Introduction to Machine Learning

This presentation provides an overview of the fundamentals of machine learning and its applications.
It also walks through the steps involved in implementing a real project in ML.

Created by : Zahra Alipour

z.alipour7034891@gmail.com

History of artificial intelligence

1940-1960

Birth of AI

Alan Turing publishes Computing Machinery and Intelligence.father of computer science"— asks the following question: "Can machines think?" From there, he offers a test, now famously known as the "Turing Test," where a human interrogator would try to distinguish between a computer and human text response.

Herbert Simon, economist and sociologist, prophesied in 1957 that the AI would succeed in beating a human at chess in the next 10 years, but the AI then entered a first winter.

1980-2004

Expert systems

Neural networks which use a backpropagation algorithm to train itself become widely used in AI applications.Stuart Russell and Peter Norvig publish Artificial Intelligence: A Modern Approach .John McCarthy writes a paper, What Is Artificial Intelligence?

2011-2015

Rise

A kind of deep neural network called a convolutional neural network to identify and categorize images with a higher rate of accuracy than the average human

2016-2023

LLM-Revolution

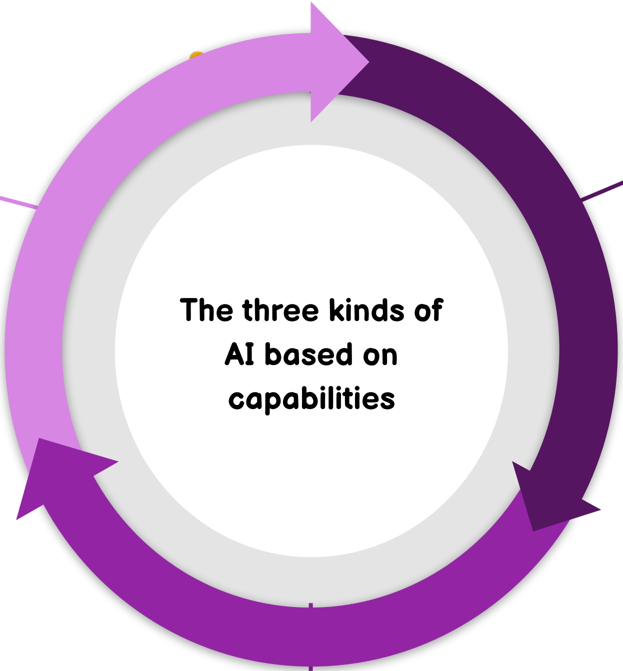
DeepMind's AlphaGo program, powered by a deep neural network. Google purchased DeepMind for a reported USD 400 million.A rise in large language models, or LLMs, such as ChatGPT, create anenormous change in performance of AI and its potential to drive enterprise value. With these new generative AI practices, deep-learning models can be pre-trained on vast amounts of raw, unlabeled data.

Introduction to Artificial Intelligence

Artificial intelligence, or AI, is technology that enables computers and machines to simulate human intelligence and problem-solving capabilities.
On its own or combined with other technologies (e.g., sensors, geolocation, robotics) AI can perform tasks that would otherwise require human intelligence or intervention. Digital assistants, GPS guidance, autonomous vehicles, and generative AI tools (like Open AI's Chat GPT) are just a few examples of AI in the daily news and our daily lives.

Super AI

Super AI is commonly referred to as artificial superintelligence and, like AGI, is strictly theoretical. If ever realized, Super AI would think, reason, learn, make judgements and possess cognitive abilities that surpass those of human beings. The applications possessing Super AI capabilities will have evolved beyond the point of understanding human sentiments and experiences to feel emotions, have needs and possess beliefs and desires of their own.



Artificial Narrow AI

Artificial Narrow Intelligence, also known as Weak AI, what we refer to as Narrow AI is the only type of AI that exists today. Any other form of AI is theoretical. It can be trained to perform a single or narrow task, often far faster and better than a human mind can. However, it can't perform outside of its defined task. Instead, it targets a single subset of cognitive abilities and advances in that spectrum. Siri, Amazon's Alexa and IBM Watson are examples of Narrow AI. Even OpenAI's ChatGPT is considered a form of Narrow AI because it's limited to the single task of text-based chat.

General AI

Artificial General Intelligence (AGI), also known as Strong AI, is today nothing more than a theoretical concept. AGI can use previous learnings and skills to accomplish new tasks in a different context without the need for human beings to train the underlying models. This ability allows AGI to learn and perform any intellectual task that a human being can.

Examples of narrow AI

01	**Image and facial recognition systems**	• These systems, including those used by social media companies like Facebook and Google to automatically identify people in photographs, are forms of weak AI.
02	**Chatbots and conversational assistants**	• This includes popular virtual assistants Google Assistant, Siri and Alexa. Also included are simpler, customer-service chatbots, such as a bot that assists customers in returning an item to a retail store.
03	Self-driving vehicles	• Autonomous or semiautonomous cars, such as some Tesla models and autonomous drones, boats and factory robots, are all applications of narrow AI.
04	**Predictive maintenance models**	• hese models rely on data from machines, often collected through sensors, to help predict when a machine part may fail and alert users ahead of time.
05	**Recommendation engines**	• These systems that predict content a user might like or search for next are forms of weak AI.

What is machine learning?

In general, machine learning algorithms are used to make a prediction or classification. Based on some input data, which can be labeled or unlabeled, your algorithm will produce an estimate about a pattern in the data.

Machine learning (ML) is a branch of artificial intelligence (AI) and computer science that focuses on the using data and algorithms to enable AI to imitate the way that humans learn, gradually improving its accuracy.

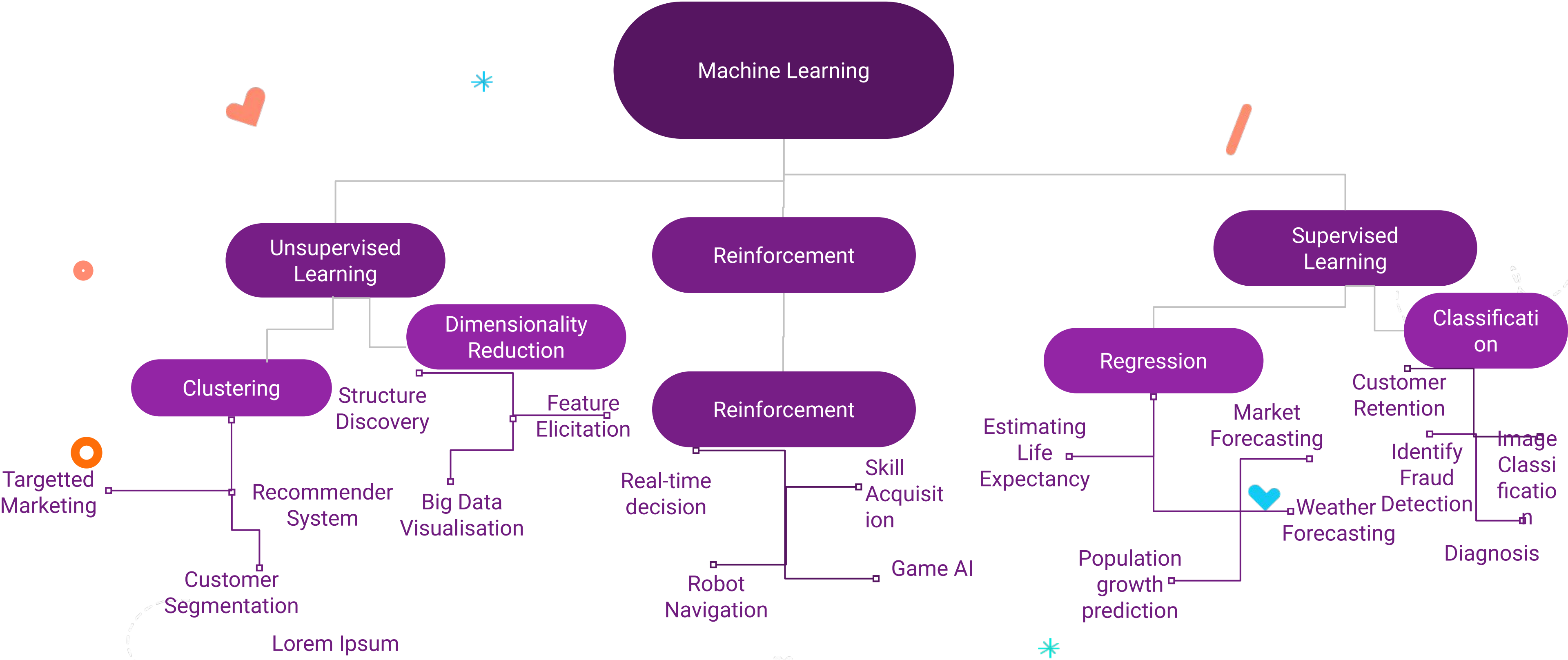
A Decision Process

A Model Optimization Process

An Error Function

An error function evaluates the prediction of the model. If there are known examples, an error function can make a comparison to assess the accuracy of the model.

If the model can fit better to the data points in the training set, then weights are adjusted to reduce the discrepancy between the known example and the model estimate. The algorithm will repeat this iterative "evaluate and optimize" process, updating weights autonomously until a threshold


Machine Learning
Unsupervised Learning
Reinforcement
Supervised Learning
Dimensionality Reduction
Clustering
Structure Discovery
Feature Elicitation
Big Data Visualisation
Targetted Marketing
Recommender System
Customer Segmentation
Lorem Ipsum
Reinforcement
Real-time decision
Skill Acquisition
Game AI
Robot Navigation
Regression
Classification
Estimating Life Expectancy
Market Forecasting
Weather Forecasting
Population growth prediction
Customer Retention
Identify Fraud Detection
Image Classification
Diagnosis

Pattern Recognition

Classify Object (instances, examples) into Categories (classes, labels):

- Probability theory
- Statistics
- Machine learning
- Linear algebra
- Image processing
- Algorithms
- ...

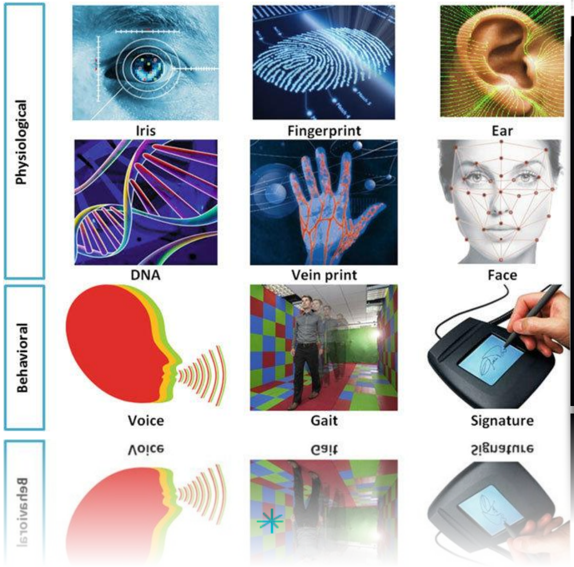
What is pattern?

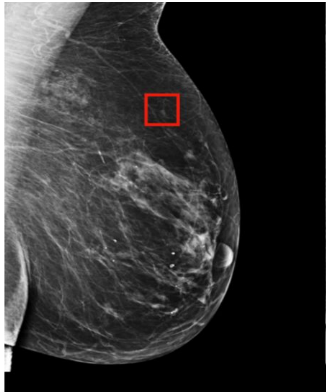
A pattern could be an object or event.

- Typically, represented by a vector x of numbers

Types of common patterns

- Handwriting Recognition
- Biometric Recognition
- Face Detection
- Autonomous System
- Medical Applications(breast cancer, skin cancer)
- Land Classification





From sensors to models

1. **Sensors**
   - Noisy
2. **Signal**
   - Medical_signals
   - Electricity_signals
3. **Features**
   - Base of dataset
4. **Modeling**
   - Algorithms
   - Model evaluation

Machine learning methods

Supervised machine learning

also known as supervised machine learning, is defined by its use of labeled datasets to train algorithms to classify data or predict outcomes accurately. As input data is fed into the model, the model adjusts its weights until it has been fitted appropriately.

- neural networks,
- naïve bayes,
- linear regression,
- logistic regression,
- random forest, and
- support vector machine (SVM).

Unsupervised machine learning

also known as unsupervised machine learning, uses machine learning algorithms to analyze and cluster unlabeled datasets (subsets called clusters). These algorithms discover hidden patterns or data groupings without the need for human intervention.

- neural networks,
- k-means clustering, and
- probabilistic clustering

Semi-supervised learning

Semi-supervised learning offers a happy medium between supervised and unsupervised learning. During training, it uses a smaller labeled data set to guide classification and feature extraction from a larger, unlabeled data set. Semi-supervised learning can solve the problem of not having enough labeled data for a supervised learning algorithm. It also helps if it’s too costly to label enough data.

What is supervised learning?

Supervised learning is a machine learning approach that's defined by its use of labeled datasets.
These datasets are designed to train or "supervise" algorithms into classifying data or predicting outcomes accurately.
Using labeled inputs and outputs, the model can measure its accuracy and learn over time.

Supervised learning

Classification

Regression

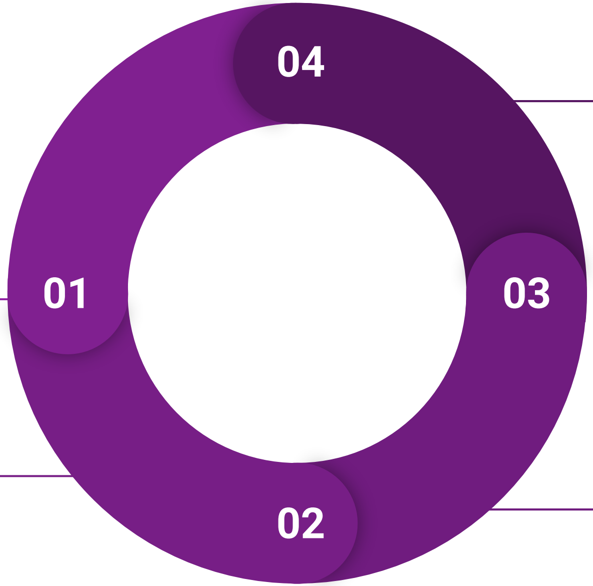
Classification

Supervised Learning

The Classification algorithm is a Supervised Learning technique that is used to identify the category of new observations on the basis of training data.

Types of Classification

- Binary Classifier: two possible outcomes
- Multi-class Classifier:more than two outcomes



Data type

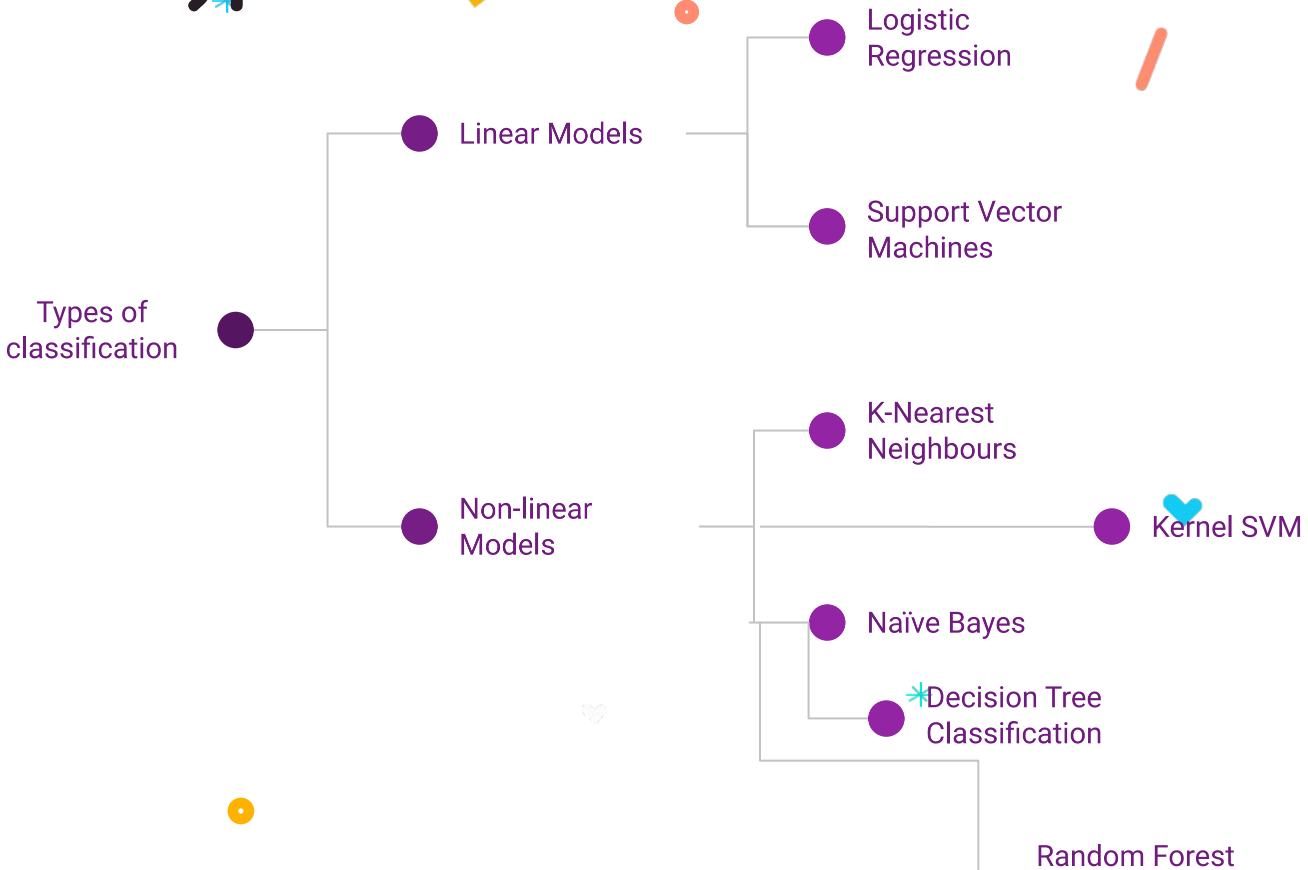
In Classification, a program learns from the given dataset or observations and then classifies new observation into a number of classes or groups. Such as, Yes or No, 0 or 1, Spam or Not Spam, cat or dog, etc. Classes can be called as targets/labels or categories.

Function

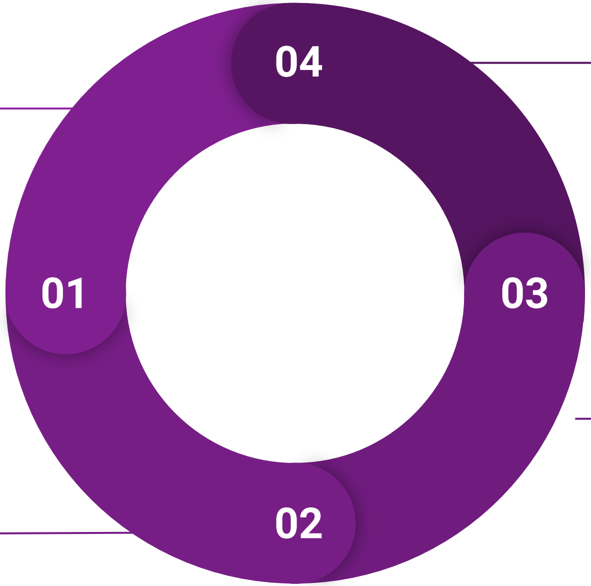
In classification algorithm, a discrete output function(y) is mapped to input variable(x).

y=f(x), where y = categorical output

Types of classification



Regression



Supervised learning

Regression is a supervised learning technique which helps in finding the correlation between variables and enables us to predict the continuous output variable based on the one or more predictor variables.

Function

Y= aX+b

Data type

Regression is a supervised learning technique which helps in finding the correlation between variables and enables us to predict the continuous output variable based on the one or more predictor variables.More specifically, Regression analysis helps us to understand how the value of the dependent variable is changing corresponding to an independent variable when other independent variables are held fixed. It predicts continuous/real values such as temperature, age, salary, price, etc.

Types of Regression

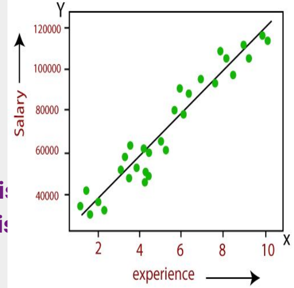
- Linear Regression
- Logistic Regression
- Polynomial Regression
- Support Vector Regression
- Decision Tree Regression
- Random Forest Regression
- Ridge Regression
- Lasso Regression

Types of Regression

01

Linear Regression

- Linear regression shows the linear relationship between the independent variable (X-axis) and the dependent variable (Y-axis), hence called linear regression.
- If there is only one input variable (x), then such linear regression is called simple linear regression. And if there is more than one input variable, then such linear regression is called multiple linear regression.



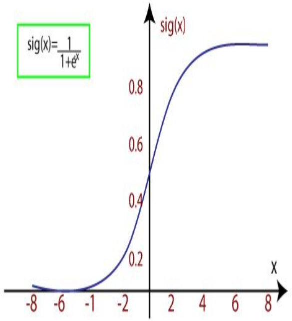
02

Logistic Regression

- Logistic regression algorithm works with the categorical variable such as 0 or 1, Yes or No, True or False, Spam or not spam, etc.
- Regression Analysis in Machine learning

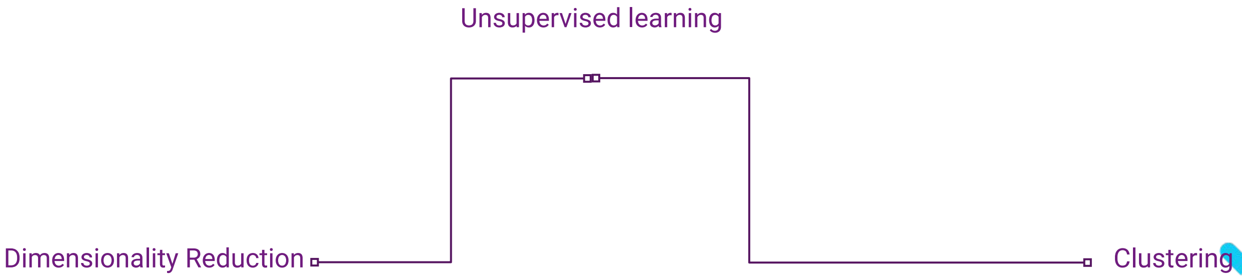
f(x)= Output between the 0 and 1 value.
x= input to the function
e= base of natural logarithm.

$$f(x)=\frac{1}{1+e^{-x}}$$



What is Unsupervised learning?

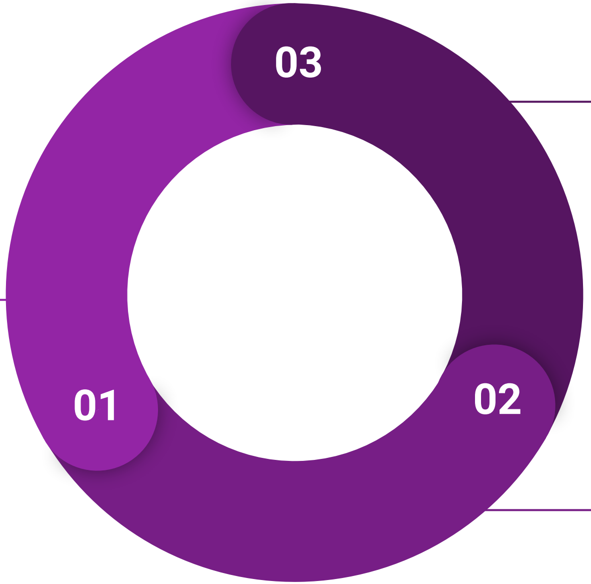
Unsupervised learning uses machine learning algorithms to analyze and cluster unlabeled data sets. These algorithms discover hidden patterns in data without the need for human intervention (hence, they are "unsupervised").



Dimensionality Reduction Technique

What is Dimensionality Reduction?

The number of input features, variables, or columns present in a given dataset is known as dimensionality, and the process to reduce these features is called dimensionality reduction.



Data

It is a way of converting the higher dimensions dataset into lesser dimensions dataset ensuring that it provides similar information

Dimensionality Reduction types

- Feature selection
- Components/Factors based
- Projection based

Dimensionality Reduction

01	**Feature Selection** : Feature selection is the process of selecting the subset of the relevant features and leaving out the irrelevant features present in a dataset to build a model of high accuracy. In other words, it is a way of selecting the optimal features from the input dataset.	• Filters Methods • Wrappers Methods • Embedded Methods
02	**Feature Extraction:** Feature extraction is the process of transforming the space containing many dimensions into space with fewer dimensions. This approach is useful when we want to keep the whole information but use fewer resources while processing the information.	• Principal Component Analysis • Linear Discriminant Analysis • Kernel PCA • Quadratic Discriminant Analysis
03	Principal Component Analysis (PCA)	• converts the observations of correlated features into a set of linearly uncorrelated features with the help of orthogonal transformation • PCA works by considering the variance of each attribute because the high attribute shows the good split between the classes, and hence it reduces the dimensionality

What is Model Evaluation?

Model evaluation is the process of using different evaluation metrics to understand a machine learning model's performance, as well as its strengths and weaknesses. Model evaluation is important to assess the efficacy of a model during initial research phases, and it also plays a role in model monitoring.

Hold-Out

In this method, the mostly large dataset is randomly divided to three subsets:Training set is a subset of the dataset used to build predictive models.

- Validation set is a subset of the dataset used to assess the performance of model built in the training phase. It provides a test platform for fine tuning model's parameters and selecting the best-performing model. Not all modeling algorithms need a validation set.
- Test set or unseen examples is a subset of the dataset to assess the likely future performance of a model. If a model fit to the training set much better than it fits the test set, overfitting is probably the cause.

Cross-Validation

When only a limited amount of data is available, to achieve an unbiased estimate of the model performance we use k-fold cross-validation. In k-fold cross-validation, we divide the data into k subsets of equal size. We build models k times, each time leaving out one of the subsets from training and use it as the test set. If k equals the sample size, this is called "leave-one-out".

Classification Evaluation

1. Confusion Matrix
2. F1 Score
3. Gain and Lift Charts
4. Kolomogorov Smirnov Chart
5. Area Under the ROC Curve (AUC – ROC)

Regression Evaluation

1. Relative Squared Error
2. Mean Absolute Error
3. Root Mean Squared Error (RMSE)
4. Relative Absolute Error
5. Coefficient of Determination
6. Standardized Residuals (Errors) Plot

Confusion Matrix

A confusion matrix shows the number of correct and incorrect predictions made by the classification model compared to the actual outcomes (target value) in the data. The matrix is NxN, where N is the number of target values (classes). Performance of such models is commonly evaluated using the data in the matrix. The following table displays a 2x2 confusion matrix for two classes (Positive and Negative)

Confusion Matrix		**Target**			
		Positive	Negative		
Model	Positive	a	b	*Positive Predictive Value*	a/(a+b)
	Negative	c	d	*Negative Predictive Value*	d/(c+d)
		Sensitivity	*Specificity*	**Accuracy** = (a+d)/(a+b+c+d)	
		a/(a+c)	d/(b+d)		

1

Accuracy

the proportion of the total number of predictions that were correct.

$$\text{Accuracy}=\frac{TP+TN}{TP+FP+FN+TN}$$

2

Misclassification rate

It is also termed as Error rate, and it defines how often the model gives the wrong predictions. The value of error rate can be calculated as the number of incorrect predictions to all number of the predictions made by the classifier.

$$\text{Error rate}=\frac{FP+FN}{TP+FP+FN+TN}$$

3

Precision

the number of correct outputs provided by the model or out of all positive classes that have predicted correctly by the model

$$\text{Precision}=\frac{TP}{TP+FP}$$

4

Sensitivity or Recall

the proportion of actual positive cases which are correctly identified

$$\text{Recall}=\frac{TP}{TP+FN}$$

5

F-measure

If two models have low precision and high recall or vice versa, it is difficult to compare these models. So, for this purpose, we can use F-score. This score helps us to evaluate the recall and precision at the same time.

$$\text{F-measure}=\frac{2*Recall*Precision}{Recall+Precision}$$

Steps to Implement a Real Project in ML

1 Define the problem statement

2 Collect and preprocess data

3 Select the appropriate machine learning algorithm

3 Train the model & Evaluate and finetune the model

3 Deploy the model

Thank you for your time 😊

Zahra Alipour