MACHINE LEARNING - U1

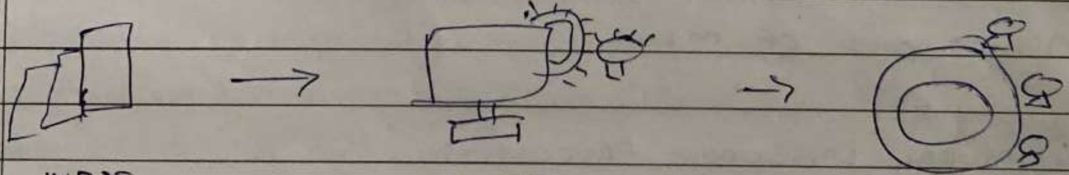
Q. Machine Learning & It's Concepts

ANS. Machine Learning is an branch of Artificial Intelligence that allows computers to learn from data and improve their performance on a task over time, without being explicitly program with step by step rules.

It focuses on teaching computers to learn from data instead of following fixed instructions. Instead of programming everything we feed the computer with large amount of information and let it find its own pattern. Over the time, the computer improves it's ability to make predictions or decisions when it sees new data. This makes machine learning powerfull for solving problems where writing exact rules would be complex.

Machine Learning finds its applications in diverse fields such as image and speech recognition, natural language processing, recommendation system and automating tasks also recognizing faces or predicting customer behaviour

WORKING OF ML



INPUT	ML TECHQ	OUTPUT
STOCK DATA	REGRESSION	STOCK MAR PRED
CS TRAS DATA	CLUSTERING	MARKET SEGMEN
STREAM DATA	ASSOCIATION	RECOMM SYSTEM
EMAIL	CLASSIFICATION	SPAM DETECTION

TYPES OF ML

(i) SUPERVISED

A type of machine learning where the model is trained on the labeled data (both the input and correct output are provided) It learns by mapping between them to make predictions on new data

Ex: Predicting house prices using past data of houses with their actual prices

(ii) UNSUPERVISED

A type of machine learning where the model is trained on unlabeled data with no pre-defined outputs. It finds the hidden patterns, groups or structure in the data

(iii) REINFORCEMENT

A type of machine learning where an agent learn by ~~itself~~ interacting with an environment and recieving awards and penalties. It aims to find best strategy for long term sucess

APPLICATIONS OF ML.

(i) NATURAL LANGUAGE PROCESSING

NLP helps machines to understand and work with human language with the help of ML, systems can recognize speech, understand text and can even translate it between languages.

Ex Includes : Voice assistant -> SIRI, ALEXA

Chatbots for cust queries

Google Translate for language Conversion

(ii) FRAUD DETECTION

Machine learning is broadly used in finance to spot unusual or suspicious activities. By studying millions of past transaction, ML models can flag patterns that look like fraud

Ex: A sudden large withdrawl from a foriegn location on your credit card. This helps banks and companies protect customer and reduce losses

(iii) RECOMMENDATION SYSTEM.

Recommedation System uses ML to suggest items as per user might like, based on their past choices and behaviour.

Ex: Netflix reccomends you movies that you might enjoy.
Amazon suggests you products similar to what you've bought
Youtube suggest videos based on your viewing history

These systems keeps on improving as they keep on learning

Q. Classification VS Regression in Machine Learning

ANS. Classification and Regression are two main tasks in supervis machine learning, which means both use labelled data (inputs with correct answers) to learn. They key difference is the type of results they give

• Classification is used when the output's category or class

Ex: Predicting wheather the Email is spam or not or wheather the fruit is apple OR banana. The model's goal is to draw boundaries that seperate different groups.

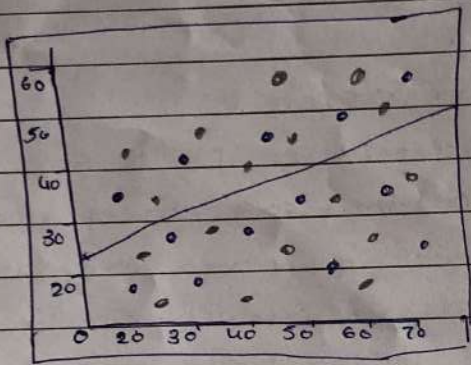
• Regression is used when the output is an continous number

Ex: Predicting the price of an house or the temperature of tomorrow. The model's goal is to find best path / fit that can estimate the value.

- REGRESSION IN ML

Regression is used when we want to predict a number or continous value based on the input data.

Ex: Predicting a person's salary from their years of experience, estimating house prices from size and location or forecasting the chance of rain, all these are used to fit the straight line



Regression

Types of Regression

(i) Simple Regression - Uses a straight line to show the relationsh between one input and one output

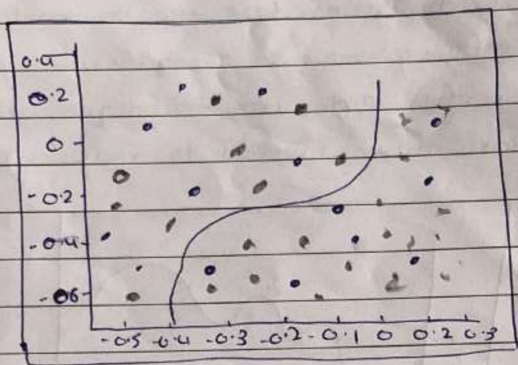
(ii) Multiple Linear Reg - uses multiple input ~~and multiple output~~ to predict single output

(iii) Polynomial Reg - fits a curves when the relationship is not a stra line

- CLASSIFICATION IN ML

Classification is used when we want to ~~use~~ put data into categori or groups instead of predicting numbers

Ex: Classfying wheather the email is spam or not spam or predict if a patient disease based on symtomps. These model ~~works~~ learns from labelled ~~data~~ examples and then assign new data to the correct class



Types of Classification
① Decision Tree - which splits data into categories
② Random Forest - which uses many trees together for higher accuracy.
③ KNN - K-Nearest neighbor, which classify data based on closest similar example.

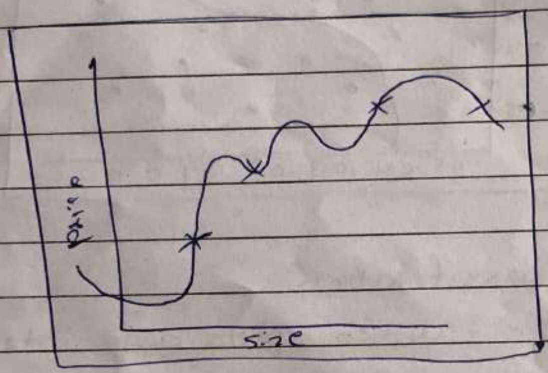
Q. Overfitting Vs UnderFitting

ANS. Overfitting

Overfitting occurs when a machine learning model learns data too closely, including its random, noise, errors or small details, while this makes the model perform extremely well on the training datasets, it struggles to make accurate perfection on new or unseen data.

overfitting happens mostly when the data model becomes too complex, such as having too many decisions rules or parameters, instead of focusing on the general trend or relationship in the

data, it tries to memorize every detail. This makes the model loose its ability to generalize the to real-world cases. It's
To prevent overfitting we can simplify the model, collect more training and usefull data, use techniques like cross-validation or apply regularization method to reduce complexity.

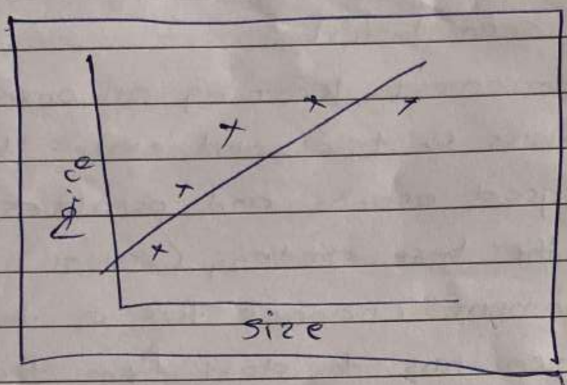


Ex: Imagine an mode training to recognize cats in photos. If the training photos mostly have cats on green grass, the model may wrongly assume that "green background" alwa means "cat". when ~~cat~~ Shown ~~ca~~ a cat indoors on a red carpet, it may fail to recognize it because it memorized details of training set instead of learning general features of cat like (ears, tail)

• UNDERFITTING :

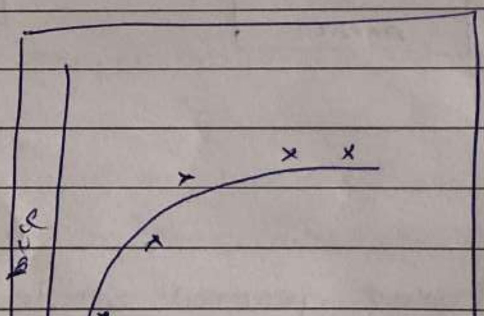
Underfitting occurs when a machine learning algorithm /model is too simple to capture the underlying patterns in the data. IT performs poorly not only on new data but also on the training data itself, showing that it has not learned enough.
This usulay occurs when the model lacks with complexity or when it hasn't been trained enough times. It fails to capture the important relationships in the datasets and ends up making

weak predictions. To fix underfitting, we can make the Model more complex, like add more features or layers, increase training time, or reduce restrictions on the model.



Ex: Suppose we want to predict house price not only use "Size of the house" as input, ignoring other important factors like location, number of bedrooms or condition, The model will be too simple and make poor predictions because it cannot capture all necessary details that affects the price.

To overcome overfitting and underfitting we need to train the model in Goodfitting



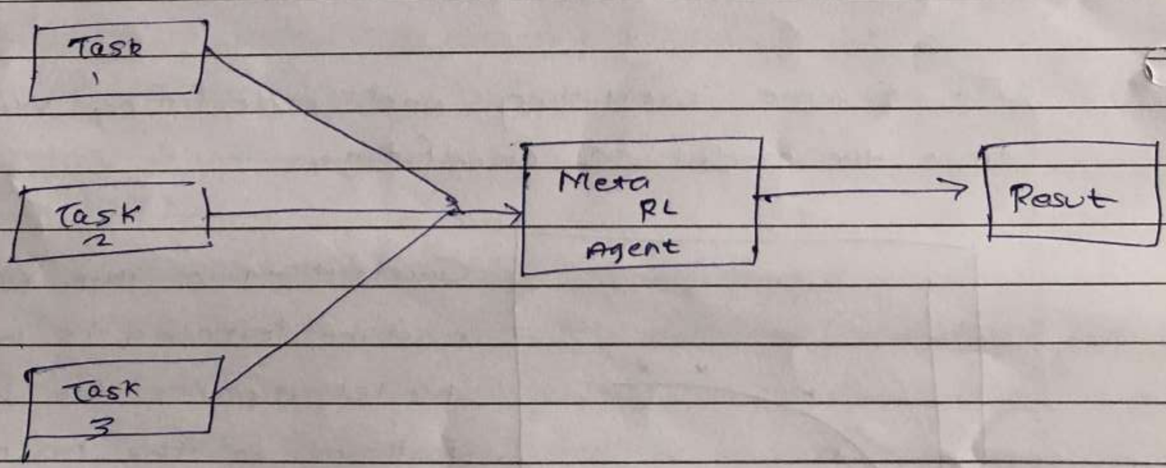
Goodfitting or the right fit in machine learning is when the model captures the important patterns in the training data without memorizing noise or being too simple.

Q. Meta ~~Learning~~ and Re-Inforcement Learning.

ANS. Meta Reinforcement Learning is an advanced form of the reinforcement learning where an agent learns how to quickly adapt new tasks by using experience from similar tasks it has seen before.

In Normal Reinforcement learning an agent learns how to perform one tasks by trial and error - it tries actions, gets rewards for good actions and penalties for bad ones, and over time the best strategy (called Policy). However if the environment changes ~~the~~ or the tasks changes, the agent often has to start from the scratch, which takes a lot of time and data.

Meta-RL Solves this problem by combining reinforcement learning with the idea of meta-learning (learning to learn), Instead of training the data only for 1 tasks, it is trained on many different by related tasks.



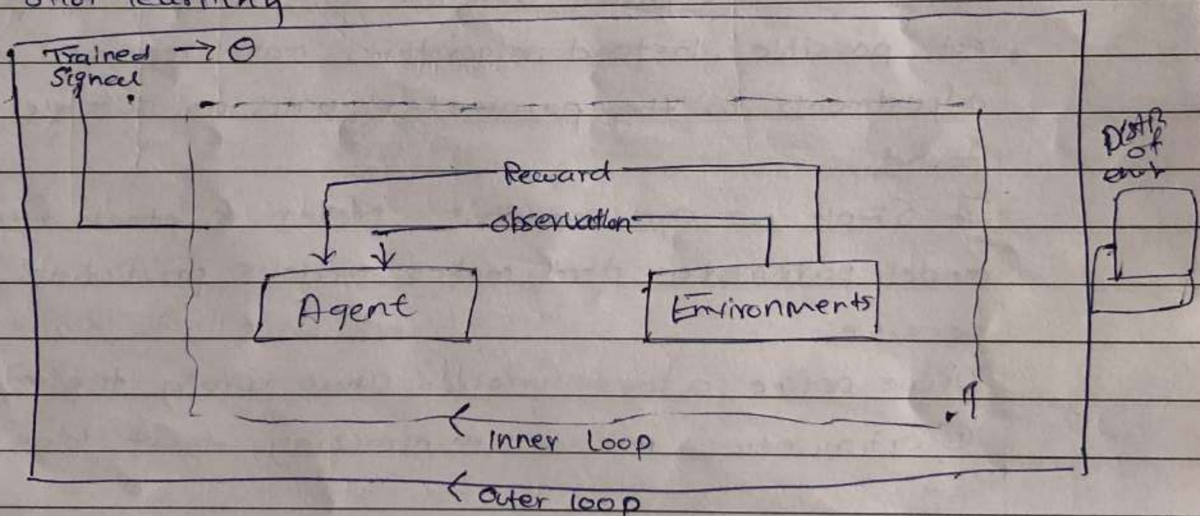
By doing this, the agent learns general strategies and patterns of learning itself, As a result, when the agent faces any brand-new tasks, it can quickly adapt and perfm well without needing a lot of extra training or learning the models.

The Typical Meta RL Involves 2 phases

(i) Meta Training phase: the agent experiences a variety of tasks drawn from a tasks distribution

(ii) Meta Testing Phase: The agent is evaluated on new, unseen tasks and should quickly adapt using minimal data.

This process helps Meta-RL agents learn new tasks faster sometime after just few tries. The agent know a few-shot learning



Ex: Imagine training a robot with regular reinforcement learning to walk on flat ground. It learns very well but struggles when asked to walk on sand or climb stairs because it hasn't seen those situation before. With Meta-RL the robot is trained on many terrains - flat grounds, sand, slopes and stairs. Overtime, it doesn't just learn how to walk in each environments but also learns the process of adapting it's walking style. So, when faced with a brand-new surface like gravel it can quickly adjust without needing to relearn everything from scratch.

Q. Numerical optimization Methods.

Ans. Numerical optimization is an core component of machine learning serving as the process by which the model "learns" from data. It is an collection of mathematical techniques used to find the minimum of function, known as loss function or cost function. By minimizing the loss, a machine learning model iteratively improves its parameters (weights & biases) to make more accurate predictions.

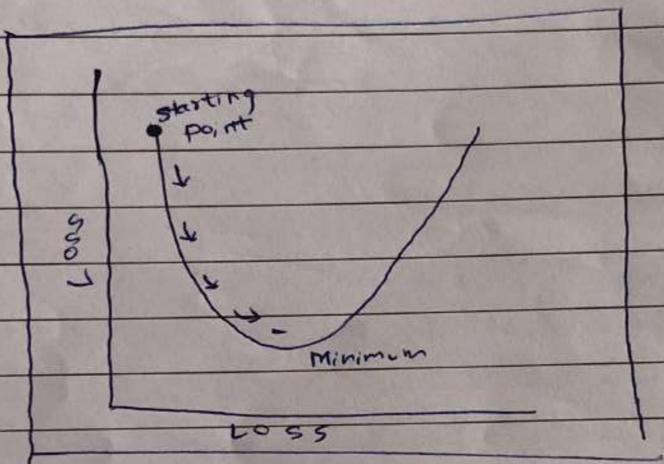
The need for numerical optimization arises because for many complex models, particularly deep learning/neural networks an direct analytical solution for the optimal parameters is not possible. Instead, algorithms make repeated, incremental adjustments to the parameter until an suitable solution is found.

In Simple we can say that NOM is about finding the best model parameter that makes model's prediction as accurate as possible

(i) we Define a loss function (how wrong the model is)

(ii) optimization means minimizing that loss

Gradient Descent is the most common optimization method used in machine learning. It's goal is to minimize the error (loss fun) of model by adjusting it's parameters step by step. You can think of it like finding the lowest point in the valley: at each step you look at the ~~step~~ slope of the ground under your feet (the gradient), and then take a small step downhill. In ML, the slope will tell us how the error changes if we tweak the parameters. By ~~repetadely~~ moving in the opposite direction of the slope, the model gradually learns better parameters and reduces its error. The size of each step is controlled by a number called the learning rate, if it's too big, you might overshoot the minimum, and if it is too small, learning will be very slow.



Ex: Minimizing $f(x) = x^2$

$f(x) = x^2$

Derivative of slope is $f(x) = 2x$

Formula: $x_{new} = x_{old} - \eta \cdot f'(x_{old})$

Steps with learning rate $\eta = 0.1$, Starting at $x = 5$

S1: Gradient $= 2(5) = 10$

update $x = 5 - 0.1(10) = 4$

S2: Gradient $= 2(4) = 8$

update $x = 4 - 0.1(8) = 3.2$

By this way each step moves x closer to 0 which is the minimum of $f(x) = x^2$

S3: Gradient $= 2(3.2) = 6.4$

update $x = 3.2 - 0.1(3.2) = 2.56$