Machine Learning

-Deeksha Reddy
ENG19CS0081

CONTENTS OF THIS PRESENTATION

1. Introduction to Machine Learning
2. Supervised, Unsupervised and Reinforcement Learning
3. Data WareHouse and Data Mining
4. Different ML algorithms
5. Application of ML in different domains - Healthcare and telecom

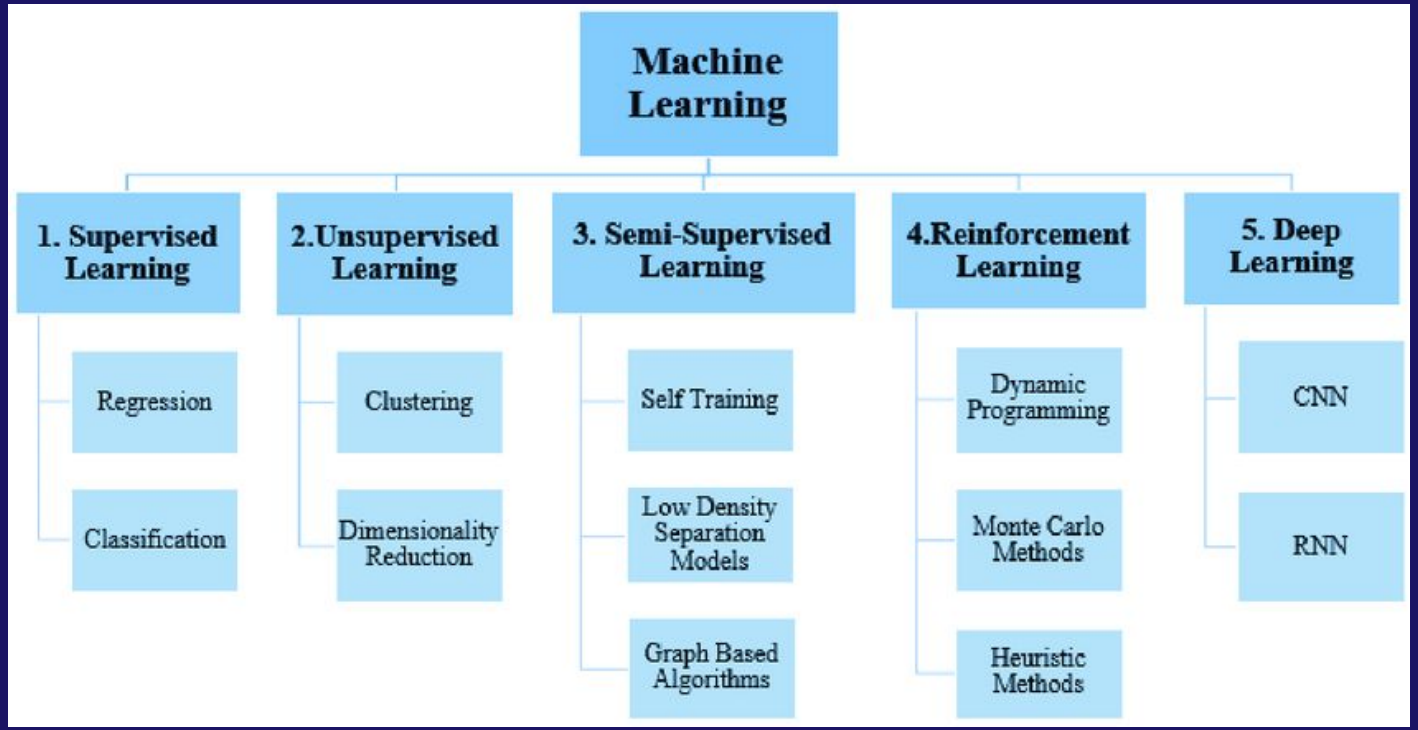
INTRODUCTION

Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning algorithms use historical data as input to predict new output values.

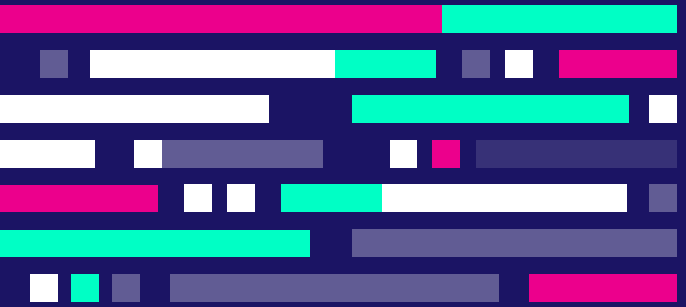
“Machine intelligence is the last invention that humanity will ever need to make.”

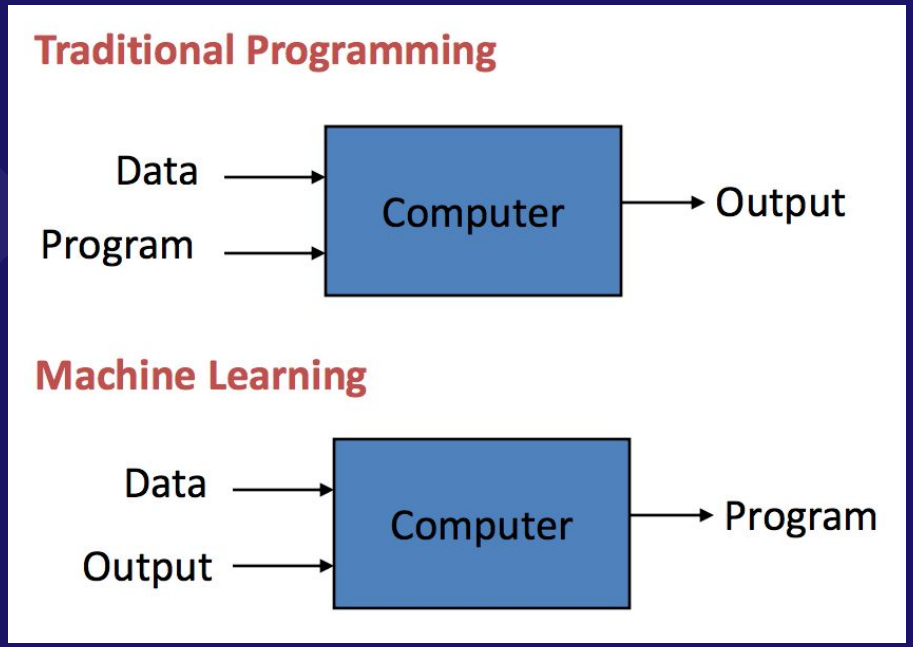
—NICK BOSTROM

TYPES OF MACHINE LEARNING



Traditional Programming vs Machine Learning

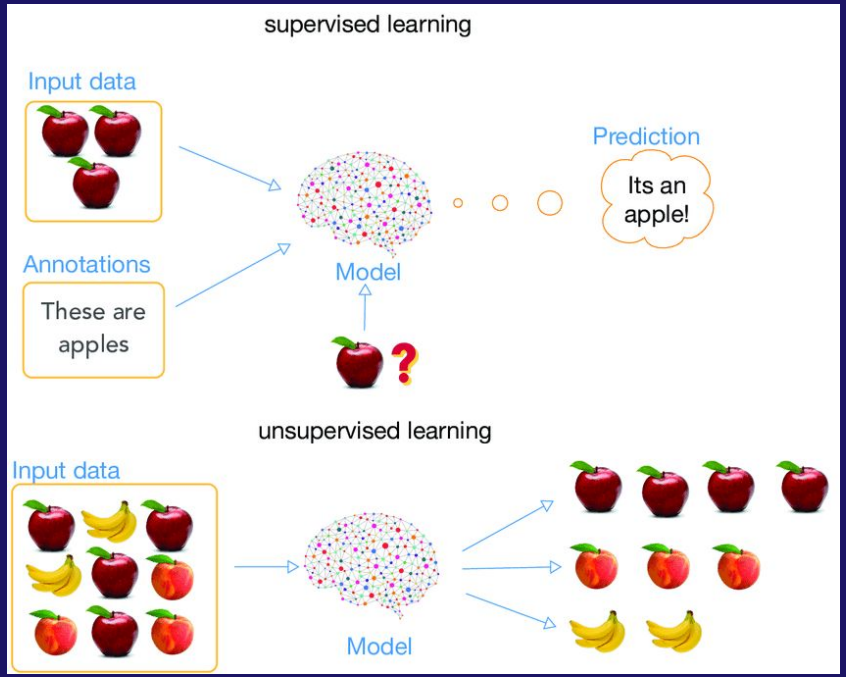




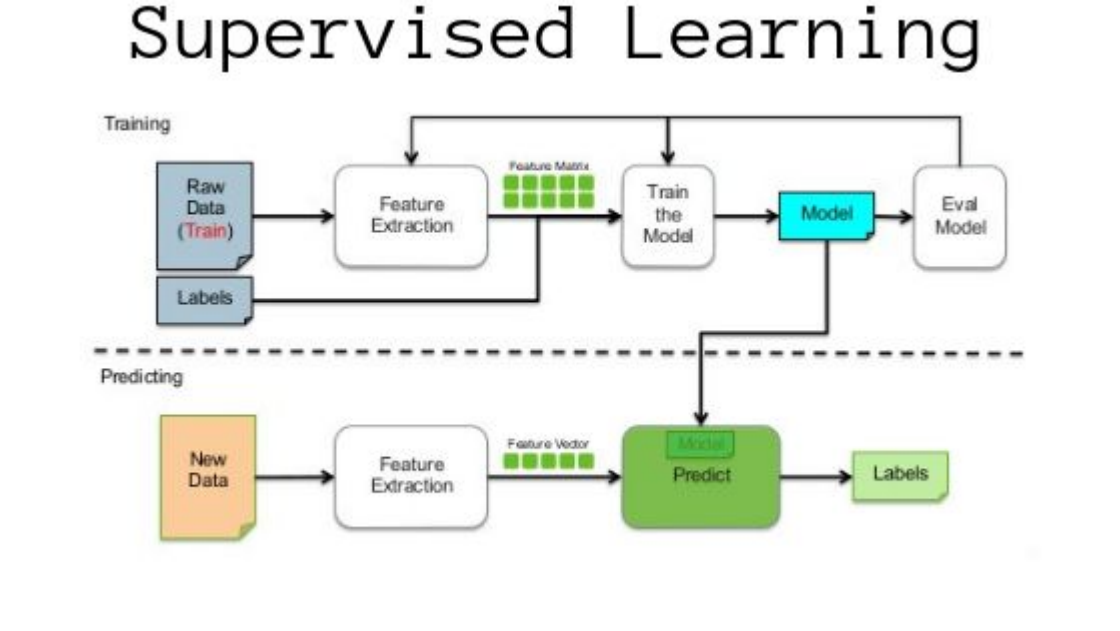
Some basic terms

- Inputs are called features.
  A feature is one column of the data in your input set. For instance, if you're trying to predict the type of pet someone will choose, your input features might include age, home region, family income, etc.
- Outputs are labels.
  The label is the final choice, such as dog, fish, iguana, rock, etc.
- Functions are models.
  A machine learning model is a file that has been trained to recognize certain types of patterns.

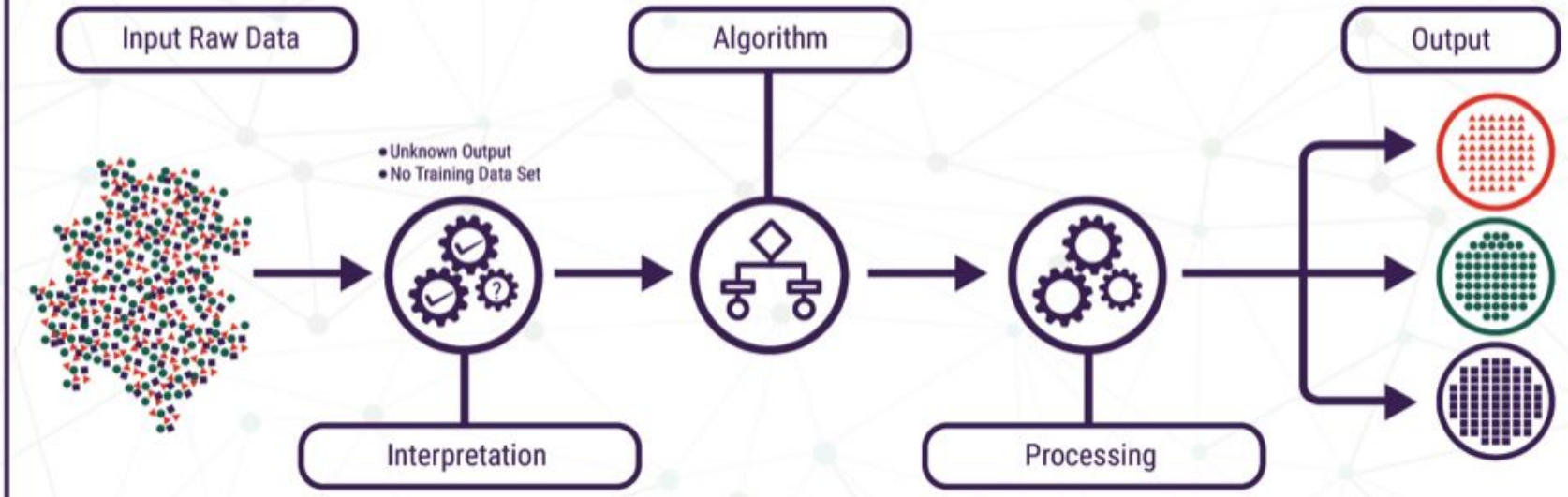
Supervised and Unsupervised Learning



SUPERVISED LEARNING

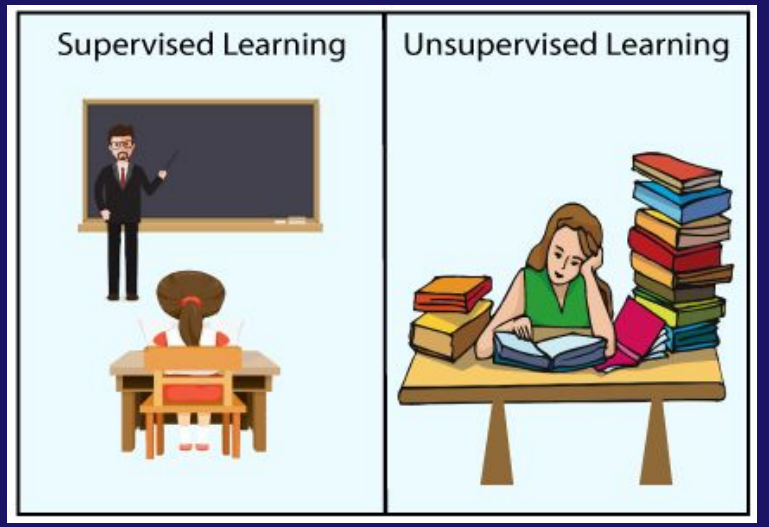


UNSUPERVISED LEARNING

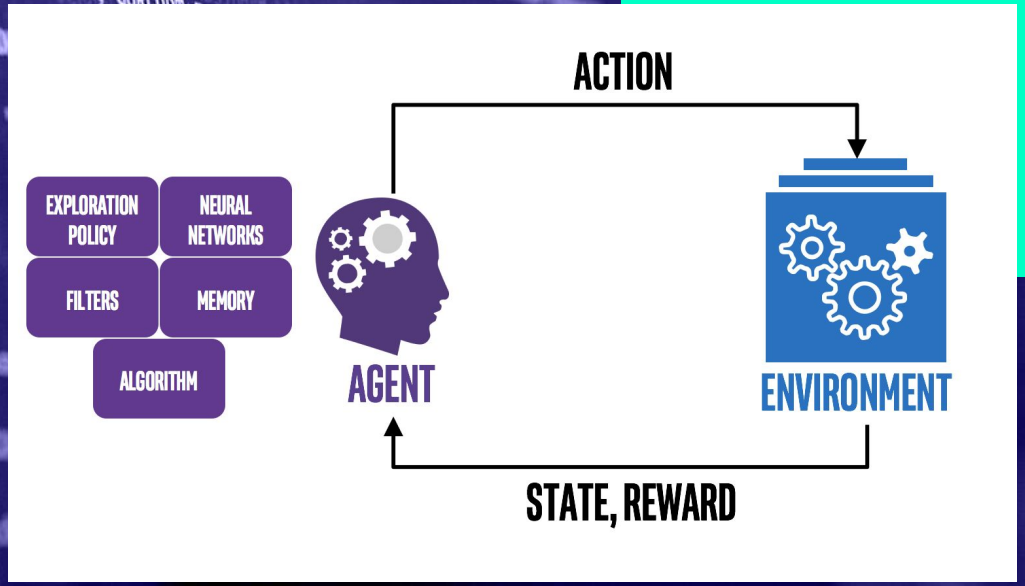


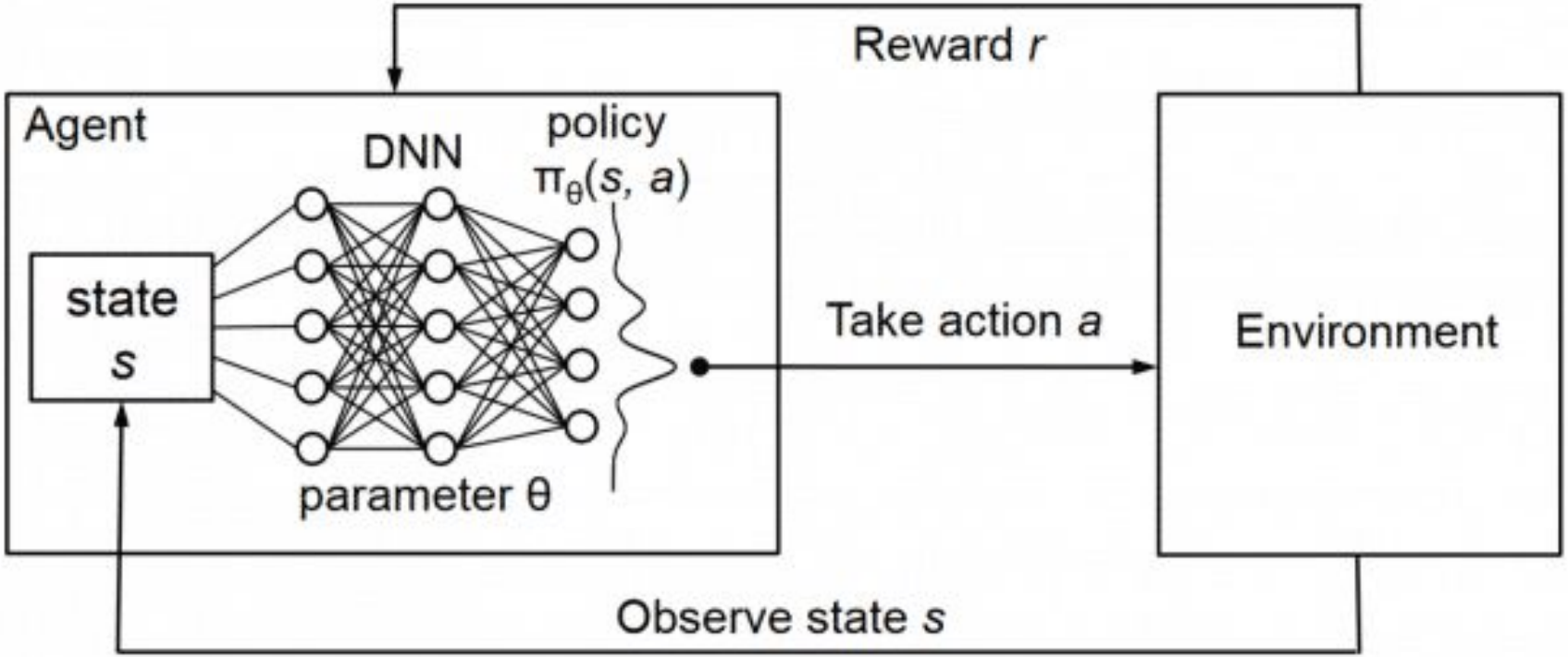
Supervised learning needs supervision to train the model, which is similar to as a student learns things in the presence of a teacher. Supervised learning can be used for two types of problems: Classification and Regression.

So unlike supervised learning, here we will not provide any supervision to the model. We will just provide the input dataset to the model and allow the model to find the patterns from the data.Unsupervised learning can be used for two types of problems: Clustering and Association.



Reinforcement Learning




Reward r
Agent
DNN
policy
$\pi_\theta(s, a)$
state
s
parameter θ
Take action a
Environment
Observe state s

Types of Machine Learning

- At a glance

Supervised Learning

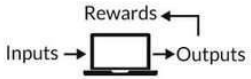
- Makes machine learn explicitly
- Data with clearly defined output is given
- Direct feedback is given
- Predicts outcome/ future
- Resolves classification & regression problems

Unsupervised Learning

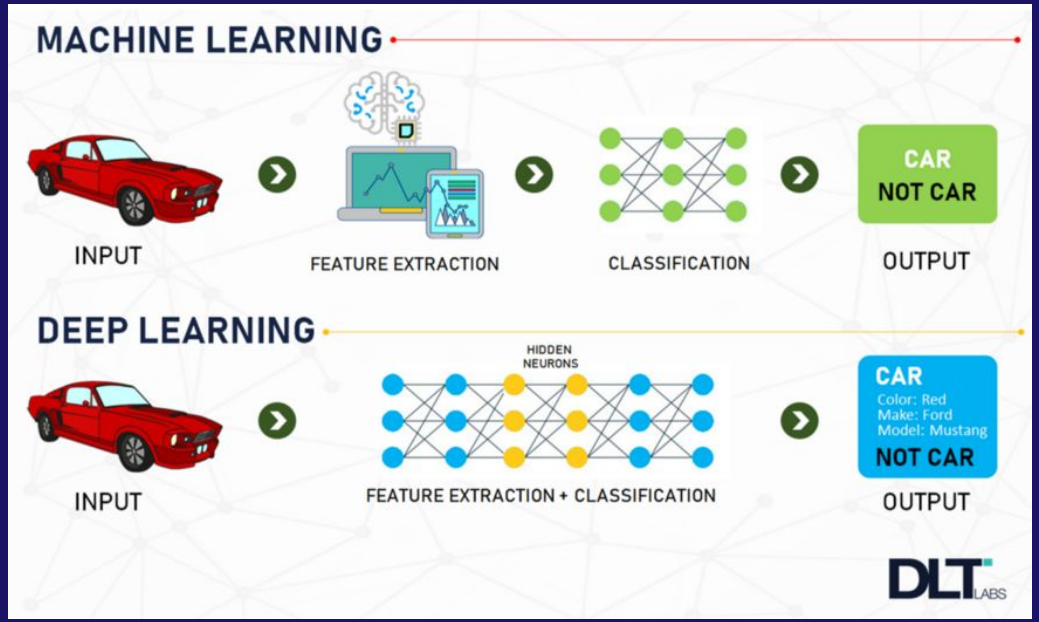
- Machine understands the data (Identifies patterns/ structures)
- Evaluation is qualitative or indirect
- Does not predict / find anything specific

Reinforcement Learning

- An approach to AI
- Reward based learning
- Learning from +ve & -ve reinforcement
- Machine learns how to act in a certain environment
- To maximize rewards



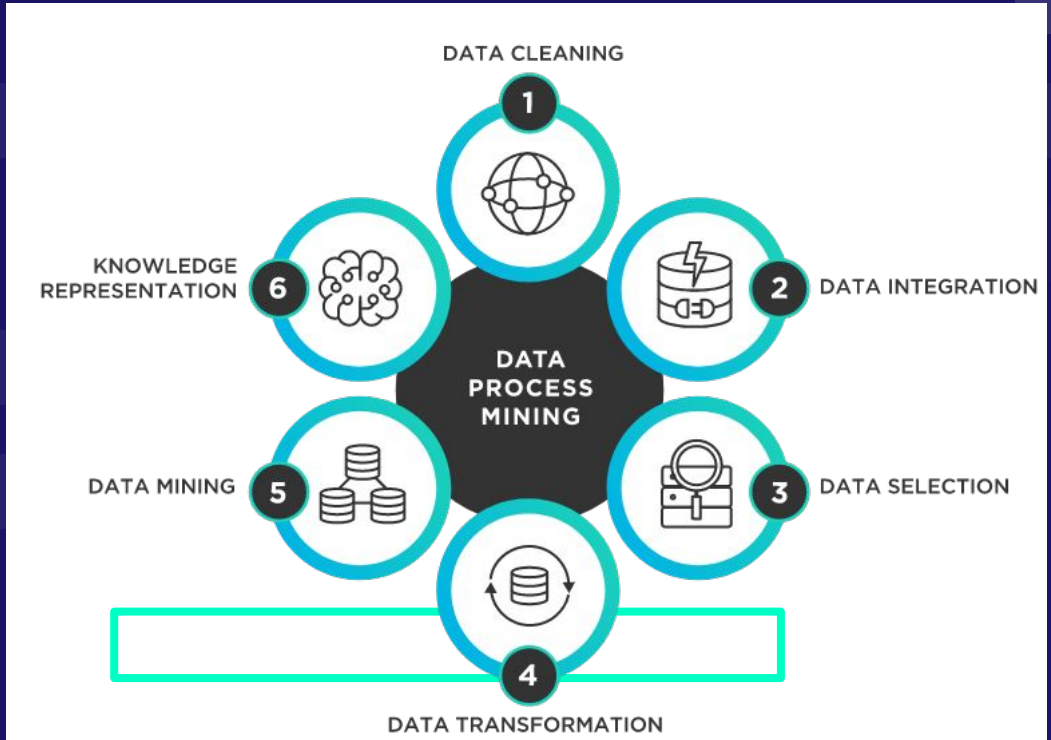
OVERVIEW DIAGRAM



Data mining

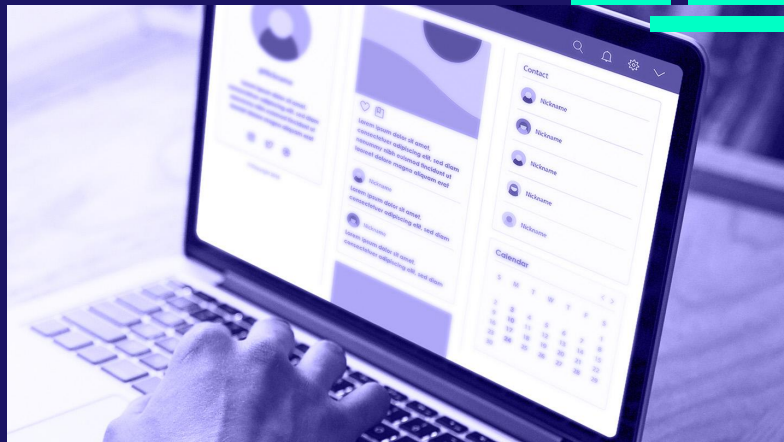
Data mining is a process used by companies to turn raw data into useful information. By using software to look for patterns in large batches of data, businesses can learn more about their customers to develop more effective marketing strategies, increase sales and decrease costs. Data mining depends on effective data collection, warehousing, and computer processing.

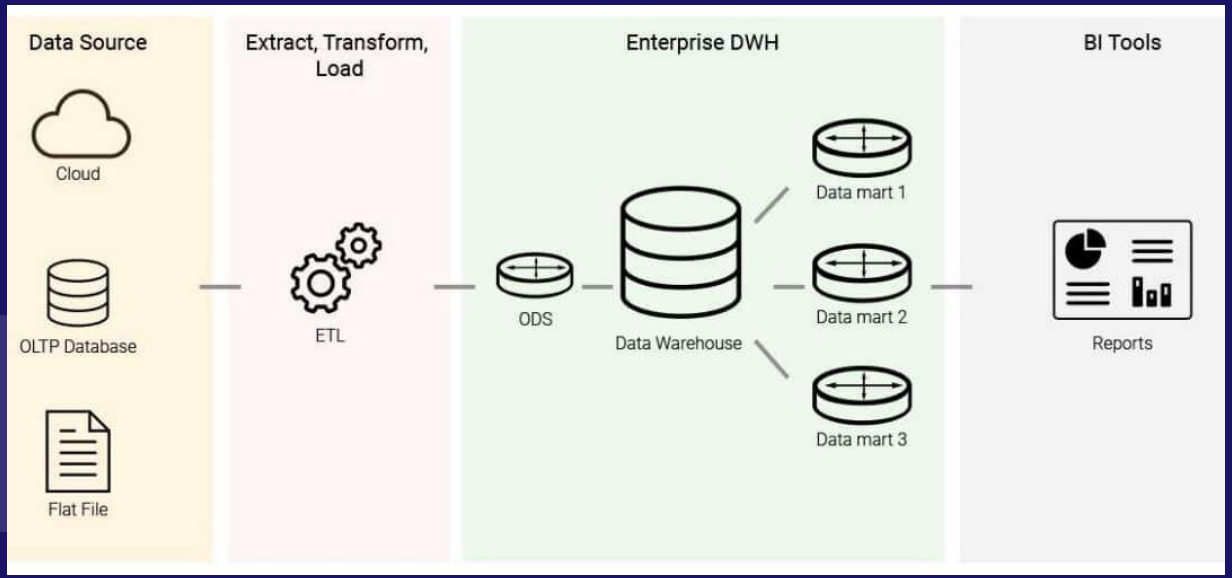
Process of data mining



Data Warehousing

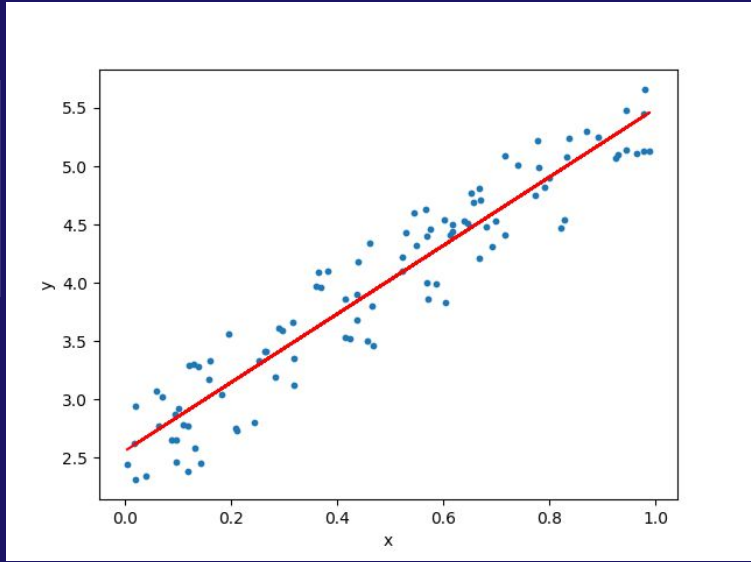
Data warehousing is **the process of constructing and using a** data warehouse. A data warehouse is constructed by integrating data from multiple heterogeneous sources that support analytical reporting, structured and/or ad hoc queries, and decision making.




Data Source
Cloud
OLTP Database
Flat File
Extract, Transform, Load
ETL
Enterprise DWH
ODS
Data Warehouse
Data mart 1
Data mart 2
Data mart 3
BI Tools
Reports

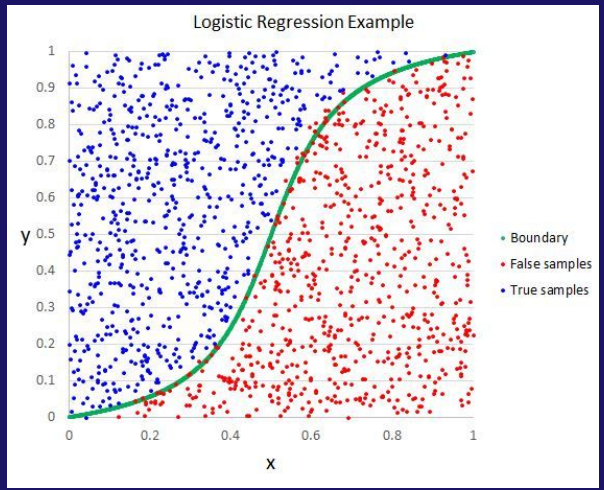
Types of ML Algorithms

1. Linear Regression - Linear Regression is a **supervised machine learning algorithm where the predicted output is continuous and has a constant slope**. It's used to predict values within a continuous range, (e.g. sales, price) rather than trying to classify them into categories (e.g. cat, dog).



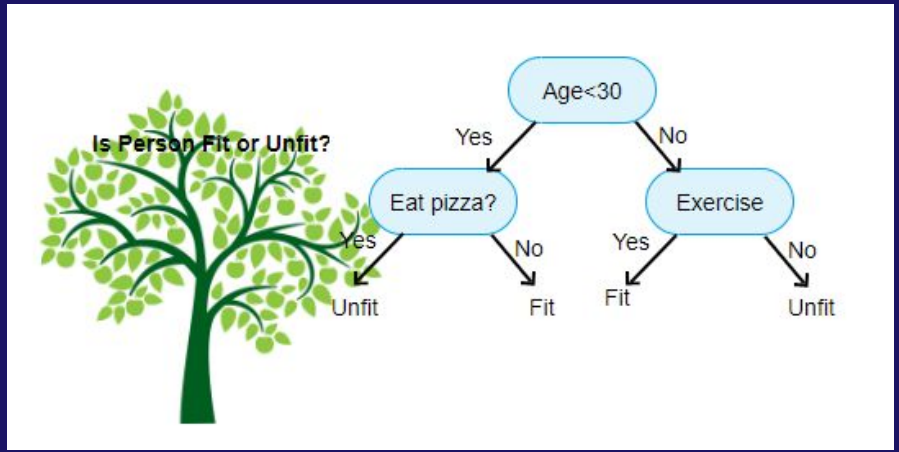
2. Logistic Regression - Logistic regression is a supervised learning classification algorithm used to predict the probability of a target variable. The nature of target or dependent variable is dichotomous, which means there would be only two possible classes.

In simple words, the dependent variable is binary in nature having data coded as either 1 (stands for success/yes) or 0 (stands for failure/no).

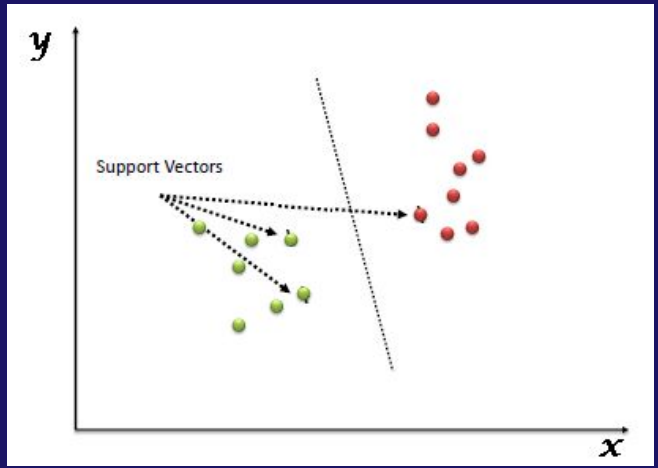


3. Decision Tree - Decision Trees are a type of Supervised Machine Learning (that is you explain what the input is and what the corresponding output is in the training data) where the data is continuously split according to a certain parameter. The tree can be explained by two entities, namely decision nodes and leaves. The leaves are the decisions or the final outcomes. And the decision nodes are where the data is split.

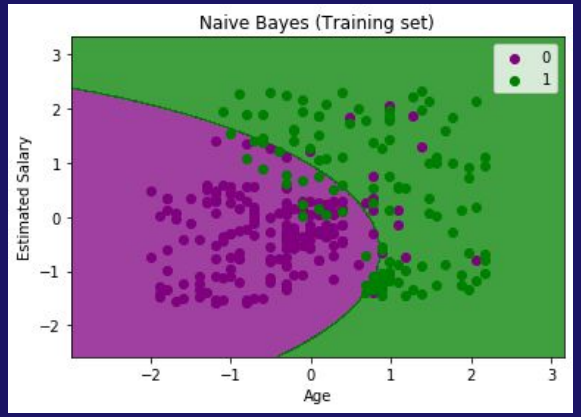
Example:



4. Support Vector Machine - “Support Vector Machine” (SVM) is a supervised machine learning algorithm that can be used for both classification or regression challenges. However, it is mostly used in classification problems. In the SVM algorithm, we plot each data item as a point in n-dimensional space (where n is a number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well (look at the below snapshot).

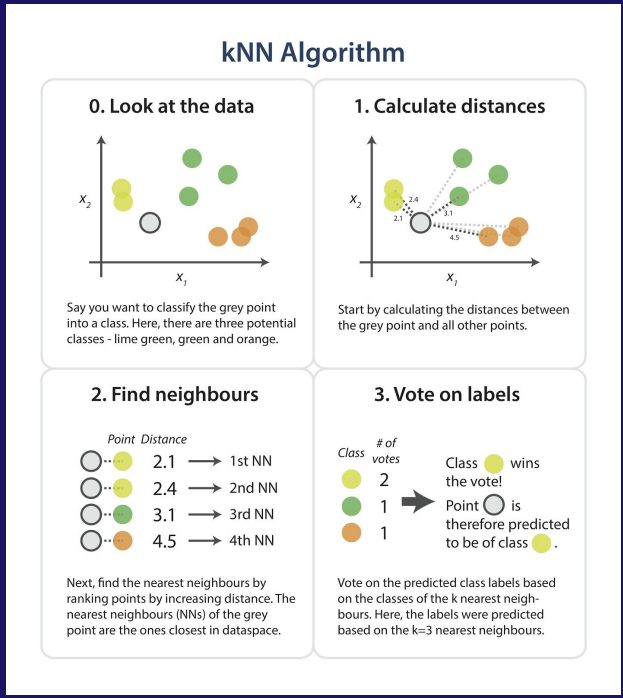


Naive Bytes - Naïve Bayes algorithm is a supervised learning algorithm, which is based on **Bayes theorem** and used for solving classification problems.It is mainly used in *text classification* that includes a high-dimensional training dataset.Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions.**It is a probabilistic classifier, which means it predicts on the basis of the probability of an object**.Some popular examples of Naïve Bayes Algorithm are **spam filtration, Sentimental analysis, and classifying articles**.



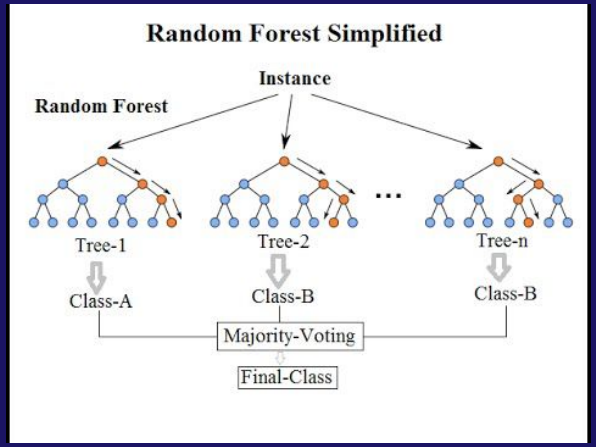
6. KNN -

- K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on Supervised Learning technique.
- K-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories.
- K-NN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm.
- K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for the Classification problems.



7. Random Forest: Random forest is a supervised learning algorithm. The "forest" it builds, is an ensemble of decision trees, usually trained with the “bagging” method. The general idea of the bagging method is that a combination of learning models increases the overall result.

Put simply: random forest builds multiple decision trees and merges them together to get a more accurate and stable prediction.



Applications of ML in different domains

HEALTHCARE - Machine learning applications can potentially improve the accuracy of treatment protocols and health outcomes through algorithmic processes.
For example, deep learning, a type of complex machine learning that mimics how the human brain functions, is increasingly being used in radiology and medical imaging. Using neural networks that can learn from data without any supervision, deep learning applications can detect, recognize and analyze cancerous lesions from images.

Faster processing speeds and cloud infrastructures allow machine learning applications to detect anomalies in images beyond what the human eye can see, aiding in diagnosing and treating disease.

Machine learning in health informatics can streamline recordkeeping, including electronic health records (EHRs). Using AI to improve EHR management can improve patient care, reduce healthcare and administrative costs, and optimize operations.

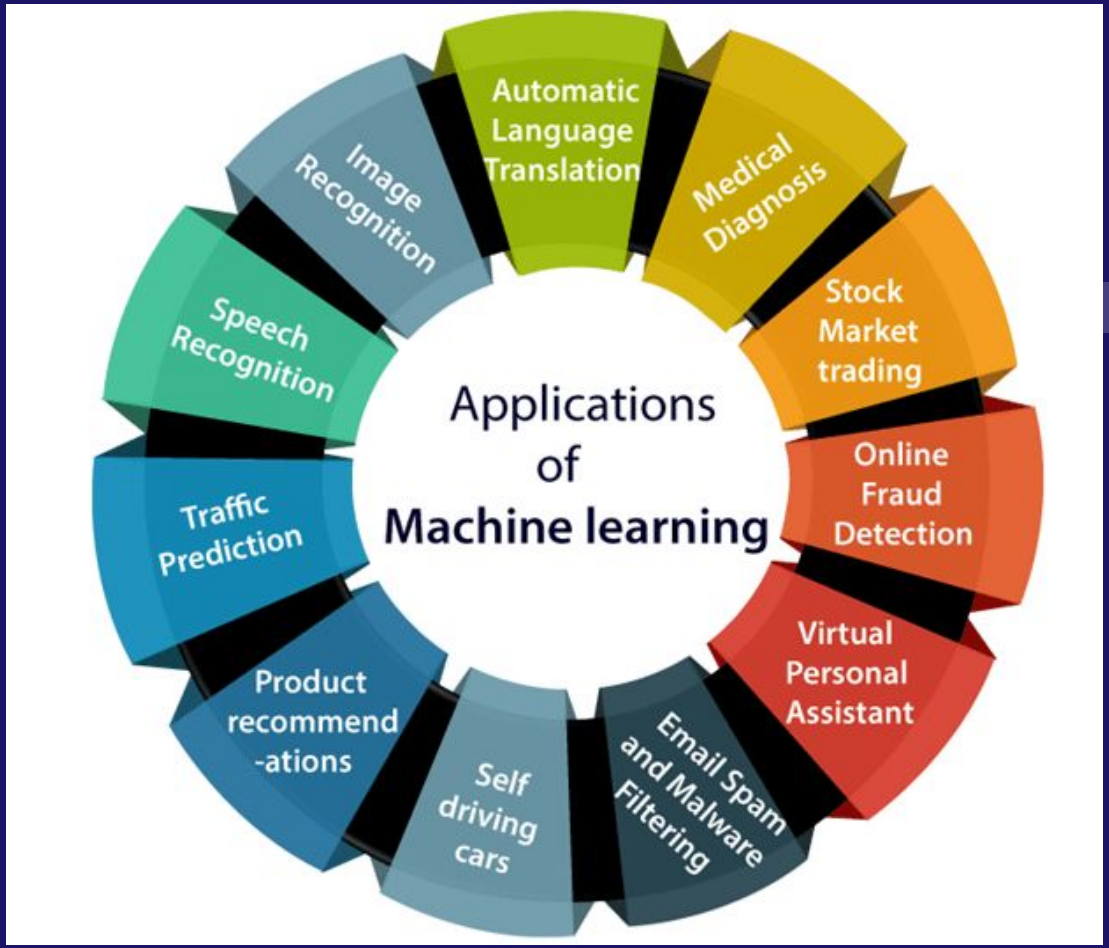
The combination of machine learning, health informatics and predictive analytics offers opportunities to improve healthcare processes, transform clinical decision support tools and help improve patient outcomes.

TELECOM -

There are numerous potential use cases for ML in telecommunications networks.

In the area of system monitoring, anomaly detection systems are crucial for identifying performance issues and problematic network behavior. Proactively predicting the degradation of key performance indicators, and identifying the likely root cause, can help reduce and prevent outages.

In the area of managed services, ML models can improve trouble ticket management by effectively classifying, prioritizing, and escalating incidents. Capacity planning and customer retention can be improved through explainable churn prediction. Furthermore, in the area of intelligent networks, the incorporation of ML tools can enable self-healing radio networks, which automatically detect issues and take corrective actions. New technologies such as deep learning and reinforcement learning can be used to automate the network design process and optimize network performance in real time.


Applications
of
Machine learning
Automatic Language Translation
Medical Diagnosis
Stock Market trading
Online Fraud Detection
Virtual Personal Assistant
Email Spam and Malware Filtering
Self driving cars
Product recommend -ations
Traffic Prediction
Speech Recognition
Image Recognition

THANKS !

eng19cs0081.deekshareddy@gmail.com
+91 9980722277