# SNS COLLEGE OF ENGINEERING

Kurumbapalayam (Po), Coimbatore – 641 107

**An Autonomous Institution**

Accredited by NBA – AICTE and Accredited by NAAC – UGC with 'A' Grade
Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai

## DEPARTMENT OF COMPUTER SCIENCE AND DESIGN

## COURSE NAME : MACHINE LEARNING

III YEAR /V SEMESTER

Unit 1 - INTRODUCTION

Topic 4 : Unsupervised Learning Categories - Advantages and Disadvantages of Unsupervised Learning- Applications of Unsupervised Learning

# Unsupervised Machine Learning

- Unsupervised learning is different from the Supervised learning technique; as its name suggests, there is no need for supervision.
- It means, in unsupervised machine learning, the machine is trained using the unlabelled dataset, and the machine predicts the output without any supervision.
- In unsupervised learning, the models are trained with the data that is neither classified nor labelled, and the model acts on that data without any supervision.

# Unsupervised Learning Categories

- The main aim of the unsupervised learning algorithm is to group or categories the unsorted dataset according to the similarities, patterns, and differences.
- Machines are instructed to find the hidden patterns from the input dataset.

# Unsupervised Learning Categories

**Example**

- Suppose there is a basket of fruit images, and we input it into the machine learning model.
- The images are totally unknown to the model, and the task of the machine is to find the patterns and categories of the objects.
- So, now the machine will discover its patterns and differences, such as colour difference, shape difference, and predict the output when it is tested with the test dataset.

# Unsupervised Learning Categories

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

1. Clustering
2. Association

# Unsupervised Learning Categories

## 1) Clustering:

The clustering technique is used when we want to find the inherent groups from the data. It is a way to group the objects into a cluster such that the objects with the most similarities remain in one group and have fewer or no similarities with the objects of other groups.

An **example** of the clustering algorithm is grouping the customers by their purchasing behaviour.

# Unsupervised Learning Categories

## 2) Association:

Association rule learning is an unsupervised learning technique, which finds interesting relations among variables within a large dataset.

The main aim of this learning algorithm is to find the dependency of one data item on another data item and map those variables accordingly so that it can generate maximum profit. This algorithm is mainly applied in Market Basket analysis, Web usage mining, continuous production, etc.Some popular algorithms of Association rule learning are Apriori Algorithm, Eclat, FP- growth algorithm.

# Unsupervised Learning

**Advantages of Unsupervised Learning:**

- These algorithms can be used for complicated tasks compared to the supervised ones because these algorithms work on the unlabelled dataset.
- Unsupervised algorithms are preferable for various tasks as getting the unlabelled dataset is easier as compared to the labelled dataset.

# Unsupervised Learning

**Disadvantages of Unsupervised Learning:**

- The output of an unsupervised algorithm can be less accurate as the dataset is not labelled, and algorithms are not trained with the exact output in prior.
- Working with Unsupervised learning is more difficult as it works with the unlabelled dataset that does not map with the output.

# Applications of Unsupervised Learning:

- **Network Analysis:** Unsupervised learning is used for identifying plagiarism and copyright in document network analysis of text data for scholarly articles.
- **Recommendation Systems:** Recommendation systems widely use unsupervised learning techniques for building recommendation applications for different web applications and e-commerce websites.

# Applications of Unsupervised Learning:

- **Anomaly Detection:** Anomaly detection is a popular application of unsupervised learning, which can identify unusual data points within the dataset. It is used to discover fraudulent transactions.
- **Singular Value Decomposition:** Singular Value Decomposition or SVD is used to extract particular information from the database. For example, extracting information of each user located at a particular location.

# Thank you