Introduction to Machine Learning

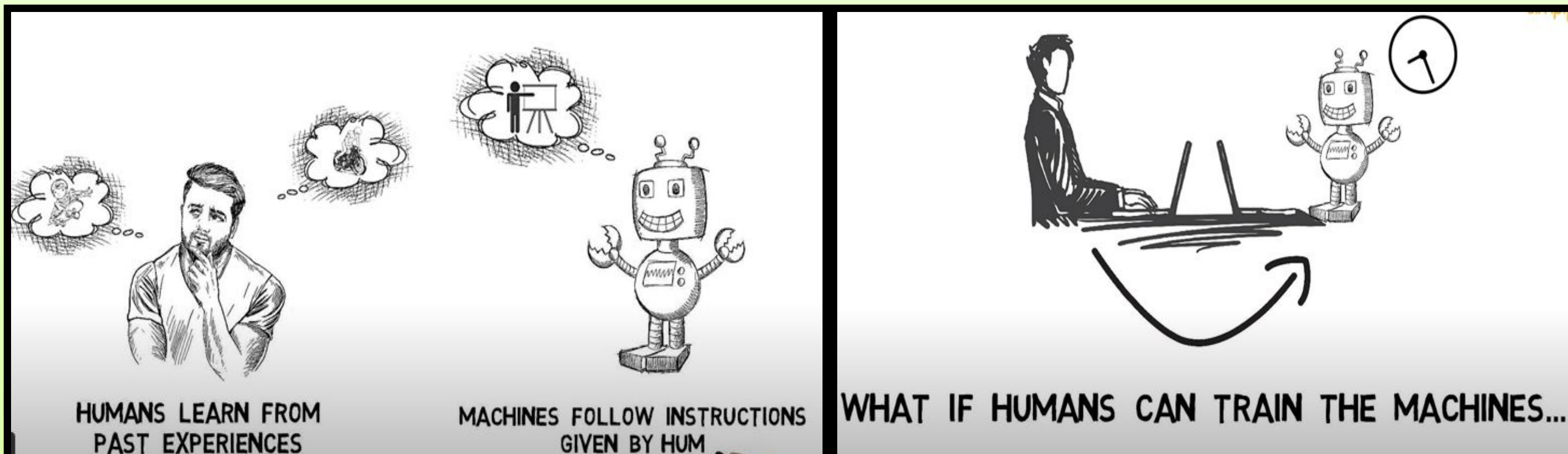
OUTLINE:

- ❖ Machine Learning
- ❖ Need for Machine Learning
- ❖ Types of Machine Learning
- ❖ How ML is relatable to AI & Deep Learning
- ❖ Applications of ML
- ❖ Modules in ML

MACHINE LEARNING

- Machine Learning:

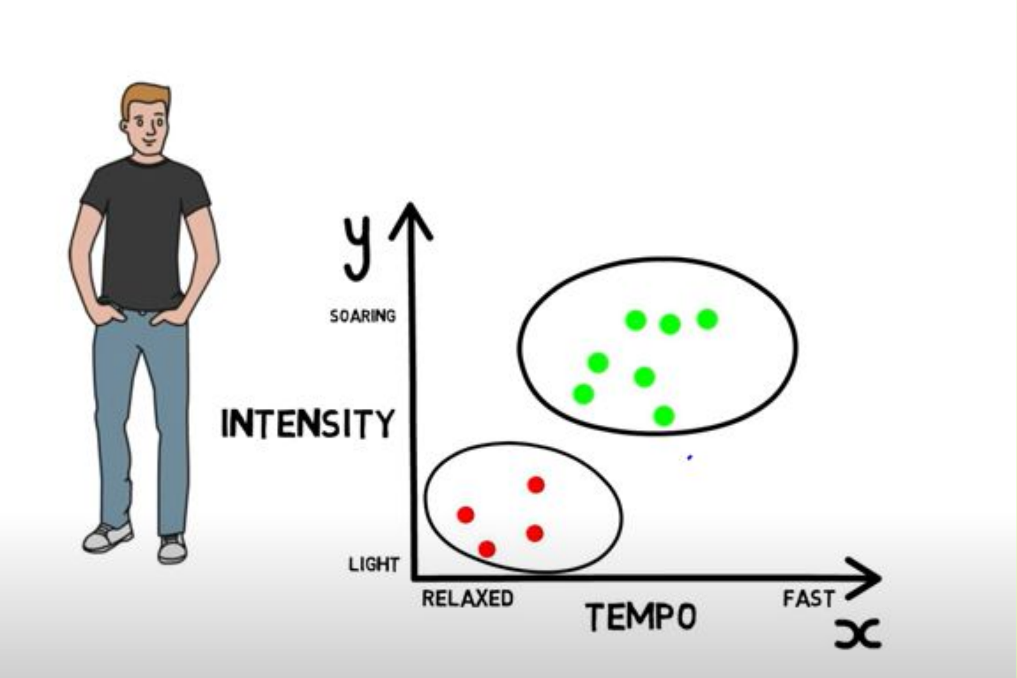
  Machine learning is **a method of data analysis that automates analytical model**. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention.



NEED OF MACHINE LEARNING

For Example, Paul's loves listening to the music, He either likes the song or dislike the song. Paul's decides this based on the

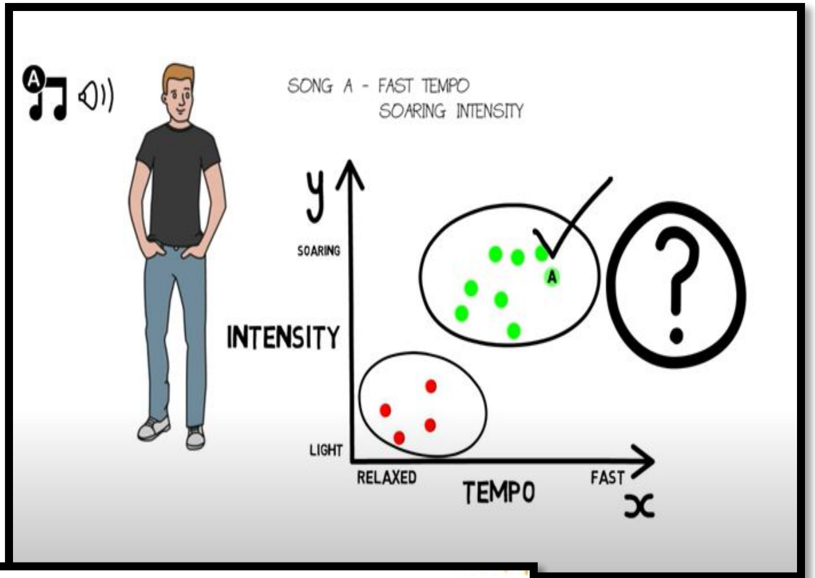
1. Tempo 2.Genre 3. Intensity 4. Gender of Voice

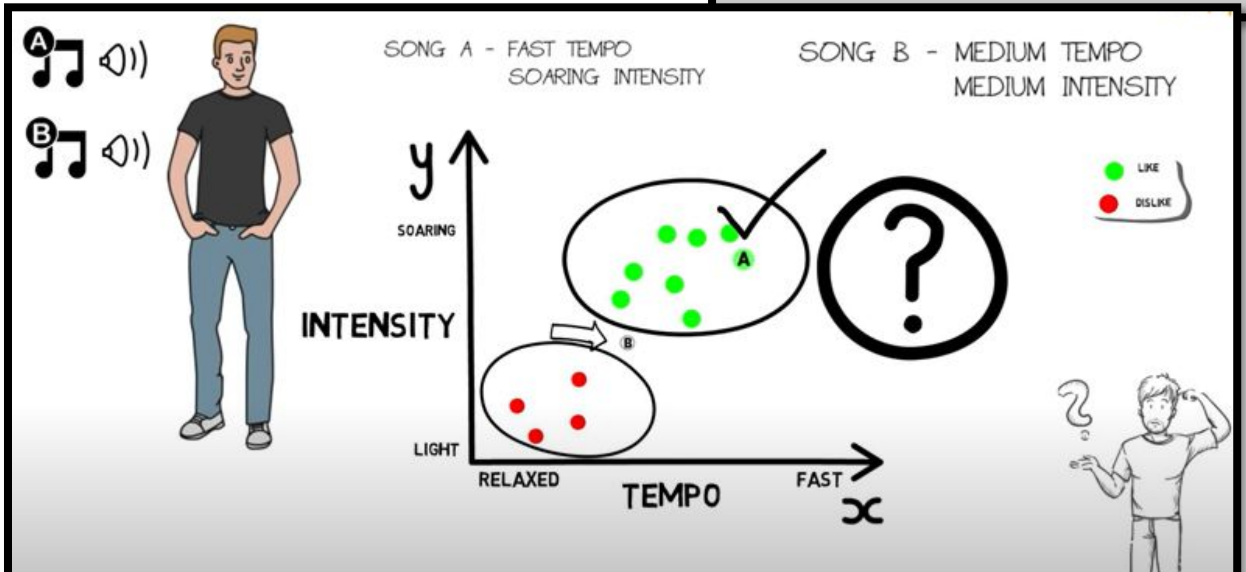


NEED OF MACHINE LEARNING

Now Paul listening to the new song (A)

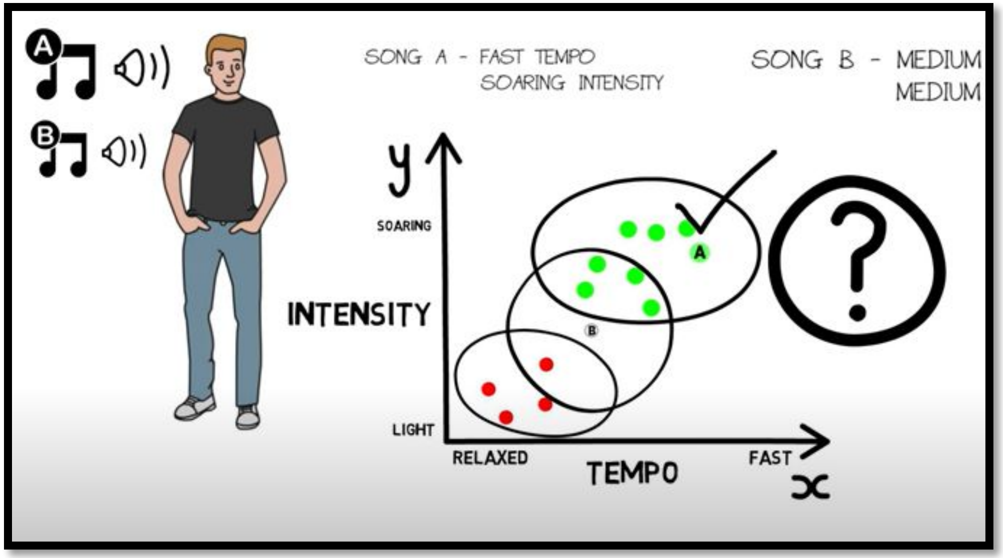
Now Paul listening to the new song (B)





NEED OF MACHINE LEARNING

- That’s were machine learning comes in
- **Learns the data**
- **Builds the prediction model**
- **When the new data comes it will easily predicts**
- **More data high will be the accuracy**



Basic machine learning KNN algorithm

TYPES OF MACHINE LEARNING

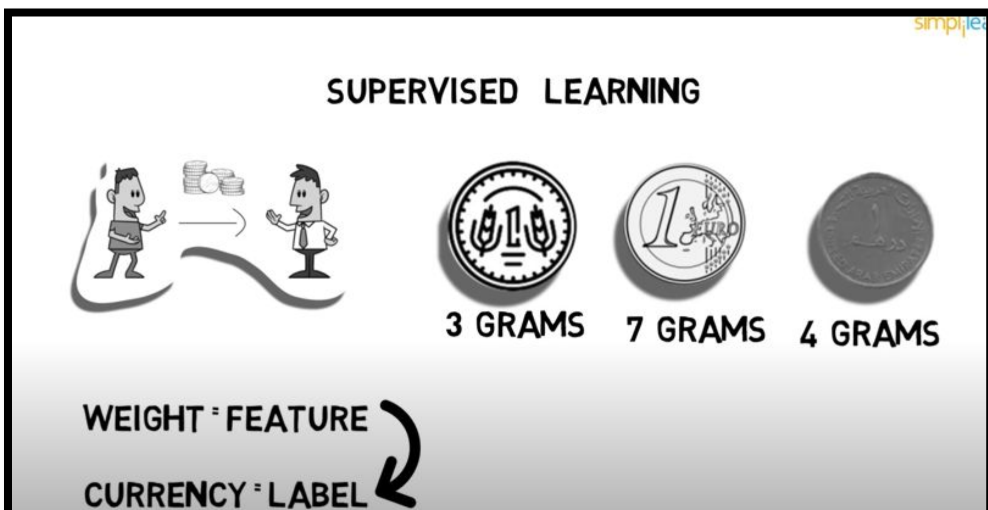
Types of Machine Learning

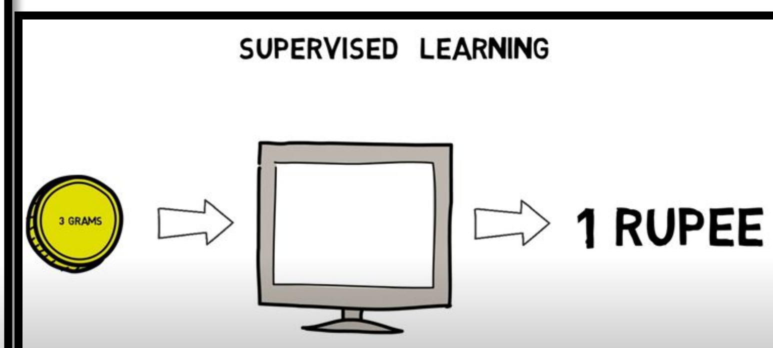
1. Supervised Learning
2. Unsupervised Learning
3. Reinforcement Learning

TYPES OF MACHINE LEARNING

1. **Supervised machine : labeled data predicted output**

- Some labled data (past data able to predict the future)
- Eg: height and weight to classify whether the patient belongs to obesity category to fit category
- Create the model train that data to predict the future

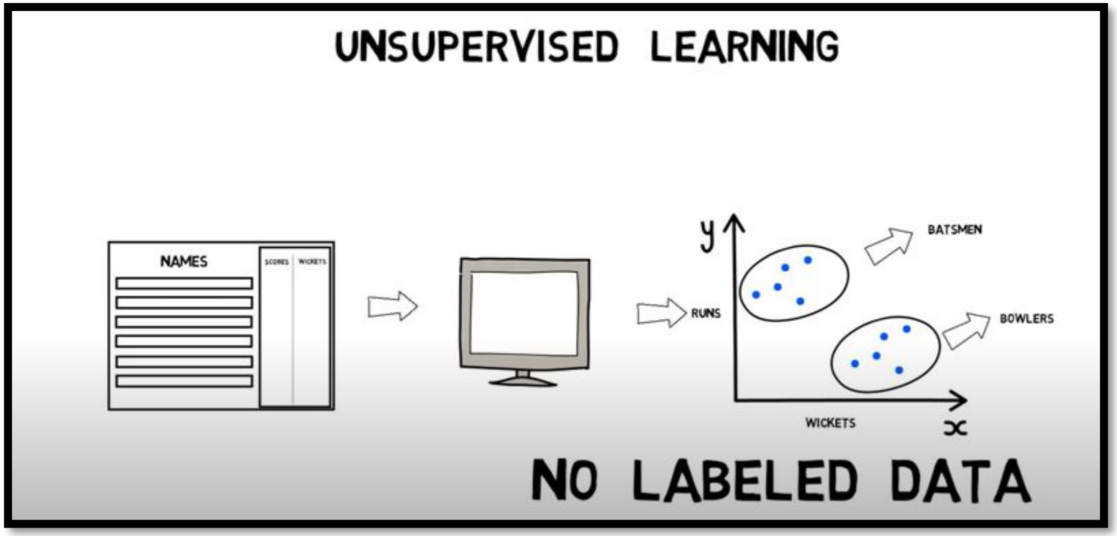




TYPES OF MACHINE LEARNING

2. Unsupervised: not labeled data

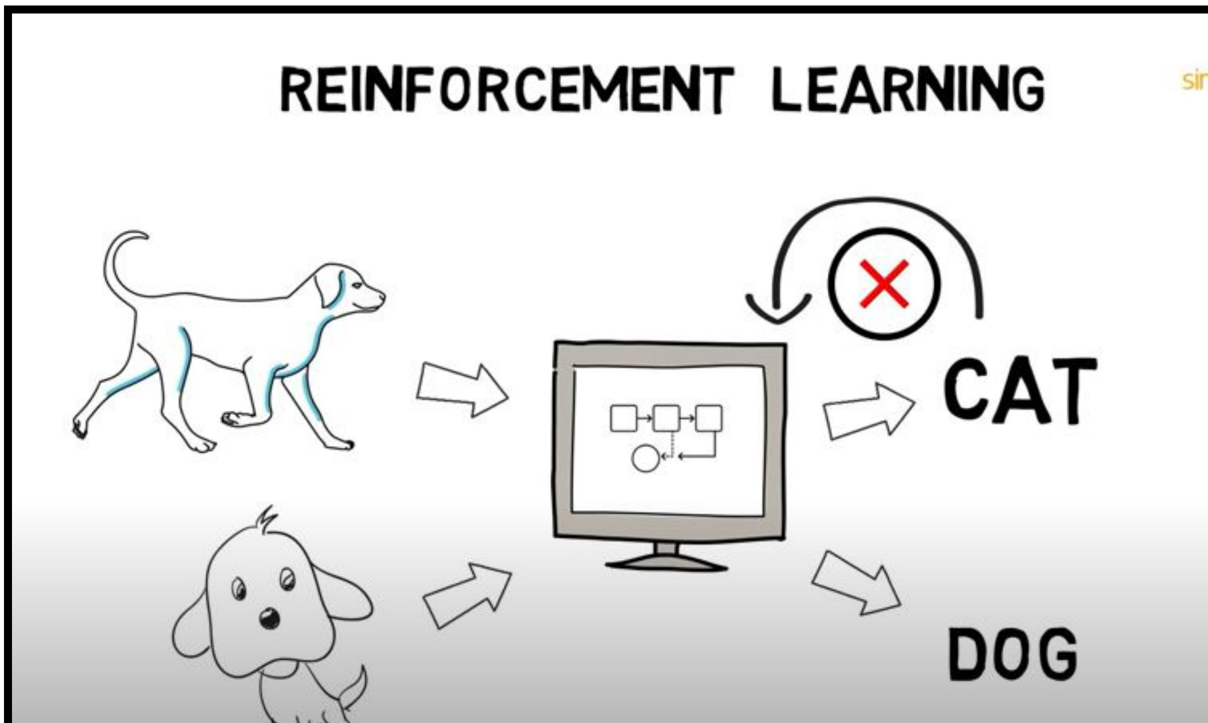
- Clustering kind of problem, k means clustering, DB clustering
- Based on the similarity of the data it will be grouped with some mathematical relations like Euclidean distance is used to group the data.
    - **DIMESNION REDUCTION ALGORITHMS:**
    - Principle Component Analysis, Independent Component Analysis.



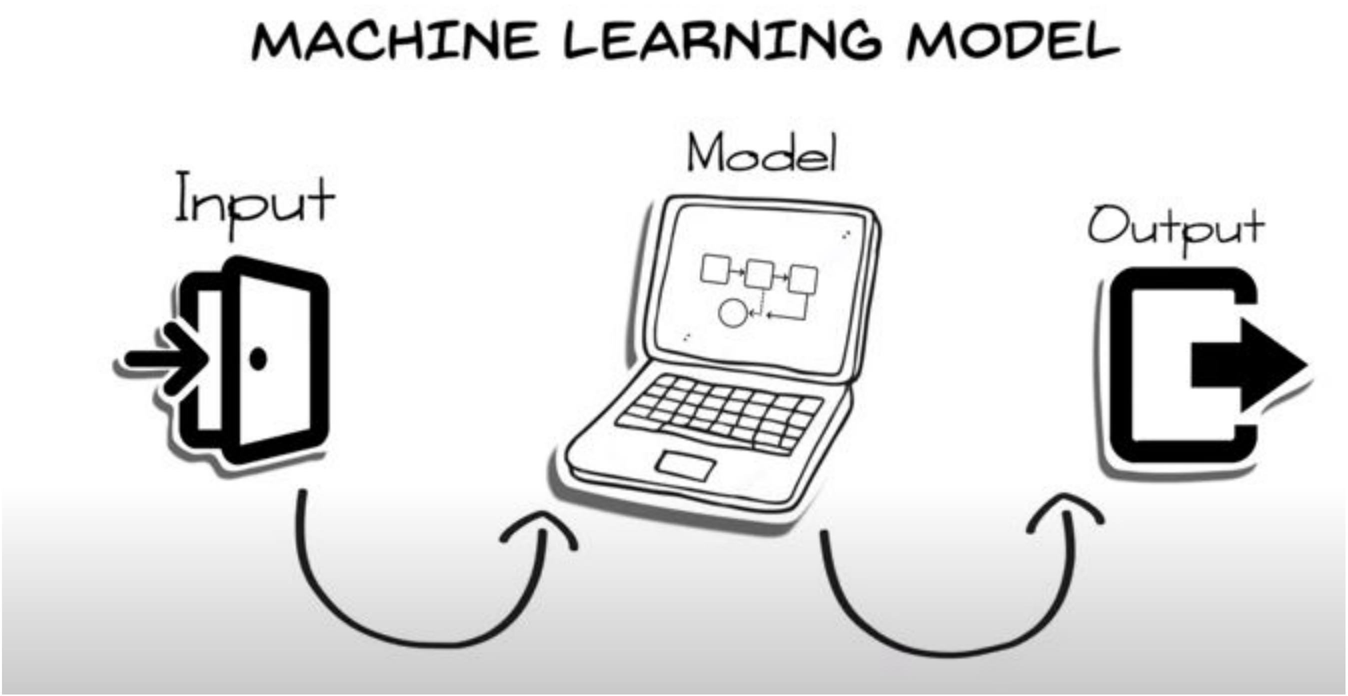
TYPES OF MACHINE LEARNING

3. Reinforcement: some data is labeled some not labeled

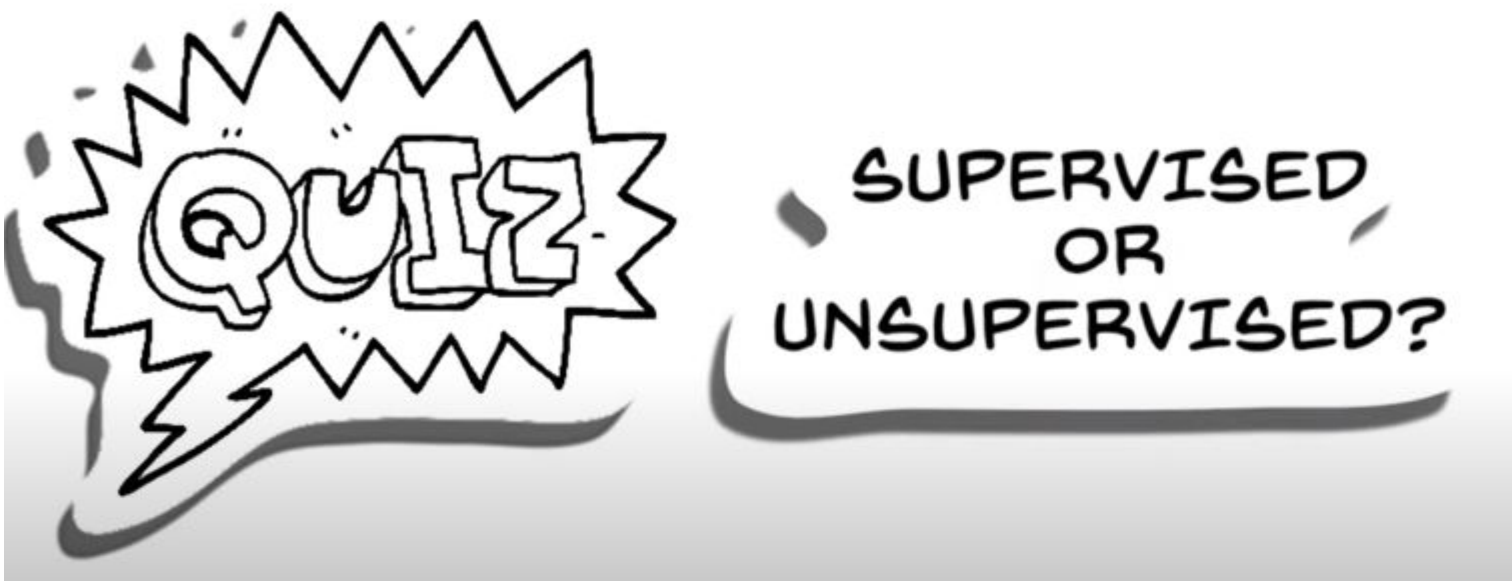
Works Based on the Negative Feedback



MACHINE LEARNING MODEL

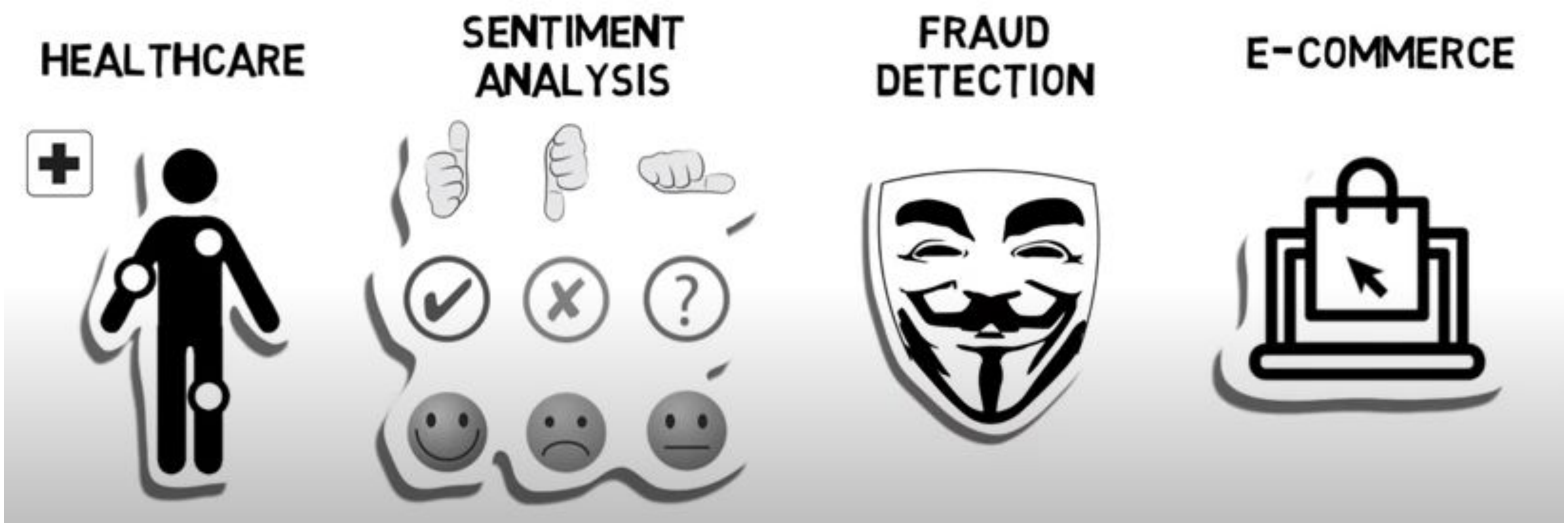


MACHINE LEARNING MODEL



MACHINE LEARNING MODEL

APPLICATIONS OF MACHINE LEARNING



APPLICATIONS OF MACHINE LEARNING

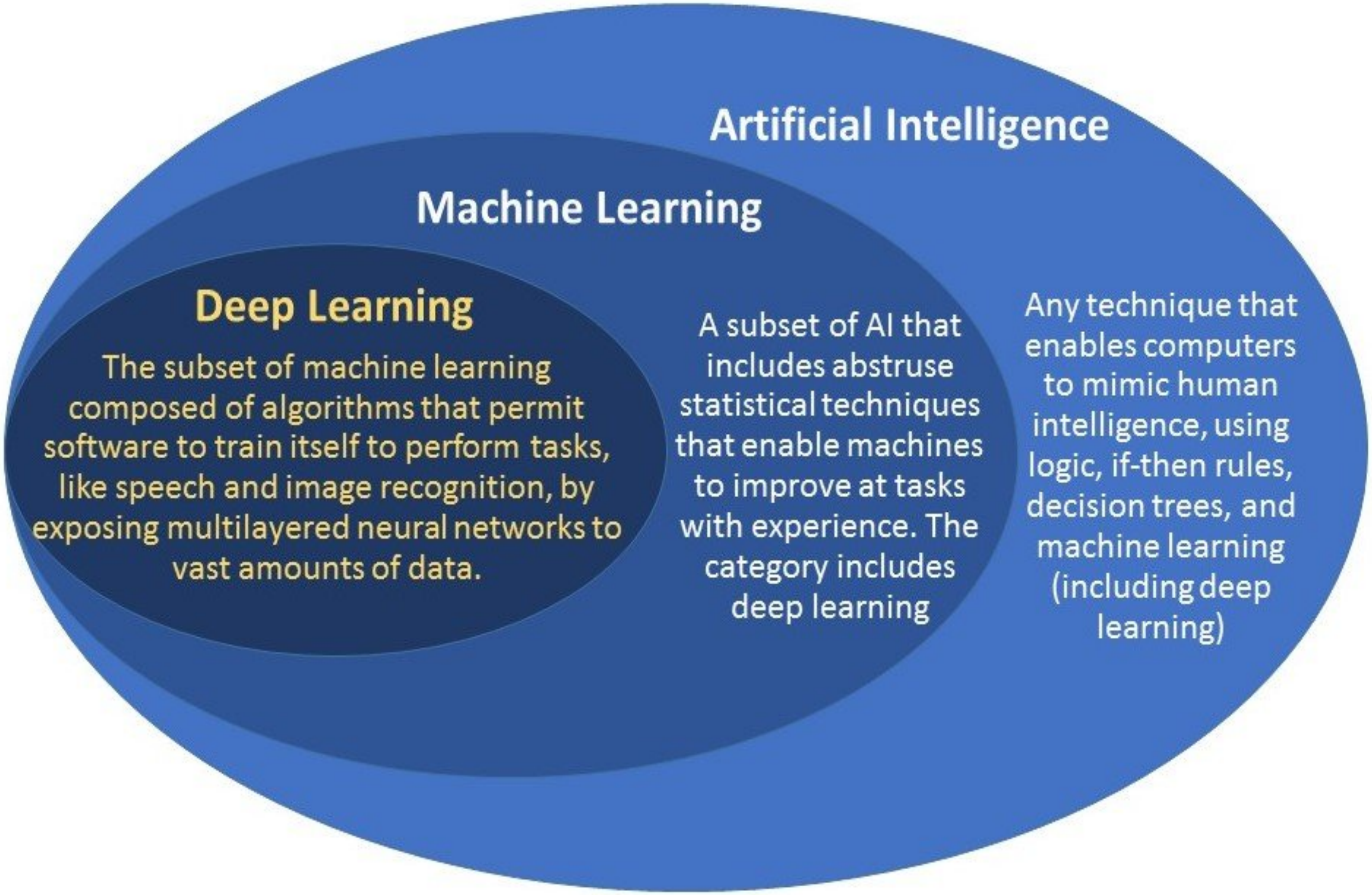
INTERESTING MACHINE LEARNING MODEL WHICH IS USED BY

UBER

DIFFERENTIAL PRICING IN REAL TIME BASED ON:

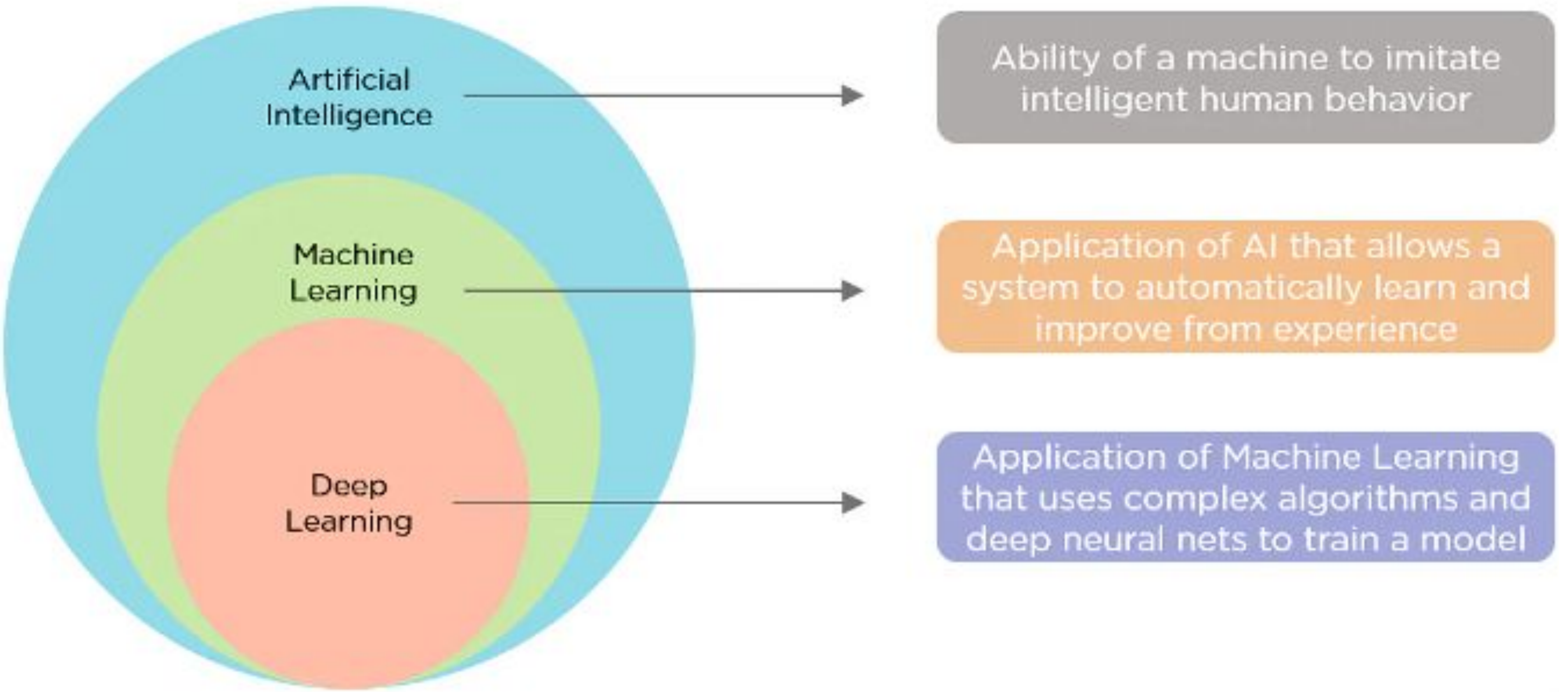
- DEMAND
- NUMBER OF CARS AVAILABLE
- BAD WEATHER

ML, DL & AI



Artificial Intelligence
Machine Learning
Deep Learning
The subset of machine learning composed of algorithms that permit software to train itself to perform tasks, like speech and image recognition, by exposing multilayered neural networks to vast amounts of data.
A subset of AI that includes abstruse statistical techniques that enable machines to improve at tasks with experience. The category includes deep learning
Any technique that enables computers to mimic human intelligence, using logic, if-then rules, decision trees, and machine learning (including deep learning)

ML, DL & AI


Artificial Intelligence
Machine Learning
Deep Learning
Ability of a machine to imitate intelligent human behavior
Application of AI that allows a system to automatically learn and improve from experience
Application of Machine Learning that uses complex algorithms and deep neural nets to train a model

ML MODULES

- Data Collection
- Data Processing
- Classification Module
- Decision Making Module

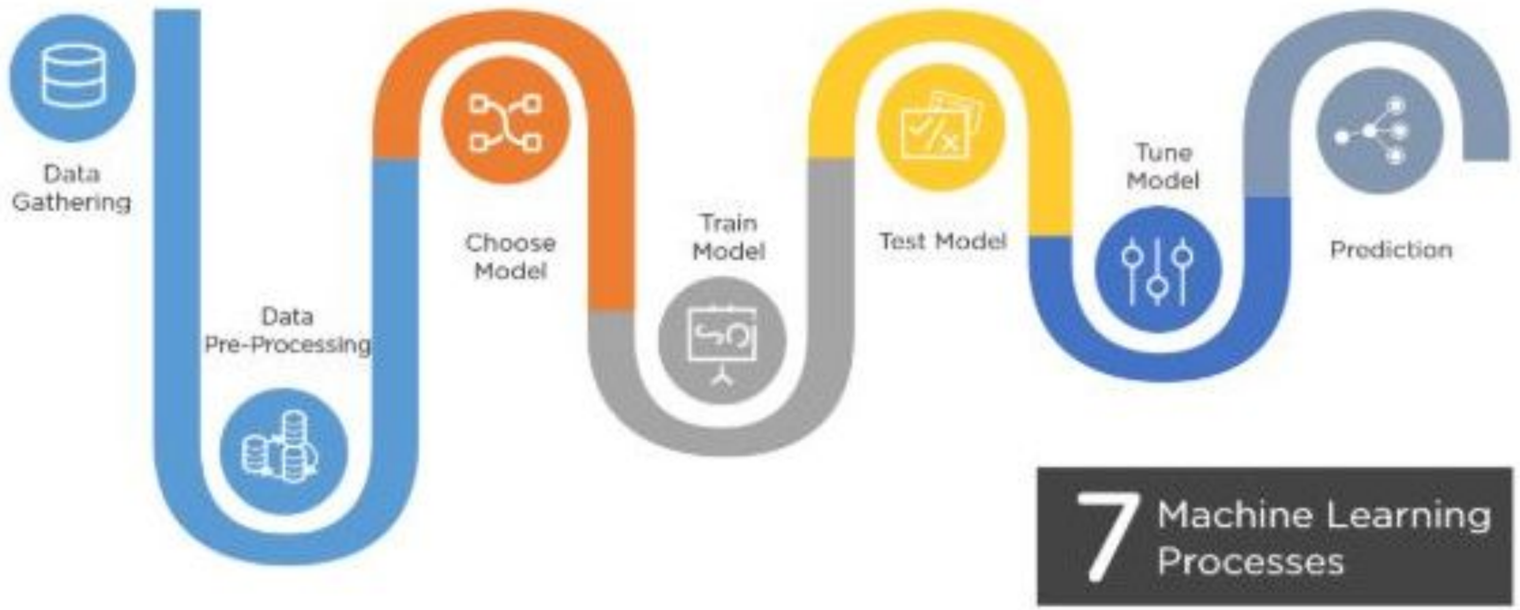
ML PROCESS

Importing the Dataset
Handling of Missing Data
Handling of Categorical Data
Splitting the dataset into training and testing datasets
Model
Performance Evaluation

APPLICATIONS

- Facial Recognition in smart phone
- Healthcare and medical diagnosis
- Smart assistants

ML PROCESS



APPLICATIONS

- Facial Recognition in smart phone
- Healthcare and medical diagnosis
- Smart assistants

SUBJECT PRESENTATION

OUTLINE:

- ❖Machine Learning
- ❖Types of Machine Learning Model

MACHINE LEARNING

Machine Learning:	Machine learning is **a method of data analysis that automates analytical model**. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention.
	❖**Learns the data** ❖**Builds the prediction model** ❖**Easily identify the new pattern** ❖**More data high will be the accuracy**

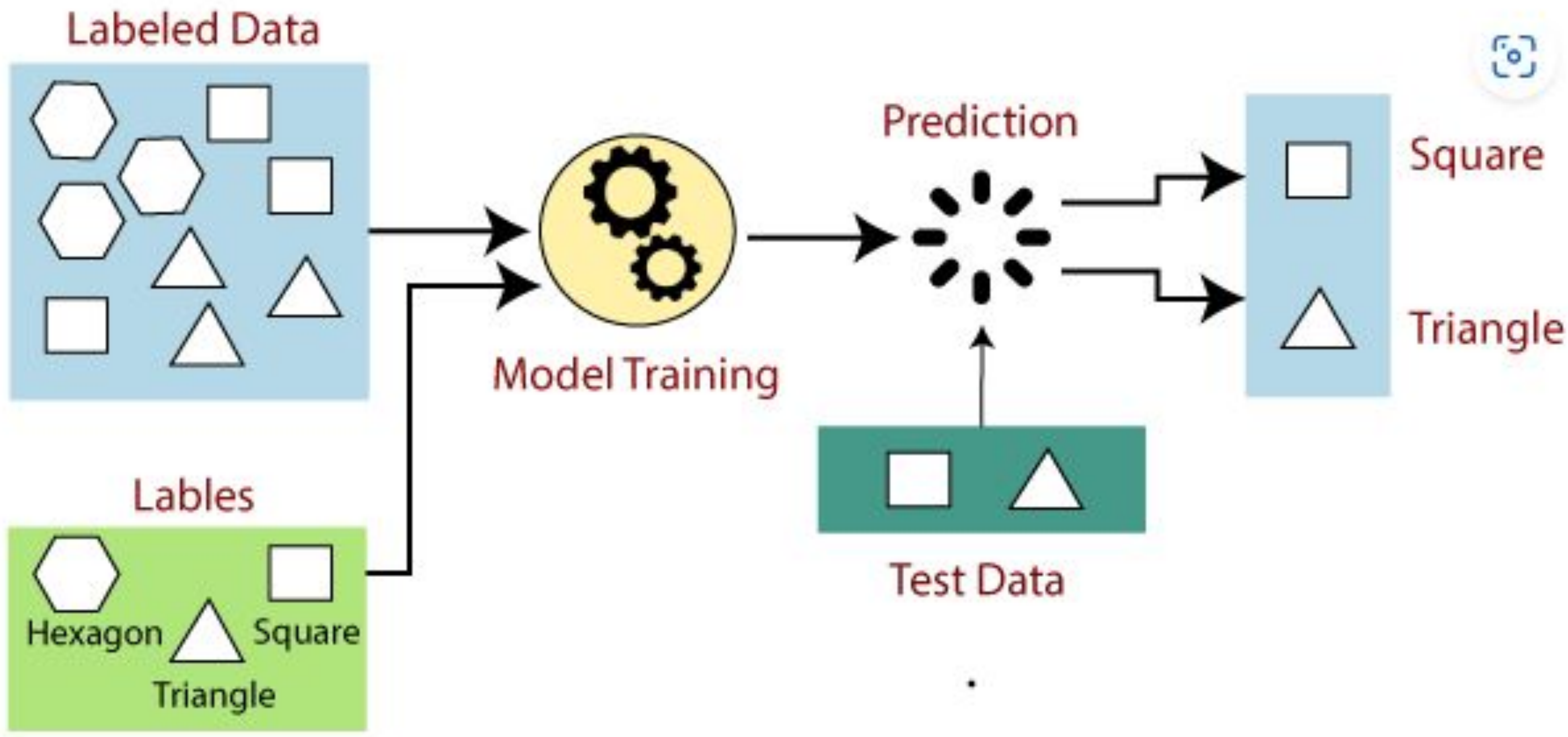
TYPES OF MACHINE LEARNING

- Based on the methods and way of learning, machine learning is divided into mainly four types, which are:
- Supervised Machine Learning
- Unsupervised Machine Learning
- Reinforcement Learning

Supervised Machine Learning

- **Supervised Machine Learning:**
- Supervised machine learning is based on supervision. It means in the supervised learning technique, we train the machines using the "labelled" dataset, and based on the training, the machine predicts the output.
- Here, the labelled data specifies that some of the inputs are already mapped to the output.

Supervised Machine Learning



Supervised Machine Learning

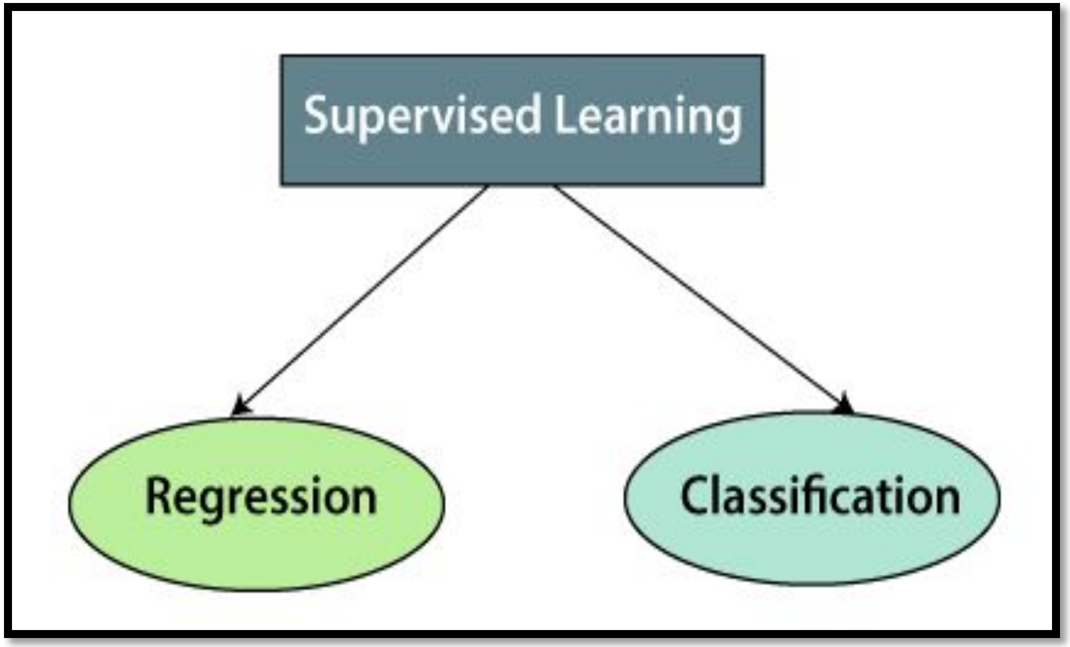
- **Supervised learning with an example**.
- Suppose we have an input dataset of cats and dog images. So, first, we will provide the training to the machine to understand the images, such as the shape & size of the tail of cat and dog, Shape of eyes, colour, height etc.
- After completion of training, we input the picture of a cat and ask the machine to identify the object and predict the output.
- Now, the machine will check all the features of the object, such as height, shape, colour, eyes, ears, tail, etc., and find that it's a cat.
- **The main goal of the supervised learning technique is to map the input variable(x) with the output variable(y).**

Supervised Machine Learning

- **Steps Involved in Supervised Learning:**
- First Determine the **type of training dataset**
- Collect/ **Gather the labelled training data**.
- Split the training dataset into **training dataset, test dataset, and validation dataset.**
- Determine the input features of the training dataset, which should have enough knowledge so that the model can accurately predict the output.
- **Determine the suitable algorithm** for the model, such as support vector machine, decision tree, etc.
- Execute the algorithm on the training dataset. Sometimes we need validation sets as the control parameters, which are the subset of training datasets.
- **Evaluate the accuracy of the model** by providing the test set. If the model predicts the correct output, which means our model is accurate.

Supervised Machine Learning

- Categories of Supervised Machine Learning
- Supervised machine learning can be classified into two types of problems, which are given below:
- **Classification**
- **Regression**



Supervised Machine Learning

- **Classification:**
- Classification algorithms are used to solve the classification problems in which the output variable is categorical, such as **"Yes" or No.**
- The classification algorithms predict the categories present in the dataset. Some real-world examples of classification algorithms are **Spam Detection, Email filtering**, etc.
- Some popular classification algorithms are given below:
- **Random Forest Algorithm**
- **Decision Tree Algorithm**
- **Logistic Regression Algorithm**
- **Support Vector Machine Algorithm**

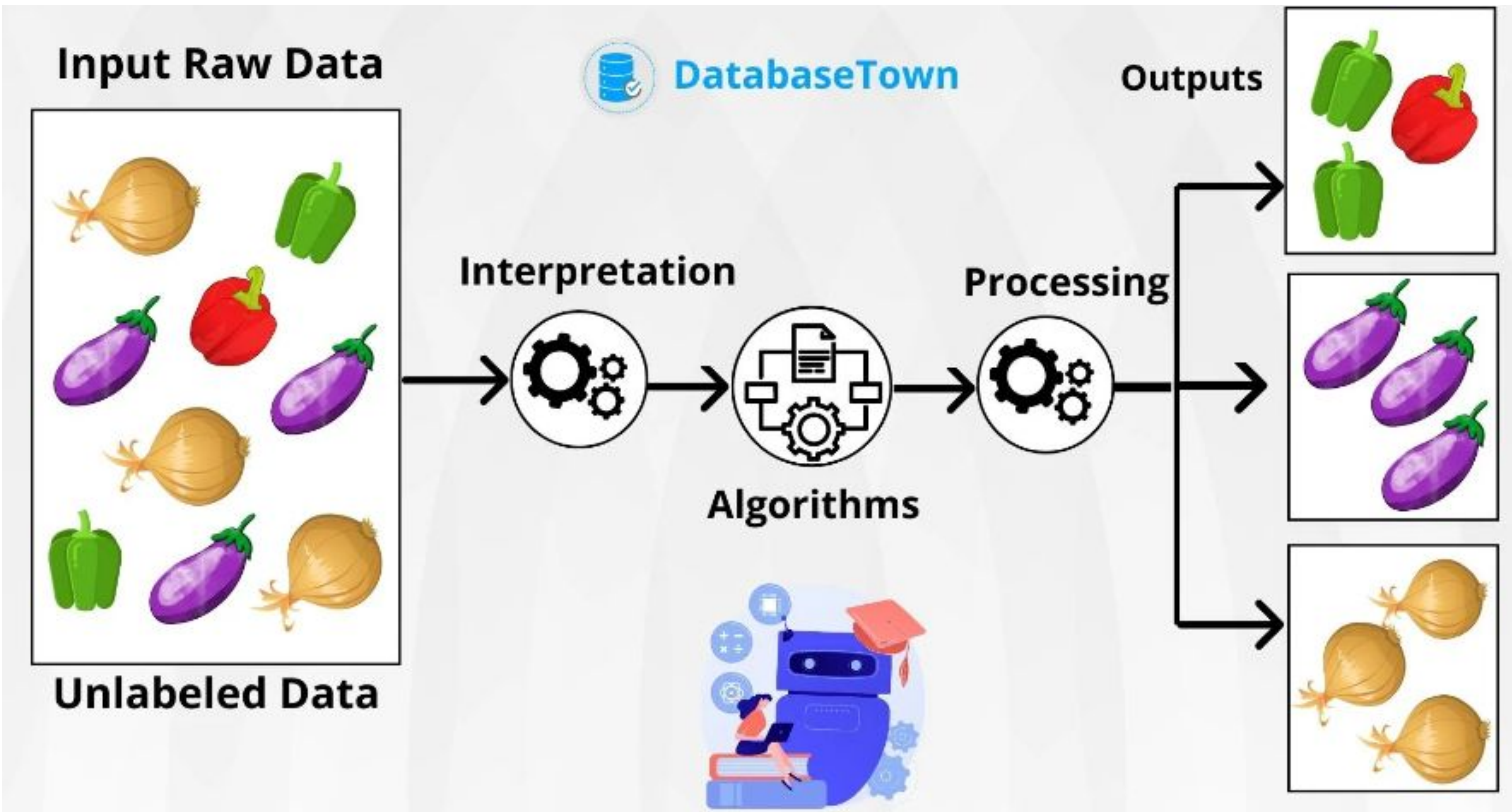
Supervised Machine Learning

- **Regression:**
- Regression algorithms are used to solve regression problems in which there is a **linear relationship between input and output variables**. These are used to predict continuous output variables, such as **market trends**, **weather prediction**, etc.
- Some popular Regression algorithms are given below:
- **Simple Linear Regression Algorithm**
- **Multivariate Regression Algorithm**
- **Decision Tree Algorithm**
- **Lasso Regression**

Unsupervised Machine Learning

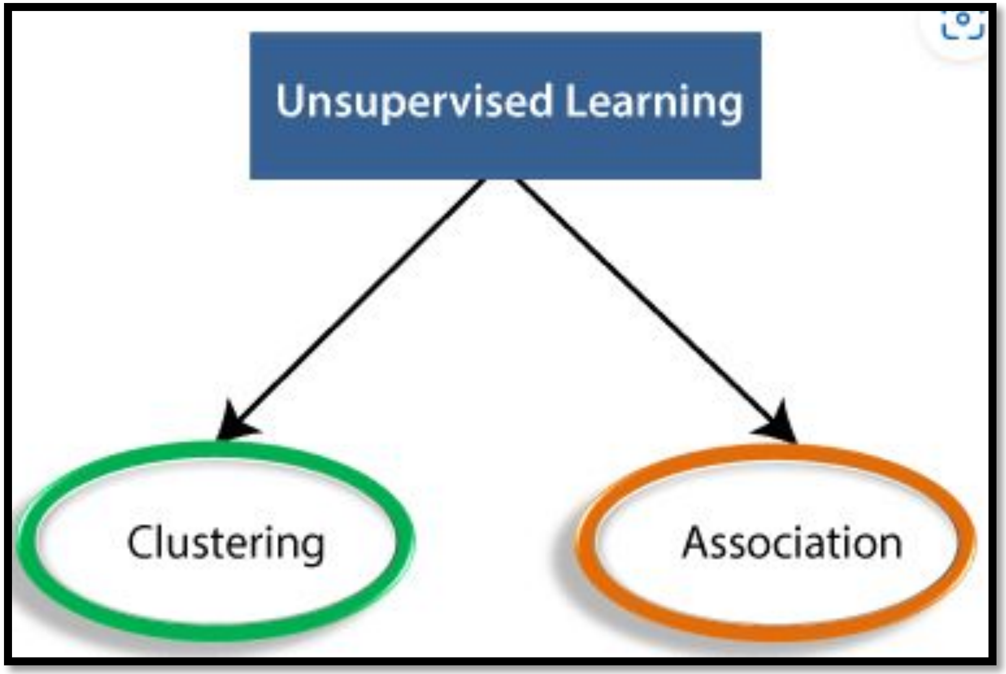
- Unsupervised learning is different from the Supervised learning technique; as its name suggests, there is no need for supervision. It means, in unsupervised machine learning, the machine is trained using the unlabeled dataset, and the machine predicts the output without any supervision.
- In unsupervised learning, the models are trained with the data that is neither classified **nor labelled, and the model acts on that data without any supervision**.
- The main aim of the unsupervised learning algorithm is to group or categories the unsorted dataset according to the similarities, patterns, and differences. Machines are instructed to find the hidden patterns from the input dataset.

Unsupervised Machine Learning



Unsupervised Machine Learning

- Categories of Unsupervised Machine Learning
- Unsupervised Learning can be further classified into two types, which are given below:
- **Clustering**
- **Association**



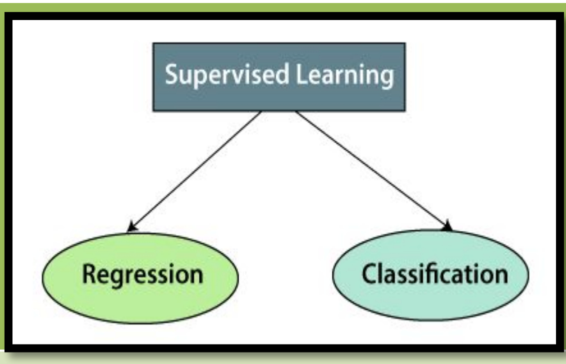
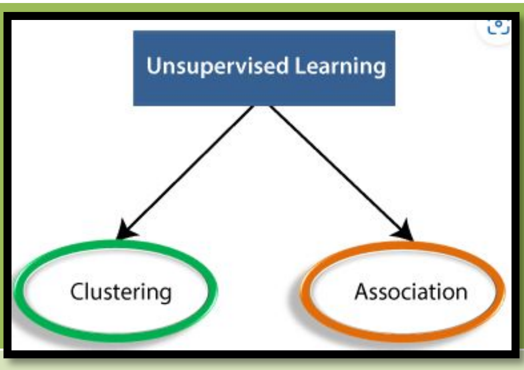
Unsupervised Machine Learning

- **Clustering:**
- The clustering technique is used when we want to find the inherent groups from the data. It is a way to group the objects into a cluster such that the objects with the **most similarities remain in one group** and have **fewer or no similarities** with the objects of other groups. **EX: Fraud detection, Fake news detection**
- Some of the popular clustering algorithms are given below:
- **K-Means Clustering algorithm**
- **Mean-shift algorithm**
- **DBSCAN Algorithm**
- **Principal Component Analysis**
- **Independent Component Analys**is

Unsupervised Machine Learning

- **Association:**
- Association rule learning is an unsupervised learning technique, which finds interesting relations among variables within a large dataset.
- The main aim of this learning algorithm is to find the dependency of one data item on another data item and map those variables accordingly so that it can generate maximum profit. This algorithm is mainly applied in **Market Basket analysis, Web usage mining, continuous product**ion, etc.
- Some popular algorithms of Association rule learning are **Apriori Algorithm, Eclat, FP-growth algorithm**.

MODULES OF MACHINE LEARNING

Model Selection:	 
Supervised Learning: Works with labeled data	**Regression:** To define the relationship between the input variable and the output variable. **EX: Linear Regression, Non-Linear Regression, Polynomial Regression** **Classification:** Output should be categorized as yes/No **EX: Random Forest, Decision Trees, SVM, Naïve Bayes**
Unsupervised Learning: Unsupervised learning works on unlabeled and uncategorized data	**Clustering:** Grouping of objects into the cluster **K-means clustering**, **K-Nearest Neighbors**, **Hierarchal clustering** **Association:** For finding the relationships between variables in the large database

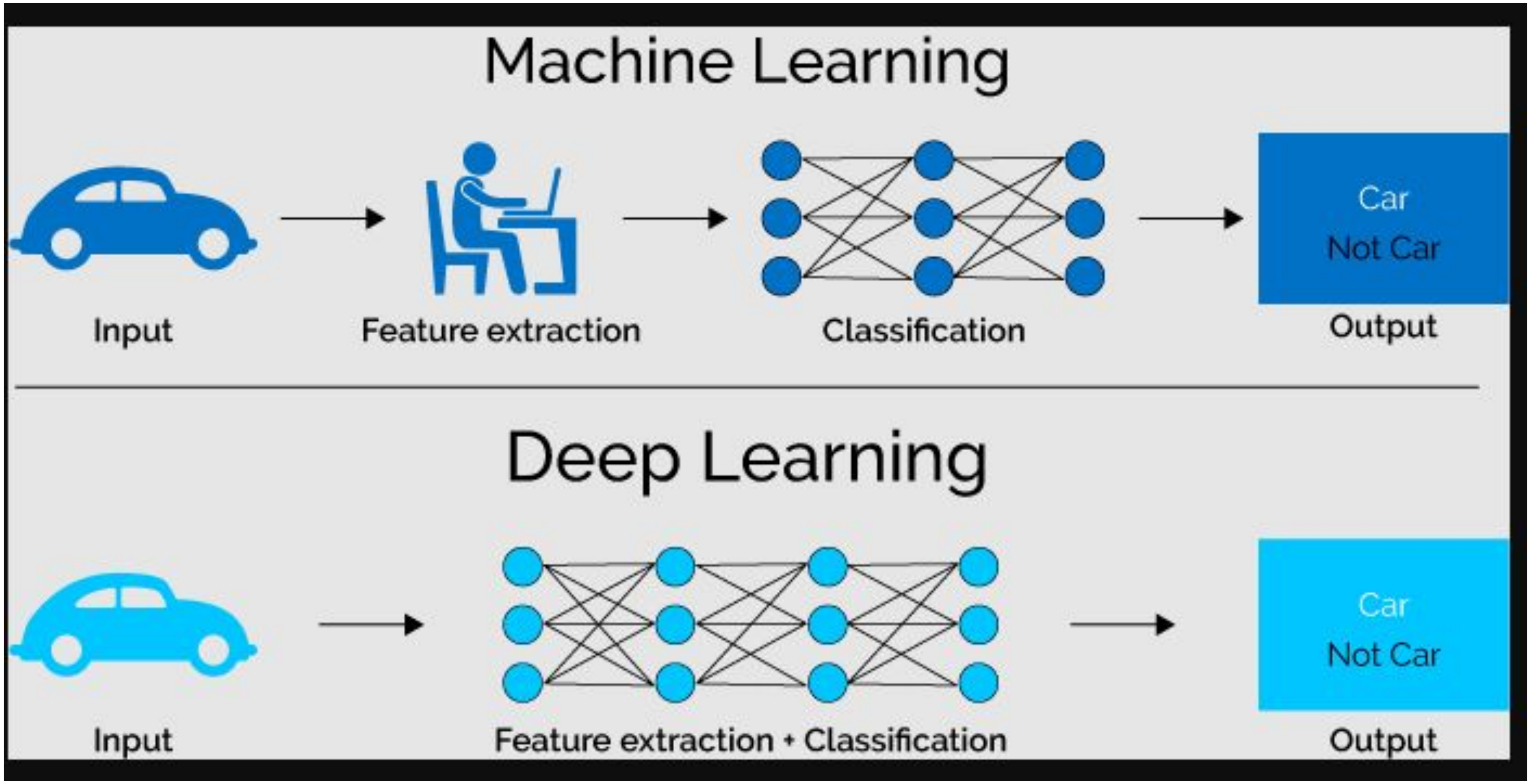
Reinforcement Machine Learning

- **Reinforcement Learning:**
- Reinforcement learning works on a feedback-based process, in which an AI agent (A software component) automatically explore its surrounding by hitting & trail, taking action, learning from experiences, and improving its performance.
- Agent gets rewarded for each good action and get punished for each bad action; hence the goal of reinforcement learning agent is to maximize the rewards.

Reinforcement Machine Learning

- **Categories of Reinforcement Learning:**
- Reinforcement learning is categorized mainly into two types of methods/algorithms:
- **Positive Reinforcement Learning:** Positive reinforcement learning specifies increasing the tendency that the required behaviour would occur again by adding something. It enhances the strength of the behaviour of the agent and positively impacts it.
- **Negative Reinforcement Learning:** Negative reinforcement learning works exactly opposite to the positive RL. It increases the tendency that the specific behaviour would occur again by avoiding the negative condition.

Difference between ML & DL



Difference between ML & DL

- Unsupervised:
- Identifying Fake News
- Marketing and Sales
- **Supervised:**
- Stock Prices
- Predicting the weather
- Classification of Genders

Unit 1 Data in Machine Learning

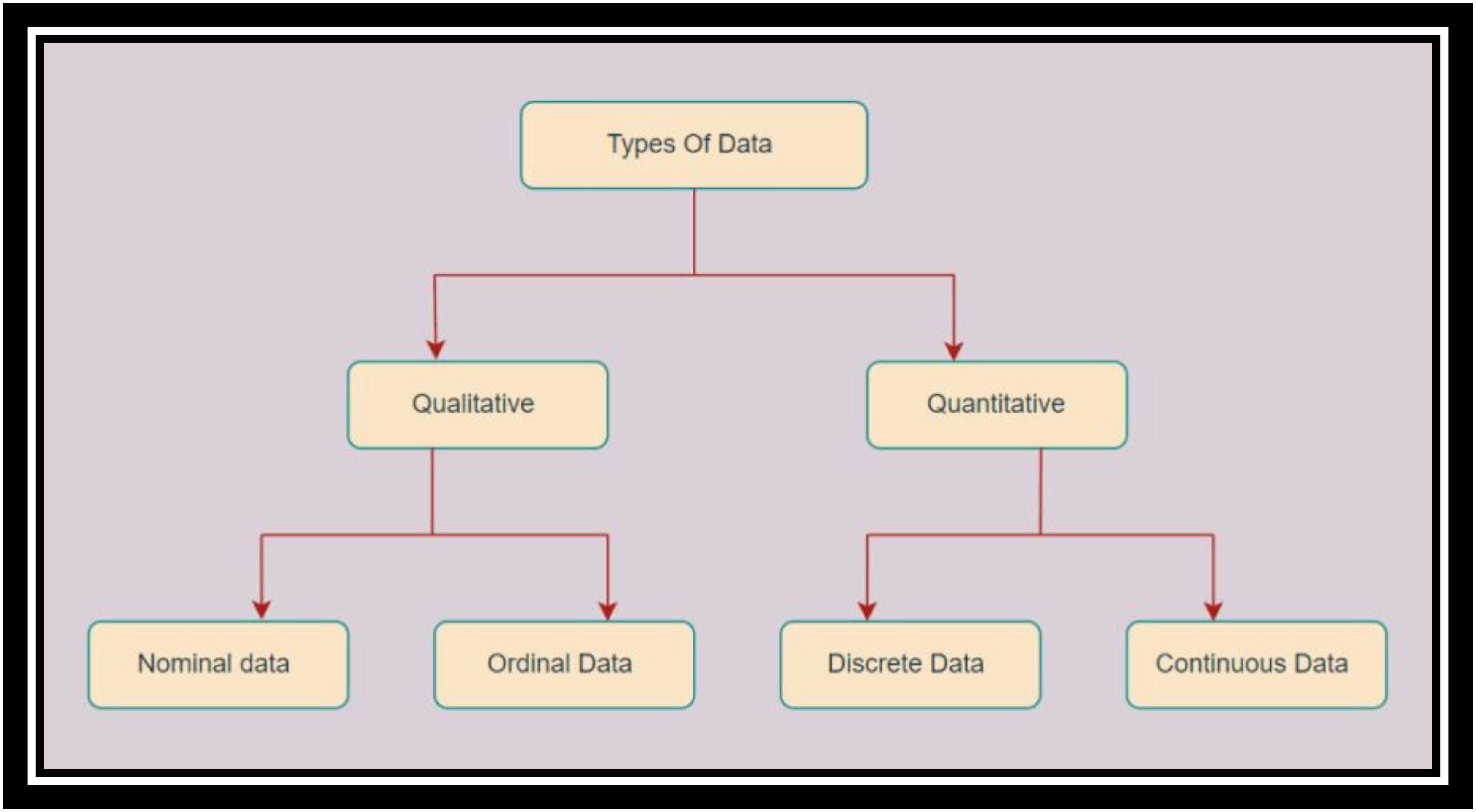
OUTLINE:

- ❖Data
- ❖Curse of Dimensionality
- ❖Over fitting & Under fitting

DATA IN MACHINE LEARNING

Data in Machine Learning:	Data is a crucial component in the field of Machine Learning. It refers to the set of observations or measurements that can be used to train a machine-learning model. The quality and quantity of data available for training and testing play a significant role in determining the performance of a machine-learning model. **Data is typically divided into two types:** •**Labeled data** •**Unlabeled data**
Data Collection:	1.Nominal data 2. Ordinal Data 3. Discrete Data 4. Continuous Data

DATA IN MACHINE LEARNING



DATA IN MACHINE LEARNING

1.Nominal data	If a feature represents a characteristic measured in numbers , it is called a numeric feature or **Quantitative Example:** Colour of hair (Blonde, red, Brown, Black, etc.) Marital status (Single, Widowed, Married) Nationality (Indian, German, American) The color of hair can be considered nominal data, as one color can't be compared with another color.
2. Ordinal Data	This denotes a nominal variable with categories falling in an ordered list . **Example:** Letter grades in the exam (A, B, C, D, etc.) Ranking of people in a competition (First, Second, Third, etc.) **Quantitative**
3. Discrete Data	The discrete data contain the values that fall under integers or whole numbers. **Example:** Total numbers of students present in a class, Days in a week
4. Continuous Data	A continuous variable can take any values. Example: Age, weight, height, length, time, and temperature.

DATA IN MACHINE LEARNING

Data Preprocessing:	**Process of transform or encoding the data to make it suitable for a machine learning model.** **1. Data Cleaning 2.Data Transformation 3.Data Integration 4. Dimension Reduction**
Data Cleaning:	**Missing Value:** Remove the missing data Filling the missing value using regression Replacing missing value with mean **Noisy Data:** Clustering, Regression
Data Integration:	Combining data residing in different sources and providing users with a unified view of these data
Data Transformation	**Normalization**: Numerical attributes are scaled up or down to fit within a specified range **[MIN-MAX normalization]**
Dimension Reduction	Reduces the dimension of the features Identify data patterns based on the features correlations Handle with large data [PCA]

DATA IN MACHINE LEARNING

handling missing data (Missing data replaced with Mean)	from sklearn. preprocessing import Imputer (estimator to fill the missing values) imputer= Imputer(missing_values ='NaN', strategy='mean',) Imputer imputer= imputer.fit(x[:, 1:3]) x[:, 1:3]= imputer.transform(x[:, 1:3])

Data Transformation :	# initialising the MinMaxScaler
	scaler = MinMaxScaler(feature_range=(0, 1))
	# learning the statistical parameters for each of the data and transforming
	rescaledX = scaler.fit_transform(X)

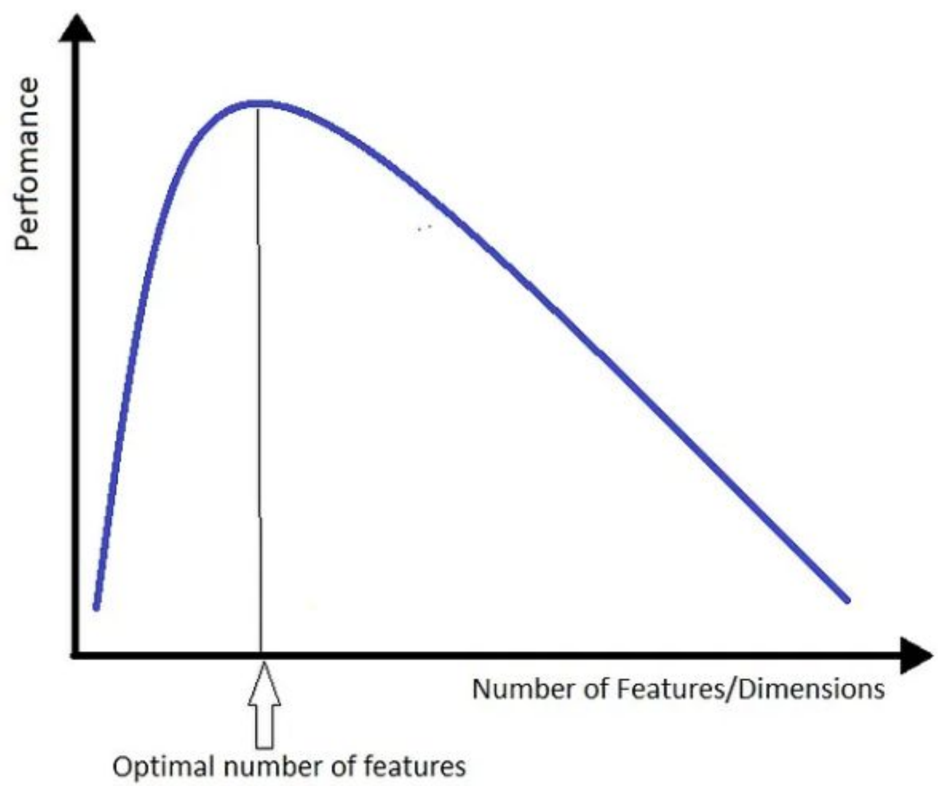
Dimension Reduction:	from sklearn.decomposition import PCA pca = PCA() X_train = pca.fit_transform(X_train) X_test = pca.transform(X_test)

Curse of dimensionality in Machine Learning

Definition:	Machine learning can effectively analyze data with several dimensions. However, it becomes complex to develop relevant models as the number of dimensions significantly increases. You will get abnormal results when you try to analyze data in high-dimensional spaces. This situation refers to the curse of dimensionality in machine learning.
Dimensions:	Dimensions are **features** that may be **dependent or independent**. Suppose we build model-1 with 3 features and model-2 with 5 features (both models have the same dataset). The model-2 has more information than model-1 because its number of features is comparatively higher. So, the accuracy of model-2 is more than that of model-1. With the **increase in the number of features**, the **model's accuracy increases**. However, after a specific threshold value, the model's accuracy will not increase, although the number of features increases. This is because a model is fed with a lot of information, making it incompetent to train with correct information. **The phenomenon when a machine learning model's accuracy decreases, although increasing the number of features after a certain threshold, is called the curse of dimensionality.**

Curse of dimensionality in Machine Learning

Why is it challenging to analyze high-dimensional data?	When more dimensions are added to a machine learning model, the processing power required for the data analysis increases.



Curse of dimensionality in Machine Learning

1. Hughes Phenomenon:

"Dimensionality Reduction" is the data conversion from a high-dimensional into a low-dimensional space.

How does Dimensionality Reduction help solve the Curse of Dimensionality?

It decreases the dataset's dimensions and thus decreases the storage space.

It significantly decreases the computation time. This is because less number of dimensions need less computing time, and ultimately the algorithms train faster than before.

It improves models' accuracy.

2. PCA:

One of the conventional tools capable of solving the curse of dimensionality is PCA (Principal Component Analysis).

One of the conventional tools capable of solving the curse of dimensionality is PCA (Principal Component Analysis).

3. Deep Learning Technique

Overfitting and underfitting

Overfitting and underfitting:

Overfitting occurs when a model is too complex and fits the training data too closely, resulting in poor generalization to new data.

How to avoid the Overfitting in Model:

- Cross-Validation
- Training with more data
- Removing features

Underfitting **occurs when a model is too simple and does not capture all the relevant patterns** in the data.

How to avoid underfitting:

By increasing the training time of the model.

By increasing the number of features.

Unit 1 Bias and Variance

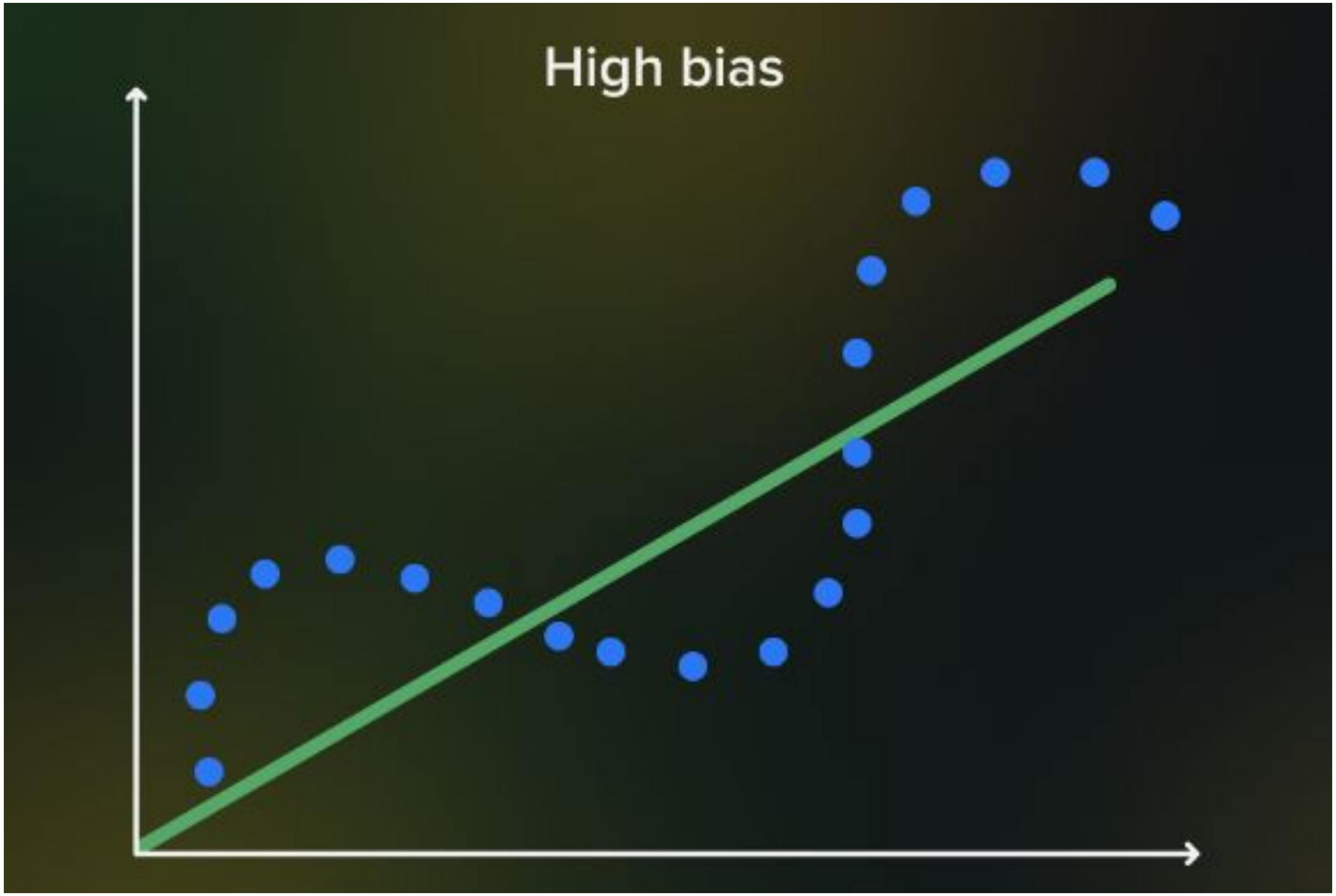
OUTLINE:

- ❖Bias
- ❖Variance
- ❖Bias and Variance Tradeoff

BIAS
IN MACHINE LEARNING

Bias:	Bias in machine learning refers to the difference between a model's predictions and the actual distribution of the value it tries to predict. The bias is known as the difference between the **prediction of the values by the Machine Learning model and the correct value**. Being high in biasing gives a large error in training as well as testing data. It recommended that an algorithm should always be **low-biased to avoid the problem of underfitting**. By high bias, the data predicted is in a straight line format, thus not fitting accurately in the data in the data set. Such fitting is known as the Underfitting of Data.

BIAS IN MACHINE LEARNING

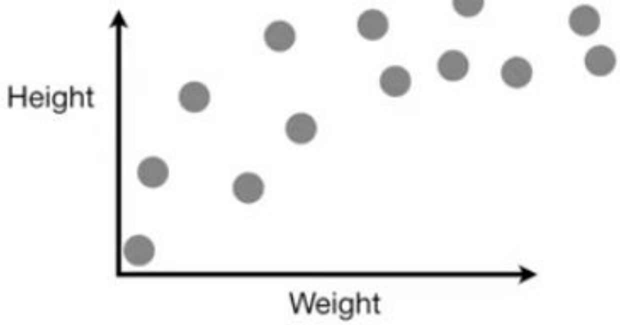


TO REDUCE BIAS PROBLEM

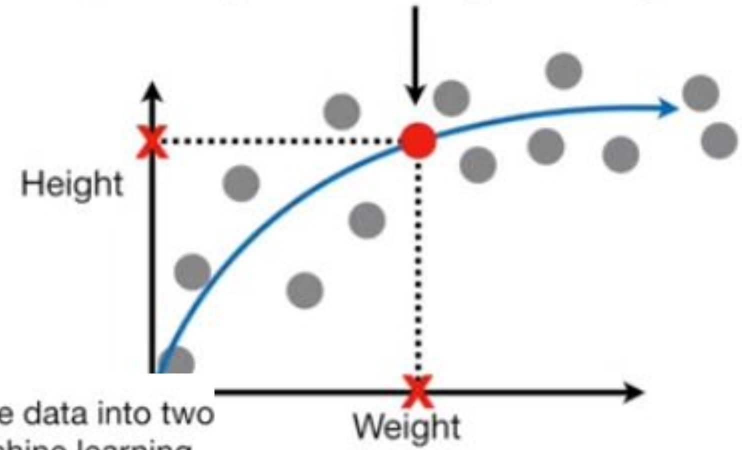
How to reduce high bias?	1. Incorporating **additional features** from data to improve the model's accuracy. 2. Increasing the **number of training iterations** to allow the model to learn more complex data. 3. **Avoiding high-bias algorithms** such as linear regression, logistic regression, discriminant analysis, etc. and instead **using nonlinear algorithms** such as k-nearest neighbors, SVM, decision trees, etc. 4. Decreasing regularization at various levels to help the model learn the training set more effectively and prevent underfitting.

EXAMPLE OF BIAS & VARIANCE

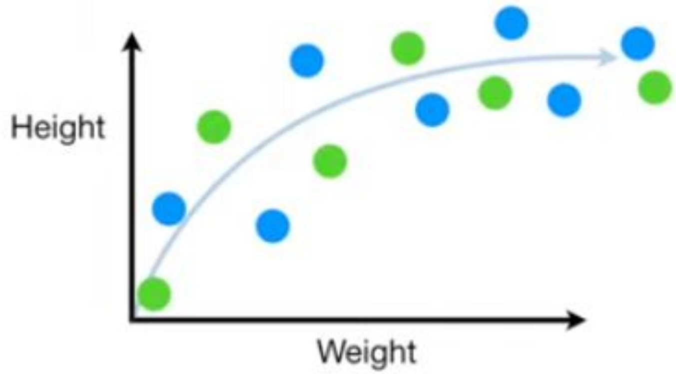
Imagine we measured the weight and height of a bunch of mice and plotted the data on a graph...



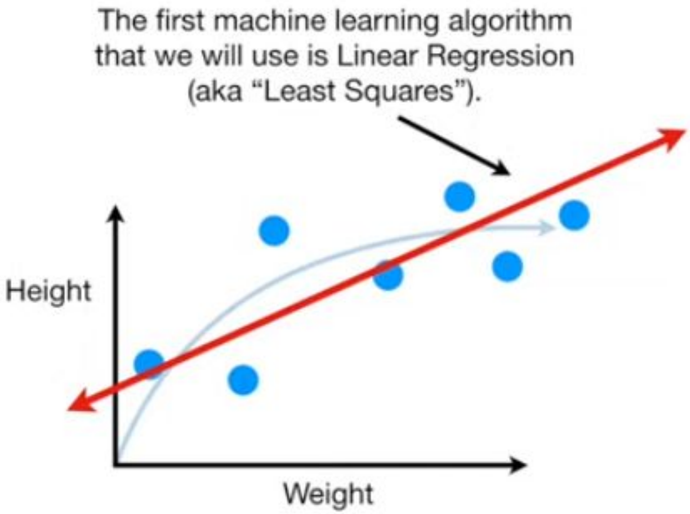
Ideally, we would know the exact mathematical formula that describes the relationship between weight and height...

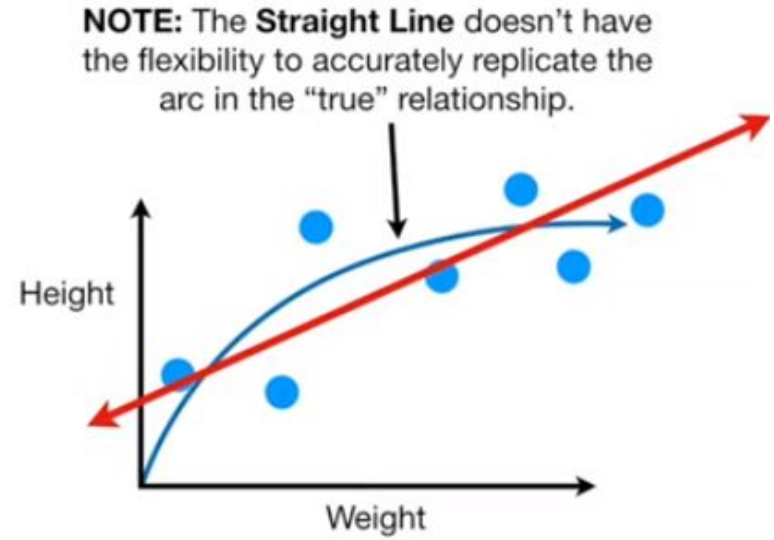


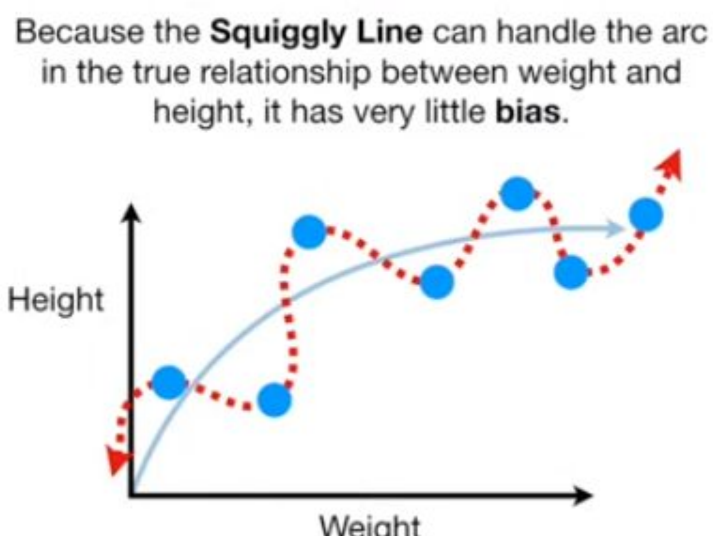
The first thing we do is split the data into two sets, one for training the machine learning algorithms and one for testing them.



EXAMPLE OF BIAS & VARIANCE

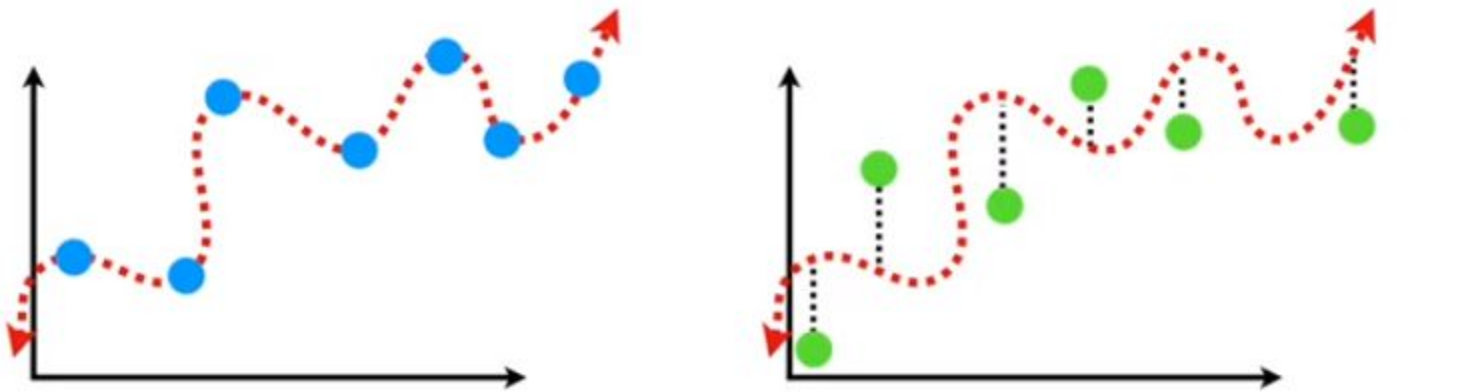




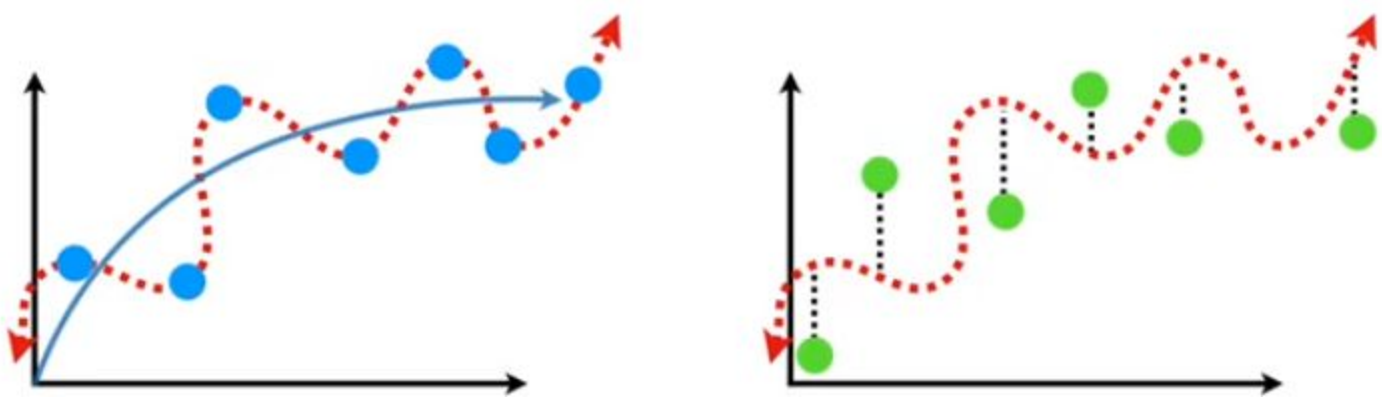


EXAMPLE OF BIAS & VARIANCE

In Machine Learning lingo, the difference in fits between data sets is called **Variance**.



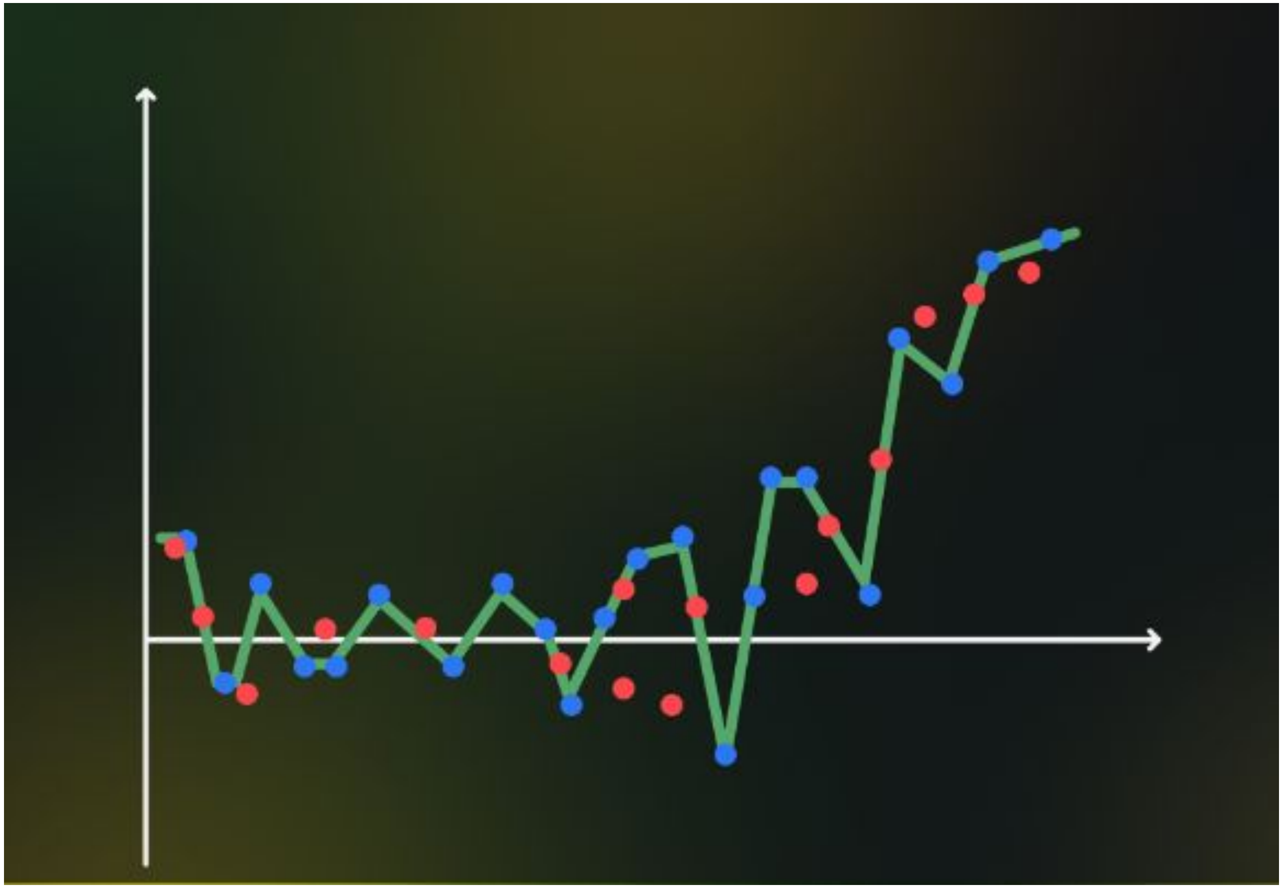
...but the **Squiggly Line** has **high variability**, because it results in vastly different Sums of Squares for different datasets.



VARIANCE
IN MACHINE LEARNING

Bias:	Variance stands in **contrast to bias**; it measures how much a distribution on several sets of data values differs from each other. The model with **high variance has a very complex** fit to the training data and thus is not able to fit accurately on the data which it hasn’t seen before. As a result, such models perform very well on training data but have high error rates on test data. **When a model is high on variance, it is then said to as Overfitting of Data**.

VARIANCE IN MACHINE LEARNING



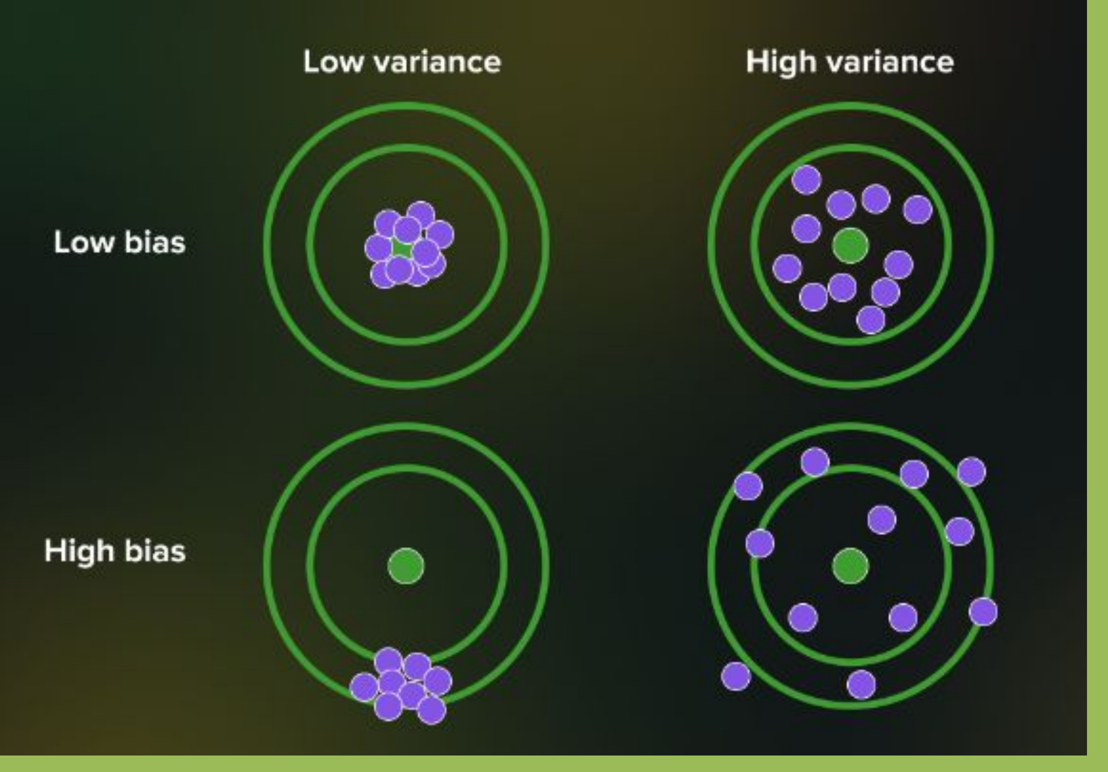
TO REDUCE VARIANCE PROBLEM

How to reduce high variance?	1. Reducing the number of features in the model. 2. Replacing the current model with a simpler one. 3. Increasing the training data diversity to balance out the complexity of the model and the data structure. 4. Avoiding high-variance algorithms (support vector machines, decision trees, k-nearest neighbors, etc.) and opt for low-variance ones such as linear regression, logistic regression, and linear discriminant analysis.

Bias-Variance Scenarios

How to reduce high variance?

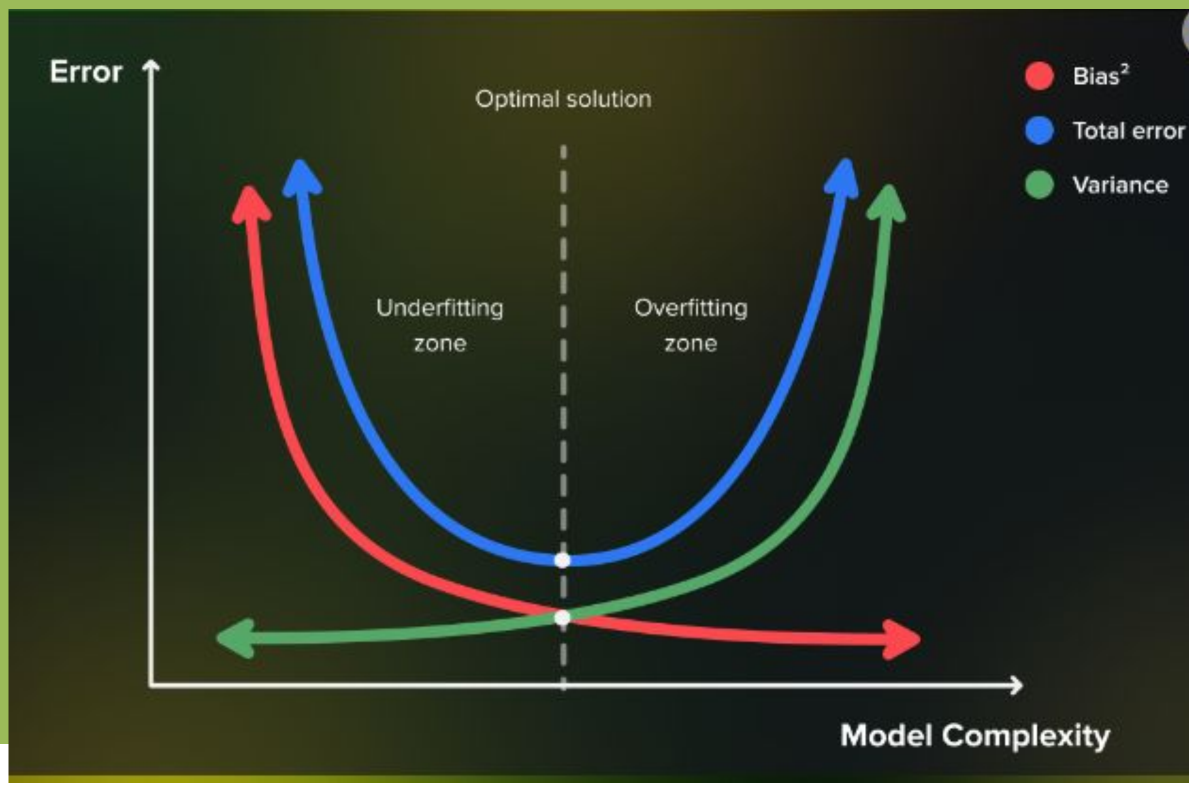
1. Low bias, low variance: ideal model
2. Low bias, high variance: results in overfitting
3. High bias, low variance: results in underfitting
4. High bias, high variance: results in inaccurate predictions



Bias-Variance Tradeoff

Bias Variance Tradeoff

If the algorithm is too simple (hypothesis with linear equation) then it may be on high bias and low variance condition and thus is error-prone. If algorithms fit too complex (hypothesis with high degree equation) then it may be on high variance and low bias. This tradeoff in complexity is why there is a tradeoff between bias and variance.



Bias-Variance Tradeoff

Bias Variance Tradeoff	•The balance between bias and variance can be adjusted in specific algorithms by modifying parameters, as seen in the following examples: •For **k-nearest neighbors**, a low bias and high variance can be corrected by increasing the **value of k**, which increases the bias and decreases the variance. •For **support vector machines**, a low bias and high variance can be altered by adjusting **the C parameter**, which increases the bias but decreases the variance.

Unit 1 Testing Cross Validation

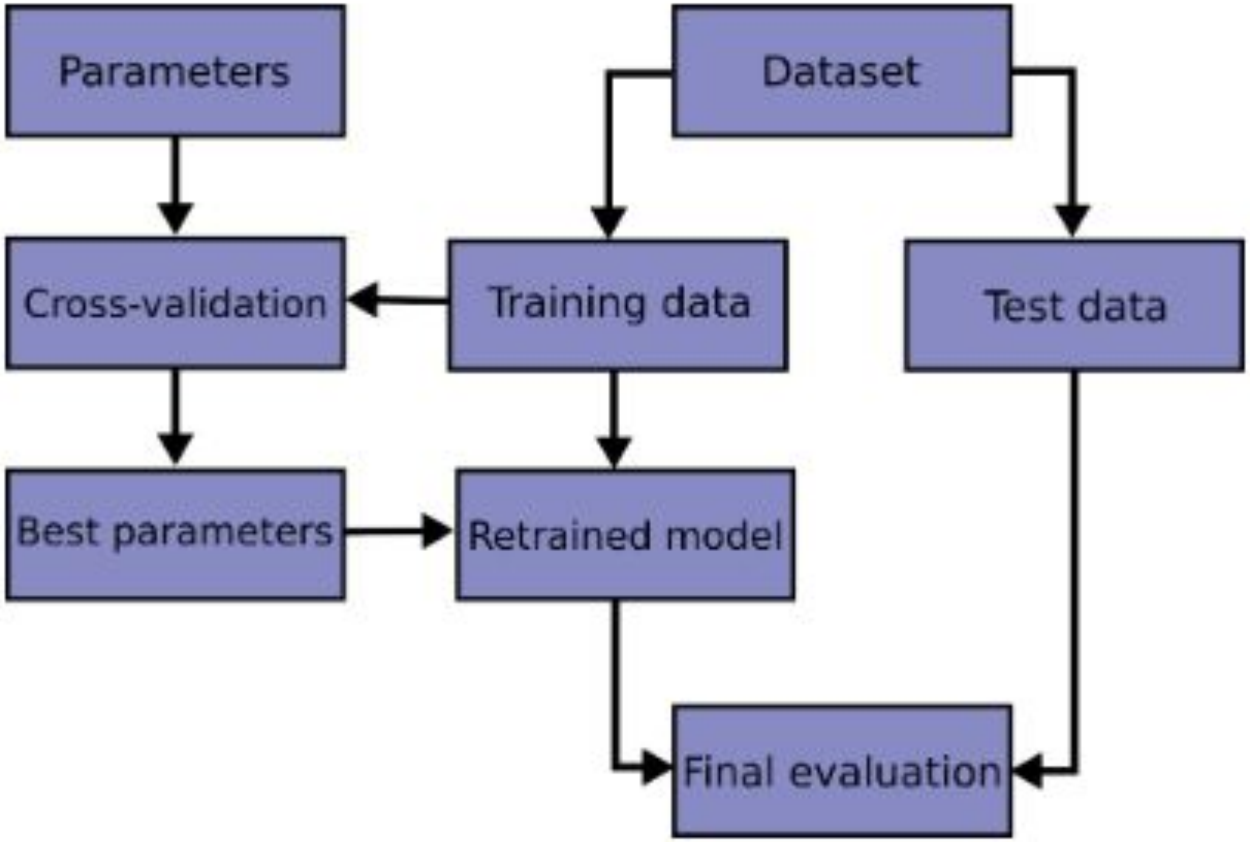
OUTLINE:

❖Cross Validation

CROSS VALIDATION IN MACHINE LEARNING

Crossvalidation:	•Cross-validation is a resampling method that uses **different portions of the data to test and train a model on different iterations.** •It is mainly used in settings where the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice. •The **80–20** is not an actual rule, and alternative ratios that range between **25~30% for testing and 70~75% for training.** • For instance, **70%** of the data is used for **training**, **20% for validation**, and the remaining **10%** is used for testing.

CROSS VALIDATION S IN MACHINE LEARNING



CROSS VALIDATION
IN MACHINE LEARNING

How does it work?

- Cross-Validation has two main steps: splitting the data into subsets (called folds) and rotating the training and validation among them.
- The splitting technique commonly has the following properties:
- **Each fold has approximately the same size.**
- **Data can be randomly selected in each fold or stratified**.
- All folds are used to train the model except one, which is used for validation.

TO REDUCE BIAS PROBLEM

How does it work? **K-fold** and **CV** are two terms that are used interchangeably. K-fold is just describing how many folds you want to split your dataset into. Many libraries use k=10 as a default value representing 90% going to training and 10% going to the validation set. The next figure describes the process of iterating over the picked ten folds of the dataset..

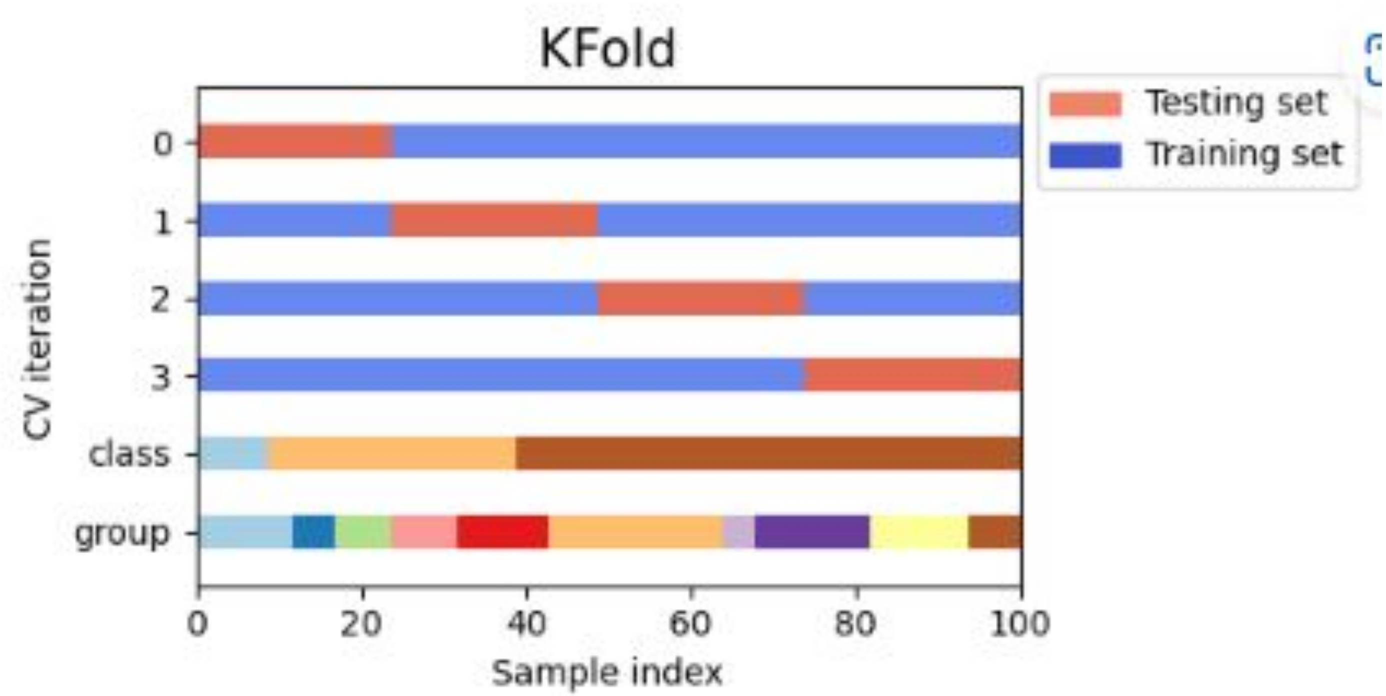
	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Fold-6	Fold-7	Fold-8	Fold-9	Fold-10
Step-1	Train	Train	Train	Train	Train	Train	Train	Train	Train	Test
Step-2	Train	Train	Train	Train	Train	Train	Train	Train	Test	Train
Step-3	Train	Train	Train	Train	Train	Train	Train	Test	Train	Train
Step-4	Train	Train	Train	Train	Train	Train	Test	Train	Train	Train
Step-5	Train	Train	Train	Train	Train	Test	Train	Train	Train	Train
Step-6	Train	Train	Train	Train	Test	Train	Train	Train	Train	Train
Step-7	Train	Train	Train	Test	Train	Train	Train	Train	Train	Train
Step-8	Train	Train	Test	Train	Train	Train	Train	Train	Train	Train
Step-9	Train	Test	Train	Train	Train	Train	Train	Train	Train	Train
Step-10	Test	Train	Train	Train	Train	Train	Train	Train	Train	Train

Figure 2: A 10-fold representation of how each fold is used in the cross-validation process.

EXAMPLE OF BIAS & VARIANCE

```
import numpy as np
>>> from sklearn.model_selection import KFold
>>> X = ["a", "b", "c", "d"]
>>> kf = KFold(n_splits=4)
>>> for train, test in kf.split(X): ...
("%s %s" % (train, test))
```

EXAMPLE OF BIAS & VARIANCE



Unit 1 Classification

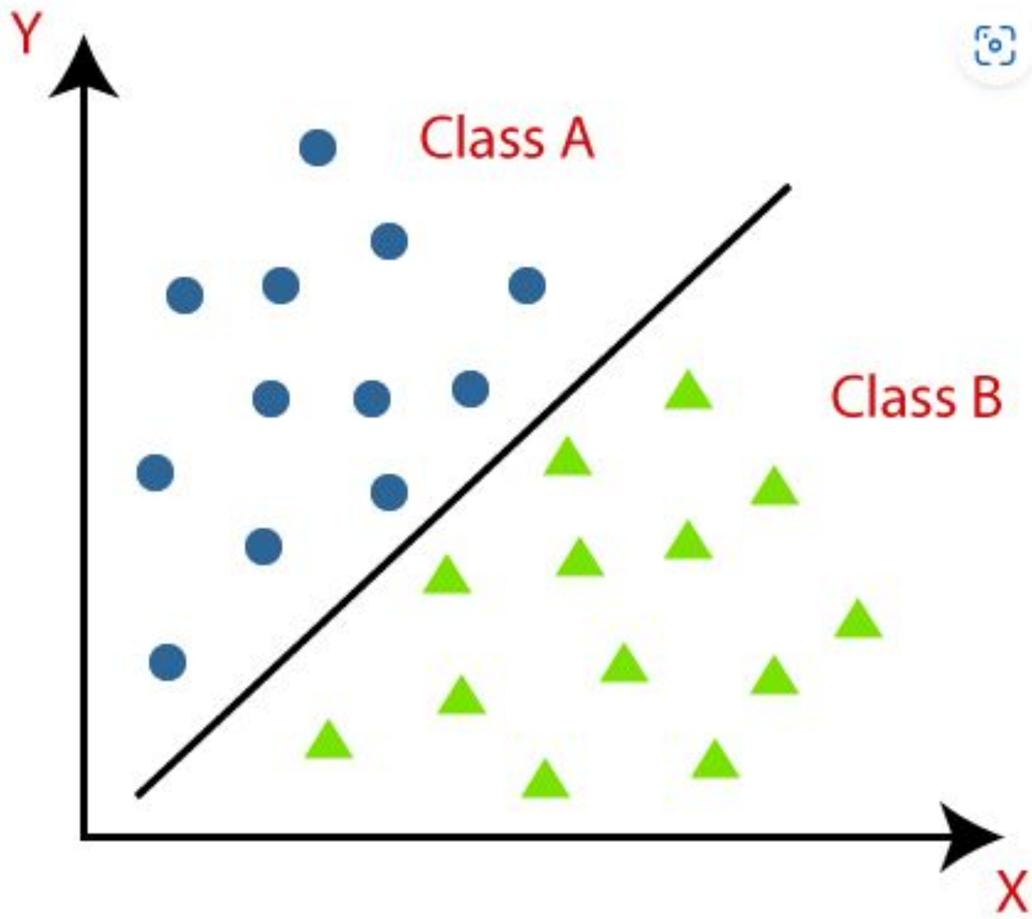
OUTLINE:

- Classification

CLASSIFICATION IN MACHINE LEARNING

Classification:	•Based on training data, the Classification algorithm is a Supervised Learning technique used to categorize new observations. •In classification, a program uses the dataset or observations provided to learn how to categorize new observations into various classes or groups. • For instance, 0 or 1, red or blue, yes or no, spam or not spam, etc. Targets, labels, or categories can all be used to describe classes. •The Classification algorithm uses labeled input data because it is a supervised learning technique and comprises input and output information. A discrete output function (y) is transferred to an input variable in the classification process (x).

CROSS VALIDATION S IN MACHINE LEARNING



Learners in Classification Problems

LEARNERS:

There are two types of learners.

Lazy Learners:

- It first stores the training dataset before waiting for the test dataset to arrive.
- When using a lazy learner, the classification is carried out using the training dataset's most appropriate data.
- Less time is spent on training, but more time is spent on predictions.
- Some of the examples are **case-based reasoning and the KNN algorithm.**

Eager Learners:

- Before obtaining a test dataset, eager learners build a classification model using a training dataset.
- They spend more time studying and less time predicting.
- Some of the examples are **ANN, naive Bayes, and Decision trees.**

Types of ML Classification Algorithms

Classification Algorithm:

Classification Algorithms can be further divided into the Mainly two category:

Linear Models:

- Logistic Regression
- Support Vector Machines

Non-linear Models:

- K-Nearest Neighbours
- Kernel SVM
- Naïve Bayes
- Decision Tree Classification
- Random Forest Classification

4 Types Of Classification Tasks In Machine Learning

Types:

Binary Classifier: If the classification problem has only two possible outcomes, then it is called as Binary Classifier.

Examples: YES or NO, MALE or FEMALE, SPAM or NOT SPAM, CAT or DOG, etc.

Multi-class Classifier: If a classification problem has more than two outcomes, then it is called as Multi-class Classifier.

Example: Classifications of types of crops, Classification of types of music.

4 Types Of Classification Tasks In Machine Learning

Types:

Binary Classifier: If the classification problem has only two possible outcomes, then it is called as Binary Classifier.

Examples: YES or NO, MALE or FEMALE, SPAM or NOT SPAM, CAT or DOG, etc.

Multi-class Classifier: If a classification problem has more than two outcomes, then it is called as Multi-class Classifier.

Example: Classifications of types of crops, Classification of types of music.

Evaluating a Classification Model

1. Confusion Matrix: The confusion matrix describes the model performance and gives us a matrix or table as an output.

The matrix appears in the following table:

	Actual Positive	Actual Negative
Predicted Positive	True Positive	False Positive
Predicted Negative	False Negative	True Negative

Evaluating a Classification Model

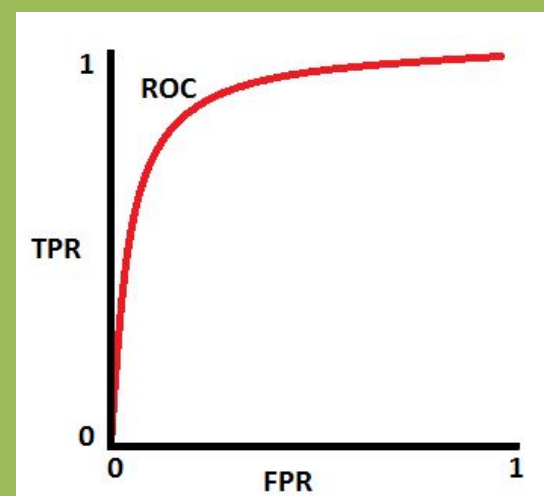
2. Log Loss or Cross-Entropy Loss:

- It is used to assess a classifier's performance, and the output is a probability value between 1 and 0.
- A successful binary classification model should have a log loss value that is close to 0.
- If the anticipated value differs from the actual value, the value of log loss rises.
- The lower log loss shows the model’s higher accuracy.
- Cross-entropy for binary classification can be calculated as:
- (ylog(p)+(1?y)log(1?p))
- Where p = Predicted Output, y = Actual output.

Evaluating a Classification Model

AUC-ROC Curve:

- AUC is for Area Under the Curve, and ROC refers to Receiver Operating Characteristics Curve.
- It is a graph that displays the classification model's performance at various thresholds.
- The AUC-ROC Curve is used to show how well the multi-class classification model performs.
- The TPR and FPR are used to draw the ROC curve, with the True Positive Rate (TPR) on the Y-axis and the FPR (False Positive Rate) on the X-axis.



Unit 1 Learning Curves

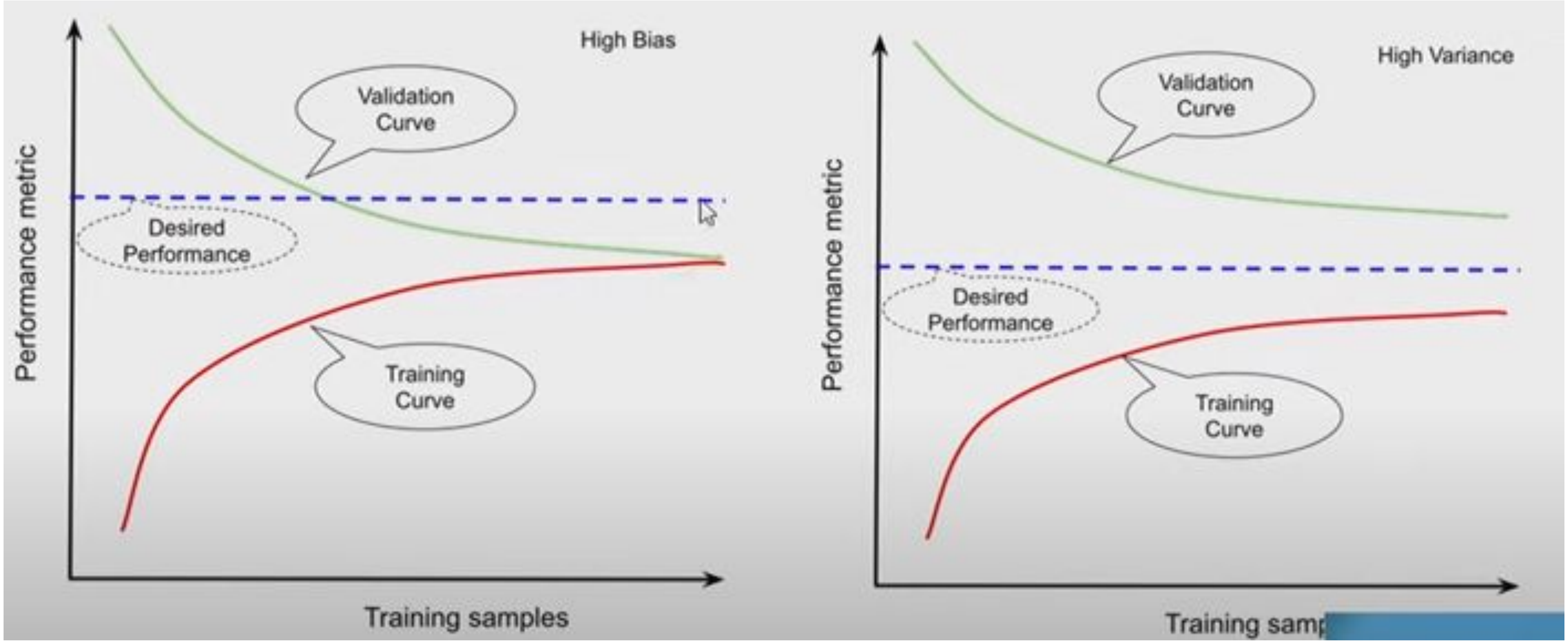
OUTLINE:

❖Learning Curves

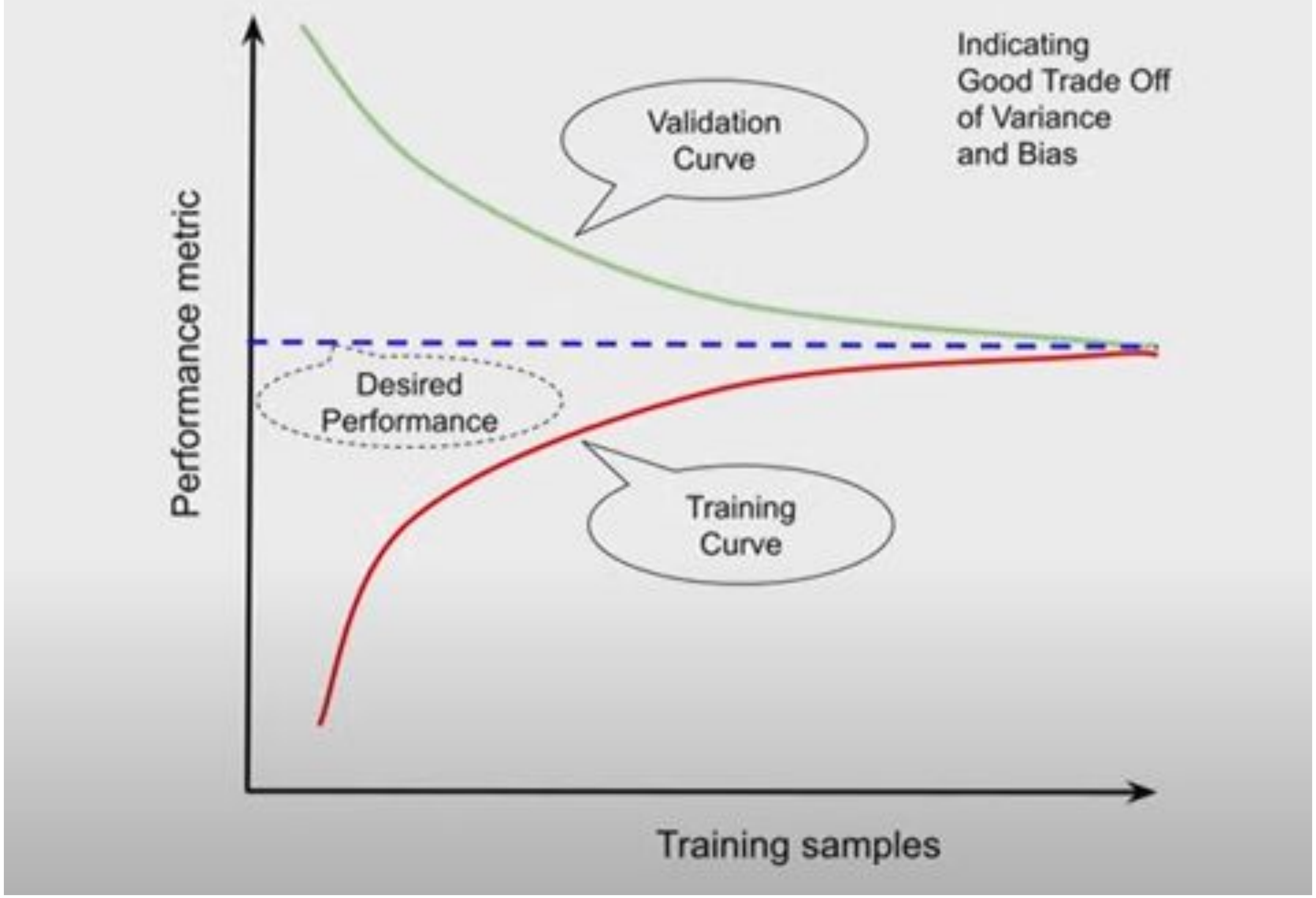
LEARNING CURVES
IN MACHINE LEARNING

Learning Curves:	•A learning curve **can help to find the right amount of training data to fit our model** with a good **bias-variance trade-off.** •Learning curves plot the **training and validation loss** of a sample of training examples by incrementally adding new training examples. •Learning curves help us in identifying whether adding additional training examples would improve the validation score (score on unseen data).

LEARNING CURVES IN MACHINE LEARNING



LEARNING CURVES IN MACHINE LEARNING



Unit 1 Regularization

OUTLINE:

- ❖Regularization

REGULARIZATION
IN MACHINE LEARNING

Overfitting:	•Overfitting is the common issues in the machine learning L1 & L2 regularizations are used the address the issues of overfitting. •Mean Square Error •Lasso Regularization •Ridge Regularization

REGULARIZATION IN MACHINE LEARNING

- **Lasso Regression** adds the *"absolute value of magnitude"* of the coefficient as a penalty term to the loss function(L).
- Lasso regression also helps us achieve feature selection by penalizing the weights to approximately equal to zero if that feature does not serve any purpose in the model.

L1 Regularization

$$ms\varepsilon = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - h_\theta(x_i)\right)^2 + \lambda \sum_{i=1}^{n}|\theta_i|$$

REGULARIZATION IN MACHINE LEARNING

Ridge Regression

A regression model that uses the **L2 regularization** technique is called **Ridge regression**. **Ridge regression** adds the "*squared magnitude*" of the coefficient as a penalty term to the loss function(L).

L2 Regularization

$$ms\varepsilon = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - h_\theta(x_i)\right)^2 + \lambda \sum_{i=1}^{n} \theta_i^2$$

Unit 1 Noise & Error

OUTLINE:

❖Noise & Error

NOISE & ERROR
IN MACHINE LEARNING

Noise:	•Humans are prone to making mistakes when collecting data, and data collection instruments may be unreliable, resulting in dataset errors. The errors are referred to as noise. •Data noise in machine learning can cause problems since the algorithm interprets the noise as a pattern and can start generalizing from it.

NOISE & ERROR IN MACHINE LEARNING

Machine learning noise detection and removal:	**Principal Component Analysis:** PCA attempts to eliminate corrupted data from a signal or picture using preservative noise while maintaining the critical features. PCA is a geometric and statistical method that reduces the input signal dimension or data by projecting it along various axes. **To better understand, imagine projecting a point in the XY dimension along the X-axis. The noise plane – Y-axis can now be removed**. The phenomenon is referred to as “dimensionality reduction.” As a result, by eliminating the axes containing the noisy data, principal component analysis can minimize the noise in input data.

NOISE & ERROR
IN MACHINE LEARNING

Machine learning noise detection and removal:	**Deep De-noising** •Auto-encoders are useful for de-noising; a stochastic variant of auto-encoder is available. Since they can be trained to recognize noise detection in a signal or data, they can be used as de-noisers by feeding them noisy data and receiving clean data as an output. •A de-noising auto-encoder does two things: it encodes the input while retaining as much detail about the output as possible. It also reverses the effects of stochastically added noise to the input data •De-noising auto-encoders' main goal is to push the secret layer to learn more robust features. The auto-encoder is then trained to reconstruct the input data from the degraded version while reducing loss. The use of auto-encoders to eliminate noise from a signal is demonstrated in one example.

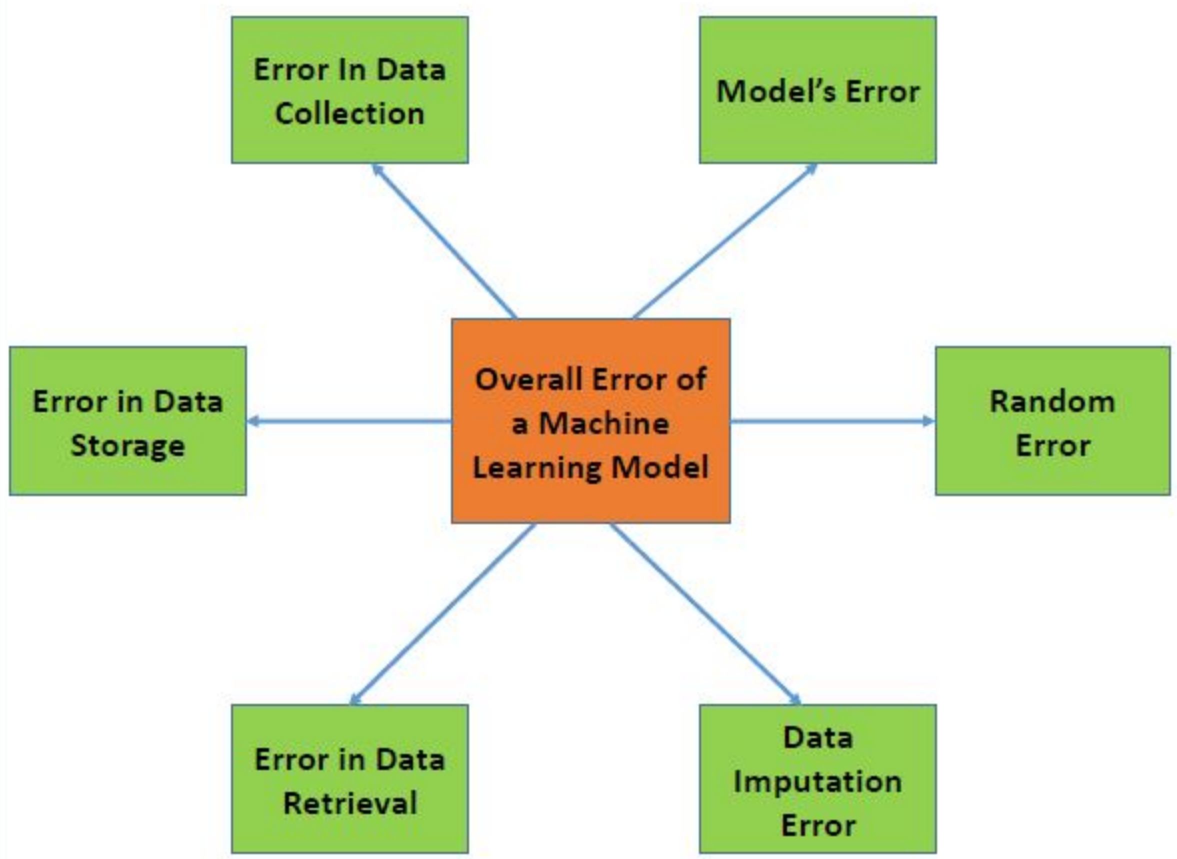
NOISE & ERROR
IN MACHINE LEARNING

Error:	•The difference between a measured value and the real value of the input. **1. Error in Data Collection:** •Data collection can produce errors at different levels. For instance, a survey could be designed for collecting data. However, individuals participating in the survey may not always provide the right information. For instance, a participant may enter the wrong information about their age, height, marital status, income, etc. **2. Error in Data Storage:** Storing data could lead to errors as some data could be saved incorrectly, or part of the data could be lost during the storage process.

NOISE & ERROR IN MACHINE LEARNING

Error:	**3. Error in Data Retrieval:** Retrieving data can also produce errors, as some part of the data may be missing or could be corrupted. **4. Data Imputation Error:** In this case, we can use different interpolation techniques to estimate the missing values from the other training samples in our dataset. One of the most common interpolation techniques is mean imputation, where we simply replace the missing value with the mean value of the entire feature column. Other options for imputing missing values are median or most frequent (mode), where the latter replaces the missing values with the most frequent values. This is useful for imputing categorical feature values. Another imputation technique.

NOISE & ERROR IN MACHINE LEARNING



Unit 1 Parametric & Non Parametric Models

OUTLINE:

- ❖Parametric & Non Parametric Models

PARAMTRIC & NON PARAMETRIC MODELS IN MACHINE LEARNING

Parametric Methods	Non-Parametric Methods
Parametric Methods uses a fixed number of parameters to build the model.	Non-Parametric Methods use the flexible number of parameters to build the model.
Parametric analysis is to test group means.	A non-parametric analysis is to test medians.
It is applicable only for variables.	It is applicable for both – Variable and Attribute.
It always considers strong assumptions about data.	It generally fewer assumptions about data.
Parametric Methods require lesser data than Non-Parametric Methods.	Non-Parametric Methods requires much more data than Parametric Methods.
Parametric methods assumed to be a normal distribution.	There is no assumed distribution in non-parametric methods.

PARAMTRIC & NON PARAMETRIC MODELS IN MACHINE LEARNING

Parametric data handles – Intervals data or ratio data.	But non-parametric methods handle original data.
Here when we use parametric methods then the result or outputs generated can be easily affected by outliers.	When we use non-parametric methods then the result or outputs generated cannot be seriously affected by outliers.
Parametric Methods can perform well in many situations but its performance is at peak (top) when the spread of each group is different.	Similarly, Non-Parametric Methods can perform well in many situations but its performance is at peak (top) when the spread of each group is the same.
Parametric methods have more statistical power than Non-Parametric methods.	Non-parametric methods have less statistical power than Parametric methods.
As far as the computation is considered these methods are computationally faster than the Non-Parametric methods.	As far as the computation is considered these methods are computationally slower than the Parametric methods.
Examples: Logistic Regression, Naïve Bayes Model, etc.	Examples: KNN, Decision Tree Model, etc.

Unit 1 Linear Algebra in Machine Learning

OUTLINE:

- ❖Linear Algebra in Machine Learning

PARAMTRIC & NON PARAMETRIC MODELS IN MACHINE LEARNING

- Machine learning has a strong connection with mathematics. Each machine learning algorithm is based on the concepts of mathematics & also with the help of mathematics, one can choose the correct algorithm by considering training time, complexity, number of features, etc.
- Linear Algebra is an essential field of mathematics, which defines the study of vectors, matrices, planes, mapping, and lines required for linear transformation.
- Linear algebra plays a vital role and key foundation in machine learning, and it enables ML algorithms to run on a huge number of datasets.

PARAMTRIC & NON PARAMETRIC MODELS IN MACHINE LEARNING

Linear Algebra in Machine Learning
Below are some popular examples of linear algebra in Machine learning:

1. **Datasets and Data Files**
2. **Linear Regression**
3. **Regression**
4. **Regularization**
5. **Principal Component Analysis**
6. **Images and Photographs**
7. **Singular-Value Decomposition**
8. **Deep Learning**
9. **Latent Semantic Analysis**

PARAMTRIC & NON PARAMETRIC MODELS IN MACHINE LEARNING

1. Datasets and Data Files:

- Each machine learning project works on the dataset, and we fit the machine learning model using this dataset.
- Each **dataset resembles a table-like** structure consisting of **rows and columns**. Where each row represents observations, and each column represents features/Variables. **This dataset is handled as a Matrix,** which is a key data structure in Linear Algebra.

2. Images and Photographs:

- In machine learning, images/photographs are used for computer vision applications. **Each Image is an example of the matrix from linear algebra** because an **image is a table structure consisting of height and width for each pixel**.
- Moreover, different operations on images, such as cropping, scaling, resizing, etc., are performed using notations and operations of Linear Algebra.

PARAMTRIC & NON PARAMETRIC MODELS IN MACHINE LEARNING

3. Linear Regression:

Linear regression is a popular technique of machine learning borrowed from statistics. I**t describes the relationship between input and output variables and is used in machine learning** to predict numerical values. The most common way to solve linear regression problems using **Least Square Optimization is solved with the help of Matrix factorization methods.**

4. Regularization:

A technique used to minimize the size of coefficients of a model while it is being fit on data is known as regularization. Common regularization techniques are L1 and L2 regularization. **Both of these forms of regularization are, in fact, a measure of the magnitude or length of the coefficients** as a vector and are methods lifted directly from linear algebra called the vector norm.

PARAMTRIC & NON PARAMETRIC MODELS IN MACHINE LEARNING

5. Principal Component Analysis:

There are several methods in machine learning that automatically reduce the number of columns of a dataset, and these methods are known as Dimensionality reduction. The most commonly used dimensionality reductions method in machine learning is Principal Component Analysis or PCA. This technique makes projections of high-dimensional data for both visualizations and training models. **PCA uses the matrix factorization method from linear algebra.**