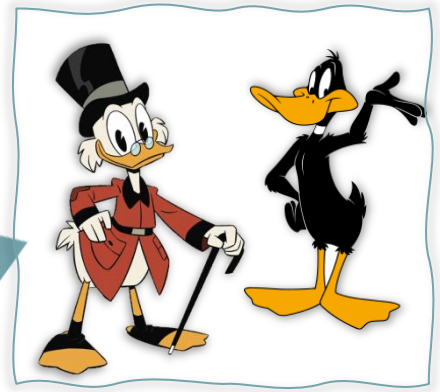
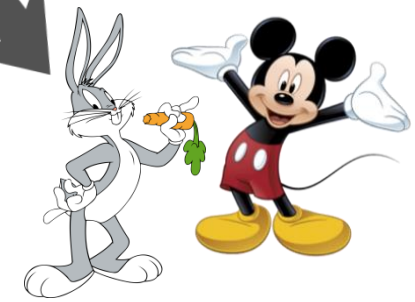
Introduction to Unsupervised Learning

Unsupervised Learning

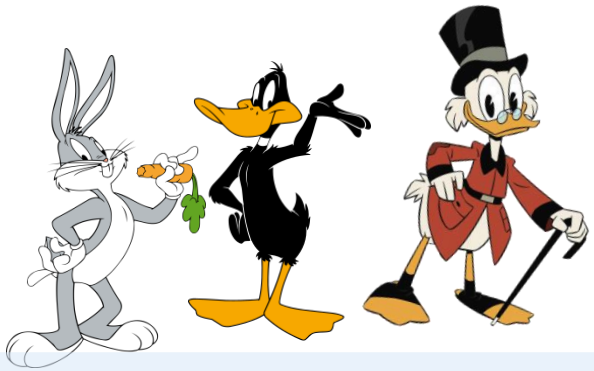
Unsupervised Learning

- Contrast to Supervised Learning only input is provided
- Agent should find patterns & divide the data into clusters
- Clusters will be formed based upon the common features in each datapoint





Individual Data Points



Unsupervised Learning

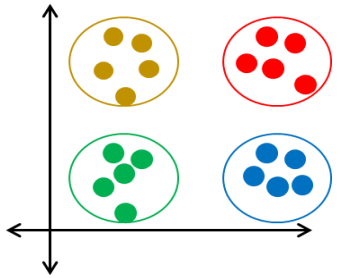
Train an algorithm to find clusters and associations in unlabelled dataset

Introduction to Unsupervised Learning

Types of
Unsupervised Learning

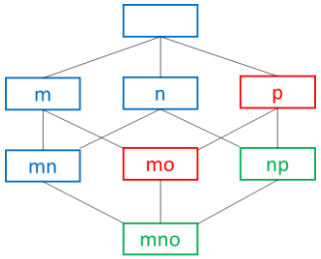
Clustering

1



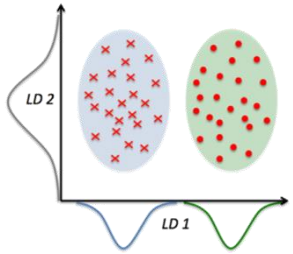
Association

2



Dimension Reduction

3



Types of Unsupervised Learning

Introduction to Unsupervised Learning

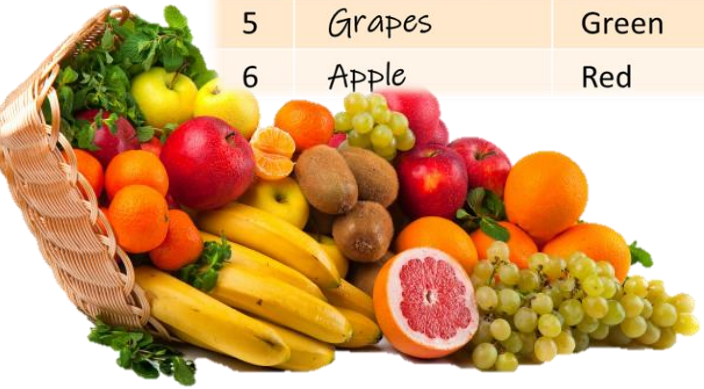
Application of Unsupervised Learning

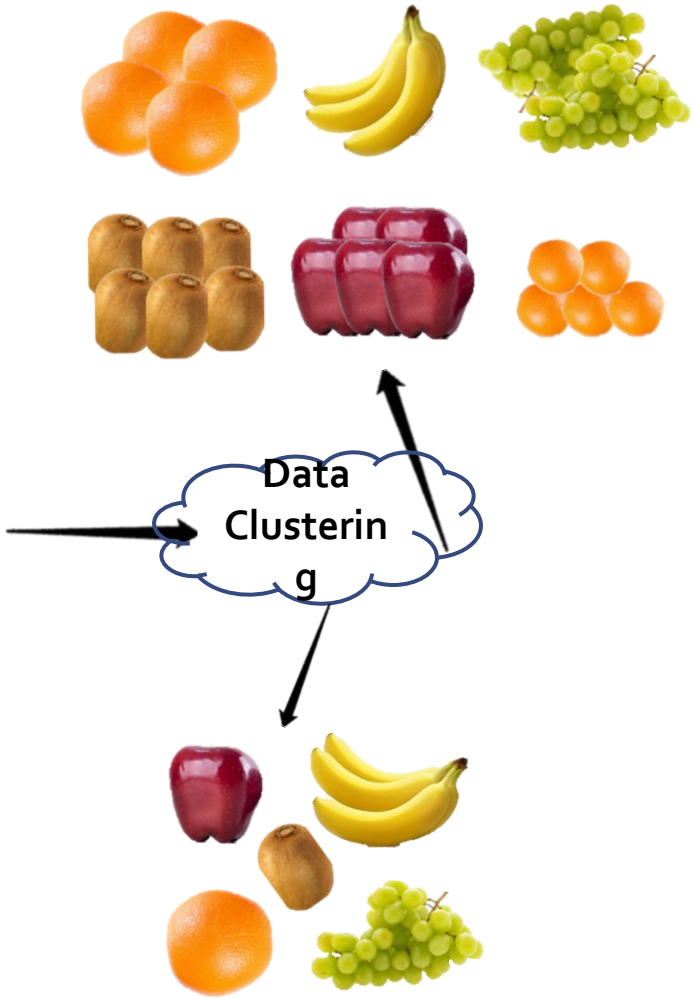
Applications of Unsupervised Learning

- Item categorization
- Clustering customers
- Anomaly Detection
- Recommendation Engines

Data Matrix

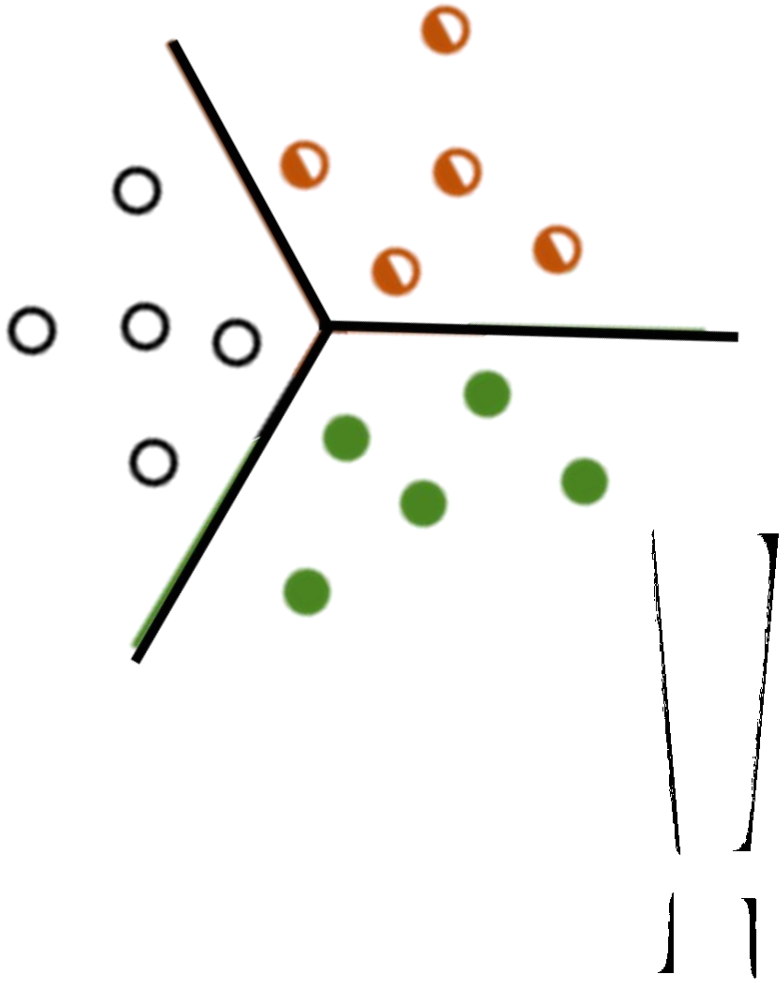
Item	Feature Type	Feature Color	Feature Weight (Ounces)
1	Banana	Yellow	5.78
2	Big Orange	Orange	4.32
3	Kiwi	Brownish	5.54
4	Baby Orange	Orange	3.65
5	Grapes	Green	1.15
6	Apple	Red	4.32





Unsupervised learning
“Clustering”

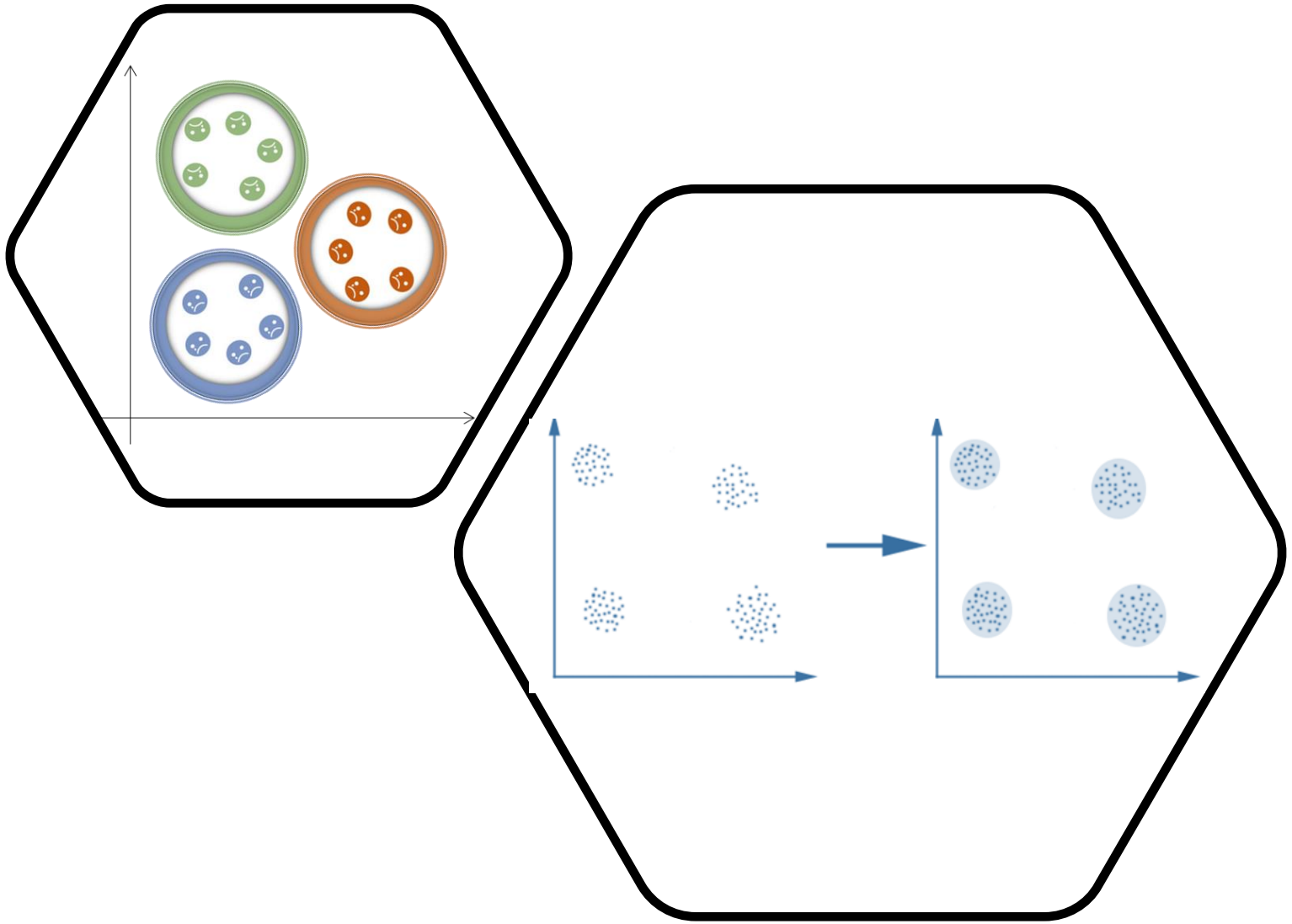
Introduction to Clustering



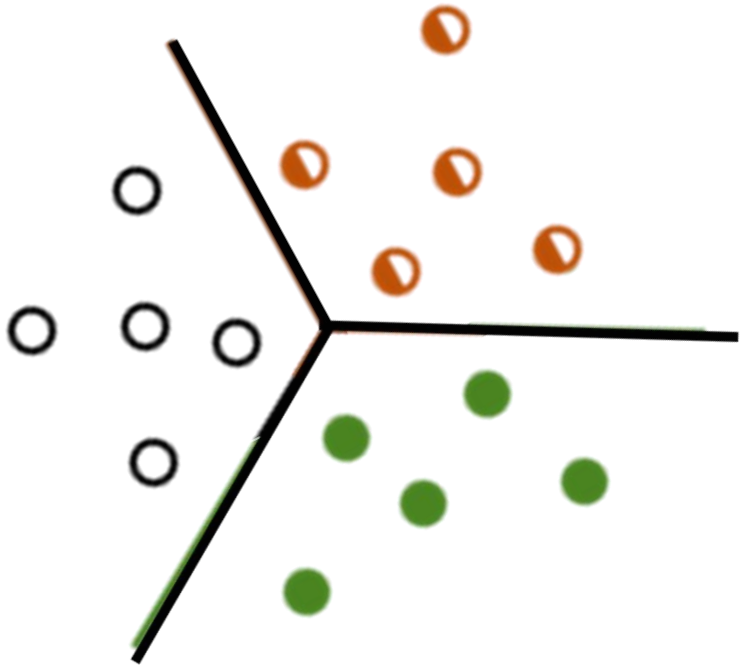
Why Clustering ?

- Determines intrinsic grouping among unlabelled data
- No criteria
- Reducing dimensions

- Organizing the objects
- Creating groups
- Collection of Objects



Introduction
to
Clustering



Challenges in Clustering

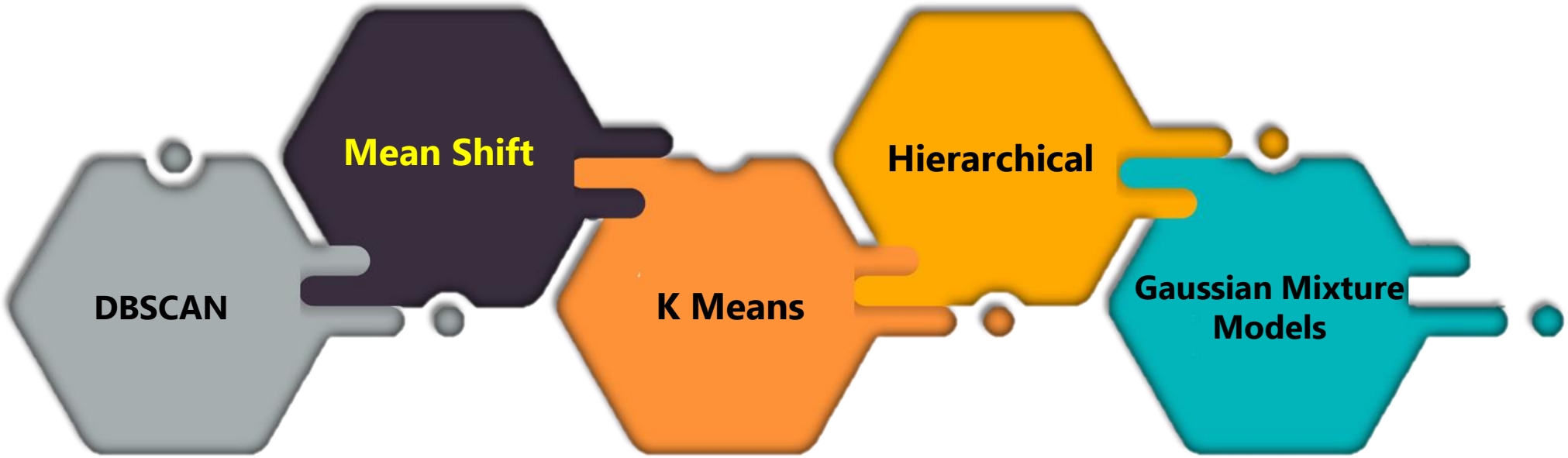
- Computational Complexity
- Risk of Inaccurate results
- Human intervention for validating outputs

Unsupervised learning
“Clustering”

Types of Clustering Algorithms

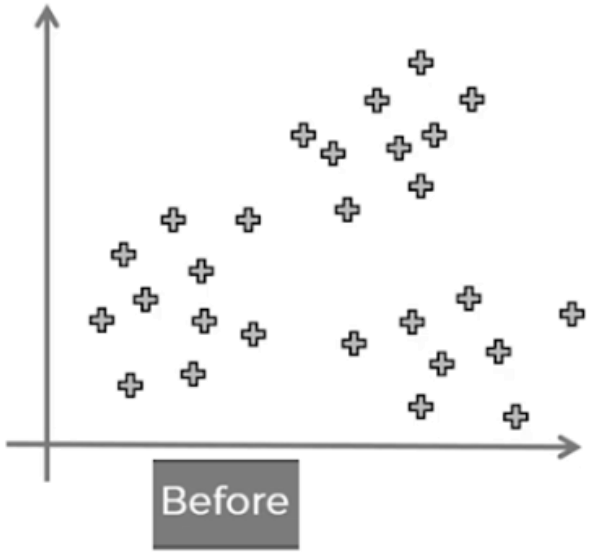
Algorithm Types

Clustering

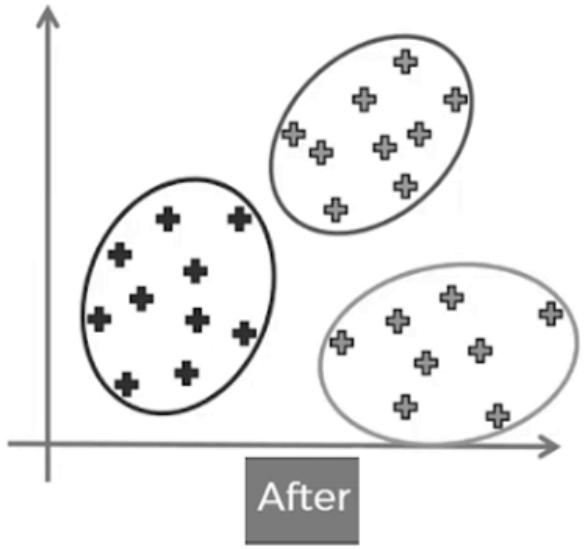


Unsupervised learning – “Clustering”

What is K-Means Clustering?



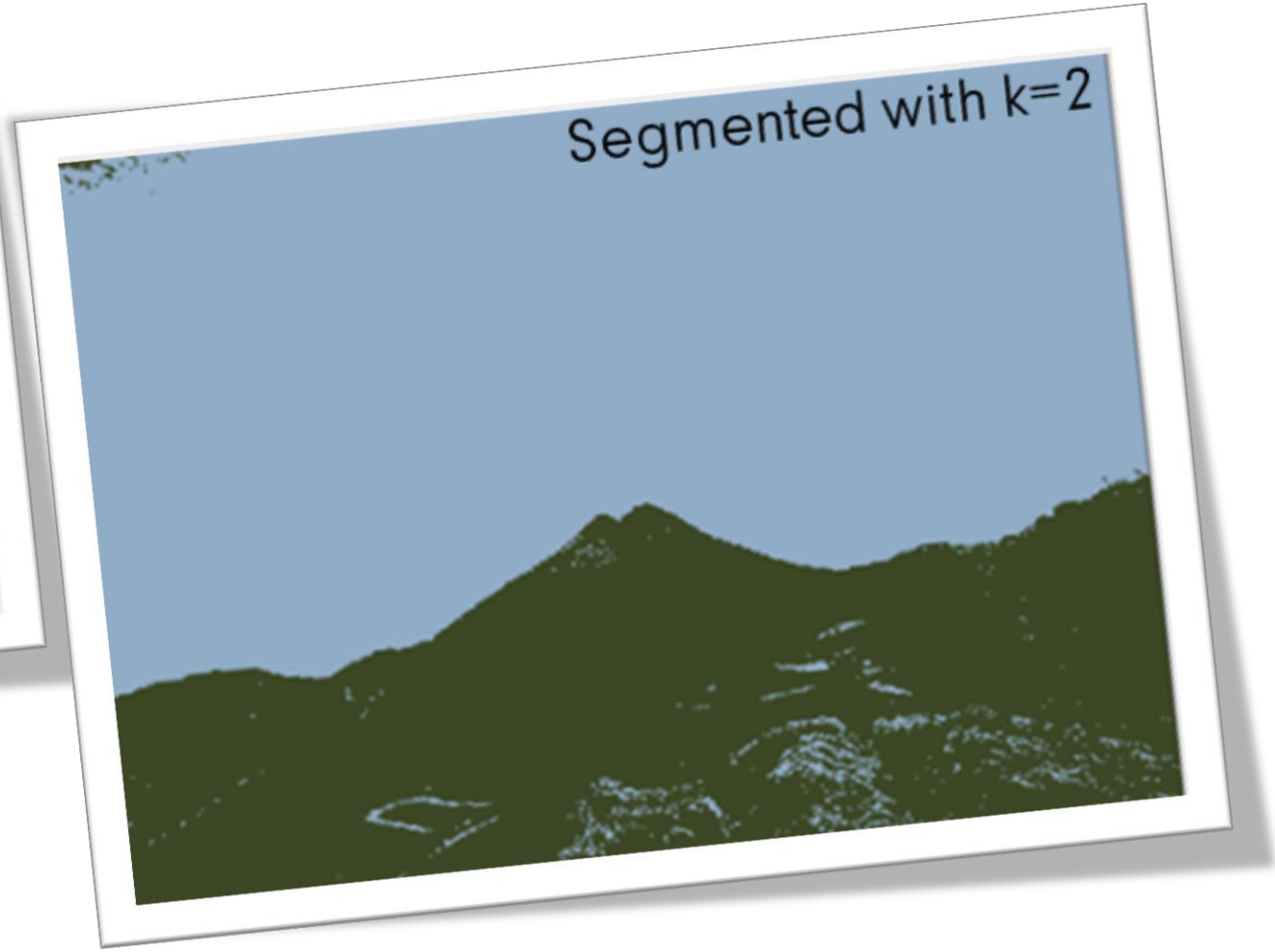
K-
Means



What is K-Means Clustering ?

- The organization of unlabelled data into groups with similarities
- Partitional Clustering
- Partitions the data into k clusters
- A part of Vector Quantization

K-Means Clustering in Image Segmentation



K Means Clustering

S_1

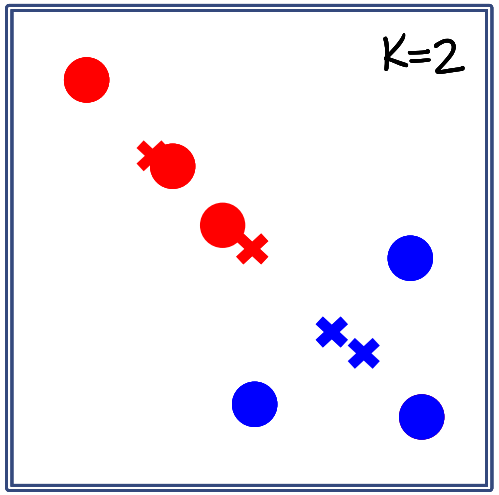
S_2

S_3

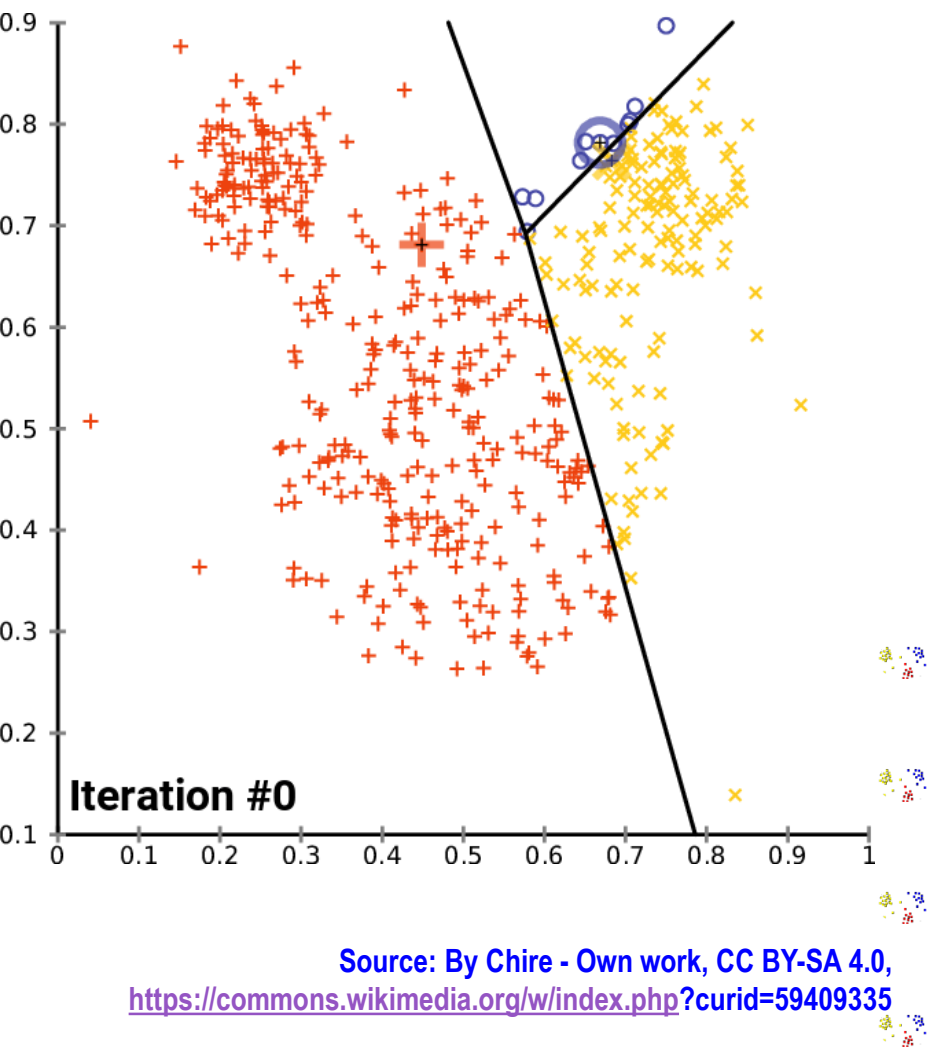
S_4

S_5

S_6



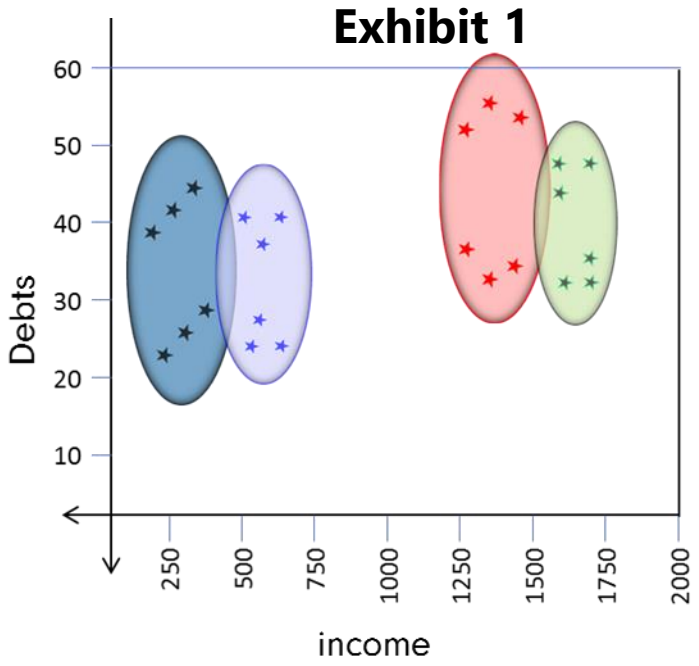
- Number of clusters K
- Randomly assigning data point to cluster
- Compute cluster centroids
- Re-assign each point to closest cluster centroid
- Re-compute cluster centroids
- Repeat step 4 & 5 till global optima is reached

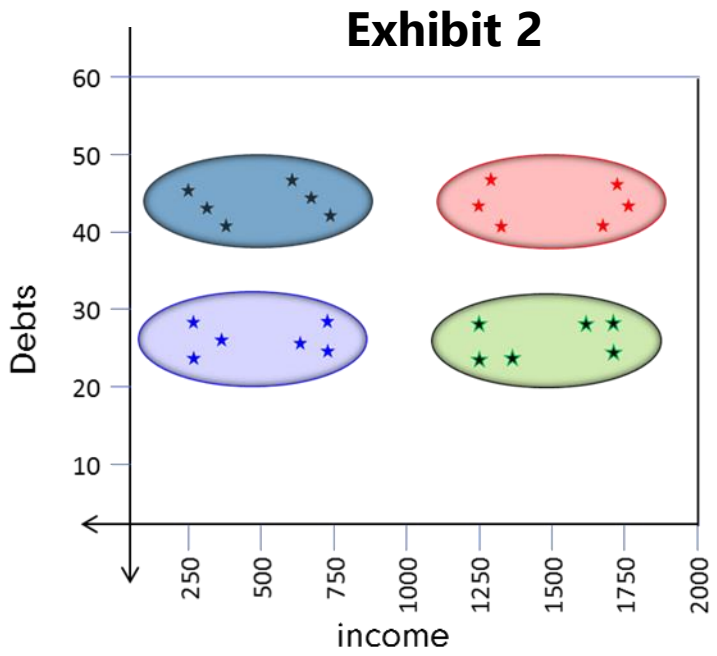


K-Means Clustering

Convergence

- Point where the algorithm stop it's process
- No or very minimum reassignments of data points to clusters
- No change in centroids
- Distance of datapoints from centroid is minimum
- Silhouette score





Properties of Clusters

Euclidean Distance

$$d(x, y) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Property 1

All data points in a cluster should be similar to each other

Property 2

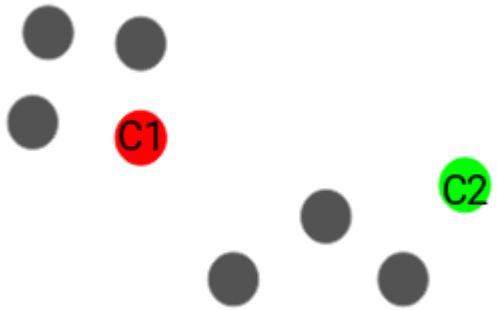
The data points from different clusters should be as different as possible

K Mean - Example

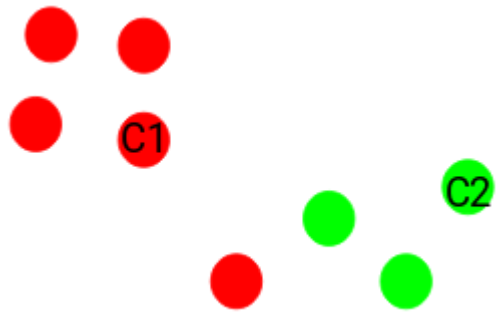
DATA POINTS

Step 1: Select the number of clusters k (For Ex K=2)

Step 2: Select k random points from the data as centroids

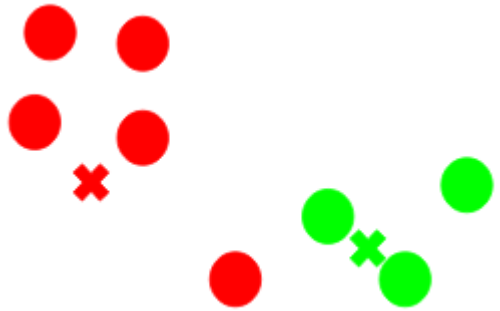


Steps 3: Assign all the points to the closest cluster centroid



Steps 4: Re-assign each point to closest cluster centroid

Steps 5: Re-compute cluster centroids



Step 6: Repeat step 4 & 5 till global optima is reached

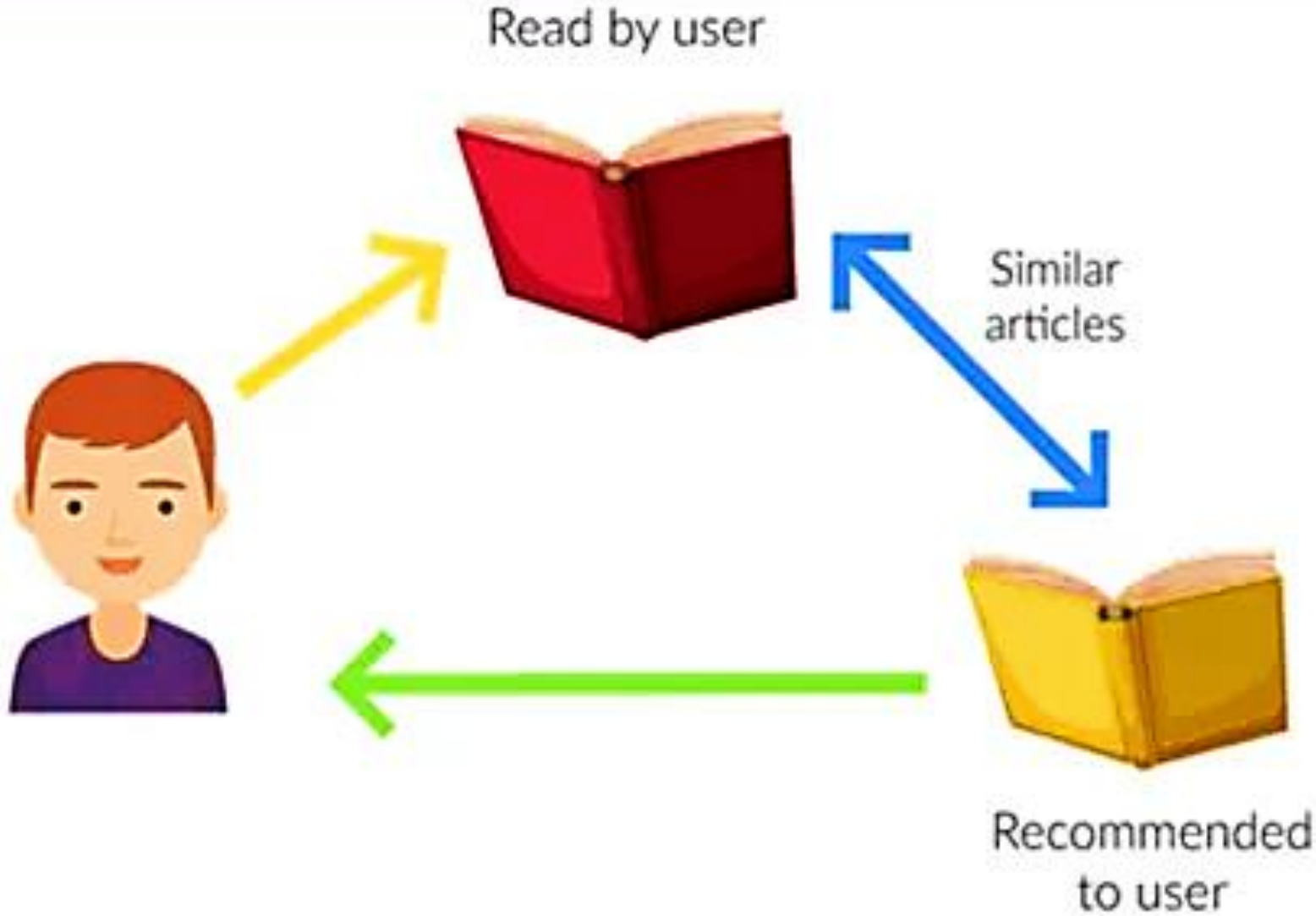
There are essentially three stopping criteria that can be adopted to stop the K-means algorithm:

- **Centroids of newly formed clusters do not change**
- **Points remain in the same cluster**
- **Maximum number of iterations is reached**

Recommender Systems

What is Recommender Systems?

- Recommender systems, also known as recommendation systems
- It is a machine learning algorithms that use data to recommend items or content to users
- based on their preferences, past behavior, or their combination.


Read by user
Similar
articles
Recommended
to user

How does a recommendation engine work?

- We can recommend items to a user which are most popular among all the users

- We can divide the users into multiple segments based on their preferences (user features) and recommend items to them based on the segment they belong to

How does a recommendation engine work?

- Making recommendations involves two main stages:
  - Candidate generation—creating a subset of products the user might be interested in.
  - Scoring—reducing and sorting a candidate list.

How To Work Recommendation System?

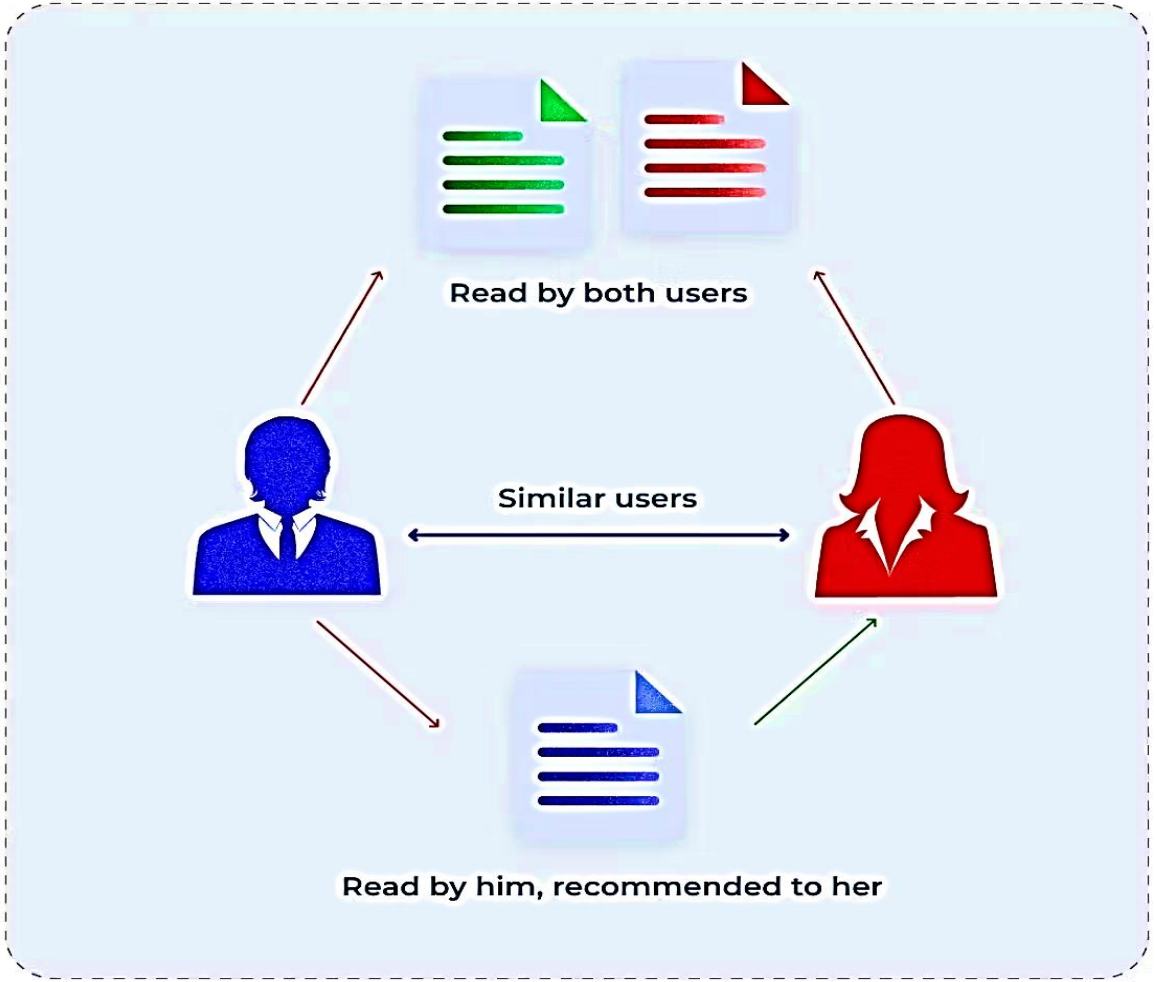
- Data collection
- Data storage
- Analyzing the data
- Filtering and recommending

Types of Recommendation System

- Collaborative filtering
- Content-based filtering
- Hybrid recommendation systems.

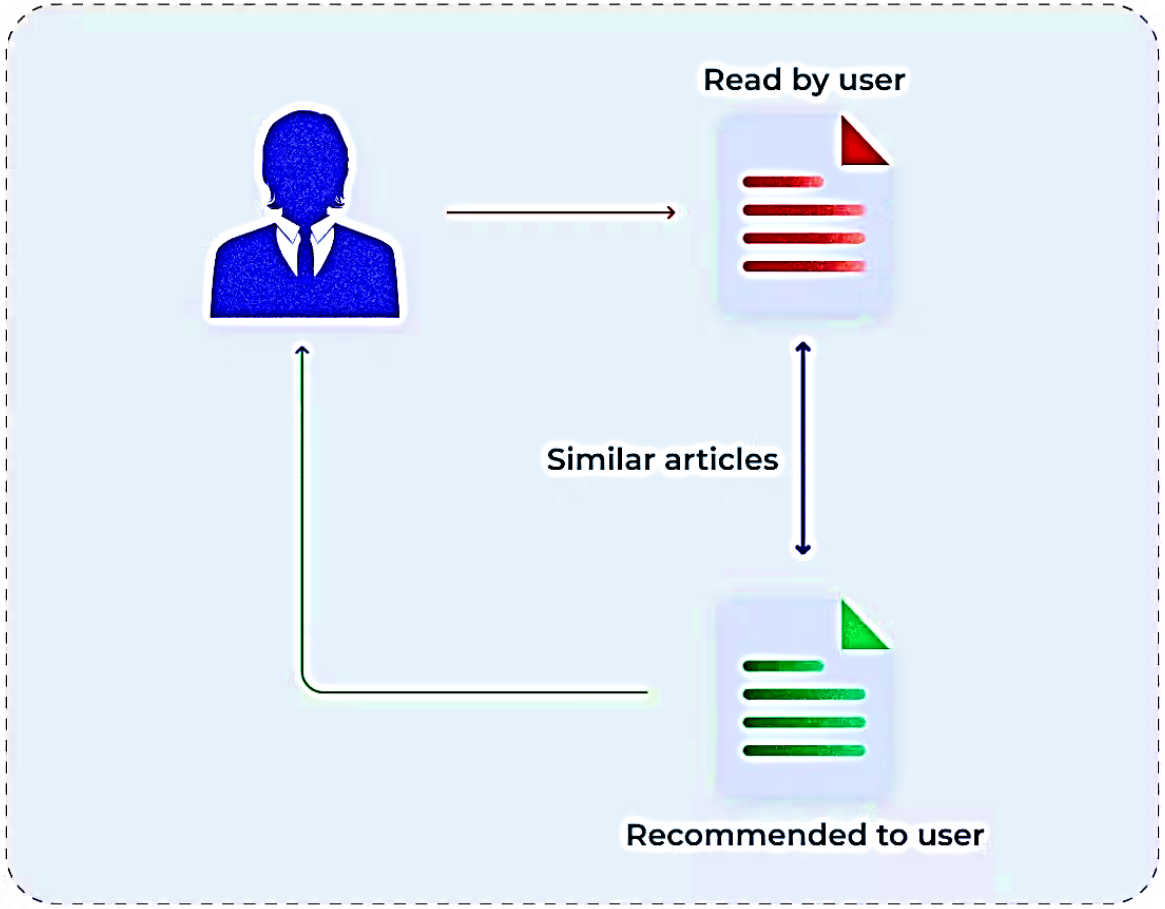
Collaborative filtering

Collaborative Filtering

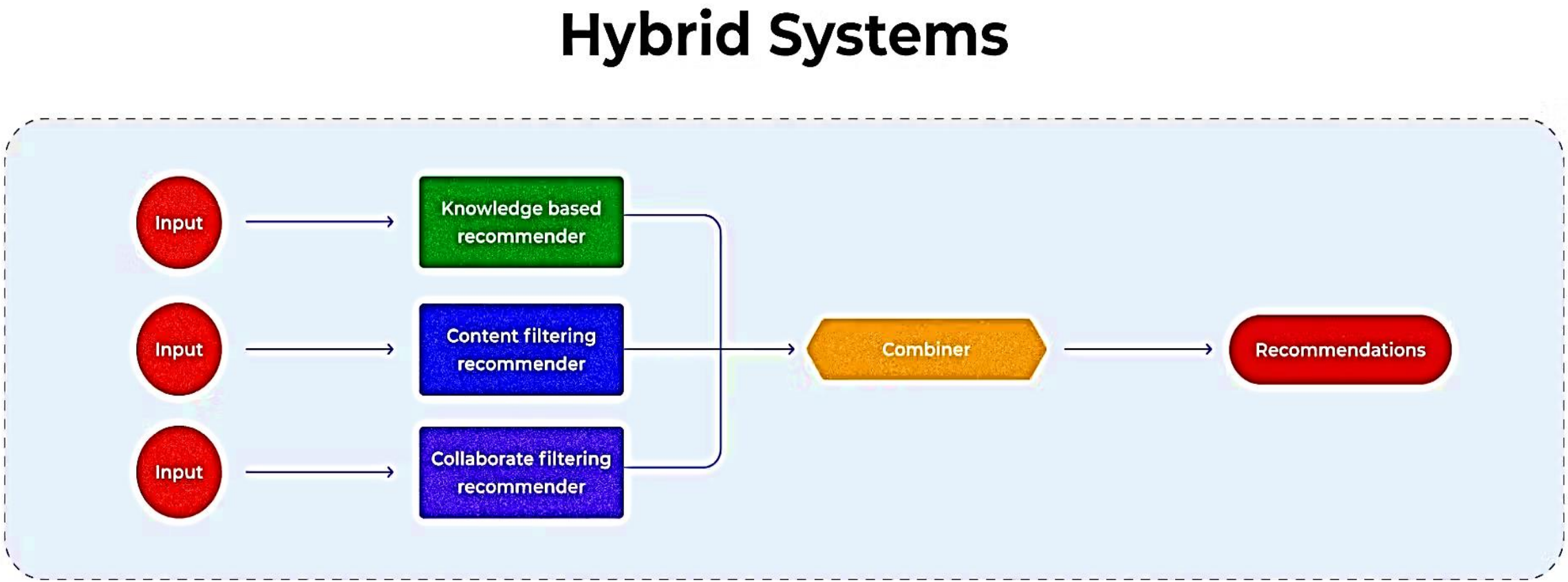


Content-based filtering

Content-Based Filtering



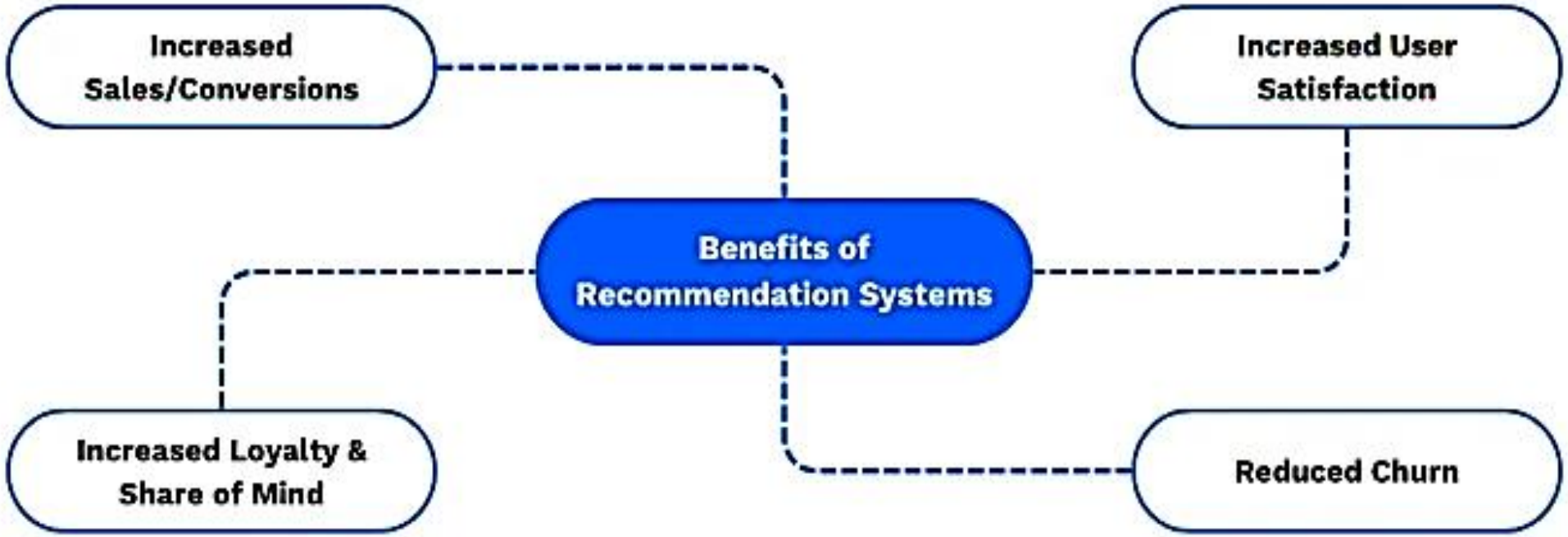
Hybrid systems



Algorithms for Recommender systems

- Clustering models,
- User-based k-nearest neighbors,
- Matrix factorization, and
- Bayesian networks

Benefits of Recommender Systems



AIMultiple

Applications of Recommender systems

- E-Commerce
- Retail
- Media
- Banking
- Telecom
- Utilities

Examples from companies that use a recommendation engine

- Amazon.com
- Netflix
- Spotify
- Linkedin

THANK YOU