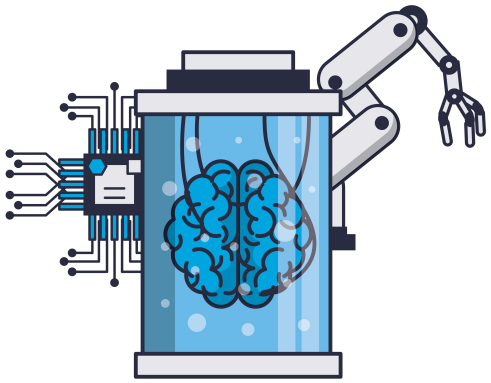
MACHINE LEARNING

Machine Learning is the process of creating models that can perform a certain task without the need for a human explicitly programming it to do something

Your smartphone, your house, your car, and your bank all use **artificial intelligence** on a daily basis.

Sometimes it's **easy to understand**, when you ask Siri, Cortana or OK Google to get you directions.

Sometimes it's **less obvious**, like when you make an abnormal purchase on your credit card and don't get a fraud alert from your bank.



AI, **ML** and **DL** are everywhere and Data Science is the interdisciplinary field of methods to extract the knowledge needed.

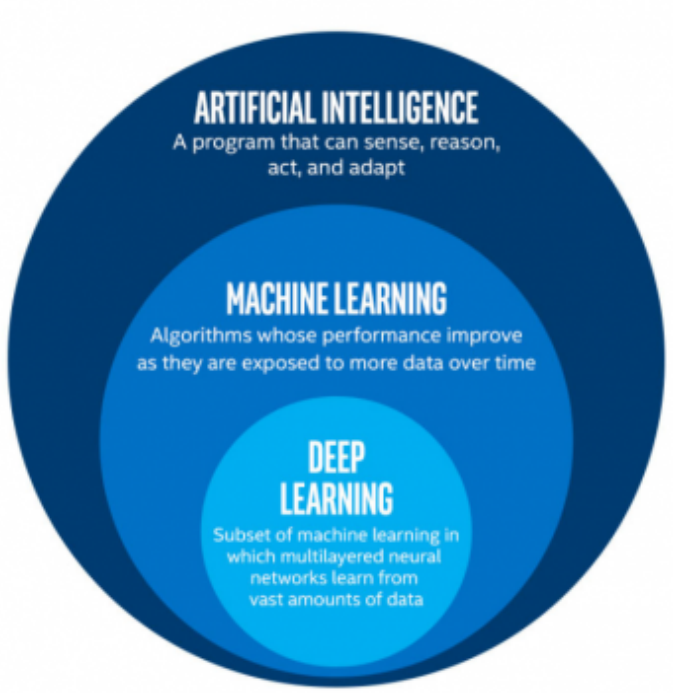
These technologies are making a huge difference in our lives every day and evolving fast by a magnitude of people working to improve them consistently.

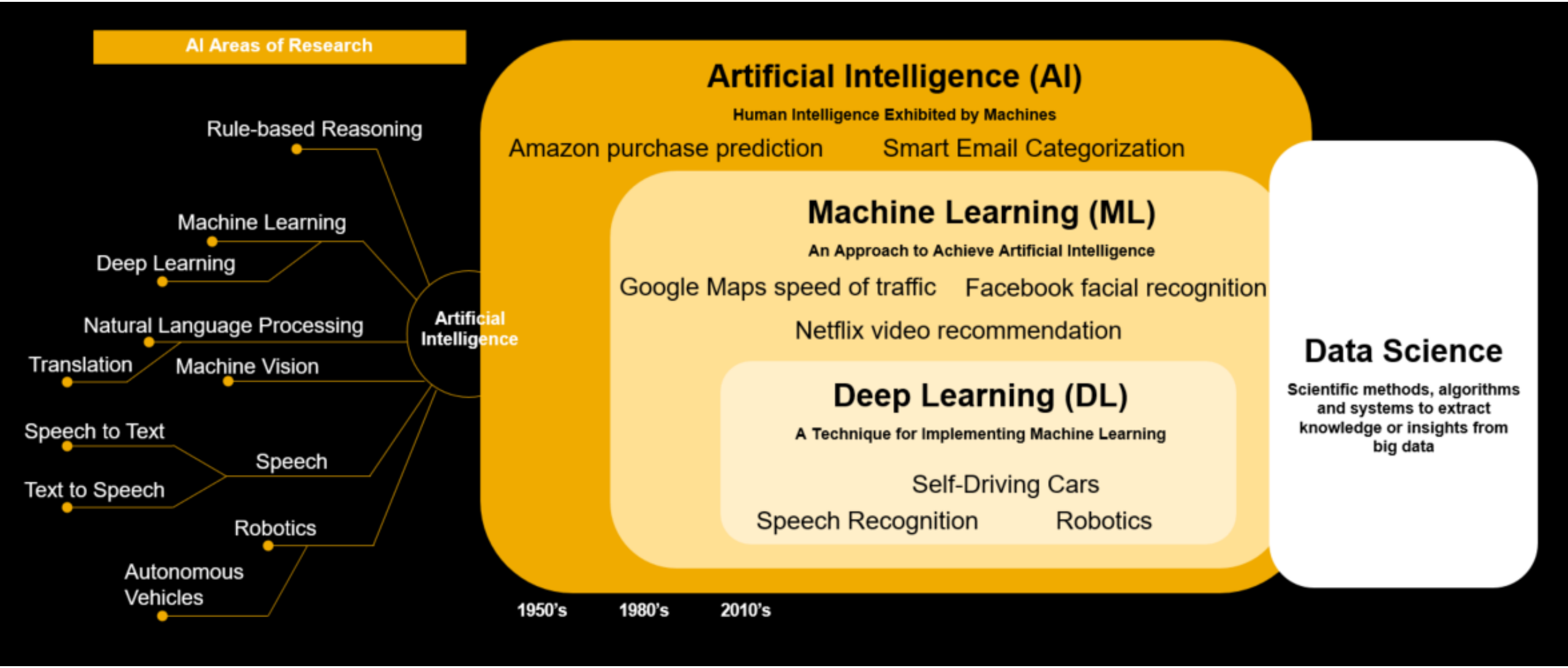
Definitions

Artificial Intelligence (*AI*) refers to the **simulation of human intelligence** processes by machines, including learning, reasoning and self-correction

Machine learning (*ML*) is an application of *AI* generating systems that can learn and **improve without being programmed**

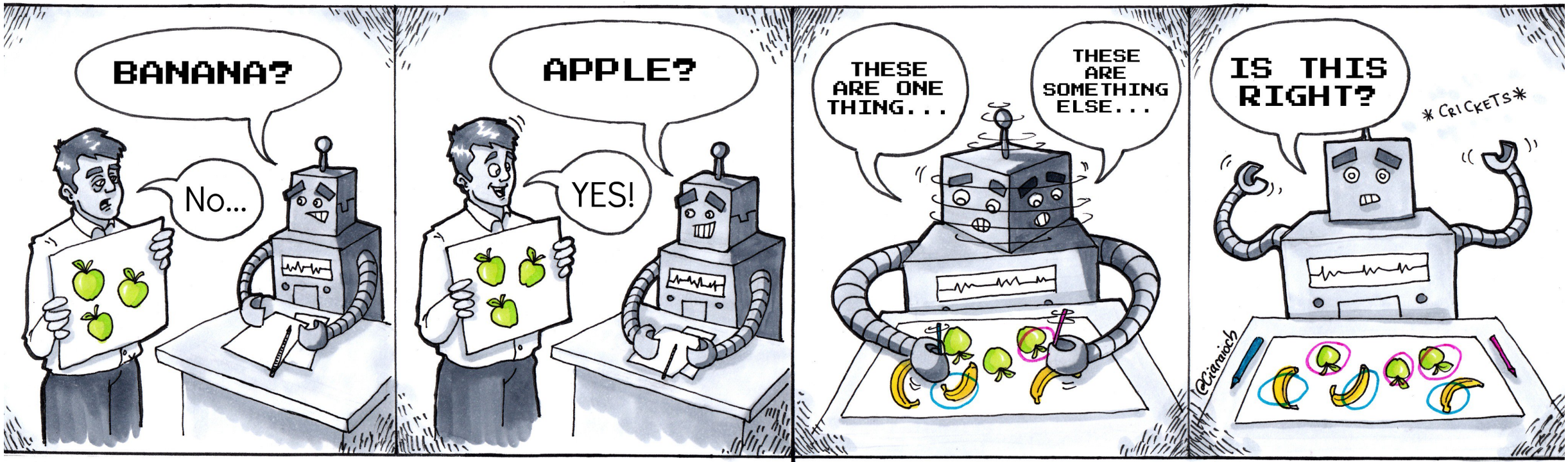
Deep Learning (DL) is a subset of Machine Learning and Artificial Intelligence. The term refers to a **particular approach** used for creating and training neural networks that are considered highly promising decision-making nodes




AI Areas of Research
Rule-based Reasoning
Machine Learning
Deep Learning
Natural Language Processing
Translation
Machine Vision
Speech to Text
Speech
Text to Speech
Robotics
Autonomous Vehicles
Artificial Intelligence
Artificial Intelligence (AI)
Human Intelligence Exhibited by Machines
Amazon purchase prediction
Smart Email Categorization
Machine Learning (ML)
An Approach to Achieve Artificial Intelligence
Google Maps speed of traffic
Facebook facial recognition
Netflix video recommendation
Deep Learning (DL)
A Technique for Implementing Machine Learning
Self-Driving Cars
Speech Recognition
Robotics
Data Science
Scientific methods, algorithms and systems to extract knowledge or insights from big data
1950's
1980's
2010's

Machine Learning algorithms

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning



Supervised Learning

Unsupervised Learning


Dynamic pricing
Predicting flight delay
01
TRAVEL
Disease prediction
Medication effectiveness
Upselling
Cross selling
Predicting lifetime valueof customer
02
MARKETING
Sentiment analysis
Digital marketing
Churn
03
HEALTHCARE
Discount offering
Demand forecasting
05
SALES
04
SOCIAL MEDIA
Self driving cars
06
AUTOMATION
Pilotless aircrafts, drones
Claims prediction
07
CREDIT & INSURANCE
Fraud & risk detection

Supervised Learning

Supervised Learning

The area of Machine Learning where you have a **set of independent variables** which helps you to analyse the **dependent variable** and the **relation between them**

Whatever you want to predict is called as **Dependent Variable**, while variables that you use to predict are called as **Independent Variables**

```
You want to predict the age of the person
based on the person's height and weight,
then height and weight will be the independent
variables, while age will be the dependent variable
```

Supervised learning is the most popular paradigm for machine learning.
It is the easiest to understand and the simplest to implement

Types of Supervised Learning

Supervised Learning can be classified into 2 types

- Regression
- Classification

Regression

Regression is the kind of Supervised Learning that learns from the Labelled Datasets and is then able to predict a continuous-valued output for the new data given to the algorithm.

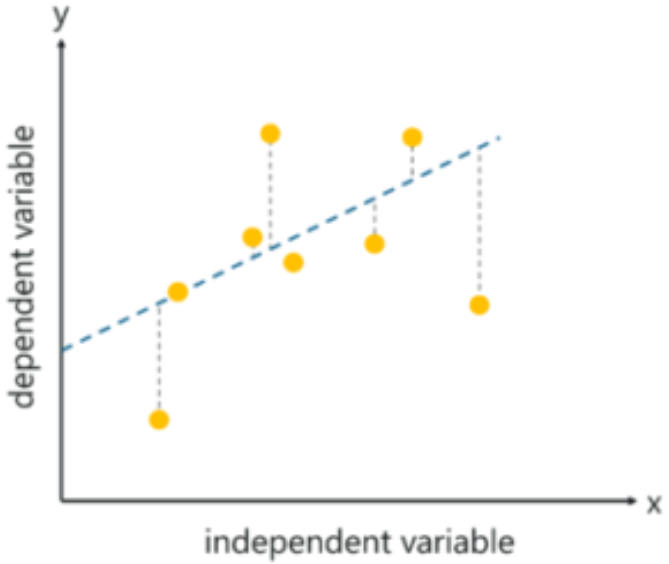
It is used whenever the output required is a number such as money or height etc.

- Linear Regression
- Logistic Regression

Linear Regression

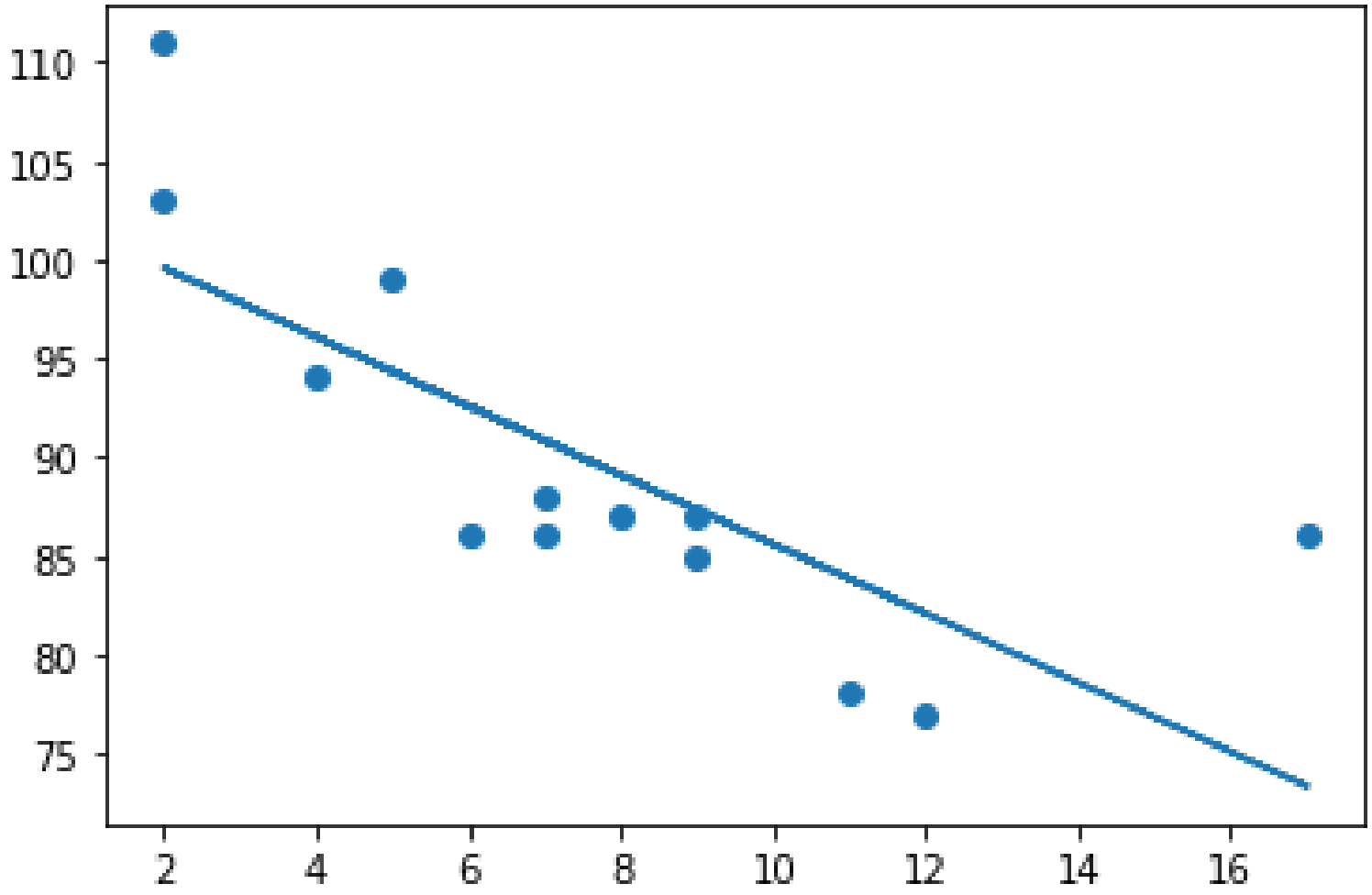
This algorithm assumes that there is a linear relationship between the 2 variables, Input (X) and Output (Y), of the data it has learnt from. The Input variable is called the Independent Variable and the Output variable is called the Dependent Variable.

When unseen data is passed to the algorithm, it uses the function, calculates and maps the input to a continuous value for the output.



Linear Regression Python

- Linear_Regression_Bad.ipynb
- Linear_Regression_Good.ipynb

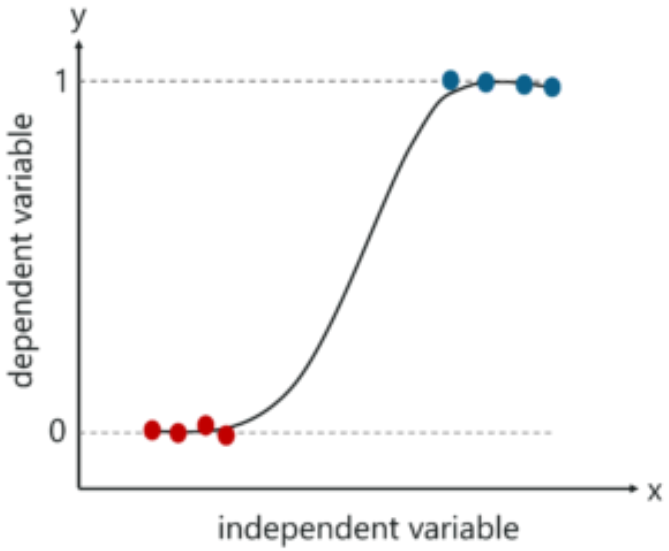


Logistic Regression

This algorithm predicts discrete values for the set of Independent variables that have been passed to it.

It does the prediction by mapping the unseen data to the logit function that has been programmed into it.

The algorithm predicts the probability of the new data and so it’s output lies between the range of 0 and 1.



Classification

Classification is the kind of machine learning where the algorithm needs to **map the new data** that is obtained to any one of the **2 classes that you have in the dataset**.

The classes need to be **mapped to either 1 or 0 **which in real-life translated to *Yes* or *No*, *Rains* or *Does Not Rain*

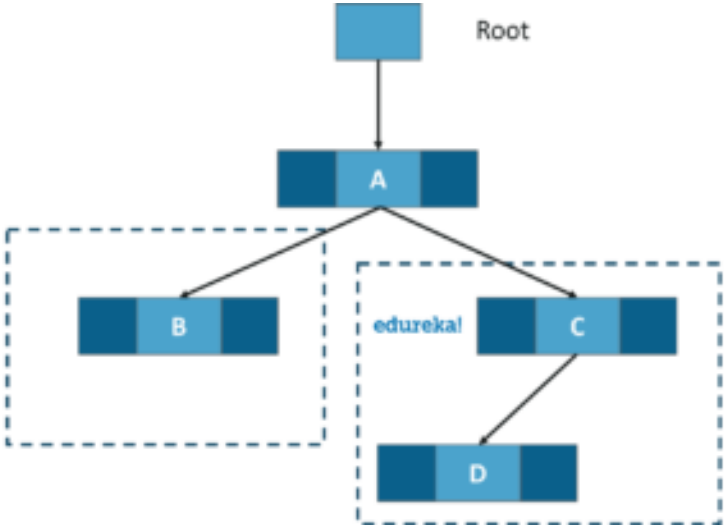
The output will be **either one of the classes and not a number** as it was in Regression.

- Decision Tree
- Naive Bayes Classifier
- Support Vector Machines

Decision Tree

Decision Trees classify based on the feature values. They use the method of Information Gain and find out which feature of the dataset gives the best of information, make that as the root node and so on till they are able to classify each instance of the dataset.

Every branch in the Decision Tree represents a feature of the dataset. They are one of the most widely used algorithms for classification.



Naive Bayes Classifier

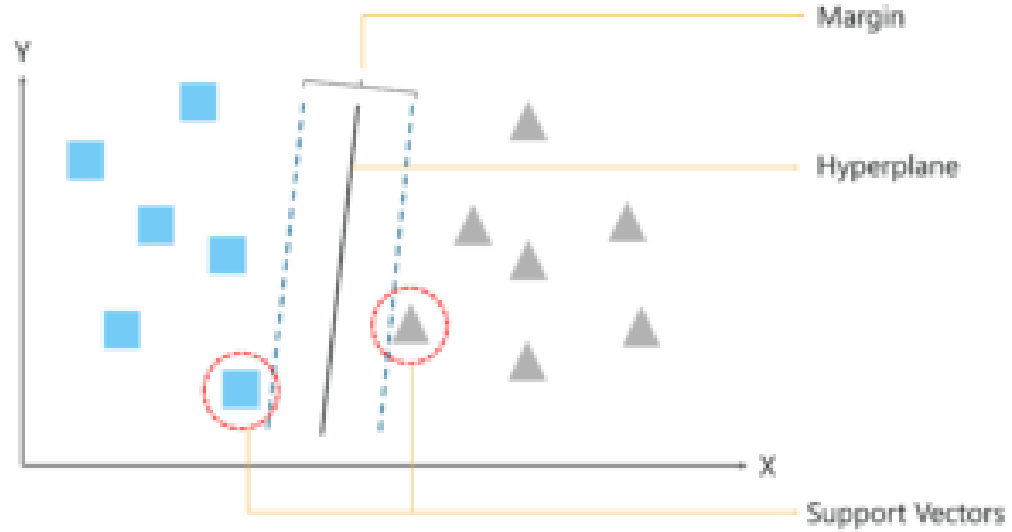
Naive Bayes algorithms assume that the features of the dataset are all independent of each other.

Works great on large datasets.

Support Vector Machines

Support Vector Machines algorithms are based on the statistical learning theory of *Vapnik*.

They use Kernal functions which are a central concept for most of the learning tasks.
These algorithms create a hyper-plane that is used to classify the two classes from each other.



Vapnik–Chervonenkis theory - https://en.wikipedia.org/wiki/Vapnik–Chervonenkis_theory

Supervised Learning - Example

Predicting house prices

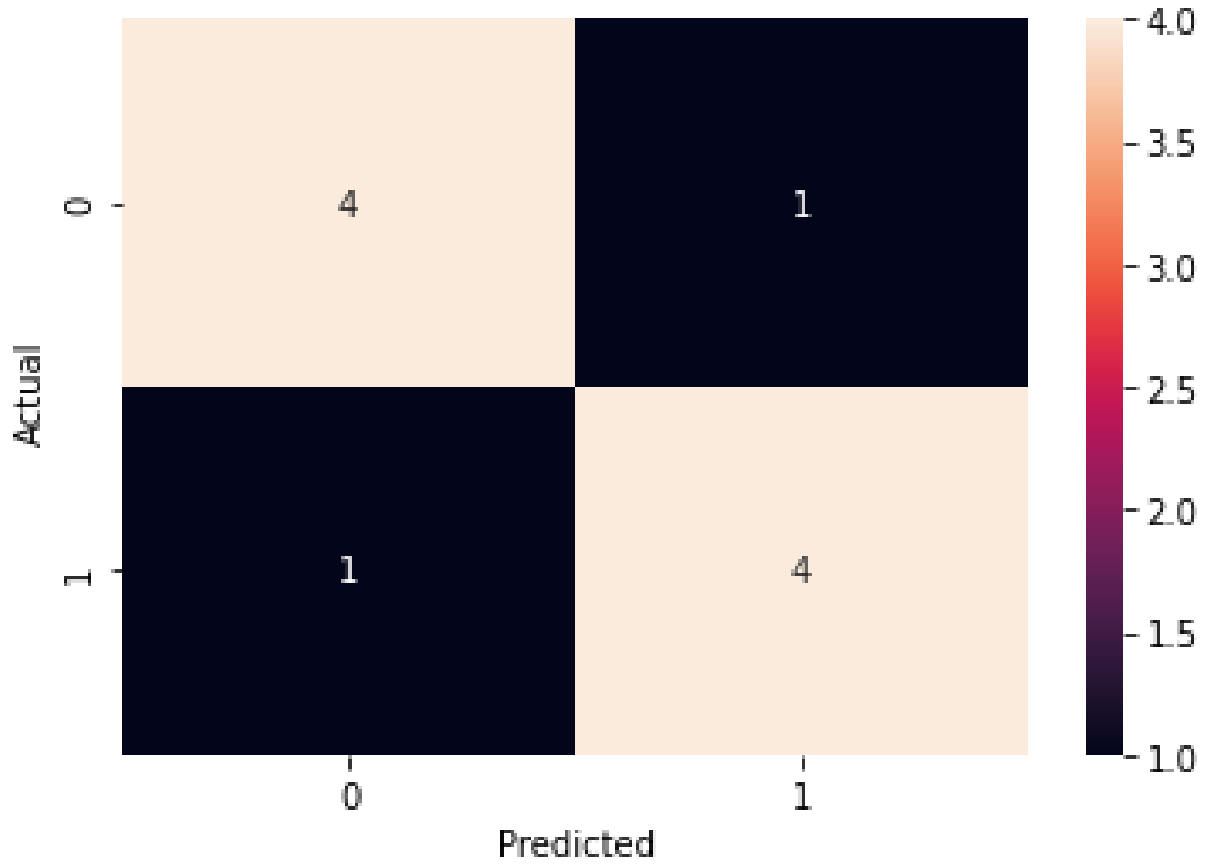
First, we need data about the houses: *square footage, number of rooms, features, whether a house has a garden or not, and so on*.

We then need to know the prices of these houses, i.e. the corresponding labels.

By leveraging data coming from thousands of houses, their features and prices, **we can now train a supervised machine learning model to predict a new house's price** based on the examples observed by the model.

Logistic Regression Python

- Logistic_Regression_Application.ipynb
- Logistic_Regression_Diabetes.ipynb



Unsupervised Learning

Unsupervised Learning

You **don’t have any dependent variable**. You just have collection of variables and try to find out similarity between them and classify them into clusters

Unsupervised Learning - Examples

- Customer segmentation, or understanding different customer groups around which to build marketing or other business strategies
- Recommender systems, which involve grouping together users with similar viewing patterns in order to recommend similar content.
- Anomaly detection, including fraud detection or detecting defective mechanical parts - *predictive maintenance*

Reinforcement Learning

Reinforcement Learning

It is the training of machine learning models to make a sequence of decisions.

The *machine* learns to achieve a goal in an uncertain, potentially complex environment.
The *machine* employs trial and error to come up with a solution to the problem.

Reinforcement Learning - Example

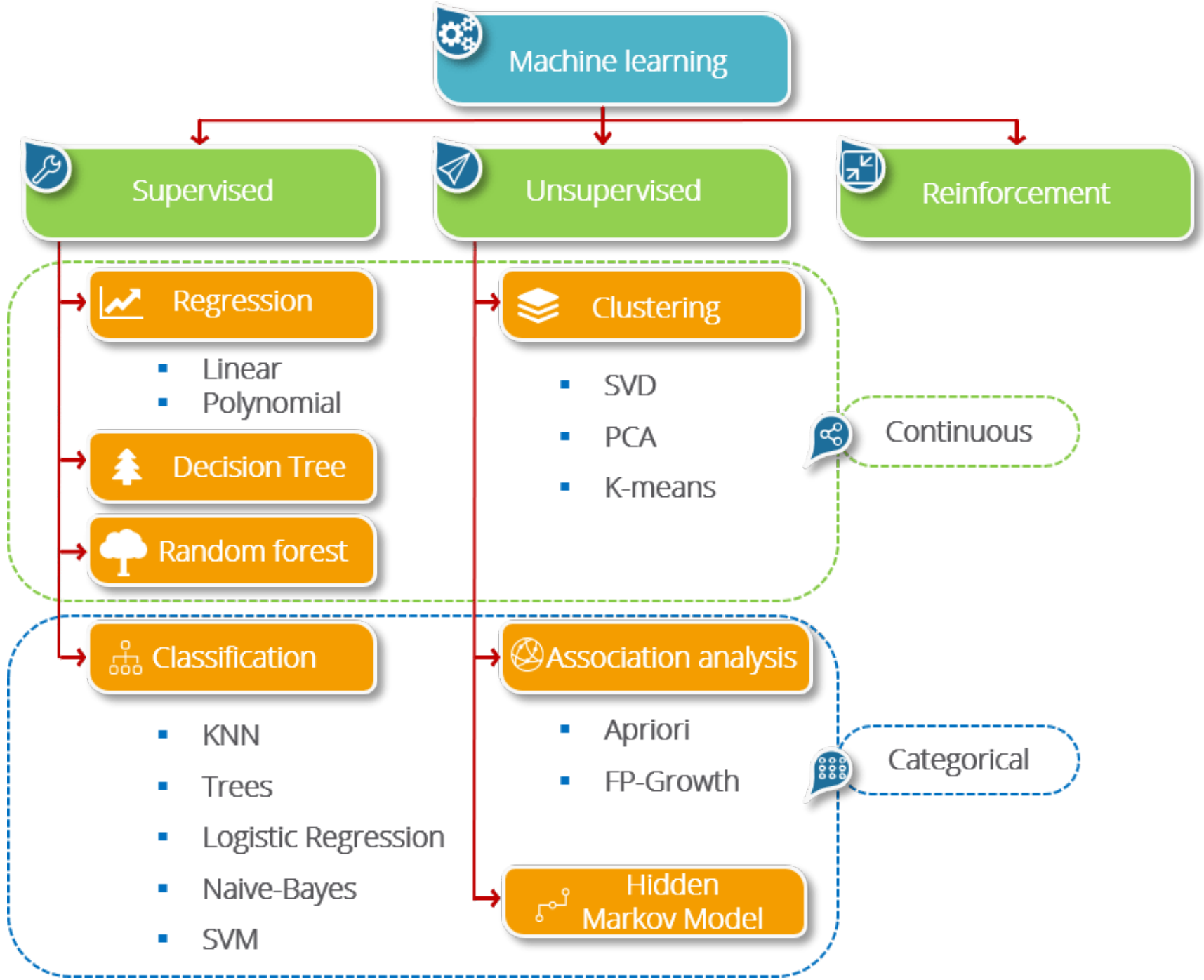
Facebook has developed an open-source reinforcement learning platform—Horizon. The platform uses reinforcement learning to optimize large-scale production systems.

Facebook has used Horizon internally

- To personalize suggestions
- Deliver more meaningful notifications to users
- Optimize video streaming quality

Read more

- https://engineering.fb.com/ml-applications/horizon
- https://research.fb.com/publications/horizon-facebooks-open-source-applied-reinforcement-learning-platform


Machine learning
Supervised
Unsupervised
Reinforcement
Regression
Linear
Polynomial
Decision Tree
Random forest
Clustering
SVD
PCA
K-means
Continuous
Classification
KNN
Trees
Logistic Regression
Naive-Bayes
SVM
Association analysis
Apriori
FP-Growth
Hidden Markov Model
Categorical

Link

- https://www.ibm.com/cloud/learn/unsupervised-learning