the best method(s) for the task.
4. Obtain the data relevant for the task of the model.

STEP 2 : MODEL INSTANTIATION

1. Create an instance of a model.
2. Define the parameters of the model.
3. Split the data for training and testing.
4. Fit the model on training data.



STEP 3 : MODEL EVALUATION AND TUNING

1. Test the model on the testing data.
2. Improve performance by altering the data and parameters of the model.
3. Evaluate model’s perfomance.

STEP 5 : DEPLOYMENT AND MAINTAINENCE

1. Deploy model to work with real life, unseen data.
2. Record and evaluate model’s performance.
3. Update and improve the model as required.

“ BETTER DATA IS BETTER THAN BETTER MODELS ”

The role of data in working of a machine learning algorithm is extremely critical, which is why machine learning engineering is only a derivative of data engineering.

Beyond the model’s behavior and its parameters, the availability and quality of data is a deciding factor in the outcome and reliability of a machine learning model.

THANK YOU !

WHILE I'M WORKING ON THE NEXT ARTICLE, I'M OVER HERE ON