01-02-2021 (187) KANAV BANSAL

LECTURE - 40

MACHINE LEARNING:

It's an application of AI (Artificial Intelligence) that provides systems the ability to learn automatically and improve from experience without being explicitly Programmed.

→ It focuses on the development of computer programs that can access the data and use it to learn for themselves.

Example: Medical Diagnosis, Image Processing, Prediction, Classification, learning association, regression etc.

→ The Intelligent systems built on ML Algorithms have the capability to learn from past

experience or historical data.

TYPES OF MACHINE LEARNING:

1. Supervised learning (TASK DRIVEN)
2. Unsupervised learning (DATA DRIVEN)
3. Reinforcement learning (LEAN FROM ERRORS)

1. SUPERVISED LEARNING:

It is a machine learning task of learni-ng a function that maps an input to an output based on example input-output pairs.

→ A supervised learning algorithm analyzes the training data and produces an inferred function, which can be used for new mapping examples.

TYPES OF SUPERVISED LEARNING:

(i) REGRESSION: A single output value is produced using training data. This value is a probabilistic interpretation, which is ascertained after considering the strength of correlation among the input variables.

Example: It can help to predict the price, of a house based on locality, size, etc.

EXAMPLES OF SUPERVISED LEARNING ALGORITHMS:

(a) LINEAR REGRESSION: It is a supervised machine learning algorithm where the predicted output is continuous and has a constant slope.

→ It is used to predict values within a continuous range, rather than trying to

classify them to categories.

Example : sales, price → CONTINUE

Example: cat, dog → CATEGORY

(b) KNN (K-NEAREST NEIGHBORS): This algorithm is a simple, supervised machine learning algorithm that can be used to solve both classification and regression problems.

→ It's easy to Implement and understand, but it has a major drawback of becoming significantly slows as the size of that data in use grows.

(C) GUASSIAN NAIVE BAYES :

* Guassian Naive Bayes is a variant of Naive Bayes that follows Guassian Normal distribution and supports continuous data.

Naive Bayes are a group of supervised machine learning classification algorithms based on the Bayes theorem.

→ It is a simple classification technique, but has a high functionality.

(d) DECISION TREES : These are a type of supervised machine learning where the data is continuously split accordingly to a certain parameters.

→ The tree can be explained by two entities, namely decision nodes and leaves.

(e) SVM (SUPPORT VECTOR MACHINE): This is a supervised machine learning algorithm which can be used for both classification or regression challenges.

→ It is mostly used in classification Problems.

→ Support vectors are simply the co-ordinates of a individual observation

(f) RANDOM FOREST : It is a supervised machine learning algorithm which is used for both classification as well as regression.

→ This algorithm creates decision trees on data samples and then gets the predicti-on from each of them and finally selects the best solution by means of voting.

(iii) CLASSIFICATION: This is a supervised learning concept which basically categorizes a set of data into classes.

→ The most common classification problems are speech recognition, face detection, handwriting recognition, document classification, etc.

EXAMPLE:

(a) LOGISTIC REGRESSION: It is a classification algorithm in machine learning that uses one or more independent variables to determine an outcome.

→ The outcome is measured with a dichotomous variable which means it will have only two possible outcomes.

→ the goal is to find a best-fitting relationship between the dependent variable and a set of independent variables.

(b) NAIVE BAYES CLASSIFIER

(c) KNN

(d) DECISION TREE

(e) RANDOM FOREST

(f) ARTIFICIAL NEURAL NETWORKS: It consists of neurons that are arranged in layers, they take some input vector and convert it into an output.

→ the process of involving each neuron taking input and applying a function which is often a non-linear function to it and then passes the output to the next layer.

(g) STOCHASTIC GRADIENT DESCENT: It is a very effective and simple approach to fit linear models.

→ Particularly used & useful when the sample data is in a large number.

→ It supports different loss functions and penalties for classification.

(h) SVM

2. UNSUPERVISED LEARNING:

It is a machine learning technique in which the users don't need to supervi-se the model.

→ Instead, it allows the model to work on its own to discover patterns and inform-ation that was previously undetected.

TYPES OF UNSUPERVISED LEARNING:

(1) CLUSTERING: Also called as cluster Analysis is a unsupervised machine learning task.

→ It involves automatically discovering natural grouping in data.

→ Unlike supervised learning (like predective modelling), clustering algorithms only interpret the input data and find natural groups or clusters in feature space.

TYPES OF CLUSTERING:

(a) HIERARCHICAL CLUSTERING: It is another unsupervised learning algorithm that is used to group together the unlabeled data points having similar characteristics

→ The hierarchy of the clusters is represented as a dendrogram or tree structure.

→ It is of two types.

↳ AGGLOMERATIVE HIERARCHICAL

↳ DIVISE HIERARCHICAL.

(b) K-MEANS CLUSTERING: It is one of the simplest and popular unsupervised machine learning algorithm.

→ It identifies 'K' number of centroids, and then allocates every data point to the nearest cluster, while keeping the centroids as small as possible.

(c) K-NN (K-NEAREST NEIGHBOR) CLUSTERING: It performs much better if all the data have the same scale but it is not true for K-means.

(d) PRINCIPAL COMPONENT ANALYSIS (PCA):

It is a unsupervised machine learning technique that attempts to derive a set of low-dimensional set of features from a much larger set while still preserving as much variance as possible.

(e) SINGULAR-VALUE DECOMPOSITION (SVD): It is a matrix decomposition method for reducing a matrix to its constituent parts in order to make certain subsequent matrix simpler.

→ For the case of simplicity we will focus on SVD for real-valued matrices and ignore the case of complex numbers.

(f) INDEPENDENT COMPONENT ANALYSIS (ICA):

It is a machine learning technique to seperate independent sources from a mixed signal.

→ Unlike PCA, which focuses on maximizing the variance of the data points, the ICA focuses on independence i.e., independent components.

(g) DIMENSIONALITY REDUCTION: It refers to techniques that reduce the number of input variables in a dataset.

→ Large number of input features can cause poor performance for machine learning algorithms.

→ This is a general field of study concerned with reducing the number of

Input features.

WHY MATH FOR ML?

* Machine learning is built on mathematical Prerequisites.

* Math is important for solving the data Science project, Deep learning use cases.

* Math defines the underlying concept behind the algorithms and tells which one is better and why.

REASON OF WHY MATH FOR ML?

↳ Selecting the right algorithm which includes giving considerations to accuracy, training time, model complexity, number of parameters and number of features.

↳ Choosing parameter settings and validation strategies.

↳ Identifying underfitting, overfitting by understanding the Bias-variance tradeoff.

↳ Estimating the right confidence interval and uncertainity.

LINEAR ALGEBRA:

It is a sub-field of math concerned with vectors, matrices and linear transforms.

→ It is a key fourdation to the field of machine learning. from notations used to describe the operation of algorithms to the implementation of algorithms in code.

GEOMETRY:

It is a new field of machine learning that can learn from complex data like graphs and multi-dimensional points.

→ It seeks to apply traditional convolutional neural networks to 3D objects, graphs and manifolds.

→ 3D object classification using Point clouds.

VECTORS:

A vector is a tuple of one or more values called scalars.

→ These are built from components, which are ordinary numbers.

→ We can think of a vector as a list of numbers and vector algebra as operations Performed on the numbers in the list.

ROW VECTOR:

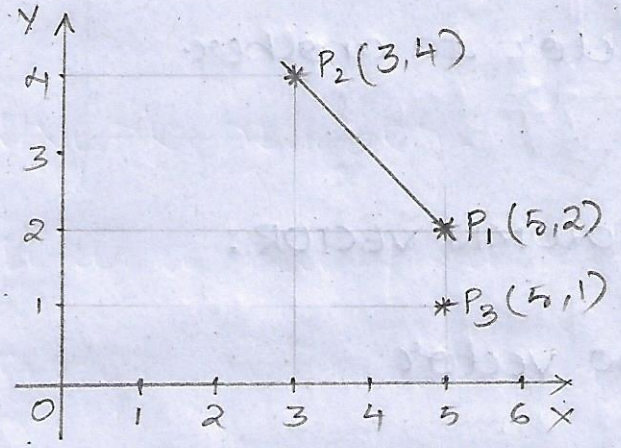
It is a "1xn matrix".

→ As it has 1 row and some number of columns.

COLUMN VECTOR:

It is a "nx1 matrix".

→ It always has 1 column and some number of rows.



X	Y
5	2
3	4
5	1

→ one row contains one data point.

→ Any point is represented as a "point on a coordinate plane which is a row in our data".

REPRESENTATION OF A VECTOR:

Vectors are usually represented by arrows with their "length representing the magnitude" and their "direction represented by

the direction the arrow points.

→ Vectors require both magnitude and direction.

→ The magnitude of a vector is a number for comparing one vector to another.

EXAMPLE OF ROW & COLUMN VECTOR:

$A = \begin{bmatrix} 3 & 4 \end{bmatrix}_{1\times 2}$ → Row vector

$B = \begin{bmatrix} 3 \\ 4 \end{bmatrix}_{2\times 1}$ → Colum vector.

NOTE

*→ We should always represent as a column vector.

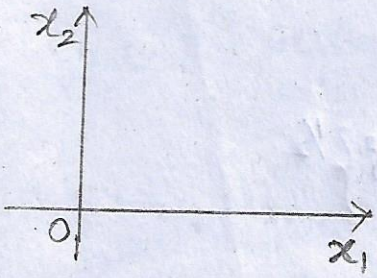
$\Rightarrow B^T = \begin{bmatrix} 3 \\ 4 \end{bmatrix}^T = \begin{bmatrix} 3 & 4 \end{bmatrix}$ → row vector.

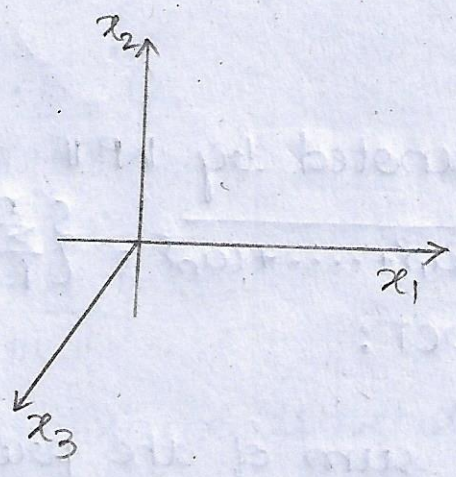
where 3 → x- component

4 → Y- component

* If there are more number of dimensions we name them as

$x_1, x_2 \ldots\ldots x_n$ respectively.





LENGTH OF A VECTOR:

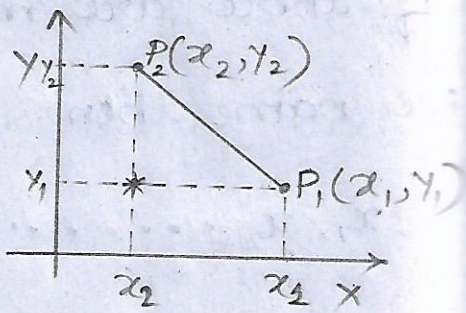
It is defined as "the square root of the sum of the squares of horizantal and vertical components".

→ If the horizantal ~~and~~ or vertical component is zero, then we don't need the vector length.

→ It is also called as "EUCLEDIAN DISTANCE" or "DISTANCE OF A POINT FROM ORIGIN".

2D $\Rightarrow d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

3D $\Rightarrow d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$



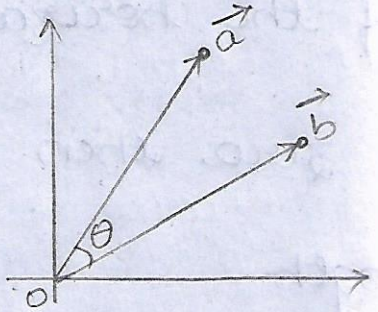
→ It is denoted by $\| P \|$

$\Rightarrow \|\vec{a}\| = \sqrt{(a_1)^2 + \ldots + (a_n)^2} = \left\{ \sum_{i=0}^{n} (a_i)^2 \right\}^{1/2}$

DOT PRODUCT :

It is the sum of the products of the corresponding entries of two sequence of numbers.

→ Geometrically, it is the product of the Euclidean magnitudes of the two vectors and the cosine of the angle between them.



$\vec{a} = [a_1, a_2, a_3, \ldots, a_d]$

$\vec{b} = [b_1, b_2, b_3, \ldots, b_d]$

$\Rightarrow \vec{a} \cdot \vec{b} = [a_1, a_2, \ldots, a_d] \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_d \end{bmatrix}$ = A single number which is a scalar type.

$\Rightarrow \vec{a} \cdot \vec{b} = \|\vec{a}\| \|\vec{b}\| \cos\theta_{a,b}$

where $\|\vec{a}\| \|\vec{b}\|$ can be stated as the length of $\vec{a}, \vec{b}$ respectively.

$$\theta_{a,b} = \cos^{-1}\left\{\frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|}\right\}$$

$$= \cos^{-1}\left\{\frac{\vec{a}^T \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|}\right\}$$

→ $\vec{a}$ & $\vec{b}$ should be in same dimensionality

NOTE:

Magnitude and direction are required in order to plot the vectors.

↳ We don't use the cross product in Machine learning.

the Angle between 2 points is

$$\cos\theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|}$$