Page No..
Date / /

Machine Learning Techniques

Unit-1 + PYQs

Topics

1. Machine learning ? Concept ? Application ? + 2 years.
2. Supervised, Unsupervised and Reinforcement learning. + 2 years.
3. Regression, classification and Clustering + 2 years.
4. Decision Tree learning.
5. Designing a learning System. + 2021-22
6. Genetic Algorithm
7. Support Vector Machine
8. Issues in Machine learning and data Science.

AKTU- 2022-23, 2020-21 [10-marks; 2 marks]

Machine Learning

Key points:

- Subset of AI
- Use algorithms and statistical models.
- Improve performance on a specific task over time.
- Without being explicitly programmed
- The primary goal of is to teach computer
- Recognize patterns in data.
- Make predictions or decisions based on patterns.

Definition:

Machine learning is a field of artifial Intelligence that involves training computer systems to learn from data and improve their performance on a specific task without being explicitly programmed. The goal of ML is to enable computers to recognize patterns in data and make predictions or decisions based on these patterns.

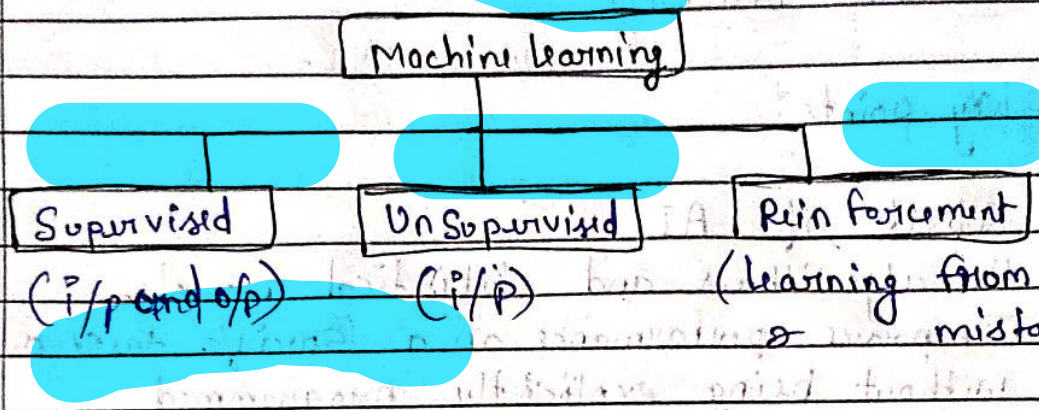
learning refers to the process of training an algorithm to recognize patterns and make predictions based on data.

e.g → M.L. used for future prediction

Page No..
Date /

AKtu - 2022-23

Types



① Supervised learning

→ It is the types of machine learning in which machines are trained using well "labelled" training data, and on basis of data, machine predict the output.

→ The labelled data means some input data is already tagged with the correct ouput.

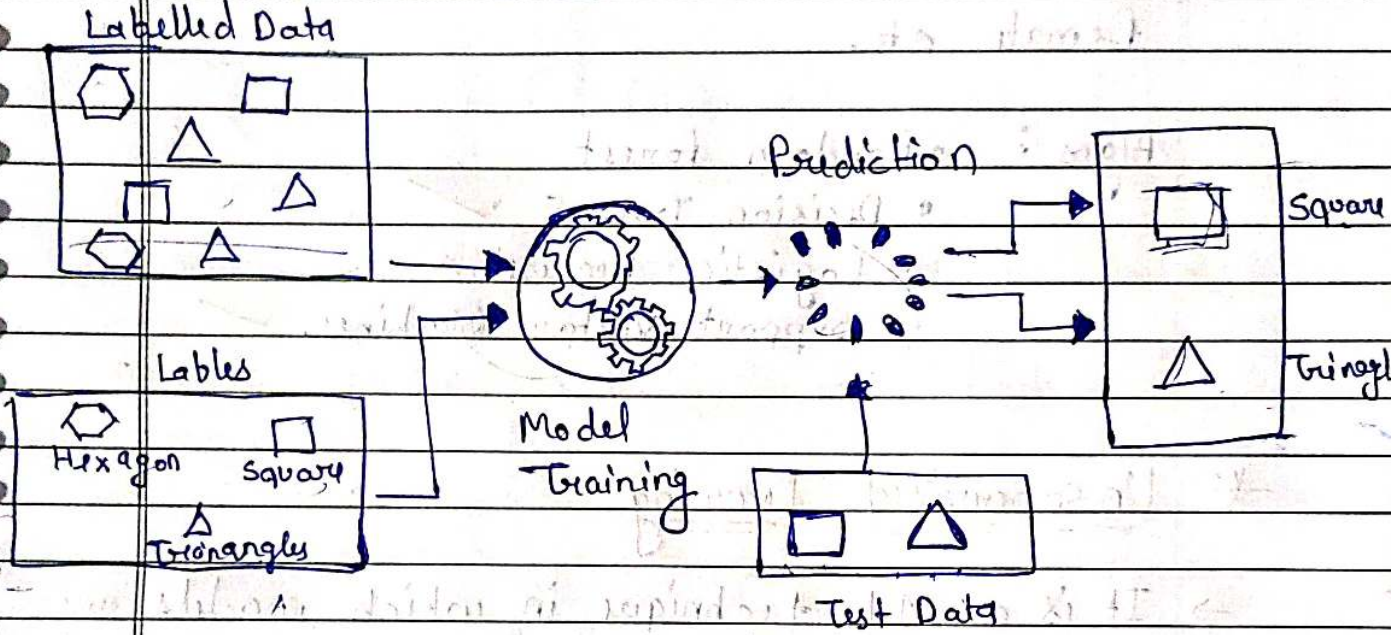
→ In real world, Supervised learning used for
* Risk Assessment,
* Image Classification,
* Fraud Detection.

How Supervised learning works?

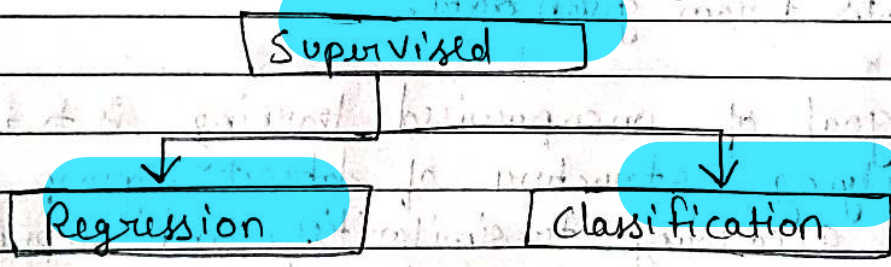
→ Suppose we have a dataset of diff. types of shapes which include square, rectangle, triangle and polygon. Now first step is that we need to train the model for each shape

Page No.:
Date: / /

6 If the given shape has four sides, and all the sides are equal ~~or~~ labelled as a Square
- If three sides labelled as triangle
- If six sides as hexagon.



Types of Supervised Learning. AKTU-2022-23



① Regression = It is used find relationship b/w the input variable and output variable. It is used for prediction of continuous variables, such as weather forecasting, Market Trends etc.

Algos :- Linear Regression, Regression Trees, Non-linear Regression, Bayesian Linear Regression.

Page No.
Date / /

② Classification

Classification algorithms are used when the output variable is categorical, which means there are two classes such as Yes-No, male-female etc.

Algos :
- Random forest
- Decision Trees
- Logistic Regression
- Support Vector Machines.

* Unsupervised Learning

→ It is a ML technique in which models are not supervised using training dataset. Instead, models itself find the hidden patterns and insights from given data.

→ The goal of unsupervised learning is to find the underlying structure of dataset, group that data according to similarities, and represent that dataset in a compressed format.

→ Clustering : It is a method of grouping the objects into clusters such that objects with most similarities remains into a group and has less or no similarities with the objects of another group.

Page No.:
Date: / /

Unsupervised Learning Algorithms :

- K-means clustering
- KNN
- Hierarchal clustering

- Neural Networks
- Anomaly detection.

Reinforcement Learning

→ A model is trained to make decisions based on trial and error.

→ It takes actions and receive feedback on those actions.

→ It receives rewards for making correct decisions and penalties for making incorrect decisions.

→ It helps to learn how to make better decision in the future.

→ Commonly used in the game playing, robotics, and Control System.

→ It is like children learn new behaviours and skills.

Page No.
Date / /

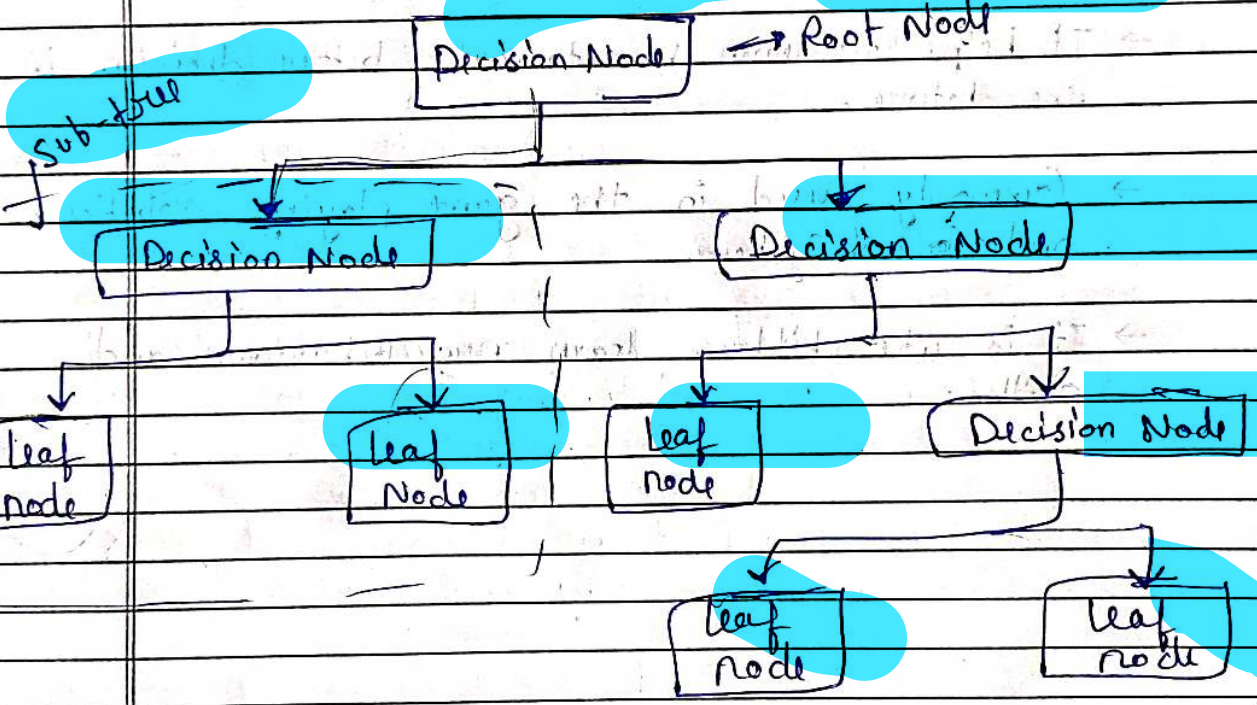
Decision Tree Learning

→ Decision Tree is a Supervised learning Technique that can be used for classification and Regression problems.

→ but mostly it is prefered for solving classification problems.

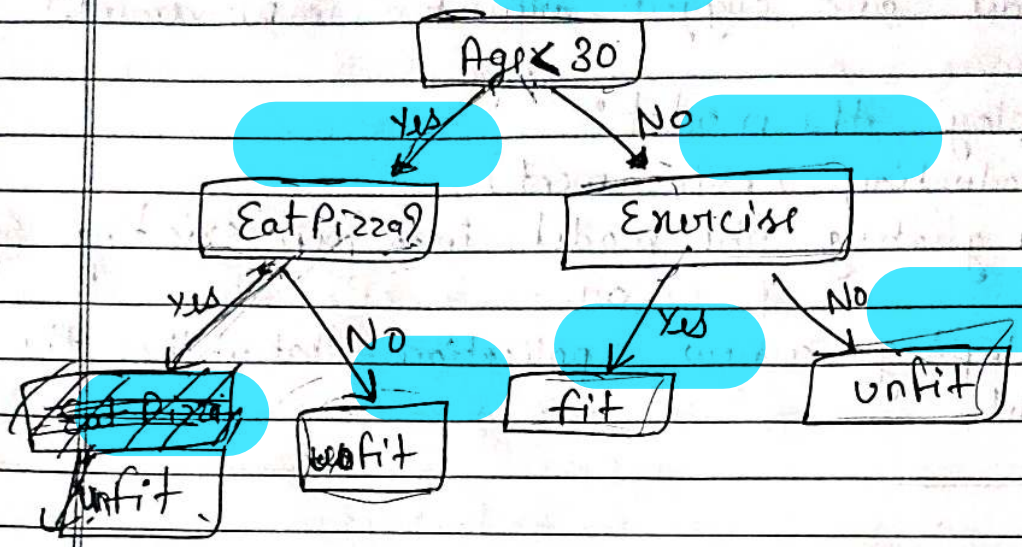
→ It is a tree-Structured classifier, where

① internal node → Features of a dataset
② branches → decision rules
③ leaf nodes → outcomes.



Page No.:
Date: / /

e.g → Is person is fit or not.



* Designing a learning system → 1 year.

Steps :

① Define the problem :
(input, output and evaluation metrics).

② Collect and pre process data :
→ (data to train and test the system).
→ (involve cleaning, transforming and normalizing the data).

③ Choose an algorithm :
(Supervised or Unsupervised or Reinforcement).

④ Train the model :
→ (on the training data)
→ Recognize pattern.

Page No.
Data / /

(5) Test the model:
(Identify any issues or limitations with the model and suggest areas for improvement)

(6) Deploy the model:
→ (production environment)
→ Integrating the model into an existing system or developing a new application that utilizes the model.

* Bayes' Theorem.

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

* Genetic Algorithm.

→ It is a type of optimization algorithm
→ It is inspired by the process of natural selection in biological evolution.
→ It starts by generating a bunch of random solutions and then tests them to see how well they perform.
→ The best solutions are then combined to create new solutions, just like new baby inherits traits from parents
→ The process is repeated until the best solution found
→ Commonly used for generate high-quality solutions for optimization problems and search problems.

Page No.:
Date: / /

* Support Vector Machine

→ SVM is a type of supervised machine learning algorithm that can be used for classification or regression tasks.

→ SVM tries to find the best possible boundry between the different classes in the data by finding a hyper plane that maxmizes the margin b/w two two classes.

→ It is often used in applications such as image recognition, text classification etc.

* Issues in Machine Learning and data Science

① Data Quality:
→ essential to have good quality data to produce quality M.L algo and model.
→ Accuracy of M.L is depend on the quality of the data.

② Transparency:
→ It is difficult to make definitive statements on how well a model is going to generalize in new environments

Date

(3) Privacy and Security :

→ M.L Algos often require access of large amount of sensitive data, such as personal info, medical records or financial data

(4) Computational Resources : M.L algos can be computationally intensive and may require high amount of processing power or specialized hardware to run efficiently.

Unit-1 Completed

Page No..
Date / /

Machine Learning Techniques

Unit-2 + PYQs.

Topics

1. Regression and its type? linear Vs logistic Regression. + 2022-23 [10 marks]
2. Bayesian learning + Baye's Thiarem + [2022-23, 2020-21]
3. Baye's optimal classifier ✓
4. Naive Bayes classifier + 2021-22 [10 marks]
5. E-M Algo - 2021-22 ✓
6. Concept learning + 2021-22 ✓
7. Support Vector Machine. + 2022-23, 2020-21 ✓
   - → Types
   - → Kernel function and its type ✓
   - → properties
   - → Issues.

Subscribe + Join Telegram

Road to 5K

* Regression - Unit-1

- It is a statistical model/method used to model the relationship b/w a dependent variable and one or more independent variables.
- The model can take various forms, such as linear regression, logistic regression, and polynomial regression, depending on the nature of the relationship b/w the variables.

* Linear Regression

↳ Same definition as Regression.

→ It makes predictions for continous/real or numeric values such as sales, age, product price, etc.

→ The goal is to find the line of best fit.

→ Simple linear equation = $y = mx + b$

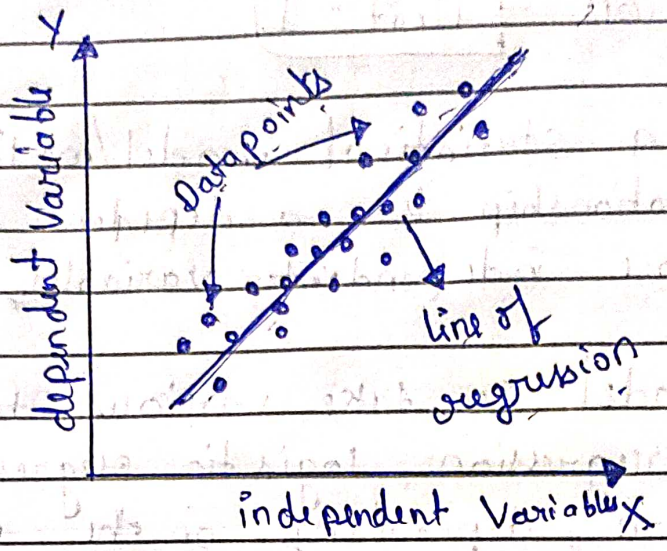
→ $Y \rightarrow$ dependent and x is independent, $m \rightarrow$ slope of line.

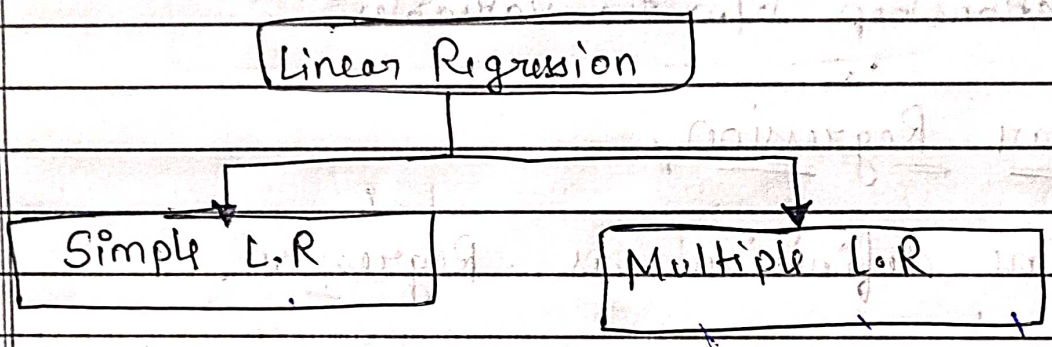
→ It can be used for both prediction and Inference (Assumption).

→ The linear regression model provides a sloped straight line representing the relationship b/w the variables.

@Multi Atoms

Page No..
Date / /





① Simple Linear Regression ⇒ If a single independent variable is used to predict the value of a numerical dependent variable.

eqv → $Y = b_0 + b_1 * x_i$

Y → dependent variable, x → independent
b1 → slope of the line.

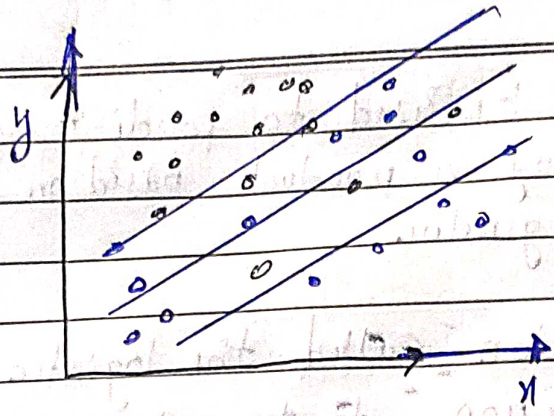
② Multiple linear Regression ⇒ If more than one independent variable is used to predict the value of numerical dependent variable.

eqv → $Y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_n * x_n$

Y → dep, x → indep, b_1, b_2 → slope.

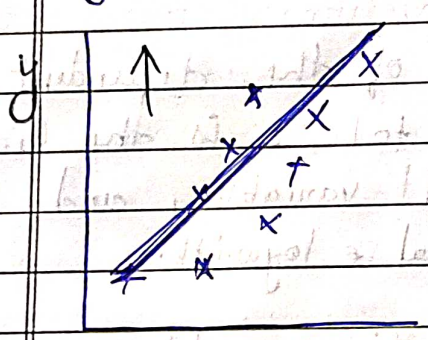
@MultiAtoms

Page No.:
Date : / /

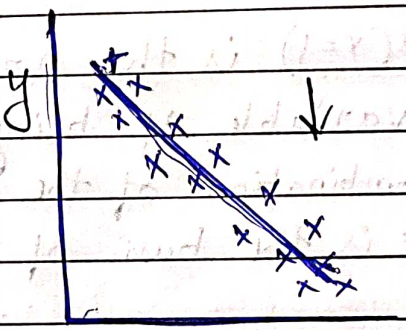


* Linear Regression line ⇒ It showing the r/p b/w the dependent and independent variables called a regression line.

→ Types of relationship:



Positive Correlation.



Negative Correlation.

⟹ Logistic Regression

→ Same as linear Regression but here the dependent variable is binary or categorical.
e.g ⇒ 010101001
0110101110

→ It is used to predict the probality of an event occurring based on the values on the independent variables.

→ Two possible outcome → True or false. or 0 or 1.

@MultiAtoms

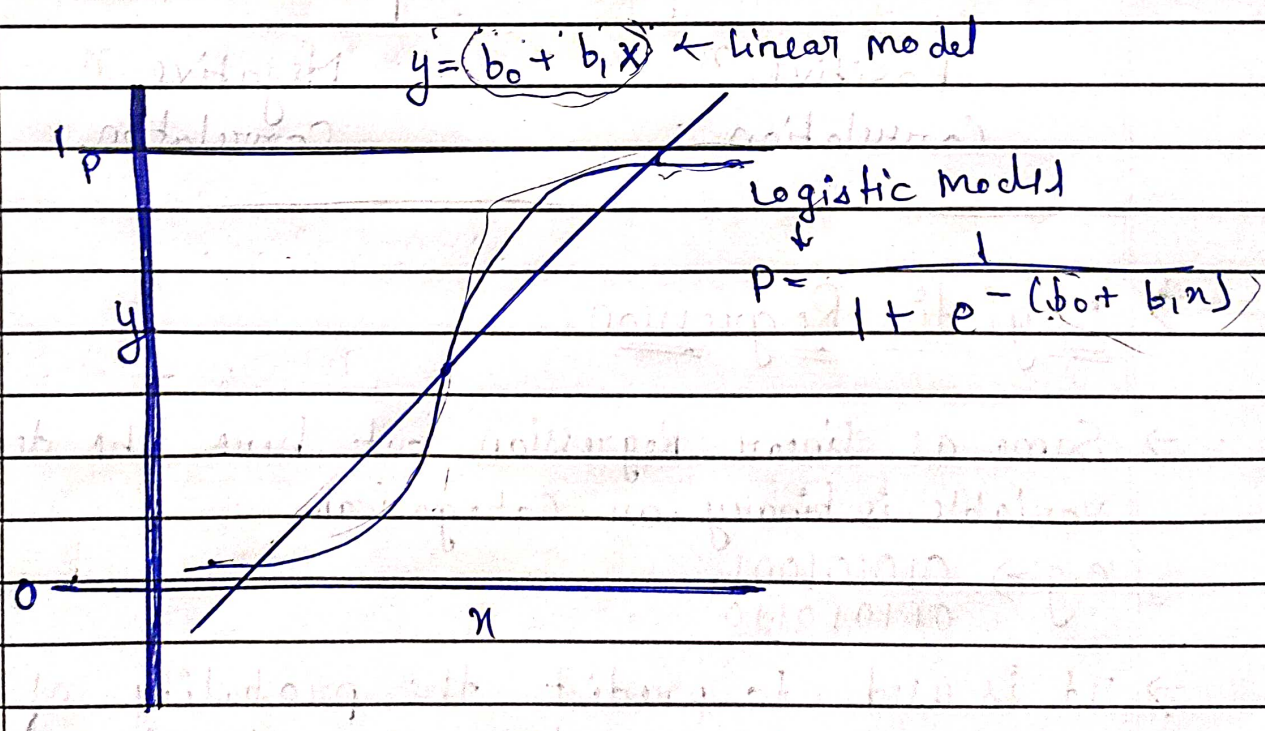
Page No..
Date / /

→ for eg → it could be used to predict whether a customer will buy a product based on their age, income, and gender.

→ It uses a function called the logistic function or sigmoid function to transform the output of a linear equation into a value b/w 0 and 1

$$\text{equ} \Rightarrow P(Y=1) = 1/(1+e^{-z})$$

P(Y=1) is the probability of the dependent variable y being equal to 1, z is the linear combination of the independent variables, and e is the base of the natural logarithm



fields ⇒ healthcare, finance and marketing.

@Multi Atoms

Page No.:
Date : / /

* Bayesian Learning ✓

→ B.L is a fundamental statistical approach to the problem of pattern classification.
→ In B.L., a model is specified with parameters that describes the probability distribution of the data.
→ These parameters are initially set using prior knowledge or assumptions, and then updated based on the observed data using Baye's theorem.
→ The updated probabilities are used to make predictions about new data.

→ It useful where is data is limited or noisy

Application ⇒ spam filtering ✓
⇒ medical diagonosis ✓
⇒ NLP ✓
⇒ image recognition etc. ✓

* Baye's Theorem - AKTU - 2022-23, 2020-21 ✓

→ It is a mathematical concept that describes the relationships b/w conditional probabilities of two events.

$$P(A/B) = \frac{P(B/A)\ P(A)}{P(B)}$$

→ It is used in A.I and M.L. to build probabilistic models and make decisions based on uncertain data.

@MultiAtoms

Page No..
Date / /

Likelihood (PRIOR)

$$P(A/B) = \frac{P(B/A) \cdot P(A)}{P(B)}$$

Posterior (The probability of A on Event B.)

(Marginalization) Probability of B,

* Concept learning - AKTU - 2021-22.

→ It refers to the process of building a probabilistic model that can recognize and generalize pattern in data and use that knowledge to classify new examples.

→ Bayesian Concept learning involves using Bayesian inference.

→ B.C.L involve

① Defining a PRIOR probability distribution.

② Collecting a set of training examples

③ Using the training examples and Baye's theorum.

④ Using the POSTERIOR distribution to make prediction.

@MultiAtoms

Page No.:
Date : / /

* Bayes optimal classifier

→ The Baye's optimal classifier is a probabilistic model that predicts the most probable outcome for a new instance using the Bayes Theorem.

* Naive Bayes Classifier AKTU-2021-22

→ A N.B.C. is a probabilistic machine learning model that's used for classification task. The base of the classifier is Bayes Theorem.

→ N.B. Cs are a collection of classification algos.

→ It is not a single algorithm but a family of algorithms where all of them share a common principle.

- Assumption = The fundamental Naive Bayes assumption is that each features makes an Independent, Equal contribution to the outcome.

- Note: The assumptions made by Naive Bayes are not generally correct in real-world situations.

→ Write Bayes theorem formula and explain it.

@Muhtiatomy

Page No..
Date / /

AKTU - 2021-22 [10 marks]

E-M Algorithm

→ The E-M (Expectation - Maximization) algorithm is an iterative optimization method used to estimate the parameters of probabilistic models when there are missing or incomplete data.

→ It is a technique to find maximum likelihood estimation when the latent variables are present.

→ It is also refferred to as the latent variable model.

→ The EM algo is combination of various unsupervised machine learning algo.

Steps.

① Initialization : start by initializing the parameters of the model with some initial values.

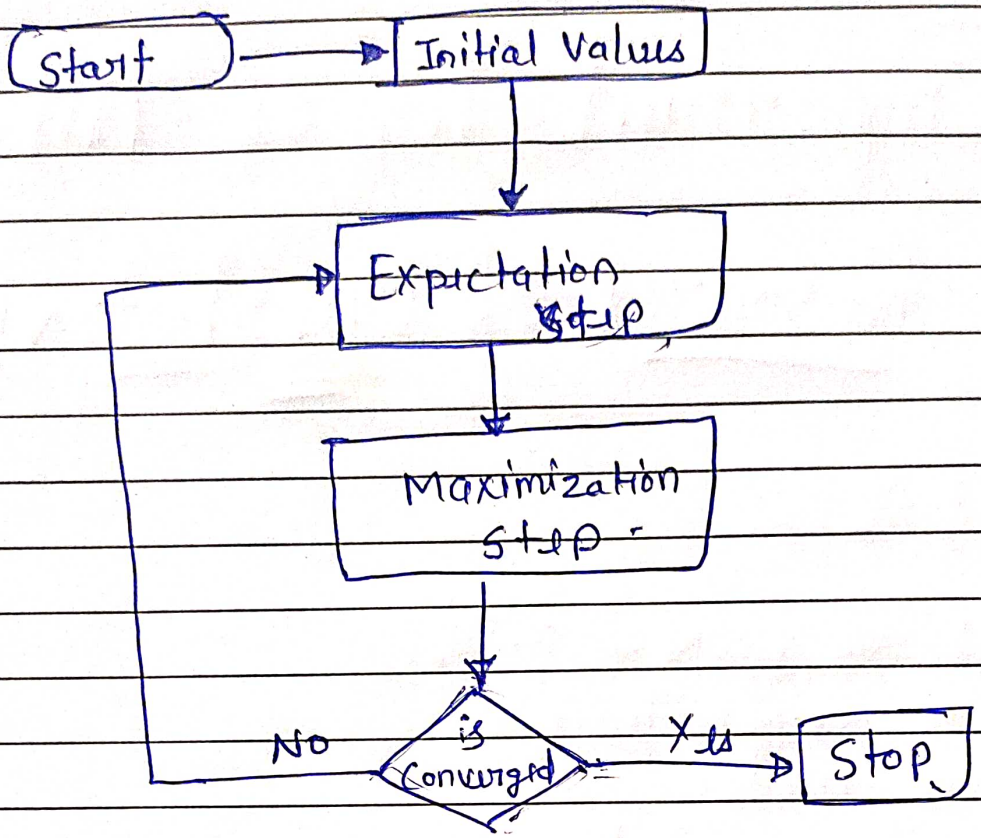
② Expectation step(E): Calculate probability.

③ Maximization step (M): Updating parameter.

④ Iteration : Repeat E-step and M-step until the convergence of the values occurs.

Page No.:

Date : / /


Start
Initial Values
Expectation Step
Maximization Step
is Converged
No
Yes
Stop

Page No.:
Date : / /

* Support Vector Machine.

→ SVM is a type of supervised machine learning algorithm that can be used for classification or regression tasks.

→ SVM tries to find the best possible boundry between the different classes in the data by finding a hyper plane that maxmizes the margin b/w two two classes.

→ It is often used in applications such as image recognition, text classification etc.

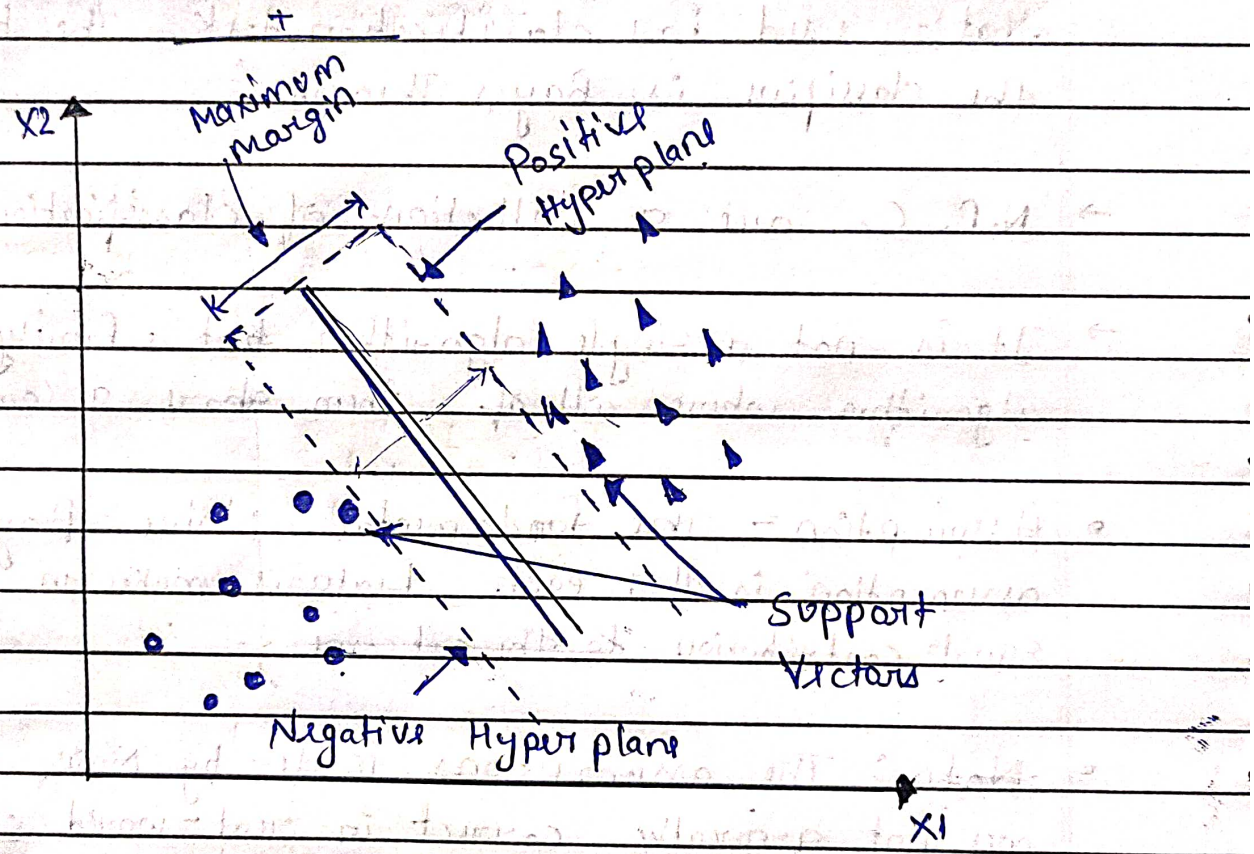
@MultiAtomy

Page No..
Date / /

* EM algorithm and its steps → Read ~~from~~ ~~Notes~~ AKTU - 2022-23, 2020-21 (2 times 10 marks).

* Support Vector Machine

→ Definition from Unit-1



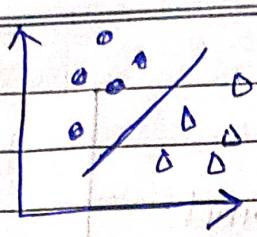
* Types of SVM

1. Linear SVM

→ It is used for linearly separable data, what means dataset divided into two classes by using single straight line.

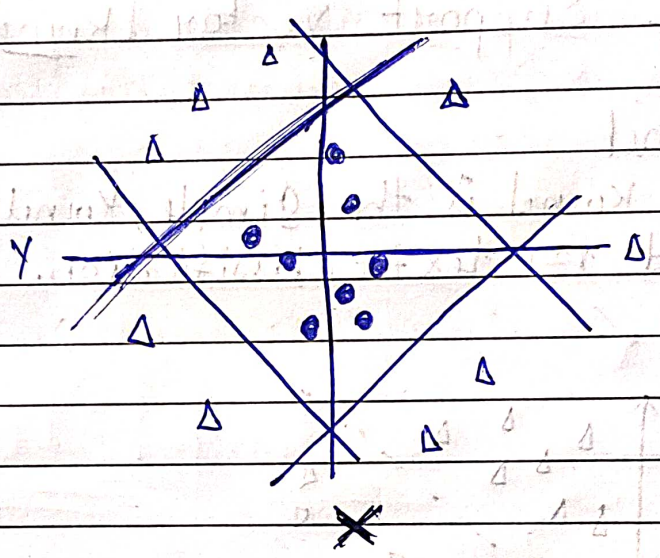
@MultiAtoms

Page No.:
Date : / /



② Non-Linear SVM

→ It is used for non-linearly separated data, which means if a dataset cannot be classified by using straight line.



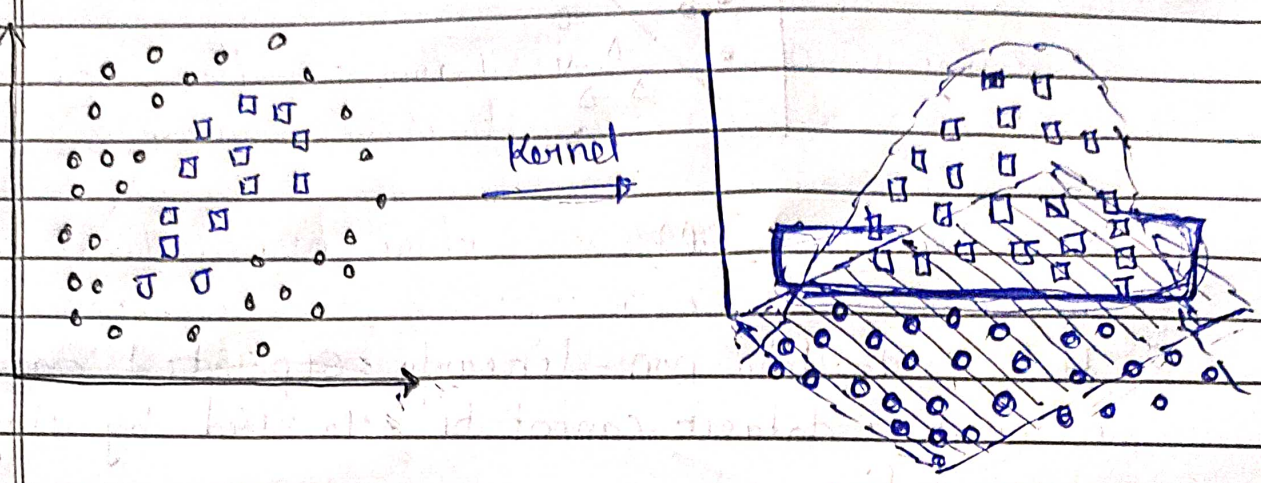
* Kernel ~ AKTU - 10 marks

→ It is function that is used to transform the original feature space into a higher-dimensional space.

→ The Kernel function enables ~~SVC~~ SVM to effectively handle non-linear classification problems. This allows SVM to find a non-linear decision boundary in the transformed feature space, even through the original data may not be linearly separable.

@MultiAtomy

Page No..
Date / /



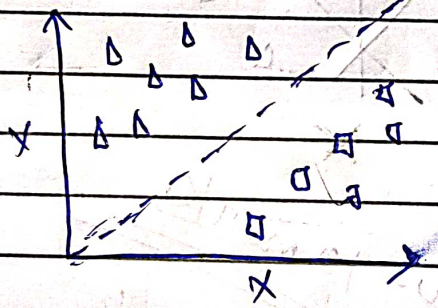
* Types of Support Vector Kernel ✓

① Linear Kernel

→ The linear kernel is the simple kernel function.

→ mostly used for text-classification.

→ straight line ✓



② Polynomial Kernel

→ It introduces new dimensions by computing all possible polynomials up to a specified degree of the original features.

→ It captures non-linear relationships b/w the data points.

Diagram

@MultiAtoms

Page No.:
Date: / /

(3) Gaussian Kernel.

→ also known as Radial Basis function (RBF) Kernel.
→ widely used and is effective for handling non-linearly separable data.
→ It maps the data into an infinite-dimensional space where each data point is represented as a Gaussian function.

Properties of SVM.

1. Maximal Margin. → distance b/w hyperplane.
2. Non-linearity with Kernels.
3. Support Kernel. → Dataset.
4. SVB also work on small number of support vectors.
5. Versatility → (use for both Regression and classifica
6. Control of Complexity → (user can control it's Complexity)

Issues in SVM

1. Scalability with large Dataset ⇒ Required more cost and memory-intensive.

Page No..
Date / /

② Lack of Probabilistic Interpretation : (does not provide probabilities outputs)

③ Imbalanced Data : where the number of samples in different classes is significantly unequal, SVMs may be biased towards the majority class.

④ Interpretability : SVMs tend to provide good predictive performance, but they may not offer direct interpretability of the learned model.

Unit-2 - Completed.

Machine learning Techniques

UNIT - 3 [One Shot]

Quick revision:

1. Decision Tree learning algorithm.
2. Inductive bias and Information gain.
3. ID-3 Algorithm
4. KNN algorithm and case-based learning.

@brevitylearning

ⓐ Decision tree learning algorithm:

- It is a popular machine learning algorithm used for classification and regression tasks.
- Process of building a decision tree:

1. Choose the best feature:
   - Use a metric (information gain or gini index) to select the feature that best splits the data
2. Split the dataset:
   - Divide the dataset into subsets based on the chosen feature
3. Create decision node or leaf node:
   - If all samples belongs to one class, create a leaf node with class lable.
   - If all samples doesn't belongs to one class, go to step 2.
4. Repeat: Continue splitting recursively until stopping criteria are met such as maximum tree depth, minimum number of samples in a node or no further information gain.

Pruning in Decision tree:

It is used to receive the size of the tree by removing nodes that provide little powere in predicting

target variables, thereby improving model performance on unseen data.

Advantages:

- Easy to understand and interpret
- Requires little data preprocessing.
- Can handle both numerical and categorical data

Disadvantages:

- Prone to overfitting
- Can be unstable with small changes in data.
- Greedy nature can be lead to suboptimal splits.

@brewilearning

Inductive bias and information gain:

* Inductive bias: It refers to the set of assumptions that a learning algorithm uses to generalize from the training data to unseen instances.

- It helps to make predictions on new data by guiding the learning process in the face of limited information.
- Inductive bias is crucial because it helps the algorithm to generalize better by avoiding overfitting to the training data, allowing it to perform well on unseen data

* Information gain:

- It is a key metric used to decide which feature to split on at each step in the construction of a decision tree.
- It measures the reduction in impurity achieved by partitioning the dataset according to a given attribute.

Calculation of Information gain:

1. **Entropy:** It is a measure of impurity or randomness in a dataset.

$$E(S) = -\sum_{i=1}^{c} p_i \log_2(p_i)$$

- S is the dataset
- c is no. of classes
- p_i is portion of instances.

2. **Conditional Entropy:** After splitting dataset based on an attribute, the conditional entropy is the weighted sum of the entropies of the resulting subsets.

$$E(S,A) = \sum_{j=1}^{n} \frac{|S_j|}{|S|} \text{Entropy}(S_j)$$

3. **Information gain:** It is the difference between the entropy of the original dataset and the conditional entropy after splitting on attribute A;

$$\text{Information gain}(S,A) = \text{Entropy}(S) - E(S,A)$$

@hrewilearning

ID-3 algorithm:

- Iterative Dichotomiser 3 is a popular decision tree algorithm used in ML.
- It aims to build a decision tree by iteratively selecting the best attribute to split the data based on information gain.
- Each node represents a test on an attribute, and each branch represents a possible outcomes of the test.
- The leaf nodes of tree represent the final classifications.

Process:

1. Calculate Entropy of the Dataset:

- Determine the proportion of each class in the dataset.

2. Compute entropy for each attribute:

- Divide the dataset into subsets and calculate the entropy for each subset.

3. Calculate Information gain for each attribute:

- Compute the weighted average entropy for the subsets created by each attribute, and calculate information gain.

4. Select best attribute:

- Choose the attribute with the highest information gain as the decision node.

5. Create decision node:

- Use the selected attribute to split the dataset into subsets.
- Create branches for each possible value of the chosen attribute.

6. Repeat for each subset:

- Apply steps 1-5 until following conditions met:
  - i) All instances in the subset belong to the same class.
  - ii) There are no more attributes to split on
  - iii) The dataset is empty

@brucilearning

#KNN Algorithm:

- K-Nearest Neighbors (KNN) algorithm is a simple, yet powerful supervised machine learning algorithm used for both classification and regression.
- It is based on the idea that similar data points are likely to be close to each other in feature space.

Steps in KNN algorithm:

1. Choose the number of neighbors (K):

- Decide the no of neighbors to consider (K).
- Common choices are small odd numbers (like 3 or 5) to avoid ties in classification.

2. Calculate the distance:
- Calculate Euclidean distance b/w given test data points.

3. Identify the K nearest neighbors:
- Sort calculated distance and select K-closest training data points.

4. Make a prediction:
- Find majority vote among the K-nearest neighbors, and assign the majority value as prediction for the the test data points.

SUBSCRIBE → @brevilearning

* Example:

X1	X2	Y
7	7	False
7	4	False
3	4	True
1	4	True
3	7	?

Step 1: let. $k = 3$ (always choose $k = 3$ or 5).

Step 2: Find Euclidean distance for each known Y, w.r.t. unknown X1 & X2.

- $ED(7,7),(3,7) = \sqrt{(7-3)^2 + (7-7)^2} = \sqrt{16} = 4$
- $ED(7,4),(3,7) = \sqrt{(7-3)^2 + (4-7)^2} = \sqrt{25} = 5$
- $ED(3,4),(3,7) = \sqrt{(3-3)^2 + (4-7)^2} = \sqrt{9} = 3$
- $ED(1,4),(3,7) = \sqrt{(1-3)^2 + (4-7)^2} = \sqrt{13} = 3.6$

Step 3: Majority check after arranging dataset in increasing order of ED, and choosing output(Y) of the 3 closest neighbors ($K=3$).

no.(True) = 2, no.(False) = 1, Hence, Y(3,7) = True.

Case-based Learning:

⇒ Used for classification and regression problem, process of solving new problems based on the solution of previous problems.

⇒ Case-based learning is an advanced instance based learning method which is used to solve more complex problem.

⇒ It does not use Euclidian Distance matrix.

⇒ When a new case arrives to classify then identical case is chick in memory if any similar cases is found then its solution is also retreived.

Steps in CBL:

- Retreive: Grather data from memory and chick previous similar solution to the problem.
- Reuse: Suggest the solution based on experience.
- Revise: Modifying or updating knowledge based on new information or feedback.
- Retain: Store this updated method in memory for future problems.

Applications:

1. Customer service help desk for diagnosis of problems.
2. Engg. and law for technical design and legal roles.
3. Medical science for patients case history and treatements.

Thanks for watching

@brwilearning YT !!

"All the best for your exams"

Page No.:
Date: /

Machine Learning Techniques

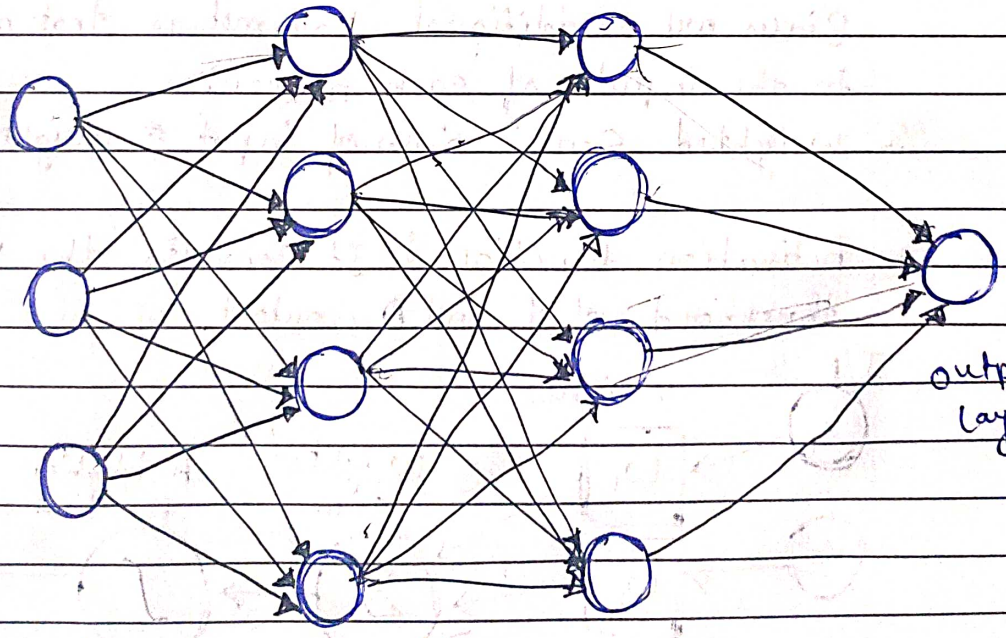
Unit-4

* Artificial Neural Network +2021-22.
* Perceptron
* Gradient Descent
* Delta Rule
* Back propagation algorithm - 2020-21
* Deep learning
* CNN - 2022-23 and its layer
* 1D - 2D CNN
* Training of Network
* Case studies.

Page No.
Date

* Artificial Neural Network (ANN)

- Inspired by structure and function of the human brain.
- Computational model that consists of interconnected nodes, called neurons, which are organized into layers.
- Each neuron receives input from other neurons, processes that input, and then sends an output signal to other neurons in the network.
- The input layer receives input data, which is then passed through one or more hidden layers, where data processed.
- The final output is produced by the output layer, which provides the result of the computation.



Page No.:
Date : / /

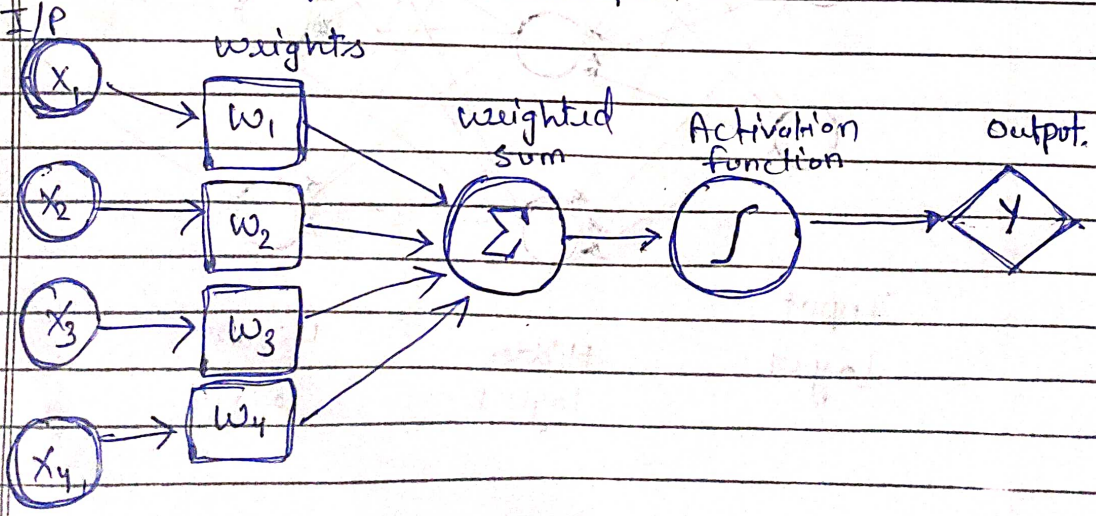
Application of ANN

- Image and speech recognition
- Natural language processing and
- Time-series prediction.

* Perceptron

→ It is the simplest form of a neural network.
→ also known as single layer neural network.
→ It contains 4 main parameters : input values, weights and Bias, net sum and an activation function.

1. Input Layer : Represent input features.
2. Weight and Bias : Strength of connection b/w neurons.
   Biases are additional parameters that are added to the input of each neuron.
3. Weighted Sum : Sum of input * weights (plus a bias term)
4. Activation function : It converts the values in threshold of 1 and 0 output.



Page No.
Date

* Gradient Descent:

→ It is an optimization algorithm used to ~~maint~~ minimize the cost function in the process of training a machine learning model.

→ The goal is to minimize the difference b/w the predicted output and the actual output.

→ The optimization process involves adjusting these parameters iteratively:

- Initialize parameters ✓
- Calculate the Gradient ✓
- Update Parameters. ✓
- Repeat step 2 & 3.

* Delta Rule: ✓

→ It is a gradient descent learning rule.

→ This rule is used to update the weights in the direction that reduces the error between the predicted output and the target output.

$$\Delta w_{ij} = \eta \cdot (t - y) \cdot x_i$$

where,

- Δw_{ij} → change in weight b/w i and j.
- η → learning rate ✓
- t → target output ✓
- y → Actual output ✓
- x_i is the input from neuron on i. ✓

Page No.:
Date: /

[2020-21]

* Back propagation algorithm

→ It is an algorithm used in the training of fed forward neural networks for supervised learning.

→ This Algo. that is designed to test for errors working back from output nodes to input nodes.

→ Back propagation efficiently computes the gradient of the loss function with respect to the weights of the network.

→ updating weights to minimize loss.

→ Diagram same as ANN with backward arrows

* Self-Organizing Map (SOM) Algorithm

→ SOM provides a data visualization technique which helps to understand high dimensional data by reducing the dimensions of data to a map.

→ represents clustering

→ also known as Kohonen Map. (world map).

Variants ⇒ (GSOM, S-SOM, I-SOM)
(Batch Self-Organizing Map, Multi layer SOM.)

Page No.
Date: / /

* Deep learning

→ It is a subfield of machine learning that involves the use of ANN to model and solve complex problems.

→ "Deep" refers to the use of multiple layers in neural network. These networks, often called deep neural networks.

→ Used in
- image and speech recognition.
- Natural language processing etc.

* Concept of Convolutional neural network 22-23

→ A CNN is a specialized type of neural network designed for tasks involving grid-like data, such as images or sequences. [matrix]. [image pixel] [2d]

→ It is a specialized type of neural network model designed for working with two dimensional image data. Although they can be used with one-dimensional and Three-dimensional data.

→ CNNs have proven to be highly effective in computer vision tasks, including [image classification, object detection, and image generation.]

AKTU - 2022-23

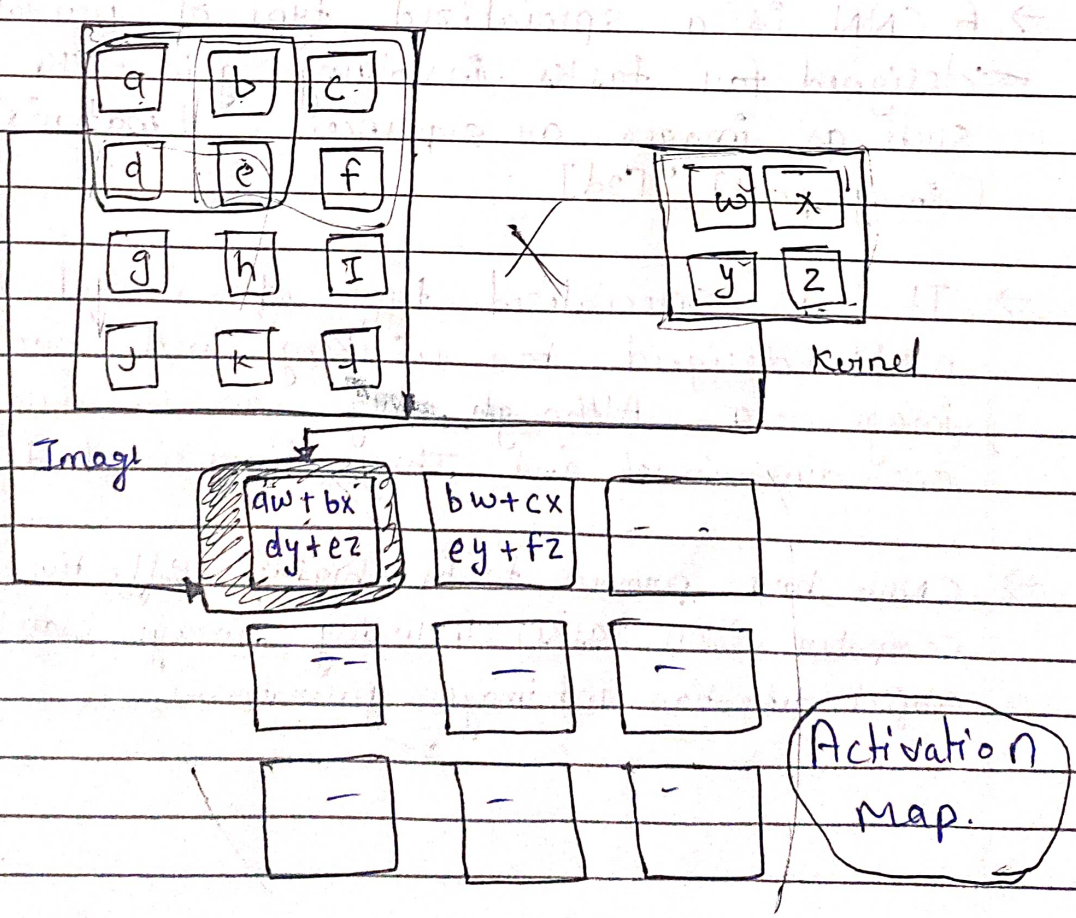
Page No.:
Date: / /

Types of Layers:

① Convolutional Layers:

→ CNNs use convolutional layers that apply convolution operations to the input data.
→ A convolution involves sliding a small filter (also called as Kernel)
→ This operation captures local patterns, such as edges or textures.

• Multiple filters: Each convolutional layer typically consists of multiple filters, and each filter captures a different aspect or features of the input.



Page No.
Date

② Activation function

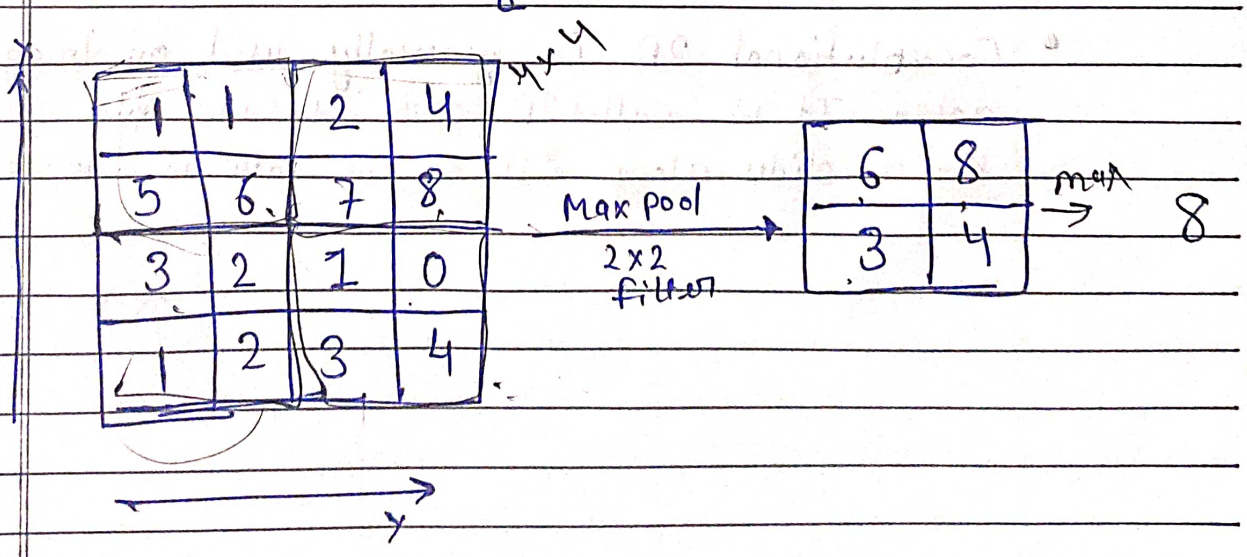
→ After the convolution operation, an activation function (commonly ReLU - Rectified Linear Unit) is applied element - wise to introduce non-linearity [This allows the network to learn more complex mappings]

③ Pooling Layers

→ It reduce the spatial dimensions of the input by down - sampling. [eg → 2D matrix to single value]

→ Common pooling operations (maximum value in a region) and (calculating the average).

• Subsampling : Pooling helps make the representation invariant to small translations and distortions in the input. It also reduces the computational load and the number of parameters in subsequent layers.



⑪ Fully Connected Layers

→ After several convolutional and pooling layers, one or more fully connected layers are often added.

→ These layers take the high-level, spatially invariant features learned by the convolutional layers and use them for classification or regression tasks.

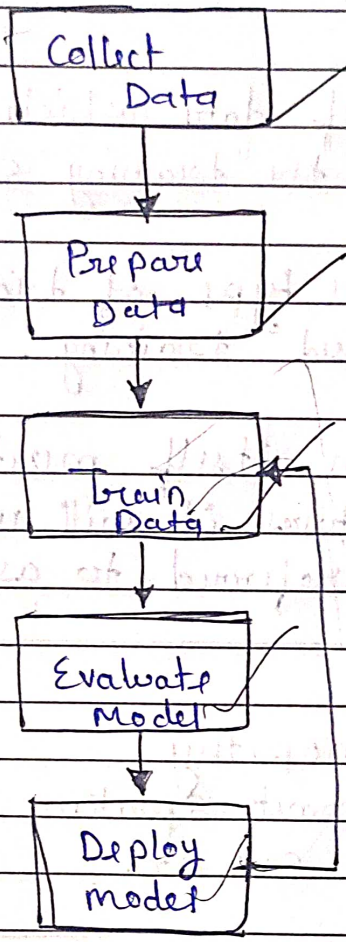
* 1D & 2D Convolutional Neural Network

- A 1D CNN is primarily used for sequential data, such as time series or text data.
  → It allows the network to learn patterns and features along a sequence.

- Convolutional 2D is generally used on image data. It is called 2D CNN because the kernel slides along 2D on the data as shown.

Page No.
Date

* Training of Network



1. → Once a network has been structured for a particular application, that network is ready to be trained.

2. → To start this process the initial weights are chosen randomly. Then, the training or learning begins.

3 → There are two approaches.

a. In Supervised training, both the input and the output provided.

Page No.:
Date: /

b. Errors are then propagated back through the System. This process occur over and over.

c. The set of data which enables the training is called the "training set".

d. The other type of training is called unsupervised training. - provide only input.

e. The system itself must then decide what features it will use to group the data. - referred to as Self-organization

Case study
* Diabetic Retinopathy
* Building a smart Speaker
* Self-driving car etc.

Describe diabetic retinopathy on the basis of deep learning.

1. Diabetic Retinopathy (DR) is one of the major causes of blindness in the western world. Increasing life expectancy, indulgent lifestyles and other contributing factors mean the number of people with diabetes is projected to continue rising.

2.Regular screening of diabetic patients for DR has been shown to be a cost-effective and important aspect of their care.

3.The accuracy and timing of this care is of significant importance to both the cost and effectiveness of treatment.

4.If detected early enough, effective treatment of DR is available; making this a vital process.

5.Classification of DR involves the weighting of numerous features and the location of such features. This is highly time consuming for clinicians.

6.Computers are able to obtain much quicker classifications once trained, giving the ability to aid clinicians in real-time classification.

7.The efficacy of automated grading for DR has been an active area of research in computer imaging with encouraging conclusions.

8.Significant work has been done on detecting the features of DR using automated methods such a support vector machines and k-NN classifiers.

9.The majority of these classification techniques arc on two class classification for DR or no DR.

Using artificial neural network how we recognize speaker.

• Objective: Develop a speaker recognition system using Artificial Neural Networks (ANNs).

• Data Collection: Collect diverse audio samples from various speakers, covering different speech patterns and recording Conditions.

• Feature Extraction: Extract relevant features, such as MFCCs, from the audio data to capture unique voice characteristics.

• Model Architecture: Design an ANN architecture with input, hidden, and output layers suitable for speaker recognition.

• Training and Tuning: Train the model using optimization algorithms and tune hyperparameters for optimal performance.

• Validation and Testing: Regularly evaluate the model on validation sets, ensuring it generalizes well to unseen speakers during testing.

• Deployment: Deploy the trained model for real-world speaker recognition tasks, integrating it into applications like security systems or voice assistants.

• Continuous Improvement: Monitor the model's performance in real-world scenarios, considering continuous learning strategies for adaptation to new speakers or changing conditions.

Artificial intelligence plays important role in self-drivingcar explain.

• Objective: Develop a self-driving car system using Artificial Neural Networks (ANNs) for autonomous navigation.

• Data Collection: Gather diverse sensor data, including images, lidar, and radar, from various driving scenarios.

• Feature Extraction: Extract features using CNNs and RNNs for image and sequential data, respectively.

• Model Architecture: Design a suitable ANN architecture with convolutional, recurrent, and fully connected layers.

• Training and Tuning: Train the model on a diverse dataset, optimize hyperparameters, and validate performance regularly.

• Safety Measures: Implement emergency braking and collision avoidance mechanisms to prioritize safety.

• Deployment: Deploy the system initially in controlled environments, expanding gradually to more complex scenarios.

• Continuous Improvement: Monitor real-world performance, collect new data, and continuously improve the model's decision-making.

Machine Learning Technique

Unit - 5 [One Shot]

Most important topics:

1. Reinforcement learning, tasks & examples
2. Markov Decision Process
3. Q-Learning (functions, algorithm)
4. Deep Q-learning (introduction)
5. Genetic Algorithm

Correct learning

① Reinforcement Learning:

- It is a type of learning in ML where an agent learns to make decision by taking actions in an environment to receive feedbacks as reward or penalty.

* Key components:

- Agent: Learner or decision-maker.
- Environment: External system through which the agent interacts.
- State (S): A representation of current situation of agent.
- Action (A): Choices available to the agent.
- Reward (R): Immediate positive feedback from environment after an action.
- Penalty (P): Immediate negative feedback.
- Policy (π): Strategy that agent follows to determine its actions.

* Tasks in Reinforcement Learning:

- Exploration and exploitation: The agent must explore the environment to find the best actions and exploit known actions that provide high rewards.
- Learning from interaction: The agent learns by continuously interacting with the environment and updating its strategy based on

* Examples:

- Game playing: like chess is played by agents that learn strategies using Reinforcement Learning.
- Robotics: RL is used by robots to learn tasks like walking, navigating, or manipulating objects.

(ii) Markov Decision Process (MDP):

- It is a mathematical framework for modelling decision-making situations where outcomes are partly random and partly under the control of the agent.

* Key components:

- Set State (S): Set of all possible situations in which the agent can be.
- Actions (a): Set of all possible actions an agent can take
- Transition model (T): The probability of transitioning from one state to another due to action a.
- Reward function (R): The immediate reward received after transition.
- Policy (π): Defines agent's behaviour.

* MDP framework:

1. The agent starts in a state 's'.
2. It chooses an action 'a' based on it's policy 'π'.
3. It receives a reward R(s,a) and transitions to a new state 's'' according to a transition model T(s,a,s').
4. The process repeats, and the agent aims to maximize the cummulative reward over time.

Q-Learning:

- It is a popular, model-free Reinforcement learning algorithm that learns the quality of actions, denoted as Q-values, which represents the expected utility of taking an action in a given state and following the optimal policy.

- Q-Function (Q):

- The Q-value $Q(s,a)$ represents the expected cumulative reward of taking action 'a' in state 's' following the optimal policy.

- The goal is to learn the optimal Q-function $Q^*(s,a)$ which gives the highest expected reward for any state-action pair.

* Algorithm:

- Initialize the Q-table $Q(s,a)$.

- For each episode:
  - Initialize the starting state 's'.
  - Choose an action 'a' using an exploration strategy.
  - Take action 'a' and observe the reward 'r' and the next state 's'.
  - Update the Q-value using the Bellman equation:

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma \max_{a'} Q(s',a') - Q(s,a)]$$

where,

α: learning rate, control the amount of updation

γ: discount factor, determines importance of future rewards

  - Update the state s to the new state s'.

Deep Q-Learning :

- It is an extension of Q-Learning that uses a neural network to approximate the Q-values, making it feasible to handle large or continuous state spaces.

Introduction:

- Neural network: Instead of a Q-table, DQL uses a neural network to approximate the Q-values.
- Experience replay: To stabilize training, DQL stores the agents experiences in a replay buffer.
- Target Network: DQL uses two neural networks: the primary network Q and the target network Q'. It helps in stablizing the training by providing consistent targets during updates.

@brainlearning

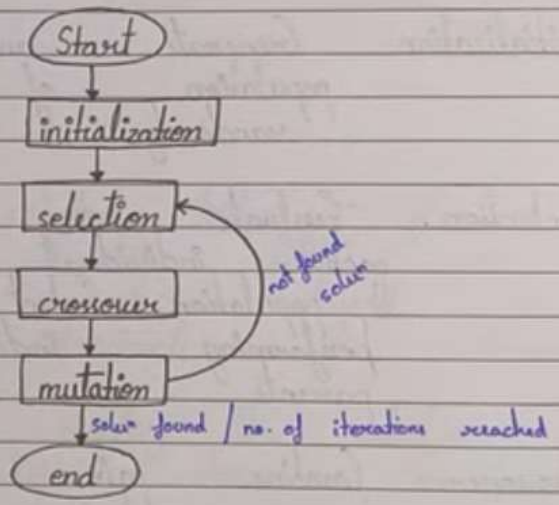
Genetic Algorithm.

- It is an optimization technique based on the principles of natural selection and genetics.

Steps:

1. Initialization: Generate an initial population of candidate randomly.
2. Selection: Evaluate the fitness of each individual in the population. Select best performing individuals to be parents.
3. Crossover: Combine pairs of parents to create new offspring.
4. Mutation: Introduce random changes to some offspring to maintain genetic diversity and explore new parts of the solution space.
5. Replacement: Replace the old population with the new generation of idividuals.

6. Iteration: Repeat the selection, crossover, mutation, and replacement steps for a no of generations or until a solution is found.



Thanks for watching !!

@bruvilearning