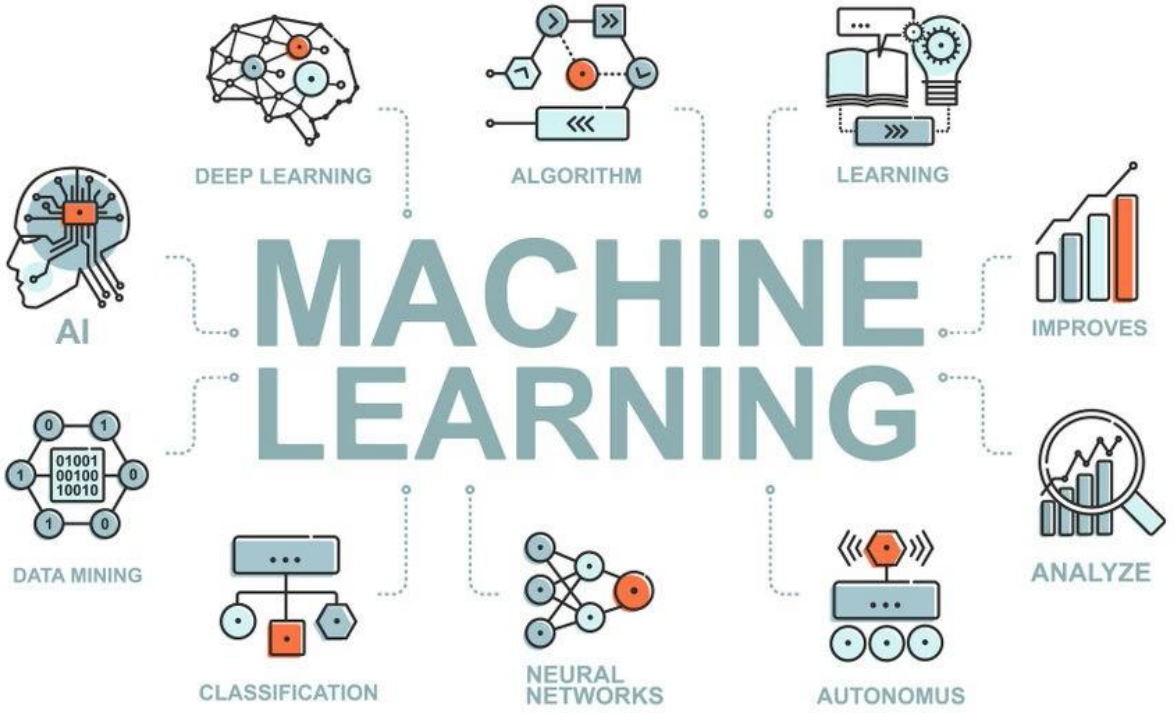


Madushanka Wadumulla
PGC in IA
Beng. IIESL, AMIE

NIBM POWERING GREAT MINDS

INTRODUCTION

Outline

- What is AI
- Areas of AI
- What is ML
- Machine Learning Process
- Types of Machine Learning
- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

What is the first thing, comes to your mind when you hear Artificial Intelligence?

What is Artificial Intelligence?

Most of us think that Al is Robotics
* Al is not robotics
* Al is a process
* Al is a study of how human brain think, learn, decide and work, when it tries to solve problems
* Robotics uses Al for bringing human intelligence to robots

Areas of Al

Artificial intelligence generally false under two broad categories.

* Narrow Artificial Intelligence

Its focus on performing specific tasks.

Ex- Chat GPT, Amazon Alexa, YouTube Recommendation, Self Driving Cars

* Artificial General Intelligence

Human level intelligence

Artificial General Intelligence (AGl)

The creation of a machine with human level intelligence that can be applied to any task is the Holy Grail for many Al researchers, but the quest for AGI has been fraught with difficulty.

* AGI, sometimes referred to as "Strong Al," is the kind of artificial intelligence we see in the movies, like the robots from Westworld or Data from Star Trek The Next Generation.

Do you know him?

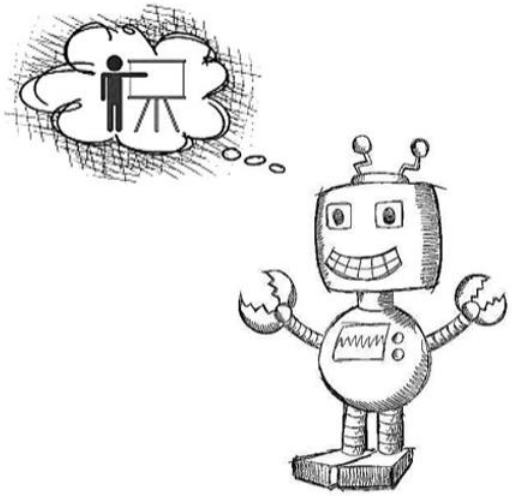
ELON MUSK
FOUNDED COMPANIES

1995 - ZIP2
1999 - PAYPAL
2002 - SPACEX
2003 - TESLA
2006 - SOLARCITY
2013 - HYPER LOOP
2015 - OPENAI
2016 - NEURALINK
2017 - THE BORING COMPANY

What is Machine Learning?

Humans Learns from past experiences

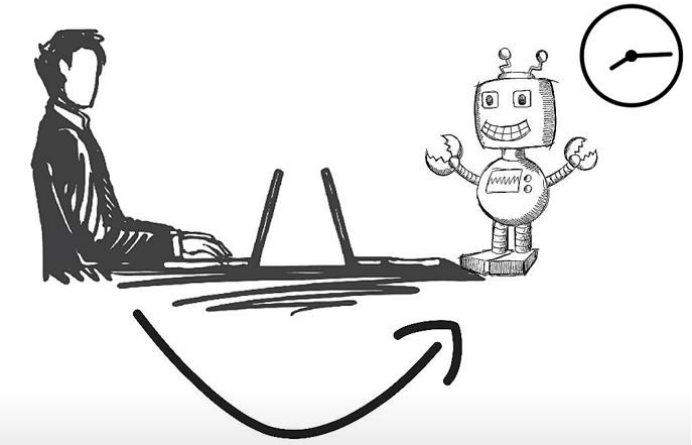
Rats Learning to Avoid Poisonous Baits:



Machines follow instructions given by humans

What is Machine Learning?

In Machine Learning,

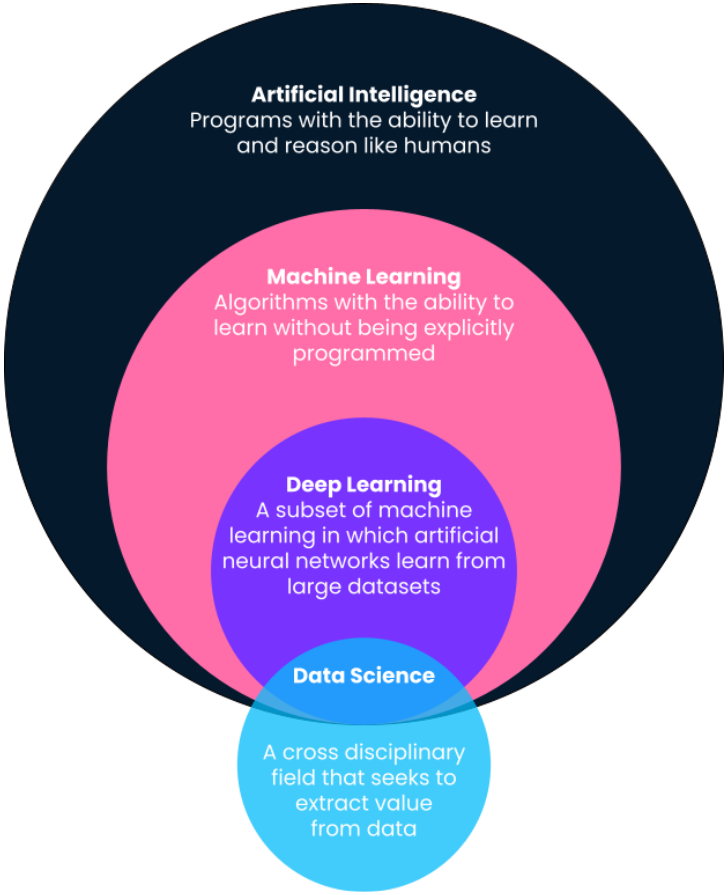


Humans train the machines based on the data

What is machine learning?

➢Machine learning (ML) is a discipline of artificial intelligence (AI) that allows machines to automatically learn from data and past experiences while identifying patterns to make predictions with minimal human intervention.

➢Machine learning methods enable computers to operate autonomously without explicit programming. ML applications are fed with new data and can independently learn, grow, develop, and adapt.



MACHINE LEARNING PROCESS

STEP 02

Data cleaning and Feature Engineering

STEP 04

Model Testing using Performance Matrices

STEP 01

Collection of Data from Various source

STEP 03

Model building using Machine Learning Algorithm

STEP 05

Model Deployment

Machine learning process

Machine learning process

Data Cleaning and Feature Engineering:
Cleanse the data by handling missing values, outliers, and inconsistencies.
Create new features (feature engineering) that enhance the model's ability to learn.
Transform data into a suitable format for ML algorithms.

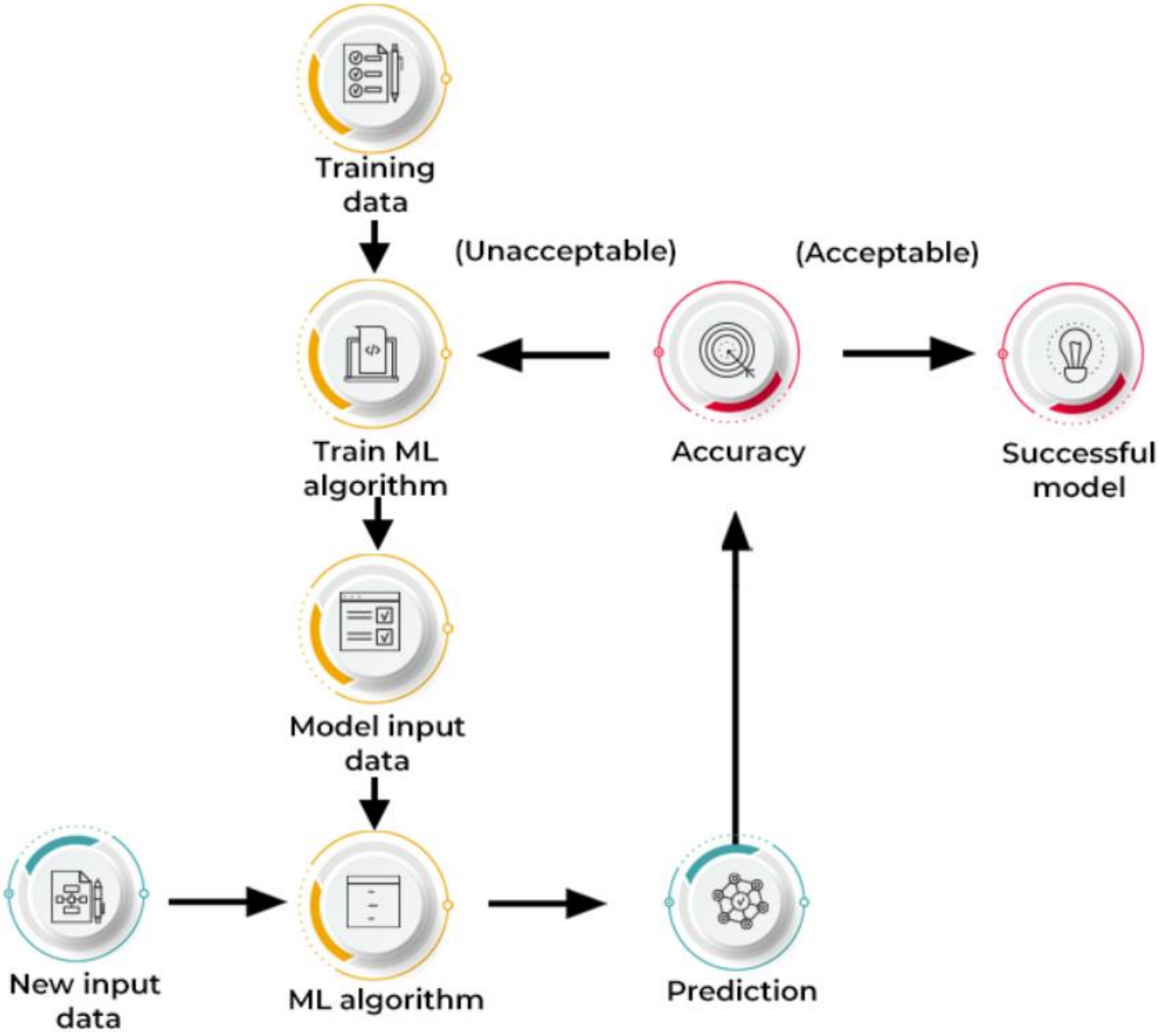
Age	Work Experience	Salary in K USD	Personal Loan Flag
25	missing	50	1
27	3	missing	1
29	5	80	0
31	7	90	0
33	9	100	1
missing	11	130	0

Machine learning process

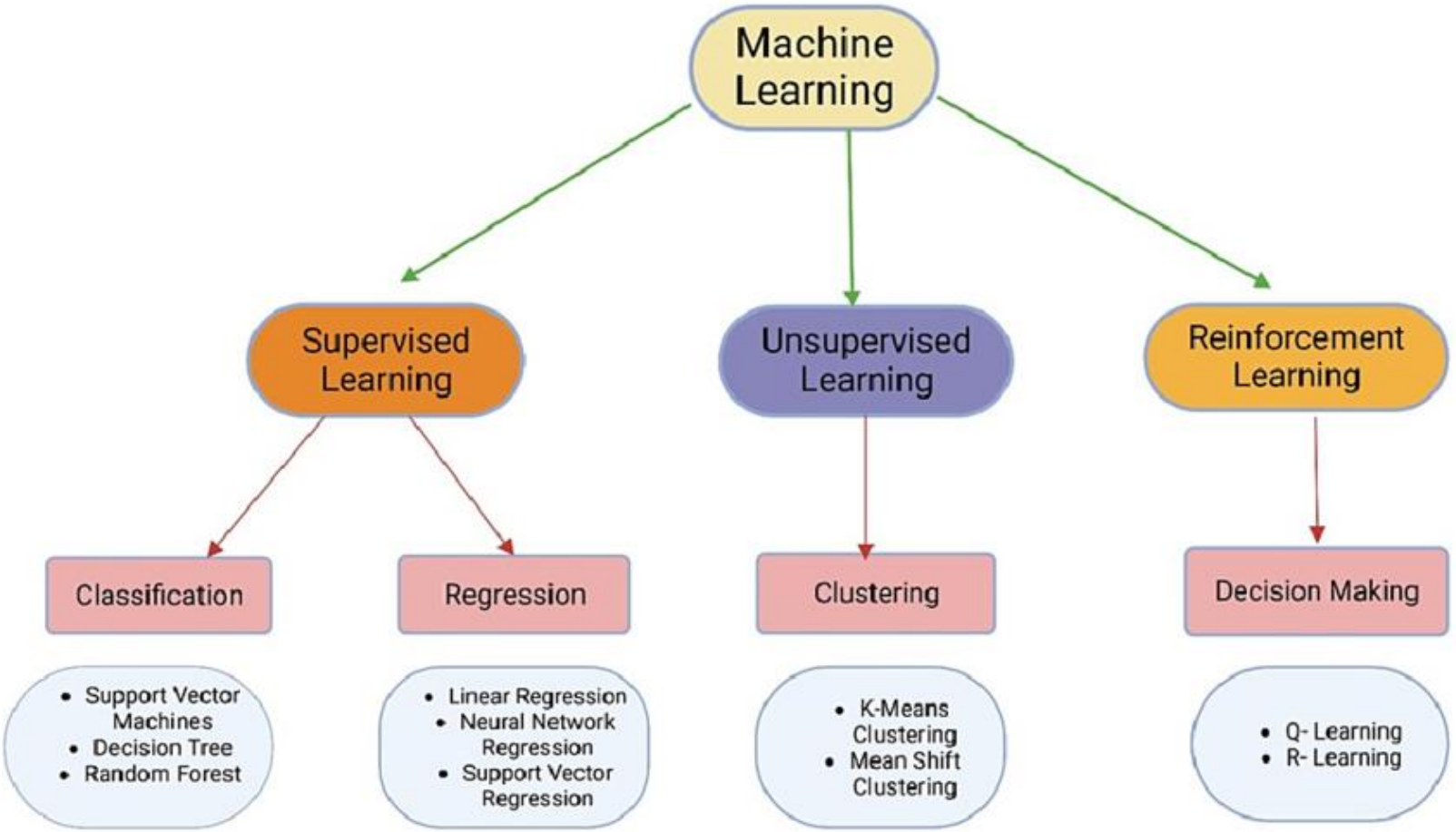
Model Building Using Machine Learning Algorithms:
Choose a machine learning algorithm (e.g., regression, classification, or clustering).
Train the model using the training dataset.

Model Testing using Performance Matrices:
Use the testing dataset to evaluate the model's accuracy.
If results are unsatisfactory, refine and improve the model.
Consider stakeholder feedback, algorithm adjustments, and parameter tuning.

Model Deployment:
Once satisfied with the model's performance, deploy it for real-world use.


Training data
(Unacceptable)
(Acceptable)
Train ML algorithm
Accuracy
Successful model
Model input data
New input data
ML algorithm
Prediction

Types of Machine Learning

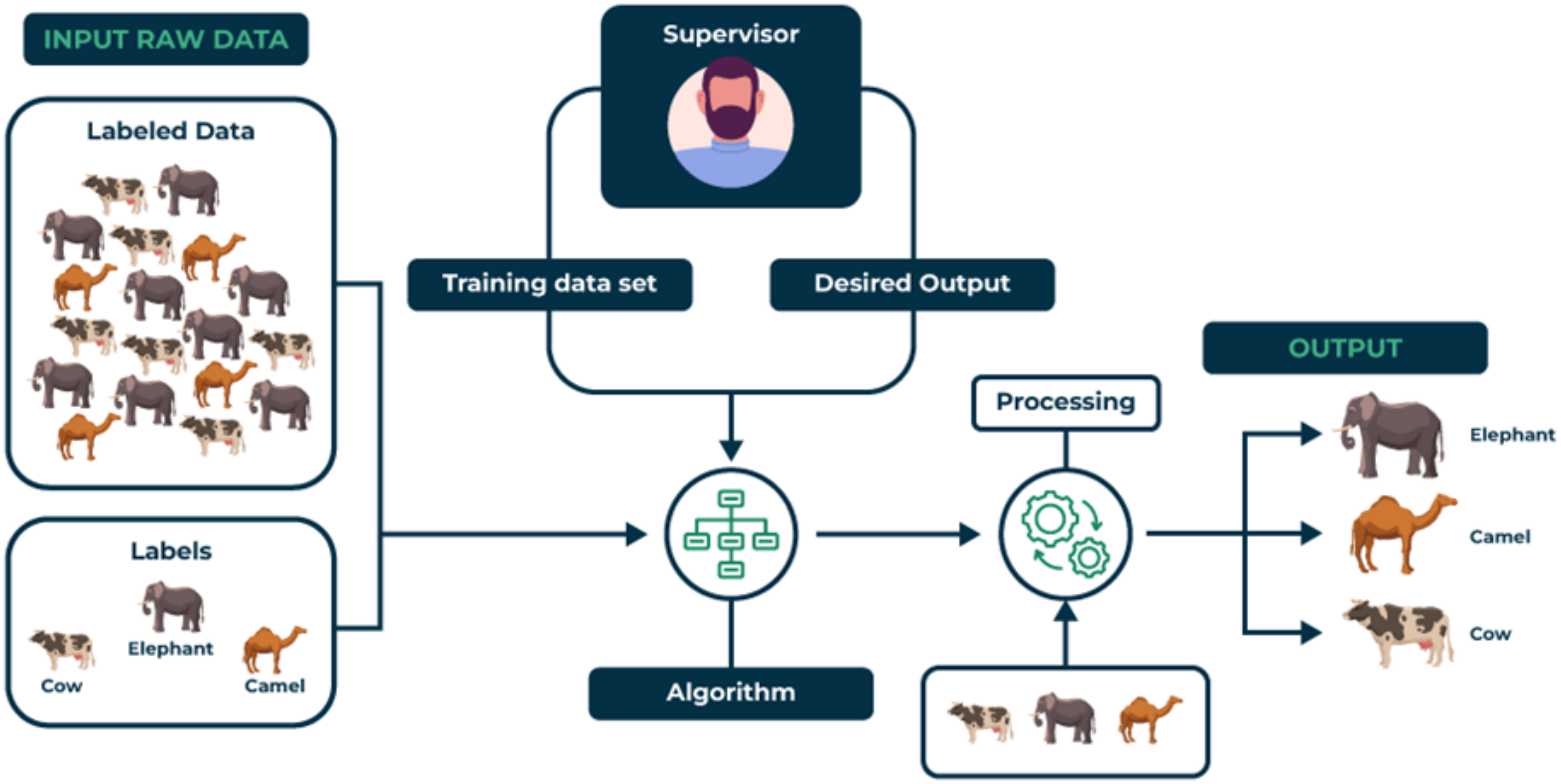


Supervised learning

Algorithm learns from labeled training data to make predictions or decisions without being explicitly programmed. This process involves training a model on a dataset that includes both the input features and the corresponding correct output, known as labels.

Ex:

- Classification: Predicting categories (e.g., spam detection)
- Regression: Predicting continuous values (e.g., house prices)


INPUT RAW DATA
Labeled Data
Labels
Cow
Elephant
Camel
Supervisor
Training data set
Desired Output
Algorithm
Processing
OUTPUT
Elephant
Camel
Cow

Supervised learning

- In Supervised Learning, teacher teaches how to do. Students learn by understanding the inputs & outputs.

- Same thing is applied here. The model will be trained with inputs and related outputs. After training it can make decisions based on predictions on unseen situations.

What happens next ?

Sometimes the actual results may not be as expectations. In supervised learning these are called the errors

Then the errors should be minimized

රමනීයයි ඒ මදුර ජවනිකා...😁🤭

> Those mosquito scenes are beautiful... 😁🤭

⚙ • Rate this translation

Supervised Learning — Overfitting & Underfitting

- **Overfitting** refers to a model that models the training data too well. This means that the noise or random fluctuations in the training data is picked up and learned as concepts by the model. The problem is that these concepts do not apply to new data and negatively impact the model's ability to generalize.

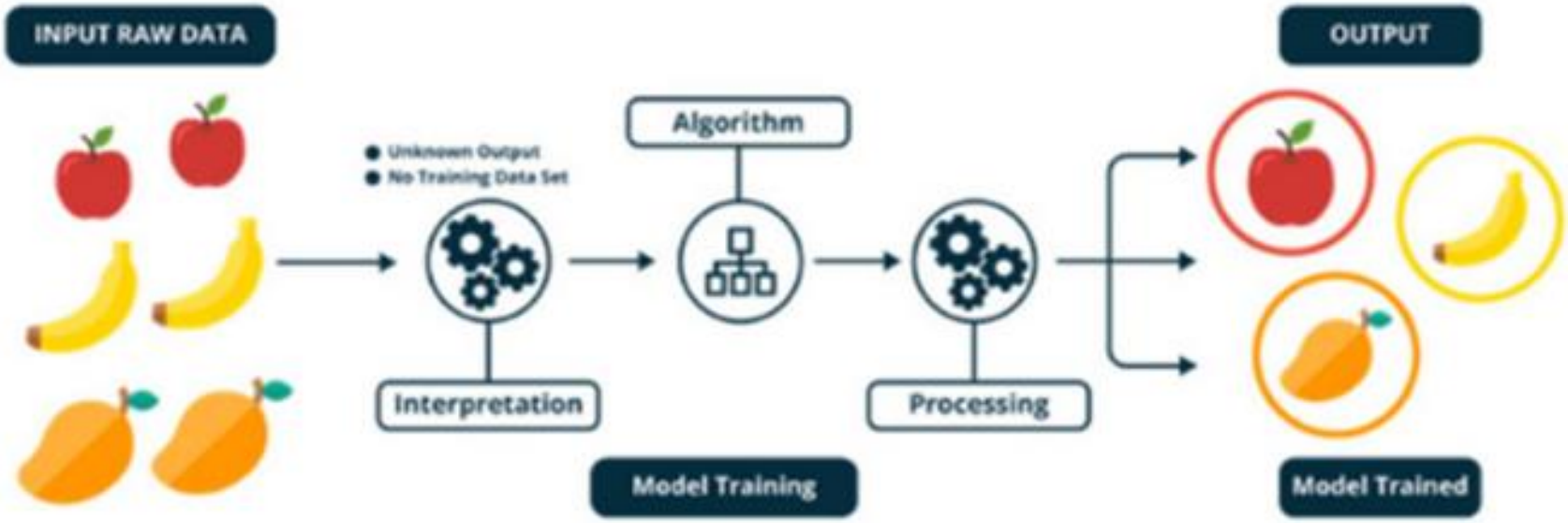
- **Underfitting** refers to a model that can neither model the training data nor generalize to new data. An underfit machine learning model is not a suitable model and will be obvious as it will have poor performance on the training data.

Unsupervised Learning

The model is trained on unlabeled data, meaning the data does not have predefined labels or outcomes. The goal of unsupervised learning is to find hidden patterns or intrinsic structures in the input data.

Unlike supervised learning, where each training example is paired with an output label, unsupervised learning works with data that has no labels. (Clustering, association rule mining, dimensionality reduction)

UNSUPERVISED LEARNING



Unsupervised Learning

- Unsupervised learning, is where you only have input data (X) and no corresponding output variables. Derive some structure or pattern from the “unlabeled data” by just looking at the relationship between the data themselves. There are few methods to check the accuracy of these models.



















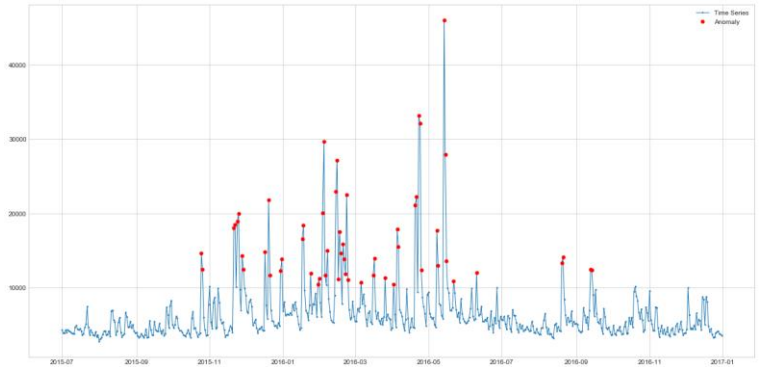






Example Applications of Unsupervised Learning

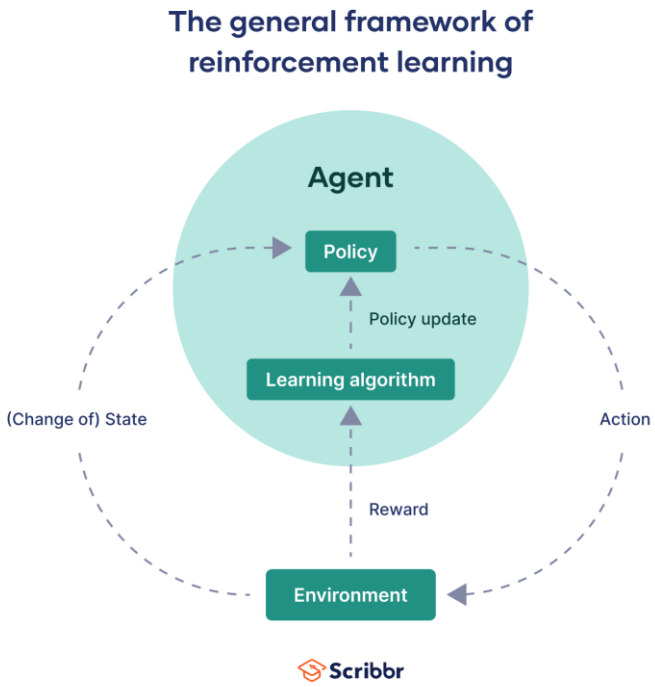
- **Customer Segmentation:** Clustering customers based on purchasing behavior.
- **Market Basket Analysis:** Identifying sets of products frequently bought together.
- **Anomaly Detection in Fraud Detection:** Identifying transactions that deviate from normal patterns.
- **Image and Video Compression:** Using techniques like autoencoders to reduce the size of image and video data.



Anomaly Detection for web traffic data

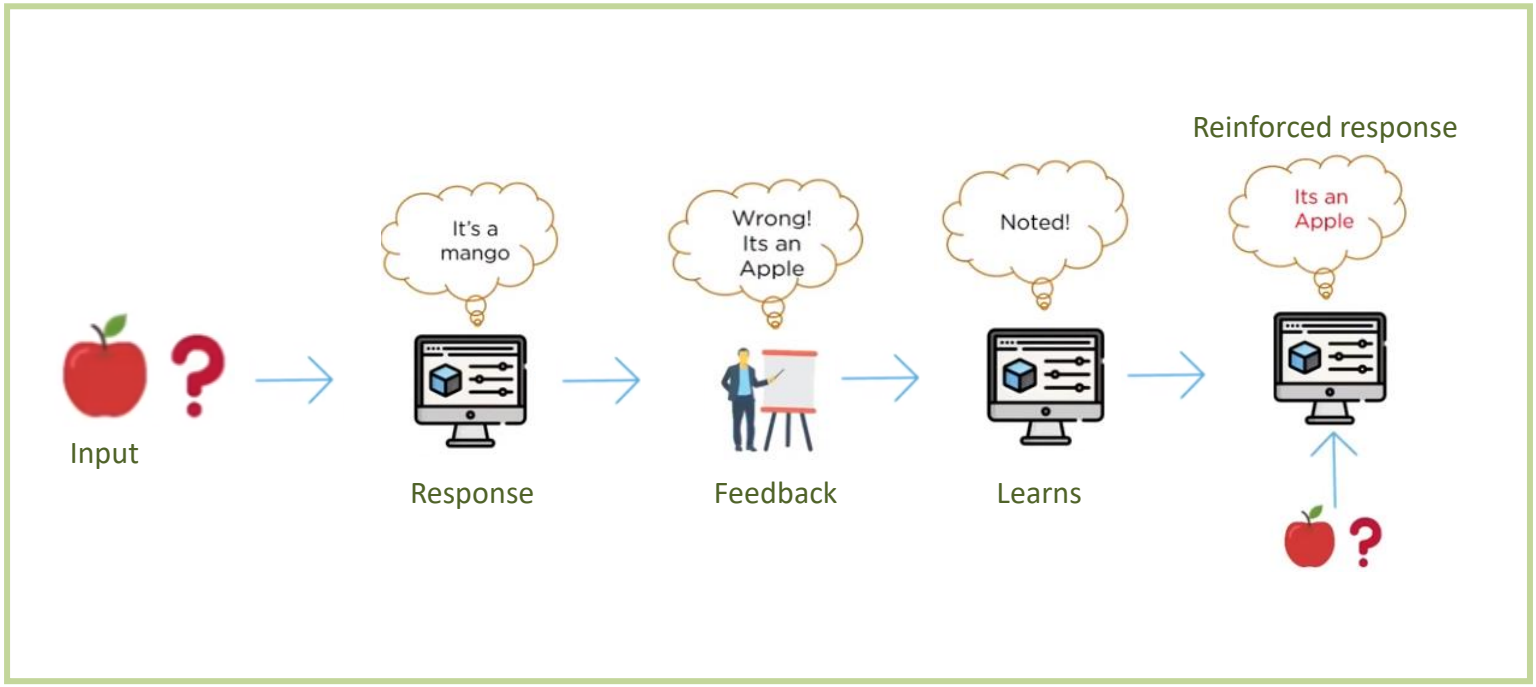
https://antonsruberts.github.io/anomaly-detection-web/
https://antonsruberts.github.io/anomaly-detection-web-2/

Reinforcement learning



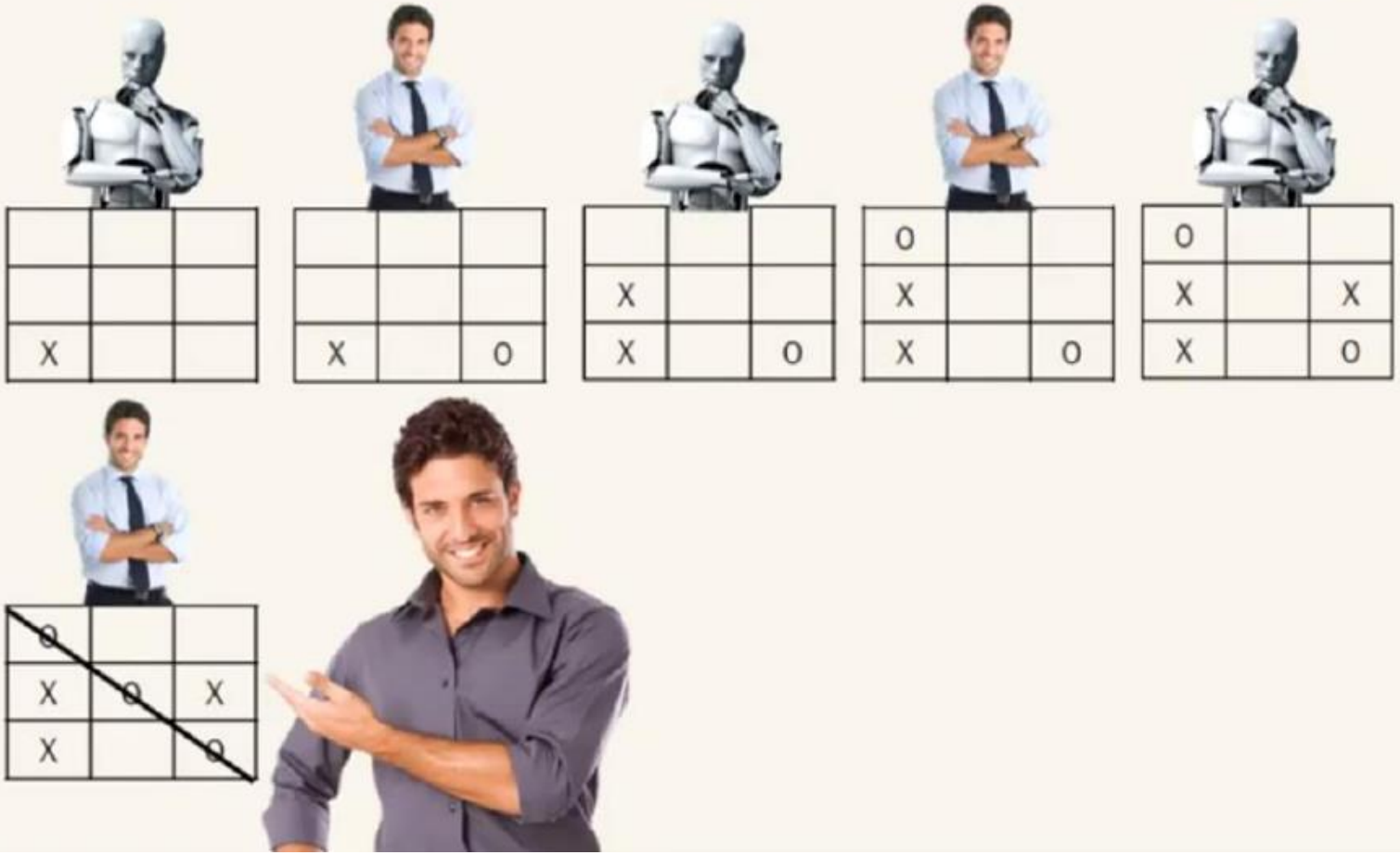
- A technique that trains software to make decisions to achieve the most optimal results. It mimics the trial-and-error learning process that humans use to achieve their goals.
- Software actions that work towards the goal are reinforced, while actions that detract from the goal are ignored.

Reinforcement learning



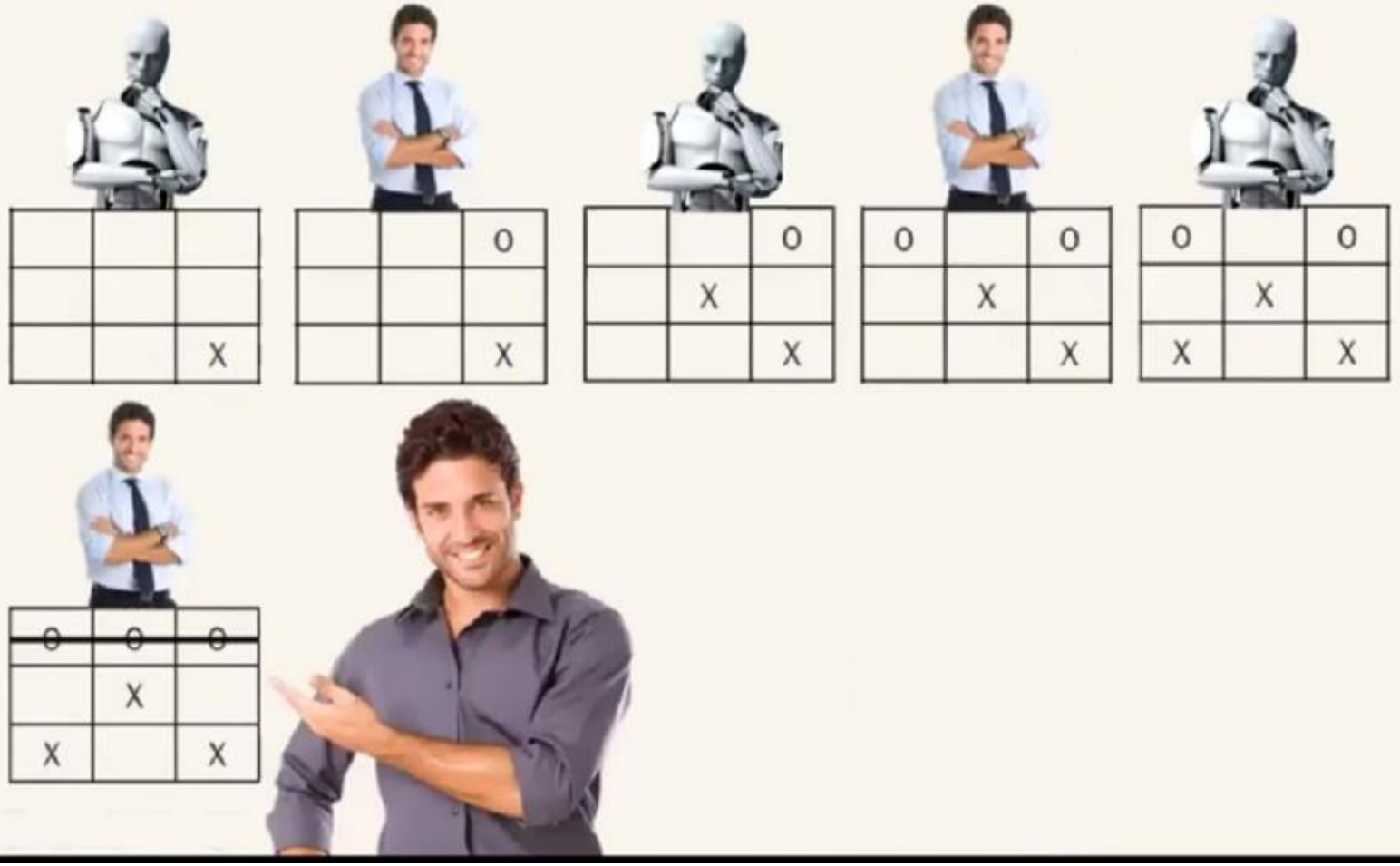

Lets Play Tic Tac Toe

Game 01



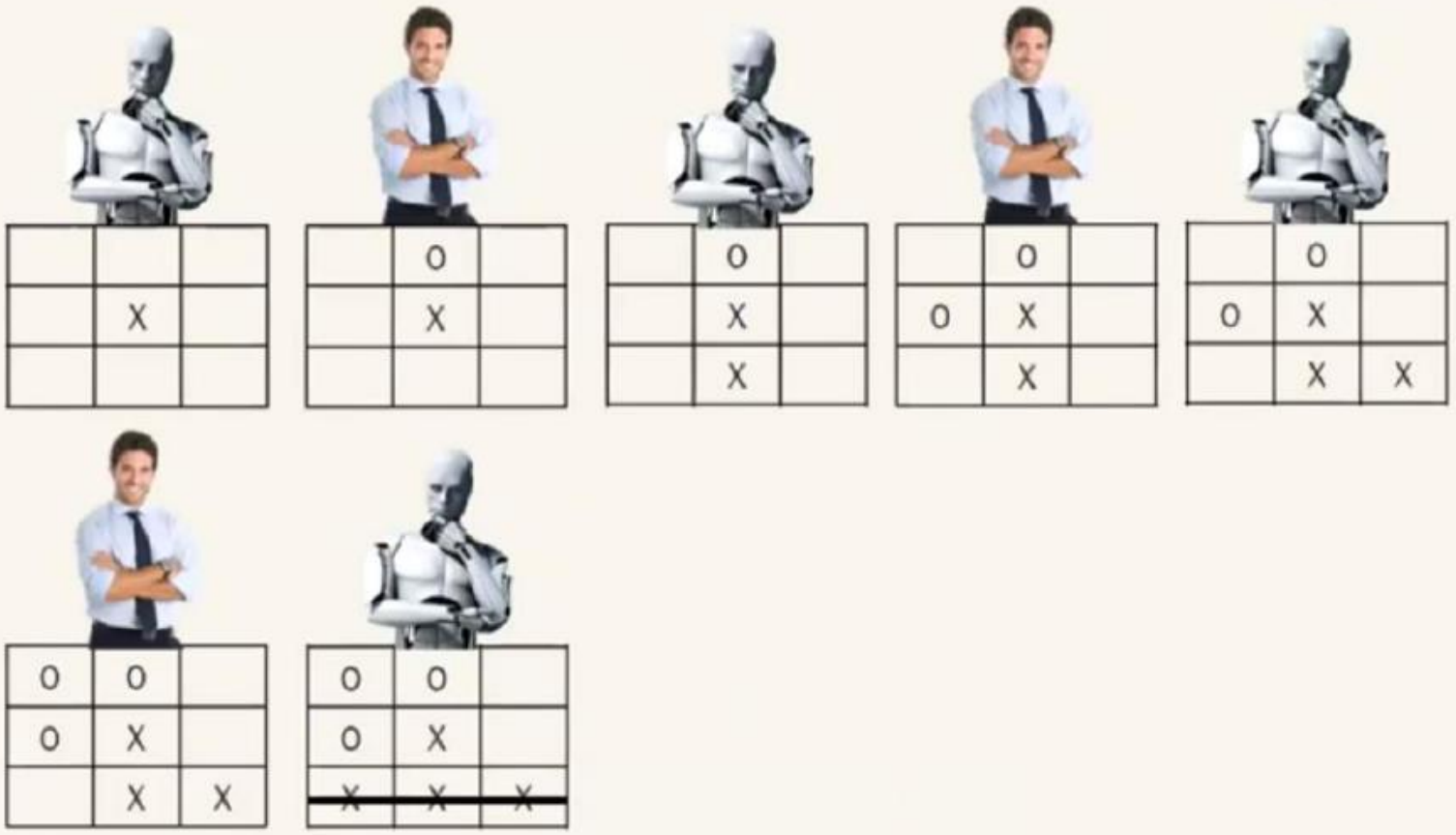

Learning through
mistakes

Game 02




Learning through
mistakes

Game 03




Never make past
mistakes

Applications of Reinforcement learning

- **Game Playing:** Notably used in AlphaGo, AlphaZero, and OpenAI's Dota 2 bots.
- **Robotics:** For autonomous control and learning complex tasks.
- **Recommendation Systems:** Personalized recommendations based on user interactions.
- **Finance:** Algorithmic trading and portfolio management.
- **Healthcare:** Personalized treatment planning and drug discovery.

https://towardsdatascience.com/a-review-of-recent-reinforcment-learning-applications-to-healthcare-1f8357600407

Data in Machine learning

- Data is considered the core of ML.
- Machine learning models learn from data, without data Machine learning is no use.
- To provide data to the machine learning model we need to create datasets for the Machine learning model.

Data in Machine learning

There are two main types of data which are:

Qualitative

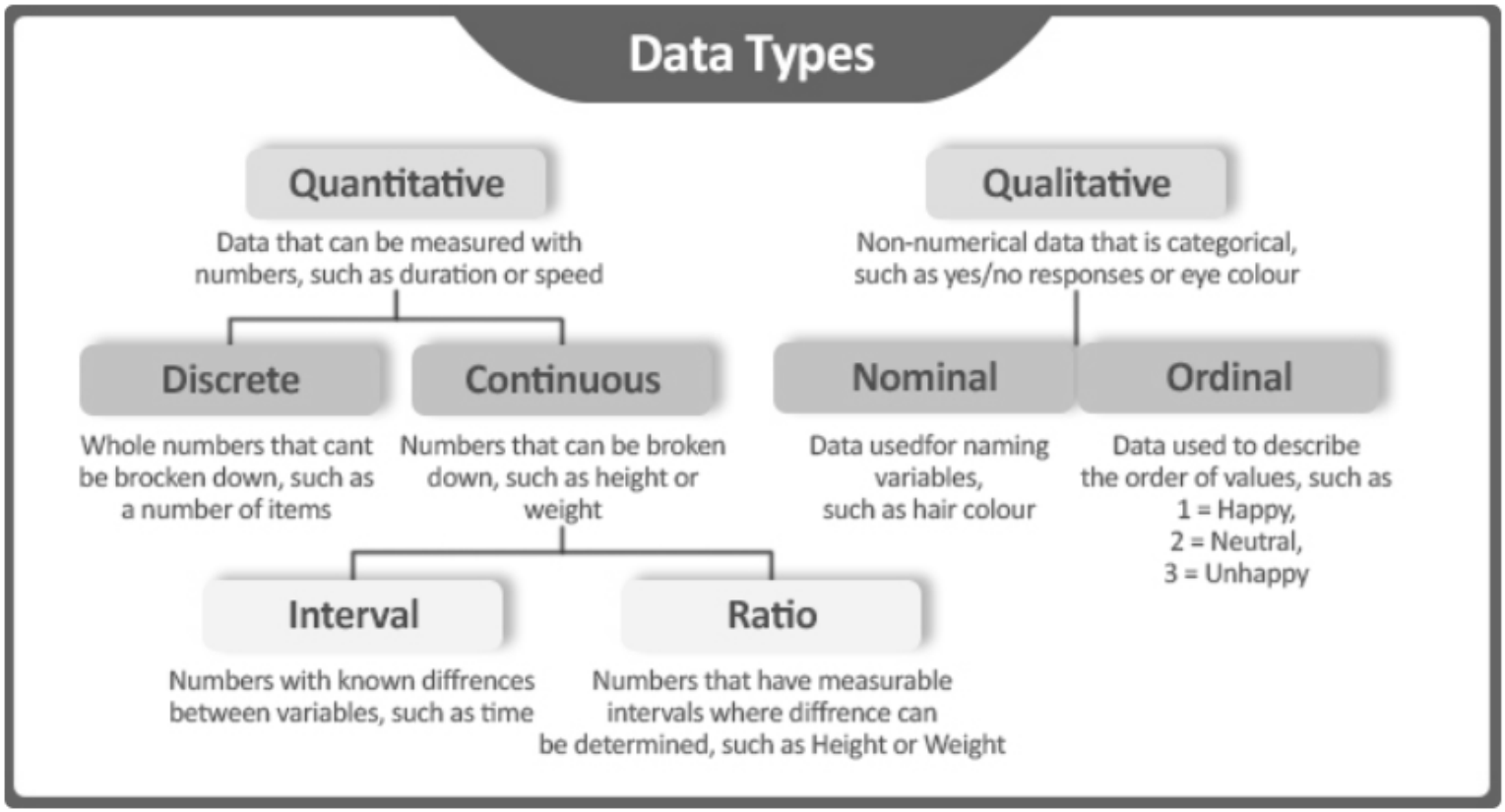
Quantitative

1. Qualitative data

It means the data of categorical type.

For example- the color of the shirt, the shape of the pizza.

It has two types: Nominal and Ordinal.


Data Types
Quantitative
Data that can be measured with numbers, such as duration or speed
Qualitative
Non-numerical data that is categorical, such as yes/no responses or eye colour
Discrete
Whole numbers that cant be brocken down, such as a number of items
Continuous
Numbers that can be broken down, such as height or weight
Nominal
Data usedfor naming variables, such as hair colour
Ordinal
Data used to describe the order of values, such as 1 = Happy, 2 = Neutral, 3 = Unhappy
Interval
Numbers with known diffrences between variables, such as time
Ratio
Numbers that have measurable intervals where diffrence can be determined, such as Height or Weight

Properties of Data

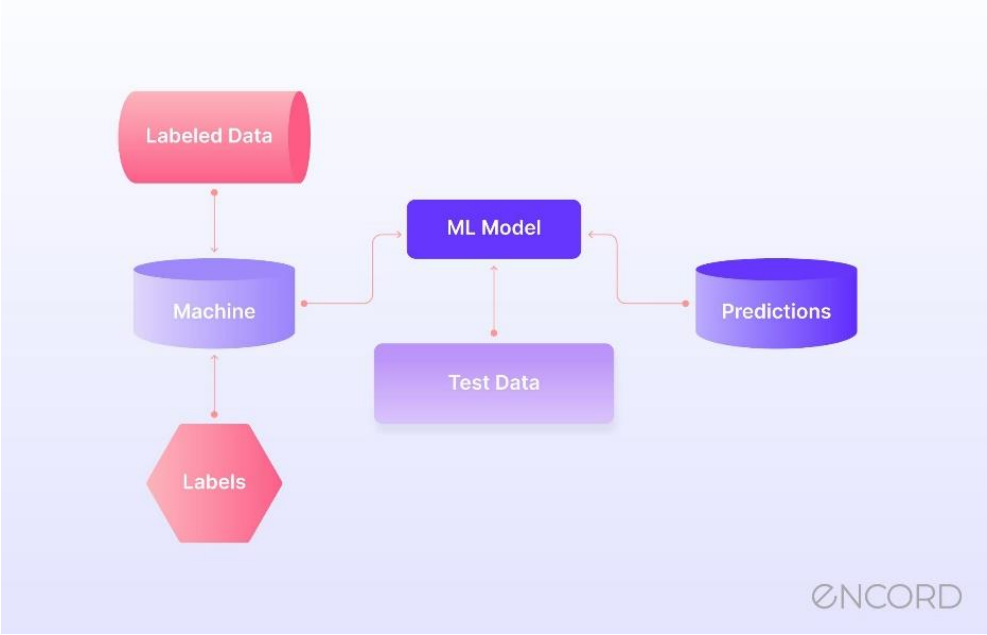
- **Volume**: Scale of Data.
- **Variety**: Different forms of data.
- **Velocity**: Rate of data streaming and generation.
- **Value**: Meaningfulness of data in terms of information.
- **Veracity**: Certainty and correctness in data.

Datasets for Machine Learning

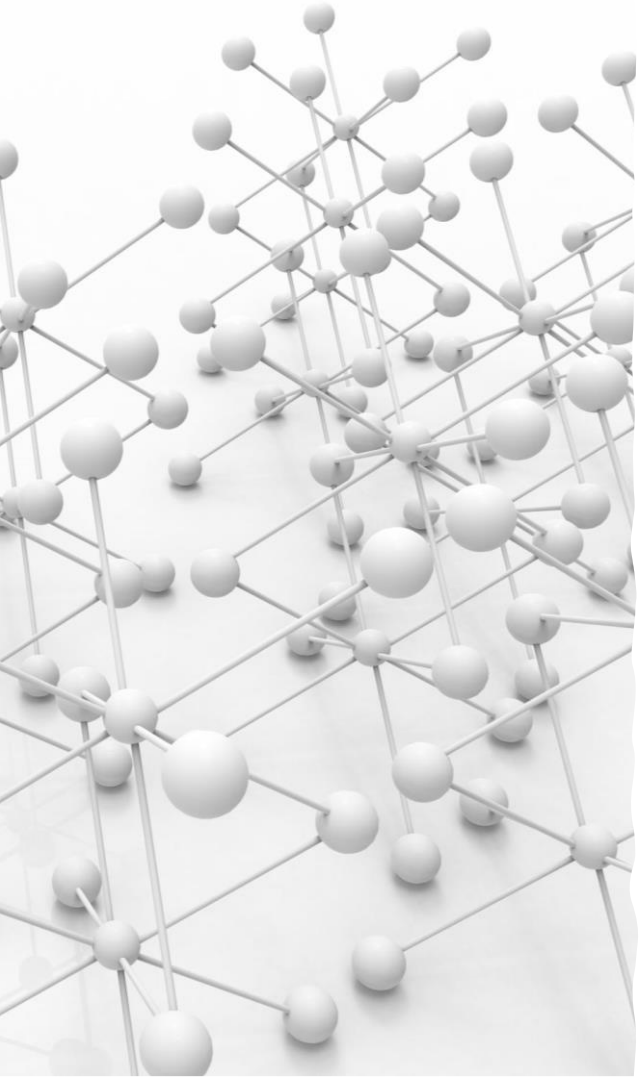
- Datasets are a collection of data.
- A dataset stores several entries for a particular type of observation with n number of columns.
- For example- if we have a dataset of movies, it contains columns- movie_name, movie_cast and movie_rating then it will have several entries/rows containing different movies.
- CSV files are the most common dataset format for Machine learning purposes.
- For Machine learning, data collection or dataset generation is an important task.
- We can create a dataset in two ways:
  - Collecting data on our own through surveys
  - Downloading the most appropriate dataset from the internet.

Key Concepts and Terminology

- **Model:** A mathematical representation of a process
- **Training:** Process of teaching a model using data
- **Testing:** Evaluating the model's performance on unseen data
- **Features:** Input variables used to make predictions
- **Labels:** Output or target variables



Popular Algorithms



- **Linear Regression** - Predicts continuous values
- **Logistic Regression** - Predicts binary outcomes
- **Decision Trees** - Tree-like model of decisions
- **Support Vector Machines (SVM)** - Classifies data by finding the best boundary
- **Neural Networks** - Models inspired by the human brain
- **K-Means Clustering** - Groups data into K clusters

Applications of Machine Learning

Healthcare - Disease diagnosis, personalized treatment

Finance - Fraud detection, algorithmic trading

Retail - Recommendation systems, inventory management

Transportation - Autonomous vehicles, route optimization

Entertainment - Content recommendation, audience analysis

Challenges in Machine Learning

- **Data Quality:** Garbage in, garbage out
- **Overfitting**: The model performs well on training data but poorly on new data
- **Interpretability**: Understanding how and why models make decisions
- **Scalability:** Handling large volumes of data
- **Bias and Fairness:** Ensuring ethical use and avoiding discrimination

Future Directions

Automated Machine Learning (AutoML): Making ML accessible to non-experts

Explainable AI (XAI): Improving model transparency

Edge Computing: Running ML models on local devices

Quantum Machine Learning: Leveraging quantum computing for ML

Conclusion

Machine Learning is transforming various industries.

Understanding the basics and different types of ML is crucial.

Real-world applications demonstrate the impact of ML.

Ongoing challenges and future directions highlight the field's dynamic nature.

THANK YOU