31/1/2023 with Sidhindetri

Interpreter

Python is complied one by one

Polymorphism ( same name, different work)

No need to declare a variable – it depends on its variables.

Paython is independent and extensible and embedded

Colab different editor for python

In python if the multiple number is saved called list

Farqakay lagal array awaya ka list atwant different data type qbol bkat

String and number single

* List
* Tuple
* Dictionary

Amanaya multiple value Agran xoyan

For example

200//3 = 66 ( wata la dway sfrawa lay abat)

# Python Program to find Total, Average, and Percentage of Five Subjects

english = float(input("Please enter English Marks: "))

math = float(input("Please enter Math score: "))

computers = float(input("Please enter Computer Marks: "))

physics = float(input("Please enter Physics Marks: "))

chemistry = float(input("Please enter Chemistry Marks: "))

total = english + math + computers + physics + chemistry

average = total / 5

percentage = (total / 500) \* 100

print("\nTotal Marks = %.2f"  %total)

print("Average Marks = %.2f"  %average)

print("Marks Percentage = %.2f"  %percentage)

odd=[1,2,3]

odd.append(7) ama wary nagret

odd.extend(7)ama wary agret

2/2/2023: (5 shama)

List ( atwanet update bkaet wo 1 data type war agret)

Tuple ( natwanret update bket wo tekalaya) wako record e data basce waya

6/2/2023:

class BankAccount: // bo drostkrdny class

# create the constuctor with parameters: accountNumber, name and balance

def \_\_init\_\_(self,accountNumber, name, balance): //bo drostkrdny constractor bam shewaya

self.accountNumber = accountNumber

self.name = name

self.balance = balance

# create Deposit() method //bo drost krdny method bam shewaya

def Deposit(self , d ):

self.balance = self.balance + d

# create Withdrawal method

def Withdrawal(self , w):

if(self.balance < w):

print("impossible operation! Insufficient balance !")

else:

self.balance = self.balance - w

# create bankFees() method

def bankFees(self):

self.balance = (95/100)\*self.balance

# create display() method

def display(self):

print("Account Number : " , self.accountNumber)

print("Account Name : " , self.name)

print("Account Balance : " , self.balance , " $")

# Testing the code :

newAccount = BankAccount(2178514584, "Shko" , 2700)

# Creating Withdrawal Test

newAccount.Withdrawal(300)

# Create deposit test

newAccount.Deposit(100)

# Display account informations

newAccount.display()

* wata method wako function drost darket wo hich farqy neya
* parameterkany konstractor atwaneret bakarbhenret la har shwenik ka btawet

NumPy is a Python library.

NumPy is used for working with arrays.

NumPy is short for "Numerical Python".

AI (7-2-2023)

Renforment learning

Supervised learning

Unsupervised learning

Machine learning -> intelligence from Data

AI is the branch of Computer Science

Chatbot ->Turing test (Alan Turing 1950) it is a test of a machine’s ability to exhibit intelligent behavior equivalent …..

ChatGPT -> Open AI (is the domain of NLP)

Open AI

Open Coder

Github

Co-palet

AWS

Github

Jenkins

Docker

Develops for AI/ML

Image development

Bigdata /AI (to manage the data) how we manage a massive data

Better algorithms

Cognitive = connected with the process of understanding

Types of AI = based on capabilities and based on functionalities

Gaming

Finance

Yahoo finance

Quandl

Real time data of the stock market

Minimum risk

Agent

Capture is used Anti turing test

structure vs unstructured data

supervised learing – score prediction

unsupervised learning – clustering

reinforcement learning – Robot Navigations

supervised learning – Regression analysis

prediction analysis

unsupervised learning is no label

PEAS types of environment

Types of agents

Agent

PEAS = performance measure, environment, actuators, sensors

Environment types: fully observable and partially observable

Goal Formulations

Cannibal

Missionary

8-2-2022

Missionaries and cannibals problem

It can be shown by tree and diagram

We are talking about Stack and Queue

In queue and de queue

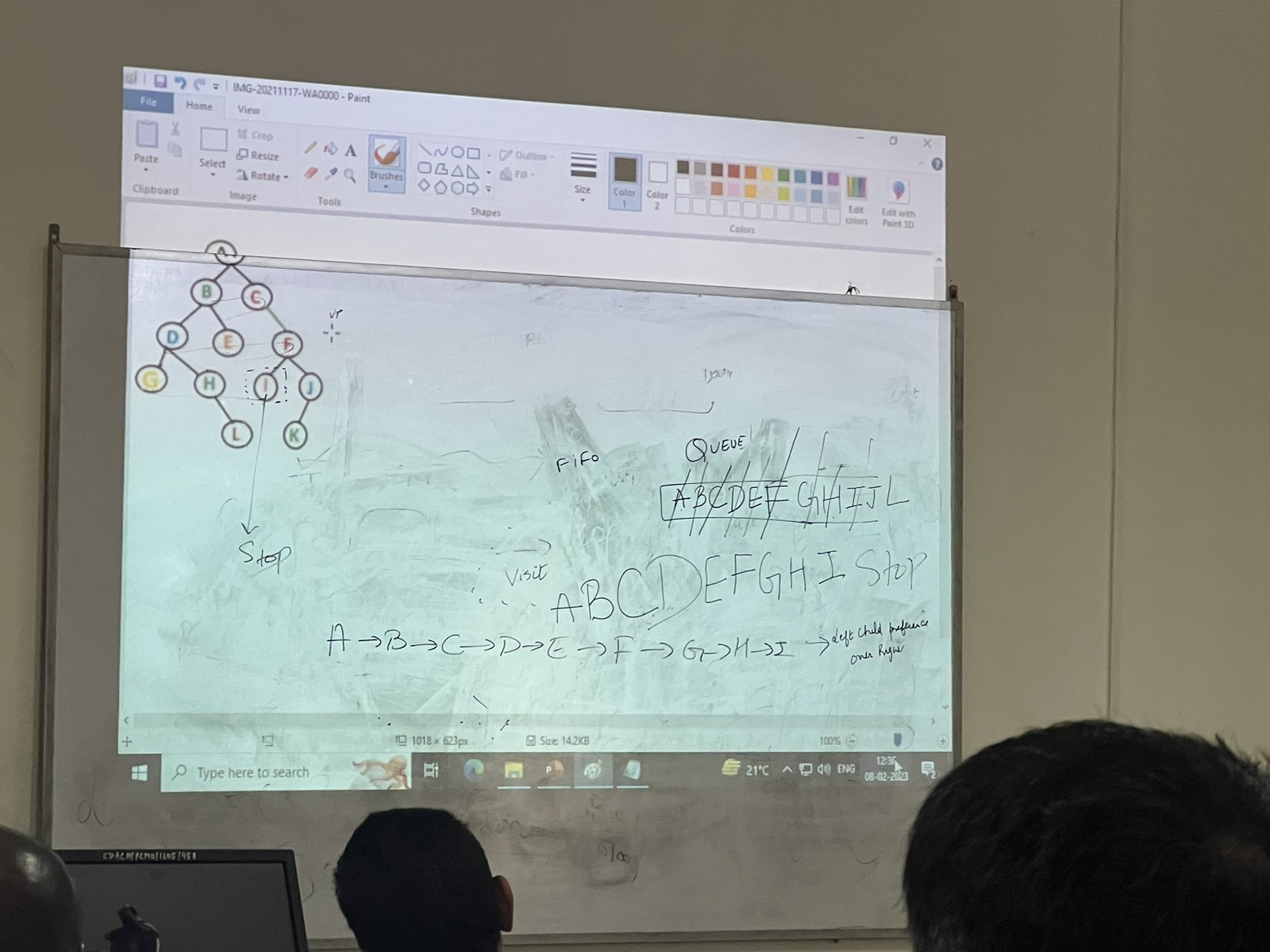
Queue is a linear data structure

Queue = FIFO

Stack =FILO ama har ba nawy xoyawa alet gir axwat

Breadth first search use FIFO

Shallow = not deep; with not much distance between top and bottom



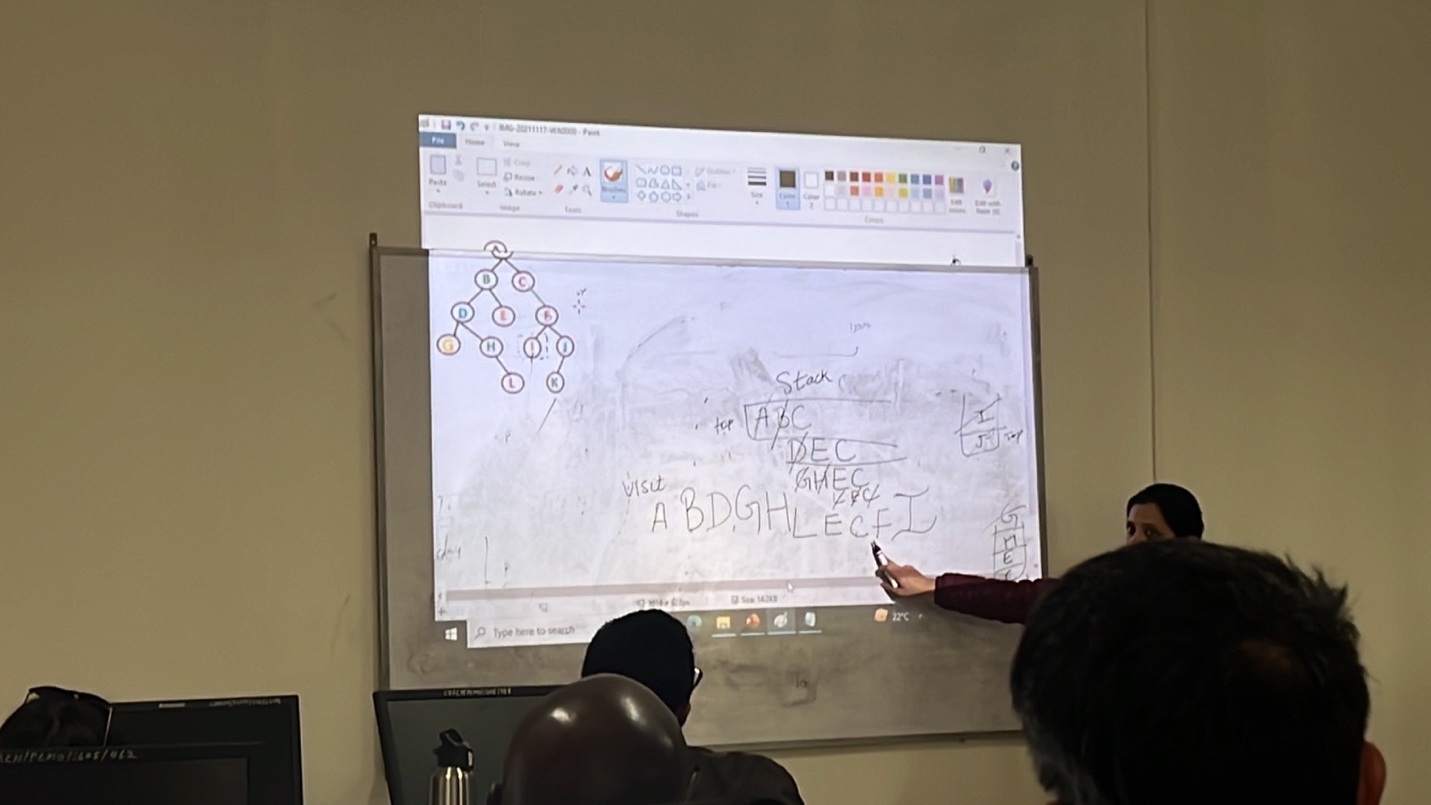
Ama zor asana tanha level by level node kan azhemner la naw stack ka da bon mona alen A child kanny chya ae nosen hata again bawo shwena ka amanawet , thanks you so much my bro

I.S = initial stat

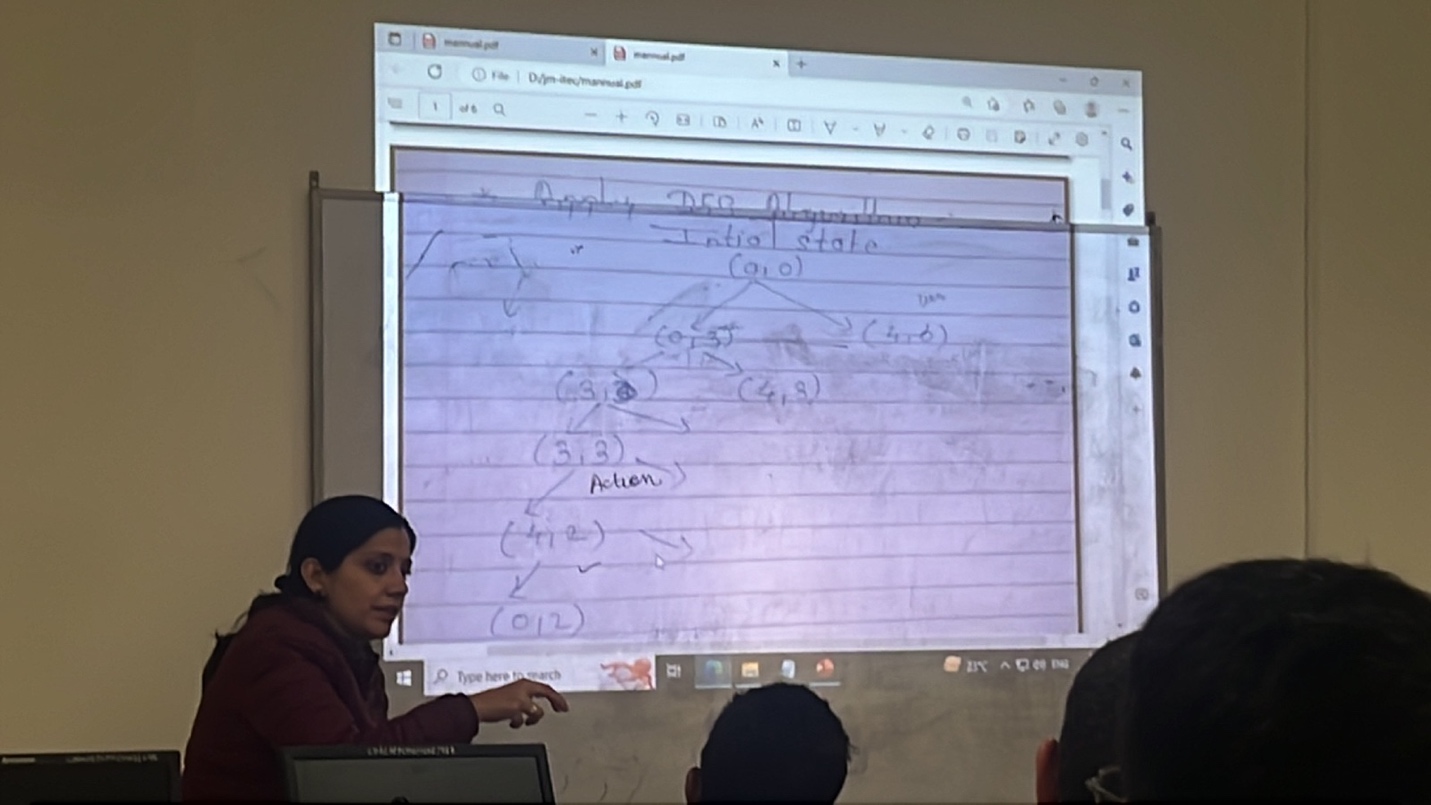
BFS

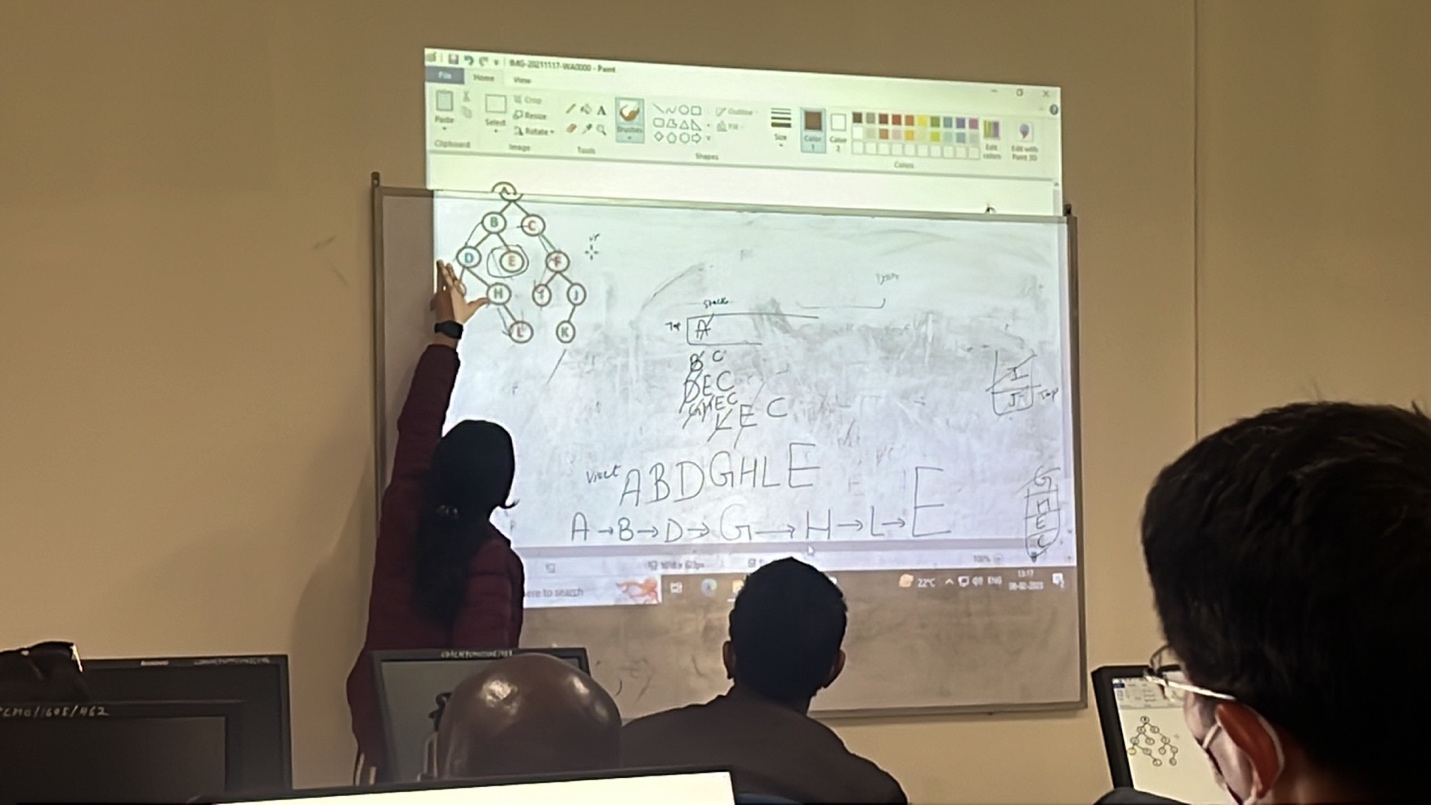


DFS =depth first search

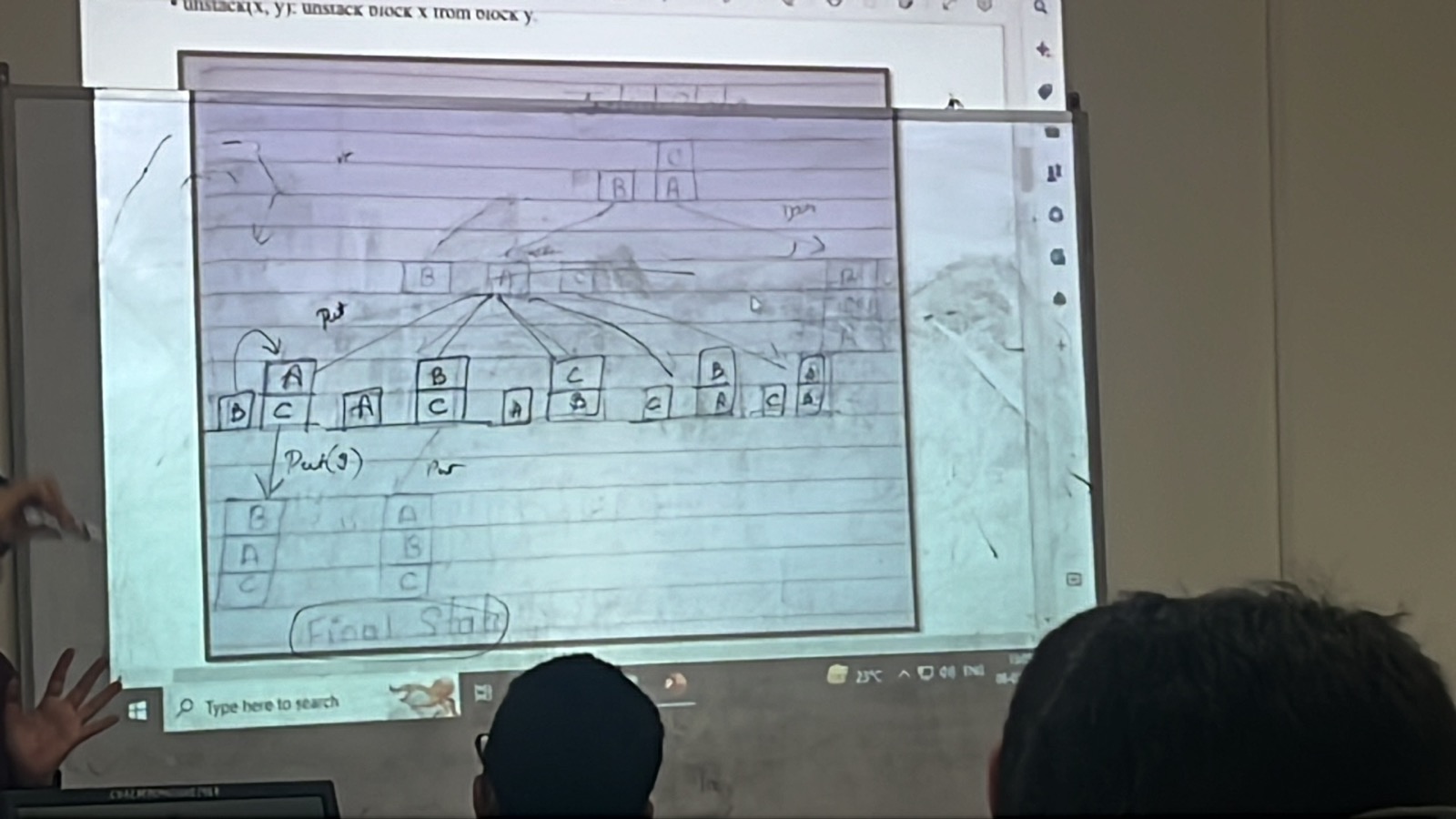


Action and achieve





BFS



Least cost unexpanded node

Uniform cost search

Use priority tools / graph expand least cost

A\* Search:

Heuristic: h(n)

Based on estimating technic

F(n) = g(n)+h(n)

Heuristic value

Inform search strategies

Categorise of search algorithms

Least evaluations

9-2-2023

Evaluation equation

F(n) = g(n) + h(n)

Tic -tac-toe

G =1 (wata level kay diary dakret bo nmona esta la level y 1 daya)

H = hamo jarik current status barawd akaenwa ba Golaka bzanin chank zhmarayan jyawaza awa county akaen (tanha jyawazakan) pey dalen misplaced

F kash yaksana ba g+h wata hardokyan ko dakaenawa wo tawao

Note: alet agar la init state kay awaka away sarata ka barawdy akaen bag golaka agar xanay batalman hamo awa hsay bo nakaen bo H ka agadar ba

Local search algorithms are usueful for solving pure optimization problem.

Adversarial Search

We have Max,Min and alfa B searning

In tic tac toe the outcome is win loss draw with the value of +1,1,0

Final optimal path that is why we use min-max search

Utility number = +10 =win

DFS

Pruning = cut down a certain nodes

It is alfa and beta pruning

Alfa the maximizer

Beta the minimizer

How to backup the values

10-2-2022

Formal and informal search

Seach for all AI categories

Problem solving

PEAS

KR = knowledge representation

Knowledge level

Logical level

Implementation

How many types

**Knowledge Representation and Reasoning** (KR, KRR) represent data from the real world. A computer can comprehend and then use this knowledge to solve complex real-world problems, such as communicating with humans in natural language.

PL , FOPL

Proposition

Use logical connector ( AND, OR , …)

P= today is Sunday

-| P = today is not Sunday

P -🡪 Q ( if it is raining, street is wet)

Modus Ponens

Modus Tollens

Hypothetical Syllogism

Download docker

Install

And then use GIT SCM

Create an account

JDK 17

Jenkins

KR input the information interims of prepositional of logical prepositions

And eliminations

<https://www.javatpoint.com/ai-knowledge-base-for-wumpus-world>

amash basy agarakan akat ka to bac hand agara atwanet bgaeta hadafaka

First-Order Logic in AI

First order predicated logic

Types of agent = knowledge based agent

Is the way of knowledge representation

FOL

Predications of

Human are intelligent -🡪 intelligent(human)

Agar woty handik (some) yan hamoyan (ALL) awa har yakayan hemayaky bo bakar ahenin

Universal quanifier

∃ = some

∀ = all, every=anyone=anything = quantifier ( ama ya3ni jamakaya hamo awakay tya)

Switiching the order of universals and existential does change meaning:

Everyone likes someone

Someone is liked by everyone

Universal instantiation

Universal generalizations

Existential instantaions

Existential generalization

Unification

CNF = Conjunctive normal form (CNF) is **an approach to Boolean logic that expresses formulas as conjunctions of clauses with an AND or OR**. Each clause connected by a conjunction, or AND, must be either a literal or contain a disjunction, or OR operator. CNF is useful for automated theorem proving.

Machine Learning 11/2/2023 Saturday Morning

Started by an example:

Interdental statical

Probability theory

Example: Bag -🡪 (3 red , 2 blue)

Q/ find all the combinations

Q / find the probabilities of each combination

Q/ Use the probabilities to estimate profit / does ..each

Possible outcomes

1/

4 blues 3B

0 red 1R

Probabilities by distributions

Binominal distributions

Ncr = it is related to factorial

Cumulative distribution function

<https://www.gstatic.com/education/formulas2/472522532/en/cumulative_distribution_function.svg>

|  |  |  |
| --- | --- | --- |
|  | = | function of X |
|  | = | real value variable |
|  | = | probability that X will have a value less than or equal to x |

SVM = support victor machine

M = mean

M=35 segma= 5

P(25<x<45)

Standard normal probabilities

zs core

**Z-Score Formula**

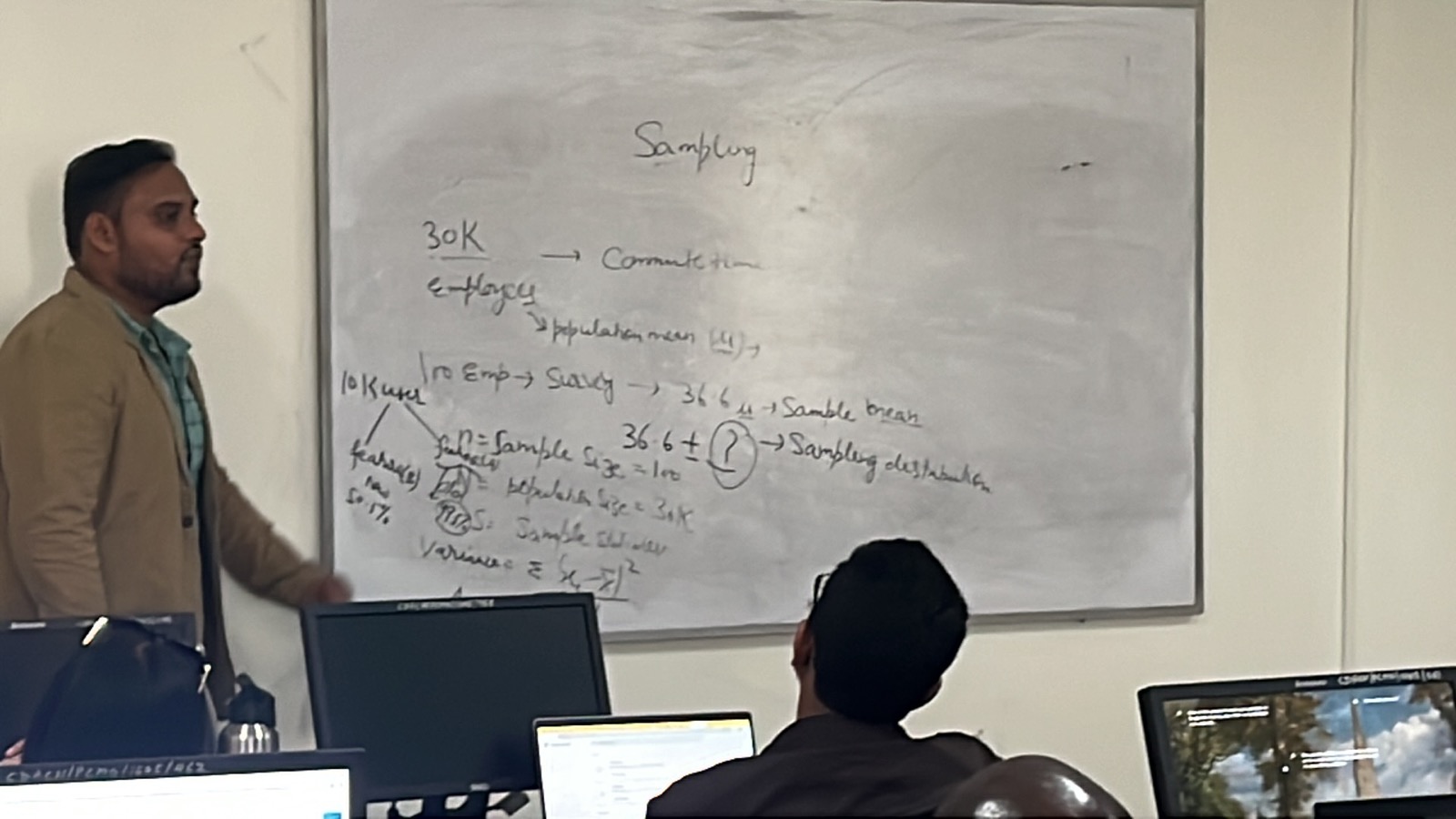
The statistical formula for a value's z-score is calculated using the following formula:

z = ( x - μ ) / σ

Where:

* z = Z-score
* x = the value being evaluated
* μ = the mean
* σ = the standard deviation

Sampling



Central limit theorem

1-Sample size n>30 normal distribution sampling distrubutuin become normal distribution

2- Sampling distribution mean (M) == population mean (M)

3- sf = segma/ root of n

Standard deviation

Confidence interval

Margin error

GPU (from google account) wako program esh akat bo eshakany python

13/2/2023

**Hypothesis Testing**

Full null hypothesis

Ho --🡪 hypothesis

Example: a person who is charged with a number case

Null hypothesis is

Status quo is true

No charge or difference in

* He is innocent
* He is guilty

Rules:

Ho => = or <……

……..

Hypothesis testing ( Samples and p-value)

LCV = lower critical value

UCV = upper critical value

<https://www.javatpoint.com/hypothesis-in-machine-learning>

2 types of hypothesis we have

Directional (one tailed)

Non directional (two tailed)

Directional tests are known as "one-tailed" tests because all of the error is one "tail" of the distribution (less than). Non-directional tests are called "two-tailed" tests because we must include the possibility that the alternative population could be less than m or greater than m.

You have to have information about extract information

**P-value method**

P value is going to helps us to measure the strengths of our evidence in support of NULL hypothesis ( p value is probabilities of null hypothesis )

P= 0.05

Y=M1X1+M2X2+……+MnXn+c

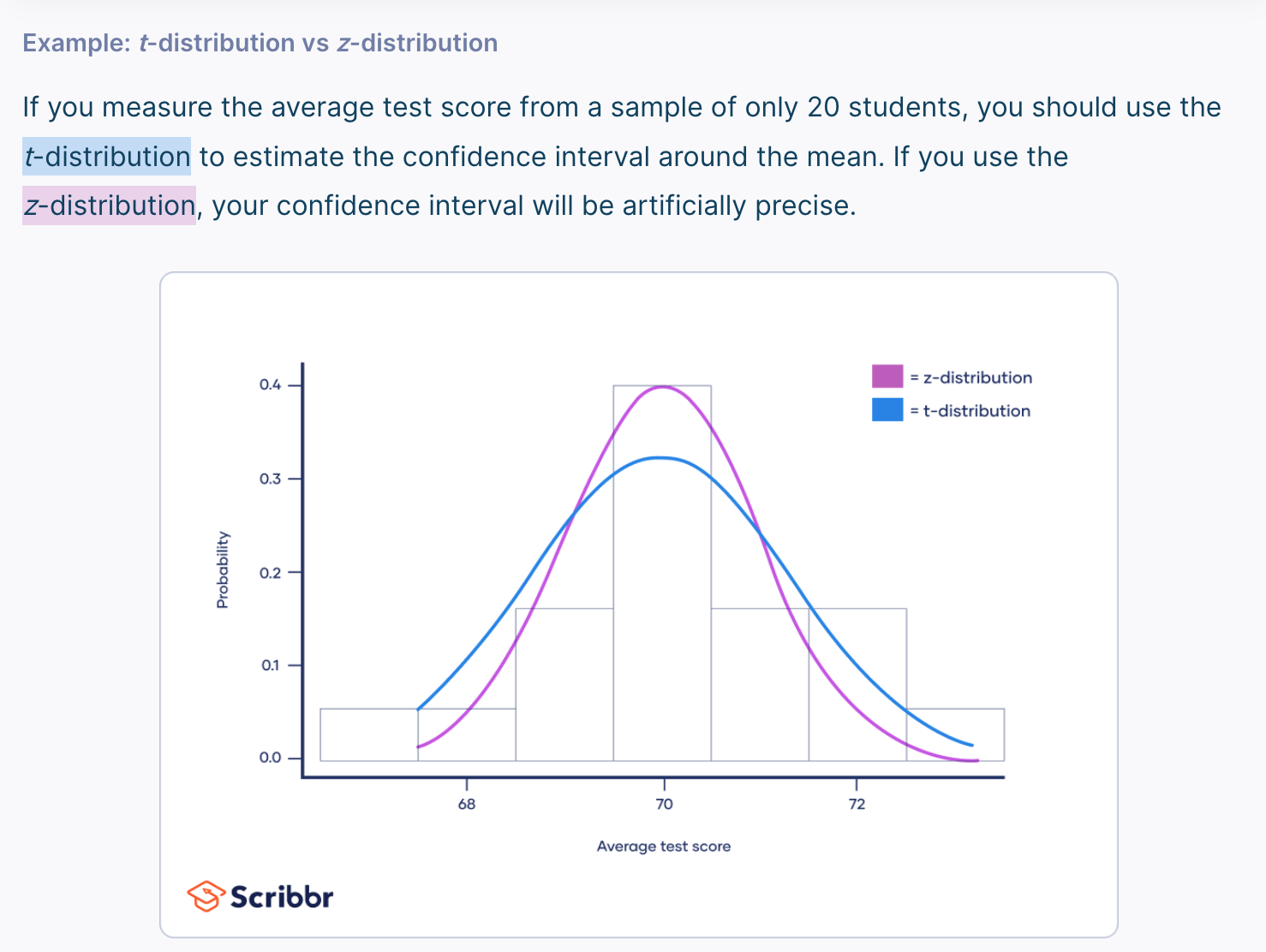
X1,….Xn features

M1,…….M2

The P-value method is **used in Hypothesis Testing to check the significance of the given Null Hypothesis**. Then, deciding to reject or support it is based upon the specified significance level or threshold. A P-value is calculated in this method which is a test statistic.

**T- distribution**

What is a t-distribution? The t-distribution is **a way of describing a set of observations where most observations fall close to the mean, and the rest of the observations make up the tails on either side**. It is a type of normal distribution used for smaller sample sizes, where the variance in the data is unknown

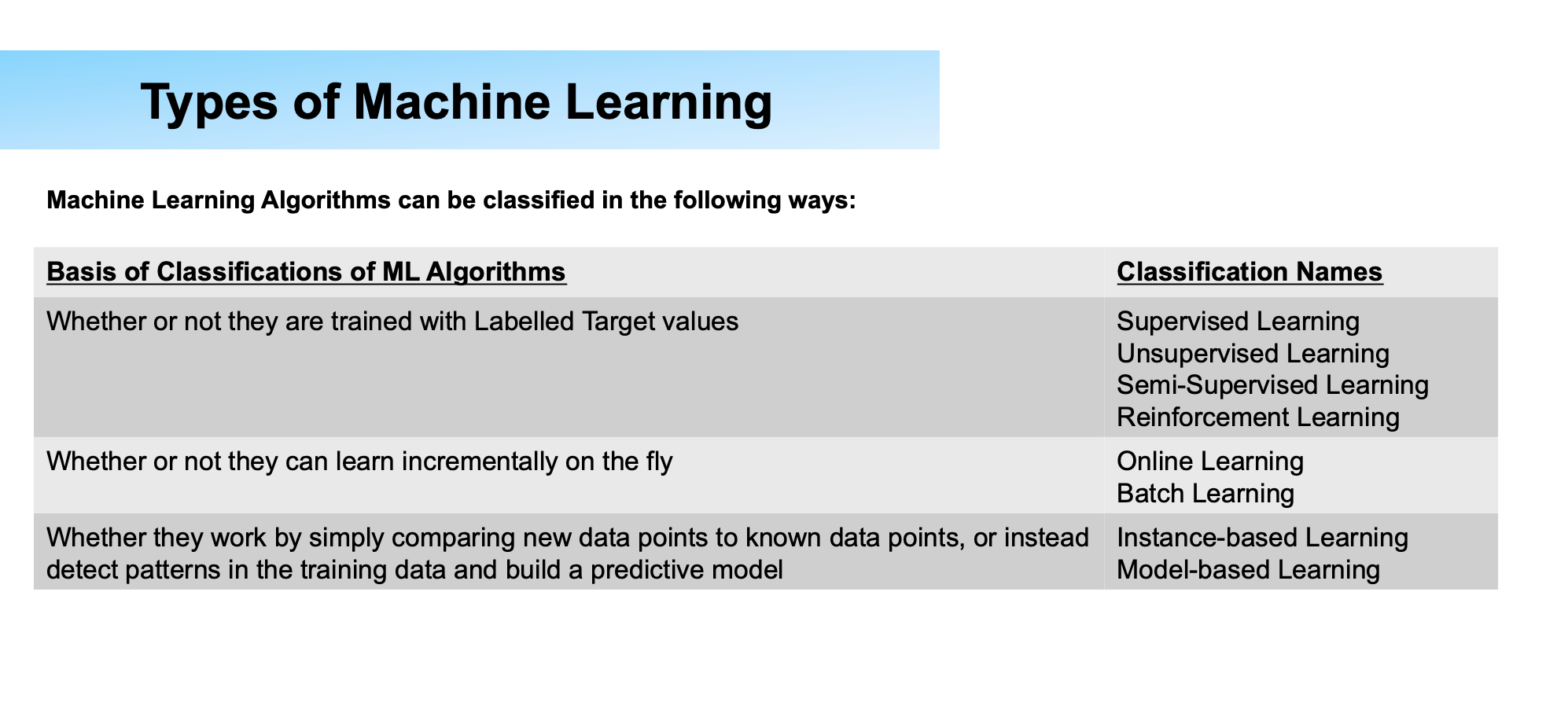


Learning ( supervised and unsupervised learning)

Data 🡪 ML 🡪 Model

GPU , CPU , TPU ( CPU is more stronger than others)

Types of machine learning:



Hugging face, <https://huggingface.co/>

Infosys :

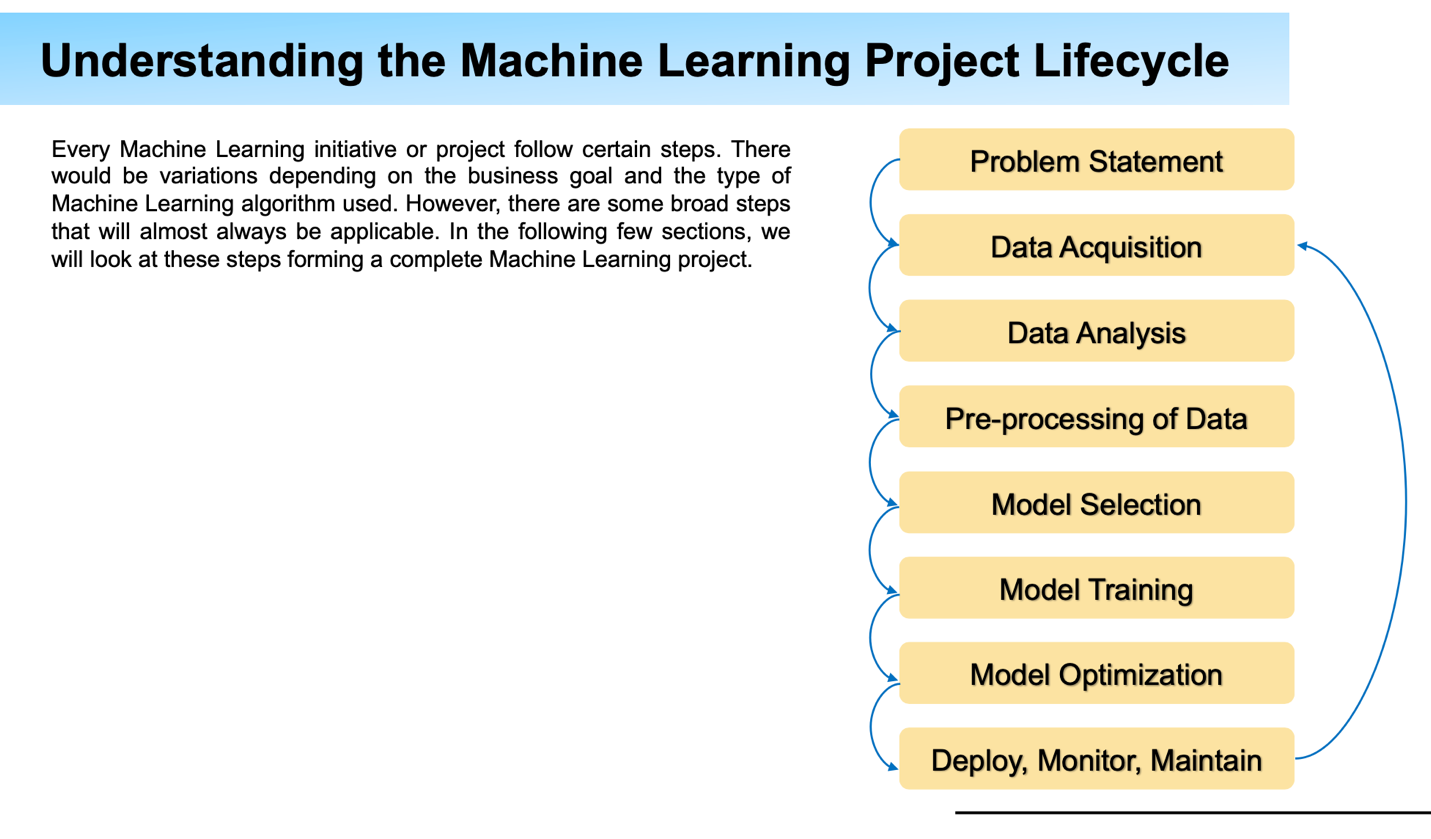
<https://www.google.com/search?q=infosys+share+price&client=firefox-b-d&ei=5PrpY7BSpL3j4Q_SrLX4Dw&oq=infosys+share&gs_lcp=Cgxnd3Mtd2l6LXNlcnAQAxgAMhAIABCABBCxAxCDARBGEPoBMgsIABCABBCxAxCDATIICAAQgAQQsQMyCwgAEIAEELEDEIMBMgUIABCABDILCAAQgAQQsQMQgwEyBQgAEIAEMgUIABCABDIFCAAQgAQyBQgAEIAEOgoIABBHENYEELADOgcIABCwAxBDOg0IABDkAhDWBBCwAxgBOhUILhDHARDRAxDUAhDIAxCwAxBDGAI6EgguEMcBENEDEMgDELADEEMYAkoECEEYAEoECEYYAVD_Cli0EWCDHGgCcAF4AIABf4gB8QSSAQMwLjWYAQCgAQHIARPAAQHaAQYIARABGAnaAQYIAhABGAg&sclient=gws-wiz-serp>

<https://www.kaggle.com/>

<https://scikit-learn.org/stable/datasets.html>

<https://archive.ics.uci.edu/ml/datasets.php>

the complete machine learning life cycle



<https://spacy.io/>

Mean Squared Error MSE

<https://seaborn.pydata.org/>

<https://matplotlib.org/>

<https://www.kaggle.com/>

14/2/2023 (Tuesday)

<https://regex101.com/>

car price prediction (linear regression -rfe) ama la link e [www.Kaggle.com](http://www.Kaggle.com) war girawa

<https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html>

<https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>

<https://www.geeksforgeeks.org/xgboost/>

Confusion Matrix:

TP = true positive

FN = false negative

TN = true negative

FT = false positive

<https://scikit-learn.org/stable/modules/model_evaluation.html#classification-report>

What is Regularization

ElasticNet

<https://machinelearningmastery.com/lasso-regression-with-python/>

<https://scikit-learn.org/stable/auto_examples/linear_model/plot_lasso_and_elasticnet.html#lasso>

<https://pandas.pydata.org/docs/reference/api/pandas.read_csv.html>

<https://machinelearningmastery.com/>

15/2/2023

**Naïve Bayes**

<https://www.analyticsvidhya.com/blog/2021/11/implementation-of-gaussian-naive-bayes-in-python-sklearn/>

<https://www.ibm.com/in-en/topics/knn>

The k-nearest neighbors algorithm, also known as KNN or k-NN, is a non-parametric, supervised learning classifier, which uses proximity to make classifications or predictions about the grouping of an individual data point. While it can be used for either regression or classification problems, it is typically used as a classification algorithm, working off the assumption that similar points can be found near one another.

For classification problems, a class label is assigned on the basis of a majority vote—i.e. the label that is most frequently represented around a given data point is used. While this is technically considered “plurality voting”, the term, “majority vote” is more commonly used in literature. The distinction between these terminologies is that “majority voting” technically requires a majority of greater than 50%, which primarily works when there are only two categories. When you have multiple classes—e.g. four categories, you don’t necessarily need 50% of the vote to make a conclusion about a class; you could assign a class label with a vote of greater than 25%. The University of Wisconsin-Madison summarizes this well with an example [here](https://sebastianraschka.com/pdf/lecture-notes/stat479fs18/02_knn_notes.pdf) (PDF, 1.2 MB) (link resides outside of ibm.com).

K= ragy 2 jay n

# ASR

**What does ASR stand for?**

ASR stands for automatic speech recognition. It refers to the ability of machines to understand and process natural speech. ASR systems are used in voice assistants, chatbots, machine translation, and more.

# CNN

**What does CNN stand for?**

In the field of artificial intelligence, CNN stands for convolutional neural network. It is a type of deep neural network often used for computer vision tasks.

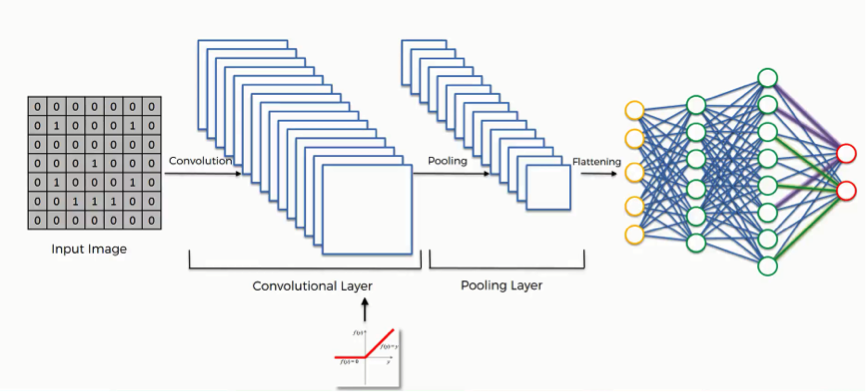


Image via superdatascience.com

Due to the effectiveness of CNNs, they have numerous applications in facial recognition, medical image analysis, and natural language processing tasks. [This beginner’s guide](https://skymind.ai/wiki/convolutional-network) goes as far to say that “CNNs are the reason why deep learning is famous.” A variation, DCNN, stands for deep convolutional neural network, which is a CNN with more layers. However, there is no standard number of layers that distinguishes CNNs from DCNNs. Therefore the two are sometimes used interchangeably.

# CSV

**What does CSV stand for?**

CSV stands for comma-separated values. It is a file format often used for AI training data. As the name suggests, CSV files use commas to separate values in the data.

# DL

**What does DL stand for?**

DL stands for deep learning and refers to machine learning tasks that use neural networks containing multiple layers. At the same time, increasing the number of layers requires more computer processing power and usually a longer training time for the model.

# FKP

**What does FKP stand for?**

In computer vision, FKP**\*** stands for facial keypoint(s). FKPs are commonly plotted around the nose, eyes, and mouth, to create a facial signature unique for each individual.



Moreover, [facial keypoint annotation](http://lionbridge.ai/services/image-annotation/#landmarkannotation), also known as landmark annotation, is used to create training data for facial recognition models.

\***alternate definition**: in biometrics, FKP stands for finger-knuckle-print. There are some machine learning approaches used to create finger-knuckle-print identification systems.

# GAN

**What does GAN stand for?**

GAN stands for generative adversarial network. In 2014, the idea was proposed in the paper [Generative Adversarial Nets](https://arxiv.org/pdf/1406.2661.pdf) and has made amazing breakthroughs since. A GAN is a neural network that can be used to generate new and unique content. Recently, it’s been used to [generate incredibly realistic human face images](https://petapixel.com/2019/02/19/this-website-generates-ai-portraits-of-people-who-dont-exist/).

# LSTM

**What does LSTM stand for?**

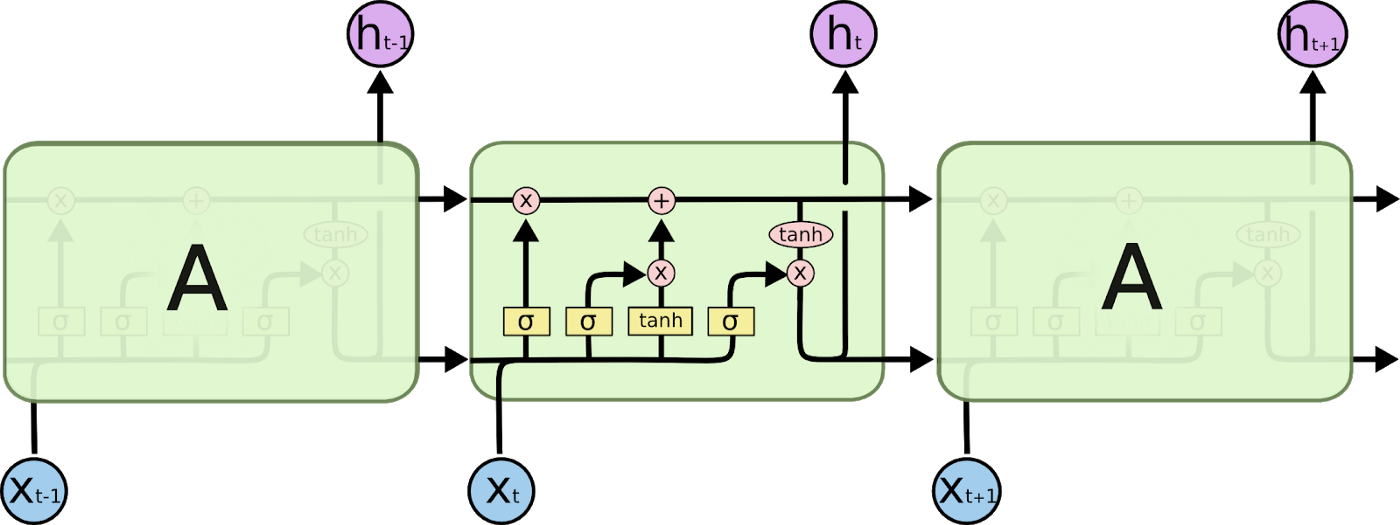


Image via colah.github.io

LSTM stands for long short-term memory and is a variant of recurrent neural networks (see definition near the bottom of this list). Basically, the strength of LSTMs is their ability to remember information for a long period of time and apply it to the present task. A good explanation of LSTMs can be found in this [beginner’s guide](https://colah.github.io/posts/2015-08-Understanding-LSTMs/).

# ML

**What does ML stand for?**

In regards to artificial intelligence (AI), ML stands for machine learning. While AI and ML are often used interchangeably, there are some differences between the two phrases. This article provides a great explanation of the [differences between AI, ML, and DL](https://lionbridge.ai/articles/whats-the-difference-between-ai-machine-learning-and-deep-learning/).

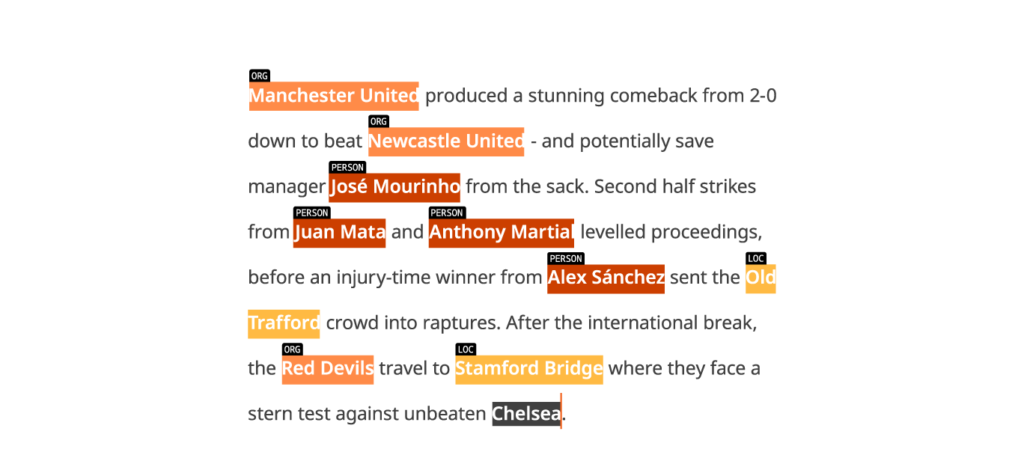
# MNIST

**What does MNIST stand for?**

MNIST stands for the Modified National Institute of Standards and Technology (database). The MNIST database is one of the most famous benchmark datasets in machine learning. Due to its usability and relatively small size, it is often used to compare the performance of various algorithms. Furthermore, it has inspired the creation of many [more datasets in the same format.](https://lionbridge.ai/datasets/mnist-datasets-for-machine-learning/)

# NER

**What does NER stand for?**



NER stands for [named entity recognition](https://lionbridge.ai/services/entity-annotation/), which is an important process in numerous natural language processing (NLP) models. Named entities refer to proper names within text, usually people, places, or organizations. By identifying and tagging these entities, you can use the data to teach the algorithm information about those entities and how to spot new named entities in the future.

# NLP

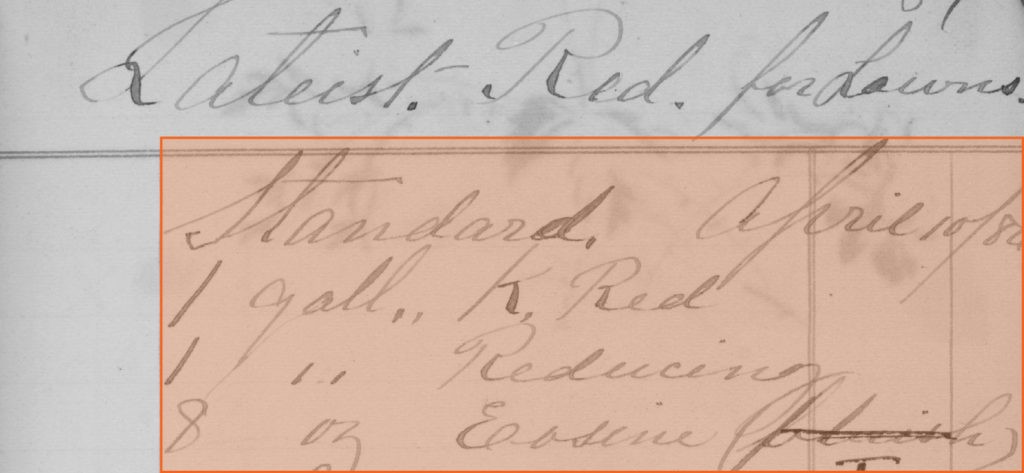
**What does NLP stand for?**

As mentioned above, NLP stands for natural language processing and is a large field of study within machine learning. It is the study of natural human language and how to create machines that fully understand that language. Some of the current applications of NLP in the real world include chatbots, [voice assistants](https://en.wikipedia.org/wiki/Virtual_assistant), and [machine translation](https://www.microsoft.com/en-us/translator/business/machine-translation/).

# OCR

**What does OCR stand for?**

OCR stands for optical character recognition, which is the process of identifying written or printed characters.



After being identified, those characters are then converted into digital data. OCR is often used for the preservation of written text documents by identifying the characters and converting the text into PDFs.

# PFE

**What does PFE stand for?**PFE stands for probabilistic facial embeddings, which is a method for facial recognition tasks in unconstrained settings.

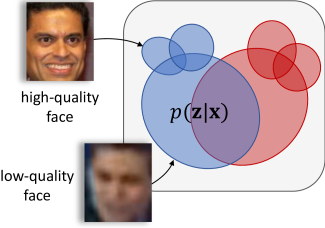


Image via groundai.com

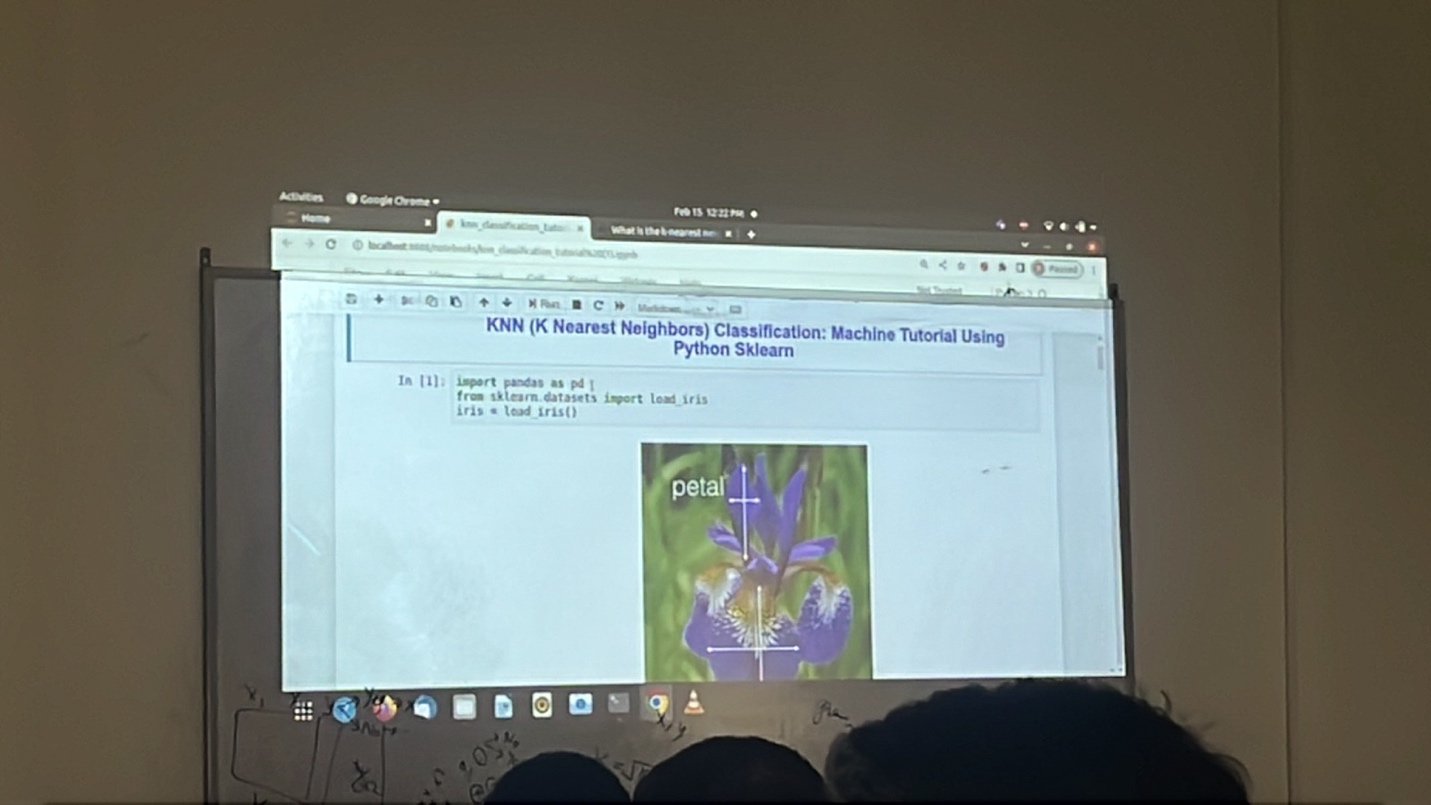
Unlike deterministic facial embeddings, PFEs convert each face image into a distribution, in order to account for uncertainty. For a deeper understanding of PFEs and facial recognition, please see these [5 essential papers on facial recognition](https://lionbridge.ai/articles/5-machine-learning-papers-on-face-recognition/).

# RNN

**What does RNN stand for?**

RNN stands for recurrent neural network and is a type of neural network which has loops. Its structure is designed to allow previously processed information to affect how the system interprets new information. A more detailed introduction to RNNs can be found [here.](https://medium.com/explore-artificial-intelligence/an-introduction-to-recurrent-neural-networks-72c97bf0912)

Hopefully, this AI abbreviations glossary helped strengthen your knowledge of common acronyms for terms used in various machine learning fields. More terms and abbreviations will constantly be added to this list, so please [follow me for further updates and other AI content.](https://hackernoon.com/@limarc2000#subscribe-embed)



<https://scikit-learn.org/stable/datasets/toy_dataset.html>

Support Vector Machine

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.

The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.

SVM chooses the extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is termed as Support Vector Machine. Consider the below diagram in which there are two different categories that are classified using a decision boundary or hyperplane:

<https://www.kaggle.com/c/mnist-handwritten-digit-recognition>

16-2-2023

OpenCV

OpenCV (Open Source Computer Vision Library) is **an open source computer vision and machine learning software library**. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products.

OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today's systems. By using it, **one can process images and videos to identify objects, faces, or even handwriting of a human**.26-Oct-2022

**Applications of OpenCV:** There are lots of applications which are solved using OpenCV, some of them are listed below 

* face recognition
* Automated inspection and surveillance
* number of people – count (foot traffic in a mall, etc)
* Vehicle counting on highways along with their speeds
* Interactive art installations
* Anomaly (defect) detection in the manufacturing process (the odd defective products)
* Street view image stitching
* Video/image search and retrieval
* Robot and driver-less car navigation and control
* object recognition
* Medical image analysis
* Movies – 3D structure from motion
* TV Channels advertisement recognition

<https://archive.ics.uci.edu/ml/index.php>

Decision Tree Classifier

Entropy

<https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html>

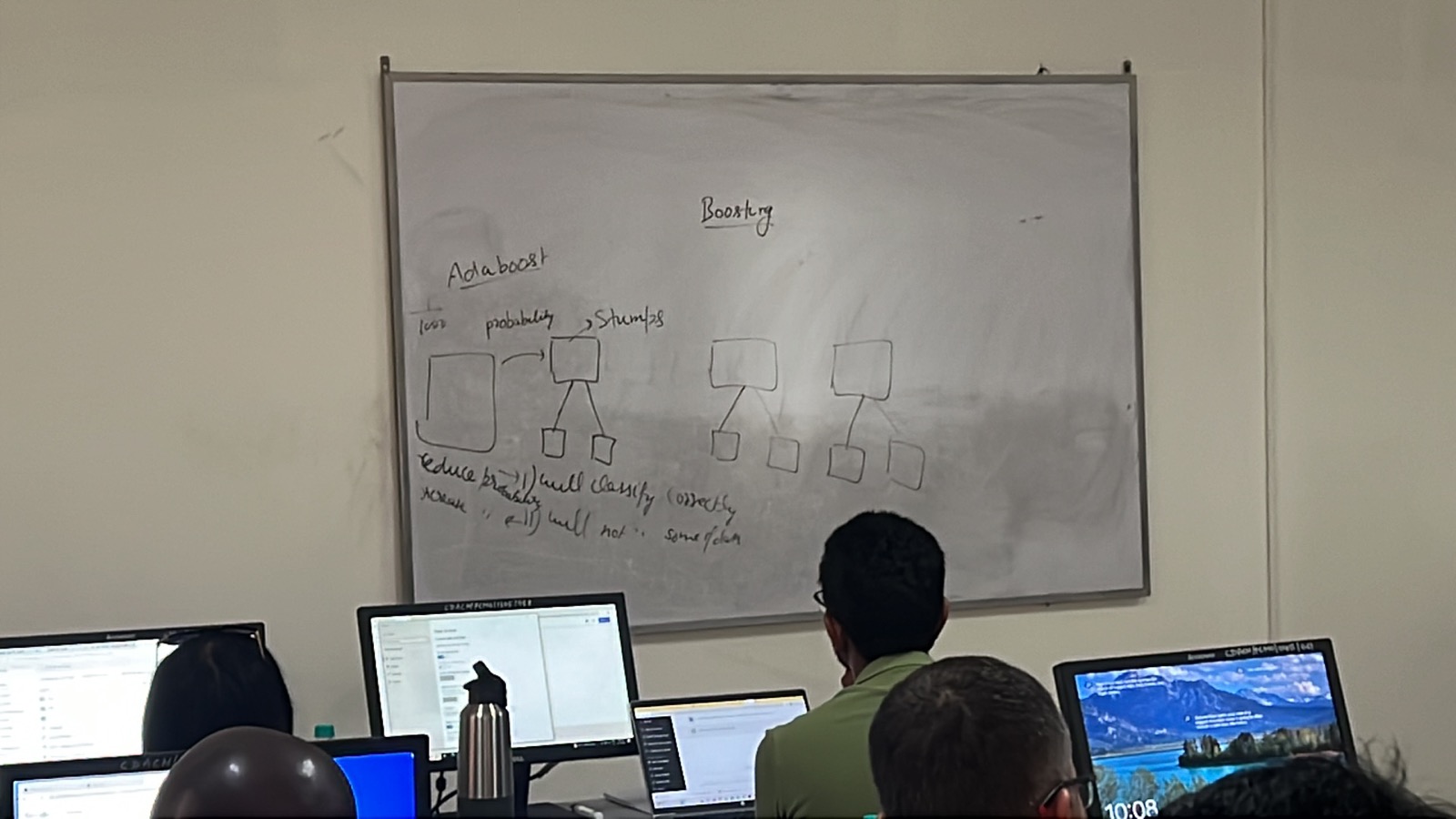
<https://scikit-learn.org/stable/modules/generated/sklearn.tree.plot_tree.html>

Bootstrap aggregating, also called bagging, is a machine learning ensemble meta-algorithm designed to improve the stability and accuracy of machine learning algorithms used in statistical classification and regression. It also reduces variance and helps to avoid overfitting.

<https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>

oob\_score = out of bag

17-2-2023 (Friday)



Boosting is a powerful technique that has been widely used in various applications, such as classification, regression, and ranking. Boosting has been shown to improve the accuracy of the prediction and to be robust to noisy data. However, boosting can be computationally expensive, and it can be sensitive to overfitting if the base models are too complex.

**AdaBoost**

AdaBoost, short for "Adaptive Boosting," is a popular ensemble learning algorithm used in machine learning for classification and regression tasks.

In AdaBoost, the algorithm creates a "strong" classifier by combining multiple "weak" classifiers. A weak classifier is a classifier that performs only slightly better than random guessing. During the training process, the algorithm assigns weights to each instance in the dataset, with the goal of emphasizing the instances that are misclassified by the weak classifier.

In each iteration of the algorithm, a new weak classifier is trained on the modified dataset. The weights assigned to each instance are updated based on the performance of the weak classifier, with misclassified instances receiving higher weights. The new weak classifier is then combined with the existing weak classifiers to create a stronger, more accurate classifier.

The final output of the AdaBoost algorithm is a weighted combination of all the weak classifiers. The weights assigned to each weak classifier are determined based on its accuracy during training, with more accurate classifiers being given higher weights.

AdaBoost is often used in classification tasks and has been shown to perform well on a variety of datasets. It has also been extended to regression tasks, where it is known as AdaBoost.R2.

Document clustering is a machine learning technique used to group similar documents together based on their content or other features. The goal of document clustering is to discover underlying patterns and structure in large collections of documents, which can be useful for tasks such as document categorization, topic modeling, and information retrieval.

In document clustering, the algorithm typically works by first representing each document as a vector of features. These features may include word frequencies, TF-IDF scores, or other types of text-based features. Other features could include metadata such as author, date, or location of the document.

Once the features are extracted, the clustering algorithm groups similar documents together based on the similarity of their feature vectors. Different clustering algorithms can be used, such as k-means clustering, hierarchical clustering, and density-based clustering.

The choice of clustering algorithm and the features used to represent the documents can have a significant impact on the performance of the clustering. It is important to carefully choose the features and tune the parameters of the clustering algorithm to achieve the best results.

Document clustering has many practical applications, such as organizing large collections of documents, identifying trends in large data sets, and enabling more efficient search and retrieval of documents.

There is no definitive answer to how many algorithms there are in machine learning, as the field is constantly evolving and new algorithms are being developed all the time. However, there are many different types of machine learning algorithms, each with their own strengths and weaknesses, and they can be broadly categorized into three main types:

1. Supervised learning algorithms: These algorithms learn from labeled data, where each data point is associated with a known output. The goal is to learn a mapping from inputs to outputs that can generalize to new, unseen data. Examples of supervised learning algorithms include linear regression, logistic regression, decision trees, support vector machines, and neural networks.
2. Unsupervised learning algorithms: These algorithms learn from unlabeled data, where there is no known output. The goal is to find underlying patterns or structure in the data, such as clusters or low-dimensional representations. Examples of unsupervised learning algorithms include k-means clustering, hierarchical clustering, principal component analysis, and autoencoders.
3. Reinforcement learning algorithms: These algorithms learn by interacting with an environment and receiving feedback in the form of rewards or punishments. The goal is to learn a policy that maximizes the cumulative reward over time. Examples of reinforcement learning algorithms include Q-learning, policy gradient methods, and deep reinforcement learning.

Within each of these categories, there are many different algorithms and variations, and the choice of algorithm depends on the specific task and the characteristics of the data.

Entropy is a concept that has multiple meanings depending on the context in which it is used. In the context of information theory and machine learning, entropy is a measure of uncertainty or randomness associated with a set of data or a probability distribution.

Specifically, entropy is a measure of the average amount of information contained in each message or data point in a set of data. The greater the entropy, the greater the uncertainty or randomness associated with the data. Entropy is often used as a criterion for selecting features or making decisions in machine learning, such as in decision trees and random forests.

The formula for entropy is:

H = -Σ(p \* log2(p))

where H is the entropy, p is the probability of a particular outcome, and log2 is the base-2 logarithm. The entropy is zero when all the outcomes are the same, and it is maximal when all the outcomes are equally likely.

Entropy has many applications in fields such as information theory, computer science, and statistical physics. It is a fundamental concept in the study of randomness, uncertainty, and information.

Adaboost (short for Adaptive Boosting) is a machine learning algorithm that can be used for classification and regression tasks. The main idea behind Adaboost is to combine several weak learners (i.e., classifiers that perform slightly better than random guessing) into a strong learner (i.e., a classifier that can make accurate predictions).

Here are the steps involved in the Adaboost algorithm:

1. Initialize the weights: Assign equal weight to each training example.
2. Train a weak learner: Train a weak learner on the training data with the assigned weights. A weak learner can be any algorithm that performs slightly better than random guessing, such as a decision tree with a small depth.
3. Compute the error: Calculate the error of the weak learner on the training data. The error is the weighted sum of misclassifications.
4. Compute the weight of the weak learner: Compute the weight of the weak learner based on its error. A weak learner with a low error will be given a higher weight than a weak learner with a high error.
5. Update the weights: Increase the weight of the misclassified examples and decrease the weight of the correctly classified examples.
6. Repeat: Repeat steps 2-5 for a predetermined number of weak learners or until the desired level of accuracy is reached.
7. Make predictions: Combine the weak learners using their weights to create a strong learner. Use the strong learner to make predictions on new data.

The final output of the Adaboost algorithm is a strong learner that is a weighted combination of the weak learners. This strong learner can be used for classification or regression tasks.

Adaboost algorithms

<https://www.analyticsvidhya.com/blog/2021/09/adaboost-algorithm-a-complete-guide-for-beginners/>

<https://www.kaggle.com/code/prashant111/adaboost-classifier-tutorial>

<https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html>

<https://scikit-learn.org/stable/modules/generated/sklearn.datasets.load_breast_cancer.html#sklearn.datasets.load_breast_cancer>

amana liburary n zor basha bo breast cancer

<https://www.geeksforgeeks.org/nlp-creating-shallow-tree/>

Shallow tree

PCA definition

Feature selection

Feature extraction

PCA = principle component model is a method of reducing the dimensionality of data and is used to improve data visualization and speed up machine learning model training.

20-2-2023

XGBOOST ( Graduanet Booship , Regression)

XGBoost is an open-source software library which provides a regularizing gradient boosting framework for C++, Java, Python, R, Julia, Perl, and Scala. It works on Linux, Windows, and macOS. From the project description, it aims to provide a "Scalable, Portable and Distributed Gradient Boosting Library".

Why we are taking y-mean

Input: training set { ( x i , y i ) } i = 1 N , a differentiable loss function L ( y , F ( x ) ) , a number of weak learners M and a learning rate α .

Algorithm:

1. Initialize model with a constant value:

f ^ ( 0 ) ( x ) = arg ⁡ min θ ∑ i = 1 N L ( y i , θ ) .

1. For *m* = 1 to *M*:
   1. Compute the 'gradients' and 'hessians':

g ^ m ( x i ) = [ ∂ L ( y i , f ( x i ) ) ∂ f ( x i ) ] f ( x ) = f ^ ( m − 1 ) ( x ) .

h ^ m ( x i ) = [ ∂ 2 L ( y i , f ( x i ) ) ∂ f ( x i ) 2 ] f ( x ) = f ^ ( m − 1 ) ( x ) .

* 1. Fit a base learner (or weak learner, e.g. tree) using the training set { x i , − g ^ m ( x i ) h ^ m ( x i ) } i = 1 N by solving the optimization problem below:

ϕ ^ m = arg ⁡ min ϕ ∈ Φ ∑ i = 1 N 1 2 h ^ m ( x i ) [ − g ^ m ( x i ) h ^ m ( x i ) − ϕ ( x i ) ] 2 .

f ^ m ( x ) = α ϕ ^ m ( x ) .

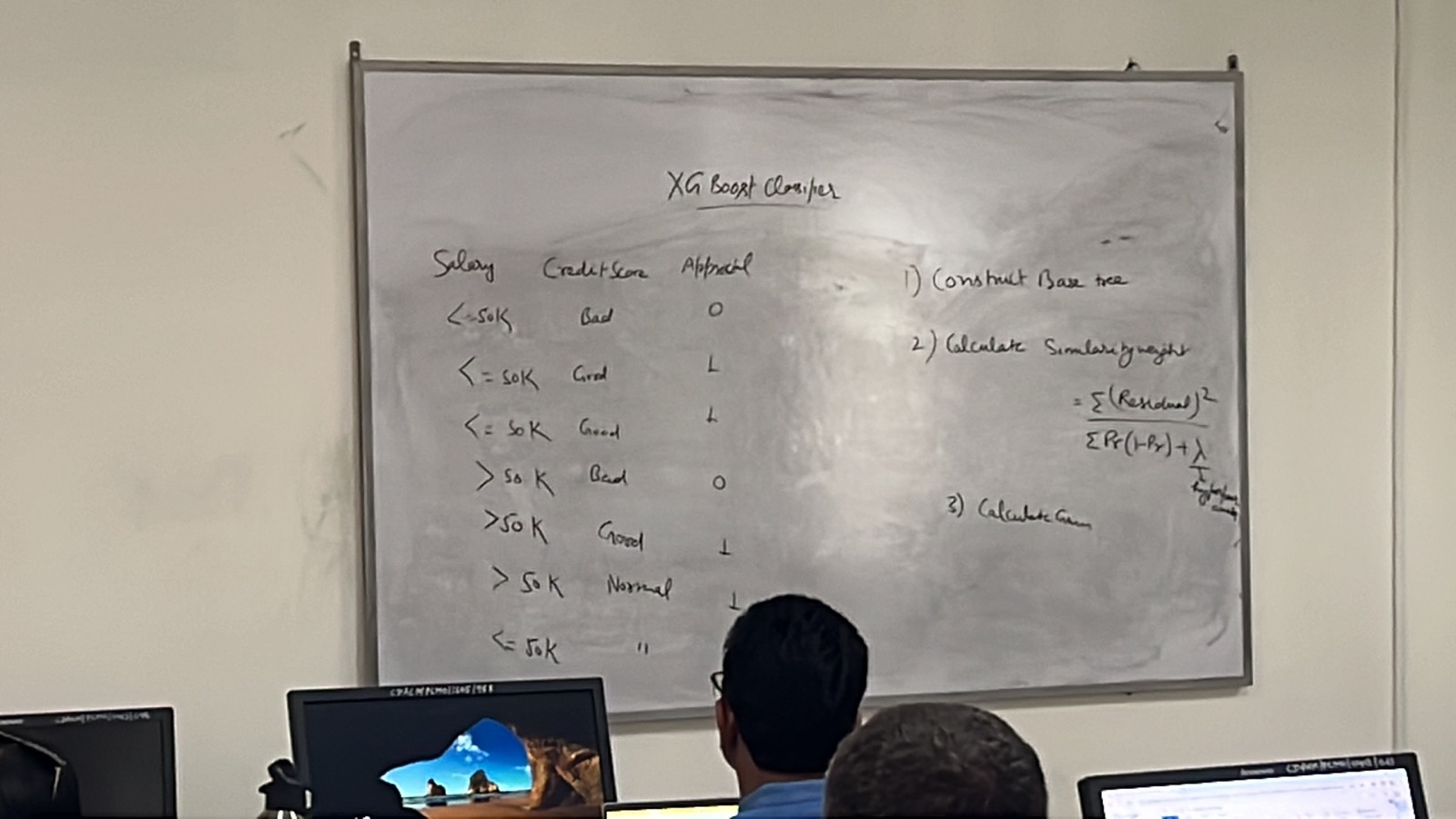
* 1. Update the model:

f ^ ( m ) ( x ) = f ^ ( m − 1 ) ( x ) + f ^ m ( x ) .

1. Output f ^ ( x ) = f ^ ( M ) ( x ) = ∑ m = 0 M f ^ m ( x ) .

<https://www.analyticsvidhya.com/blog/2021/03/gradient-boosting-machine-for-data-scientists/>

<https://www.analyticsvidhya.com/blog/2021/09/gradient-boosting-algorithm-a-complete-guide-for-beginners/>



Base model 0.5 p-value

Base model with respect of Pr=0.5

<https://towardsdatascience.com/beginners-guide-to-xgboost-for-classification-problems-50f75aac5390>

ama zor basha

<https://xgboost.readthedocs.io/en/stable/>

pip install xgboost

<https://xgboost.readthedocs.io/en/stable/parameter.html>

Linear Regression

Regression problem

Gain calculation

Calculate the error

Co2 xgboost regression

<https://sih.gov.in/>

<https://www.analyticsvidhya.com/blog/2021/08/quick-hacks-to-save-machine-learning-model-using-pickle-and-joblib/>

Pickle and joblib

<https://www.naftaliharris.com/blog/visualizing-k-means-clustering/>

zor mhma

21/2/2023

K-means clustering

<https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html>

k-means clustering customer segmentation

am file bkarawa test bka

<https://www.geeksforgeeks.org/ml-k-medoids-clustering-with-example/>

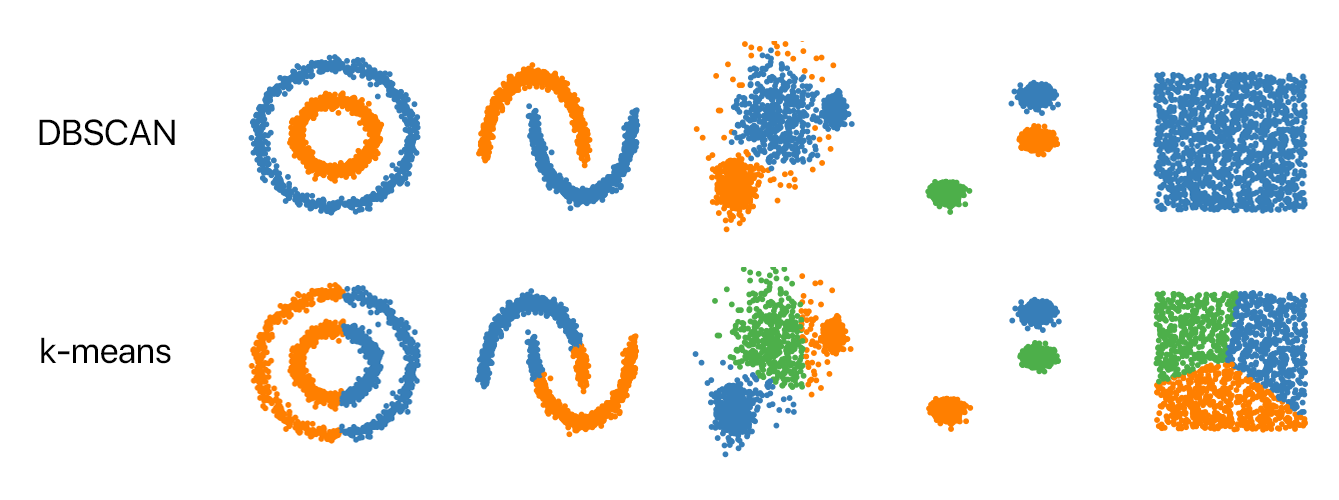
<https://www.kdnuggets.com/2020/04/dbscan-clustering-algorithm-machine-learning.html>

**Why do we need a Density-Based clustering algorithm like DBSCAN when we already have K-means clustering?**

K-Means clustering may cluster loosely related observations together. Every observation becomes a part of some cluster eventually, even if the observations are scattered far away in the vector space. Since clusters depend on the mean value of cluster elements, each data point plays a role in forming the clusters. A slight change in data points might affect the clustering outcome. This problem is greatly reduced in DBSCAN due to the way clusters are formed. This is usually not a big problem unless we come across some odd shape data.

Another challenge with k-means is that you need to specify the number of clusters (“k”) in order to use it. Much of the time, we won’t know what a reasonable k value is a priori.

What’s nice about DBSCAN is that you don’t have to specify the number of clusters to use it. All you need is a function to calculate the distance between values and some guidance for what amount of distance is considered “close”. DBSCAN also produces more reasonable results than k-means across a variety of different distributions. Below figure illustrates the fact:



Association rule mining

apriori-algorithm

<https://www.geeksforgeeks.org/apriori-algorithm/>

<https://www.w3resource.com/python-exercises/pandas/index/pandas-indexing-exercise-21.php>

This Person does not exist

<http://thispersondoesnotexist.com/>

22/2/2023

Start Deep learning

Deep\_learning.pdf

Deep\_learning\_introduction.pdf

Matrix\_math.ipynp

Gd\_and\_sgd.ipynp

Mini\_batch\_gd.ipynp

Digits\_recognition\_neural\_network.ipynp

Loss\_or\_cost\_function.ipynp

23/2/2023

<https://www.tensorflow.org/>

digits\_recognition\_neural\_network\_tensorboard(2).ipynb

tensorboard

am file:

Small\_Image\_Classification\_Using\_Simple\_Aritifical\_Neural\_Network\_GPU\_(collab\_GPU) (1).ipynb

La google Collaboratory akretwa

Dropout is a regularization technique for neural network models proposed by Srivastava et al. in their 2014 paper “Dropout: A Simple Way to Prevent Neural Networks from Overfitting” (download the PDF). Dropout is **a technique where randomly selected neurons are ignored during training**. They are “dropped out” randomly.

Deep learning is a subset of machine learning, which is a type of artificial intelligence (AI) that involves training algorithms to recognize patterns in data. Deep learning models are designed to learn from large amounts of data and can automatically discover and extract complex features and relationships from that data.

At the heart of deep learning are artificial neural networks, which are inspired by the structure and function of the human brain. These networks consist of layers of interconnected nodes, or neurons, that process and transform data as it flows through the network. Each neuron applies a mathematical operation to its inputs and produces an output, which is then fed into the next layer of neurons.

The layers of a neural network are typically organized into an input layer, one or more hidden layers, and an output layer. During training, the network adjusts the weights and biases of its neurons to minimize the error between its predictions and the true labels of the training data. This process is called backpropagation, and it allows the network to learn from its mistakes and improve its accuracy over time.

Deep learning has been successful in a wide range of applications, including computer vision, natural language processing, speech recognition, and many others. Some of the most well-known deep learning models include convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative adversarial networks (GANs).

CNN stands for Convolutional Neural Network, which is a type of deep learning model that is commonly used in computer vision applications such as image and video recognition.

CNNs are designed to process input data with a grid-like topology, such as images. The network consists of a series of interconnected layers, including convolutional layers, pooling layers, and fully connected layers.

The convolutional layers of a CNN apply a set of filters to the input data, which helps to extract features and patterns from the image. These filters are typically small, square-shaped matrices that slide over the input image and compute a dot product with each sub-region of the image. The output of this operation is a feature map that highlights regions of the image that contain the detected feature.

The pooling layers of a CNN are used to downsample the feature maps by reducing their resolution, which helps to make the network more computationally efficient and reduce overfitting. The most commonly used pooling operation is max pooling, which selects the maximum value in each sub-region of the feature map.

The fully connected layers of a CNN are used to make the final predictions based on the extracted features. These layers are similar to those in a traditional neural network, and they can be used to classify the input image into one or more categories.

CNNs have been shown to be highly effective in a wide range of computer vision tasks, such as image classification, object detection, and image segmentation. They have also been used in other applications, such as natural language processing and speech recognition, where the input data has a grid-like structure.

<https://keras.io/>

<https://www.tensorflow.org/api_docs/python/tf/function>

<https://medium.com/analytics-vidhya/haar-cascades-explained-38210e57970d>

24/2/2023

<https://towardsdatascience.com/>

mobileNet-v2

<https://www.image-net.org/>

<https://cocodataset.org/#home>

<https://www.mathworks.com/help/deeplearning/ref/mobilenetv2.html>

install cv2 juputer

<https://pixabay.com/> (good for picture)

<https://www.tensorflow.org/tutorials/images/classification>

<https://www.geeksforgeeks.org/vgg-16-cnn-model/>

TensorFlow is an open-source machine learning framework developed by Google Brain team. It provides a comprehensive set of tools for building and deploying machine learning models, including deep neural networks. TensorFlow allows developers and researchers to easily define, train, and deploy machine learning models for a wide range of applications, such as image and speech recognition, natural language processing, and even game development.

One of the key features of TensorFlow is its ability to perform computations on tensors, which are multi-dimensional arrays of data. TensorFlow uses a dataflow graph to represent the computation of a neural network, where each node in the graph represents a mathematical operation, and the edges represent the data flow between these operations. This makes it easy to parallelize computations across multiple CPUs or GPUs, making it particularly well-suited for training deep neural networks.

TensorFlow also provides a high-level API called Keras, which allows developers to quickly build and train machine learning models using a simplified interface. Additionally, TensorFlow supports a wide range of programming languages, including Python, C++, Java, and Go, making it accessible to a broad community of developers and researchers.

27-2-2023

Resent\_152CNN.ipynp

<https://www.kaggle.com/datasets/hsankesara/flickr-image-dataset>

<https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/>

<https://www.cs.waikato.ac.nz/ml/weka/book.html>

Trained data refers to the data that is used to train a machine learning model. This data is typically labeled, meaning that it has been annotated with information that the model will use to learn how to make predictions or classify new data.

The training process involves feeding the labeled data into the machine learning algorithm so that it can analyze the patterns and relationships in the data. The algorithm then uses these patterns to adjust its internal parameters and improve its accuracy at predicting or classifying new data.

The trained data is an essential component of the machine learning process. The quality and quantity of the trained data can have a significant impact on the accuracy and effectiveness of the machine learning model. Therefore, it is essential to use high-quality, relevant, and diverse training data to produce accurate and reliable results.

<https://www.tensorflow.org/api_docs/python/tf/keras/datasets/fashion_mnist/load_data>

﻿<https://www.kaggle.com/code/ahmadjaved097/multiclass-image-classification-using-cnn/data>

For the assignment

Create an account on kagle

28/2/2023

AI Computer Platforms and Trends-Apache Spark

Ms. Priti Bharadwaj

Apache Spark (Spark) is **an open source data-processing engine for large data sets**. It is designed to deliver the computational speed, scalability, and programmability required for Big Data—specifically for streaming data, graph data, machine learning, and artificial intelligence (AI) applications.

Is a framework for analyzing bigdata

Apache Hadoop is a collection of open-source software utilities that facilitates using a network of many computers to solve problems involving massive amounts of data and computation. It provides a software framework for distributed storage and processing of big data using the MapReduce programming model.

Hadoop 🡪 for analyzing data (framework)

Why DFS? Distrusted file system

Master and slave

HDFS = Hadoop distrusted file system

HDFS is **a distributed file system that handles large data sets running on commodity hardware**. It is used to scale a single Apache Hadoop cluster to hundreds (and even thousands) of nodes. HDFS is one of the major components of Apache Hadoop, the others being MapReduce and YARN.

Hadoop is fun

Fun with hadoop

Let’s lean hadoop

YARN is **the main component of Hadoop v2.** **0**. YARN helps to open up Hadoop by allowing to process and run data for batch processing, stream processing, interactive processing and graph processing which are stored in HDFS. In this way, It helps to run different types of distributed applications other than MapReduce

Hadoop is an open-source software framework used for distributed storage and processing of large data sets across clusters of computers. It was initially developed by Apache Software Foundation and is based on the MapReduce programming model, which allows for parallel processing of data across multiple machines.

Hadoop consists of two core components: Hadoop Distributed File System (HDFS) and MapReduce. HDFS is a distributed file system that stores data across multiple machines in a cluster, providing reliable and scalable storage for large data sets. MapReduce is a programming model and software framework for processing large data sets in parallel across clusters of computers.

Hadoop also includes a variety of other components and tools, including YARN (Yet Another Resource Negotiator), which manages resources in a cluster, and Hive, which provides a SQL-like interface for querying data stored in Hadoop.

Hadoop is commonly used for big data applications, such as log processing, data warehousing, and machine learning. Its distributed architecture allows it to handle large volumes of data and provides fault tolerance and scalability.

2003= GFS (google file system)

2004 = MapReduce

=Hadoop

RDD = resilient distributed dataset

rdd2 = rdd.flatMap(lambda x: x.split(" "))

x means each line

transformation = awo line code nawa ka ae noset bon mona varible

action = output

DSL = domain specific language

spark.spankcontent.text

1-3-2023

PySpark is the Python API that is used for Spark. Basically, it is a collection of Apache Spark, written in Scala programming language and Python programming to deal with data.Spark is a big data computational engine, whereas Python is a programming language. To work with PySpark, one needs to have basic knowledge of Python and Spark. The market trends of PySpark and Python are expected to increase in the next 2 years. Both terms have their own features, limitations, and differences. So, let’s check what aspects they differ.