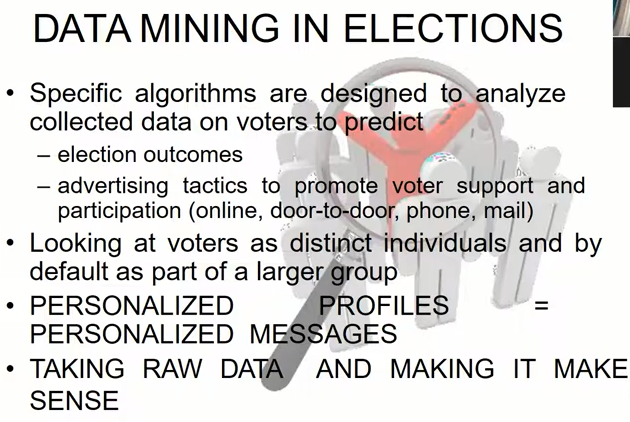
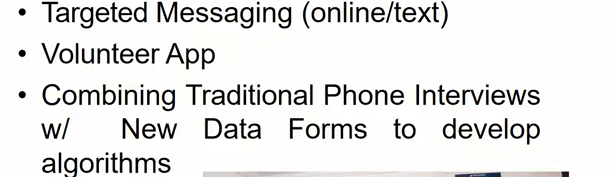
**Prediction and elections:**

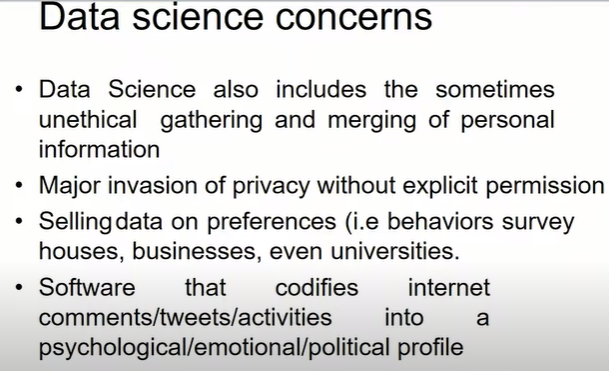












*Clustering in Machine Learning*

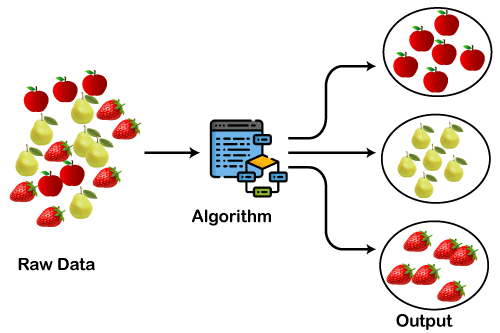
1. Clustering or cluster analysis is a machine learning technique, which groups the **unlabelled dataset.**
2. It can be defined as **"A way of grouping the data points into different clusters, consisting of similar data points.**
3. The objects with the possible similarities remain in a group that has less or no similarities with another group."
4. It is unsupervised Ml.
5. After applying this clustering technique, each cluster or group is provided with a cluster-ID.
6. ML system can use this id to simplify the processing of large and complex datasets.

Examples:

any shopping mall --🡪t-shirts are grouped in one section, and trousers are at other sections

Apart from these general usages, it is used by the **Amazon** in its recommendation system to provide the recommendations as per the past search of products.

**Netflix** also uses this technique to recommend the movies and web-series to its users as per the watch history.



## Types of Clustering Methods

1. **Hard clustering**
2. Soft clustering
3. Partitioning Clustering
4. Hierarchical Clustering

**Hard clustering** (datapoint belongs to only one group)

**Soft Clustering** (data points can belong to another group also).

### Partitioning Clustering

It is a type of clustering that divides the data into **non-hierarchical groups.**

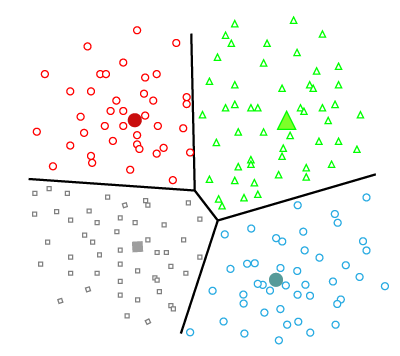
It is also known as the **centroid-based method**.

The most common example of partitioning clustering is the **[K-Means Clustering algorithm](https://www.javatpoint.com/k-means-clustering-algorithm-in-machine-learning)**

.

In this type, the dataset is divided into a set of k groups, where K is used to define the number of pre-defined groups.

The cluster center is created in such a way that the distance between the data points of one cluster is minimum as compared to another cluster centroid.



### Hierarchical Clustering

Hierarchical clustering can be used as an alternative for the partitioned clustering as there is no requirement of pre-specifying the number of clusters to be created.

In this technique, the dataset is divided into clusters to create a tree-like structure, which is also called a **dendrogram**

