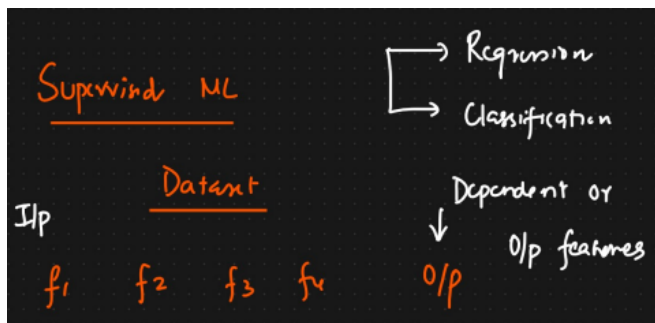**Supervised Learning**

- **Analogy:** Think of it like learning with a teacher or a supervisor. You are given examples with correct answers (labels), and you learn to predict the answers for new, unseen examples.
- **Data:** Uses **labeled data**. This means for each data point in the training set, there's a known outcome or "correct answer" associated with it.
    - Example: Emails labeled as "spam" or "not spam"; houses with their known sale prices.
- **Types of Problems:**
    - **Classification:** The output label is a category (discrete value).
        - *Examples:* Spam detection (spam/not spam), image recognition (cat/dog/car), sentiment analysis (positive/negative/neutral), predicting customer churn (yes/no).
    - **Regression:** The output label is a continuous value.
        - *Examples:* Predicting house prices, forecasting stock prices, predicting temperature, estimating student scores.
- **Common Algorithms:** Linear Regression, Logistic Regression, Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), Decision Trees, Random Forests, Naive[1] Bayes, Neural Networks (often used for complex supervised tasks).
- **Feedback Mechanism:** During training, the algorithm makes predictions, compares them to the actual labels, and adjusts its internal parameters to minimize the errors.
- 



**2. Unsupervised Learning**

- **Analogy:** Think of it like learning without a teacher. You are given a lot of data and asked to find interesting patterns, structures, or relationships within it on your own, without any predefined answers.
- **Data:** Uses **unlabeled data**. The training data consists only of input data points (X) without any corresponding output labels (Y).
- **Goal:** To model the underlying structure or distribution in the data to learn more about it. It's about discovering hidden patterns, groupings, or anomalies.
- **Types of Problems:**
    - **Clustering:** Grouping similar data points together based on their features. The algorithm discovers these groups (clusters).

- - - *Examples:* Customer segmentation (grouping customers with similar purchasing behavior), grouping similar documents, image segmentation.
  - ○ **Association Rule Learning:** Discovering rules that describe large portions of your data, like "people who buy X also tend to buy Y".
    - - *Examples:* Market basket analysis (what items are frequently bought together in a supermarket), recommendation systems.
  - ○ **Dimensionality Reduction:** Reducing the number of variables (features) while preserving important information.
    - - *Examples:* Feature extraction, data compression, noise reduction.
  - ○ **Anomaly Detection:** Identifying rare items, events, or observations which differ significantly from the majority of the data.
    - - *Examples:* Fraud detection, identifying defective products, network intrusion detection.
- **Common Algorithms:** K-Means Clustering, Hierarchical Clustering, DBSCAN, Principal Component Analysis (PCA), Singular Value Decomposition (SVD), Apriori Algorithm, Isolation Forest.
- **Feedback Mechanism:** There are no correct answers to compare against. Instead, algorithms often work by minimizing/maximizing an internal objective function (e.g., minimizing distance within clusters in K-Means) or identifying statistical patterns.
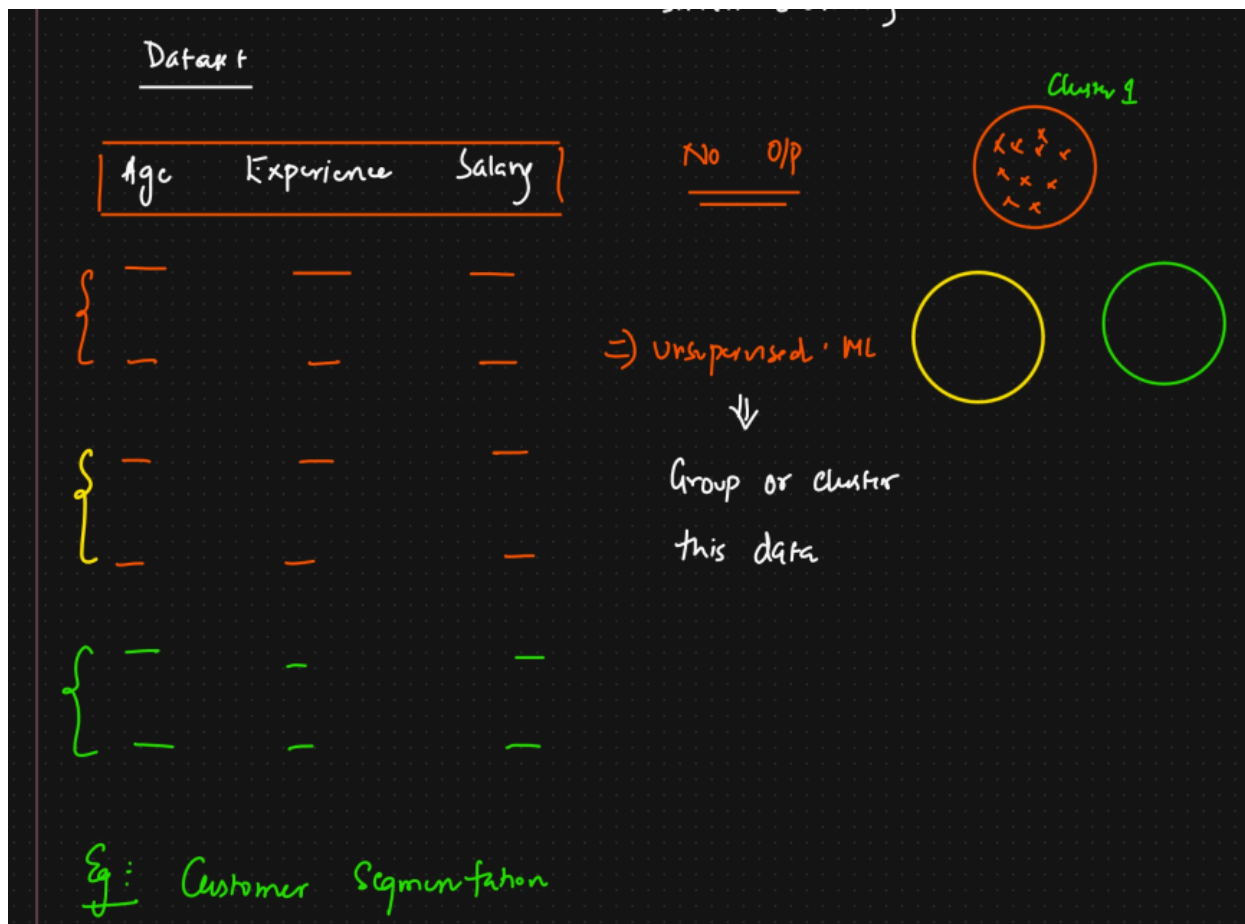
# Unsupervised ML

1. K Means Algorithm
2. Hierarchical Clustering.
3. DBScan Clustering
4. Silhoutte Scoring

Validate