Berkeley Engineering | BerkeleyHaas

PROFESSIONAL CERTIFICATE IN MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

Office Hour #6 with Matilde D'Amelio

April 21, 2022 at 9 pm UTC

TYPES of AI LEARNING METHODS

SUPERVISED LEARNING

If you're learning a task under supervision, someone is present judging whether you're getting the right answer. Similarly, in supervised learning, that means having a full set of labeled data while training an algorithm.

Fully labeled means that each example in the training dataset is tagged with the answer the algorithm should come up with on its own.

UNSUPERVISED LEARNING

Clean, perfectly labeled datasets aren't easy to come by. And sometimes, researchers are asking the algorithm questions they don't know the answer to. That's where unsupervised learning comes in.

In unsupervised learning, a deep learning model is handed a dataset without explicit instructions on what to do with it. The training dataset is a collection of examples without a specific desired outcome or correct answer.

REINFORCEMENT LEARNING

It is neither based on supervised learning nor unsupervised learning. Moreover, here the algorithms learn to react to an environment on their own. It is rapidly growing and moreover producing a variety of learning algorithms. These algorithms are useful in the field of Robotics, Gaming etc.

Source: https://www.aitude.com/supervised-vs-unsupervised-vs-reinforcement/

Unsupervised Learning Applications

CLUSTERING

Clustering is a data mining technique which groups unlabeled data based on their similarities or differences. Clustering algorithms are used to process raw, unclassified data objects into groups represented by structures or patterns in the information.

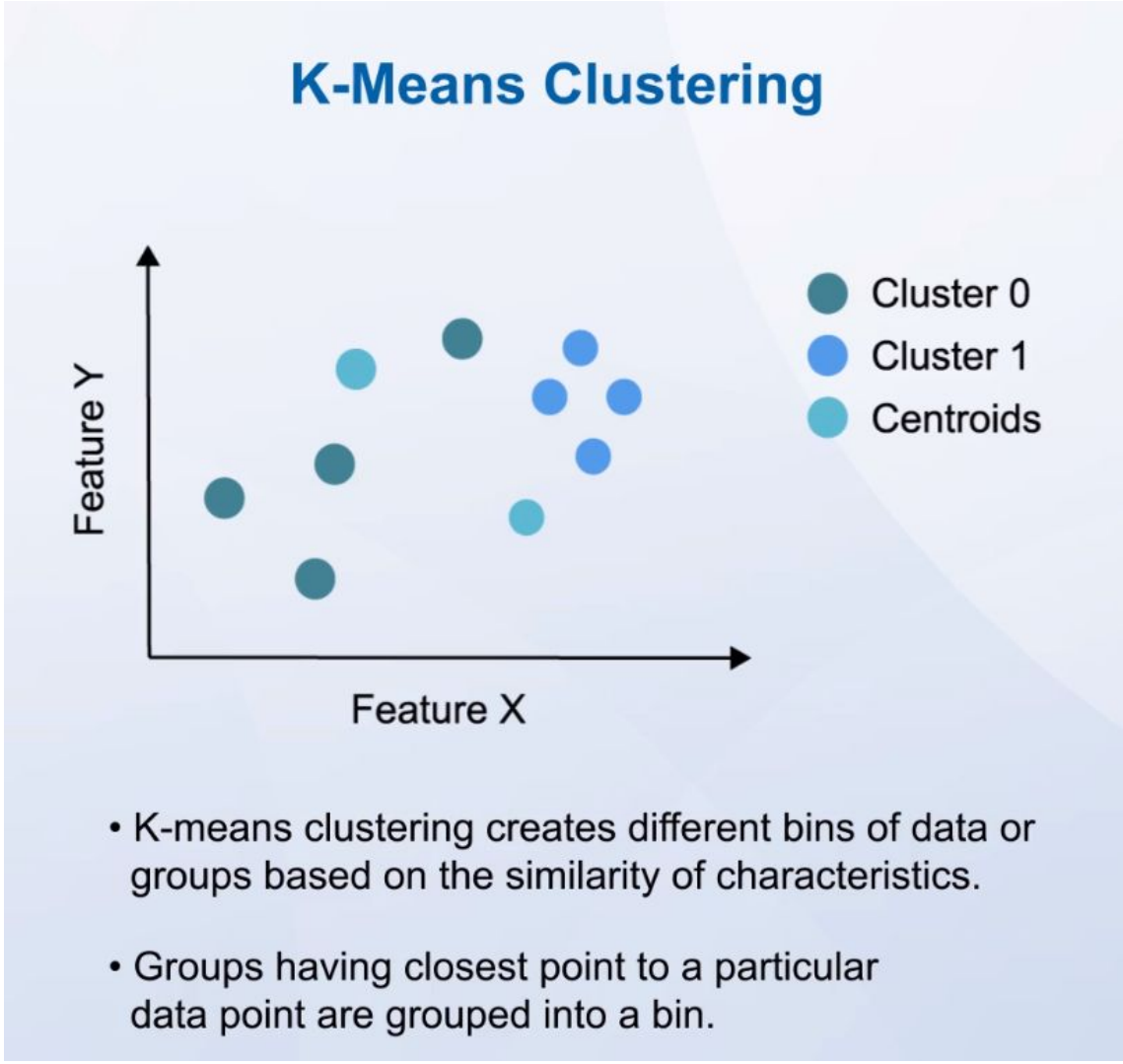
K-means
Gaussian MIxture Model

DIMENSIONAL REDUCTION

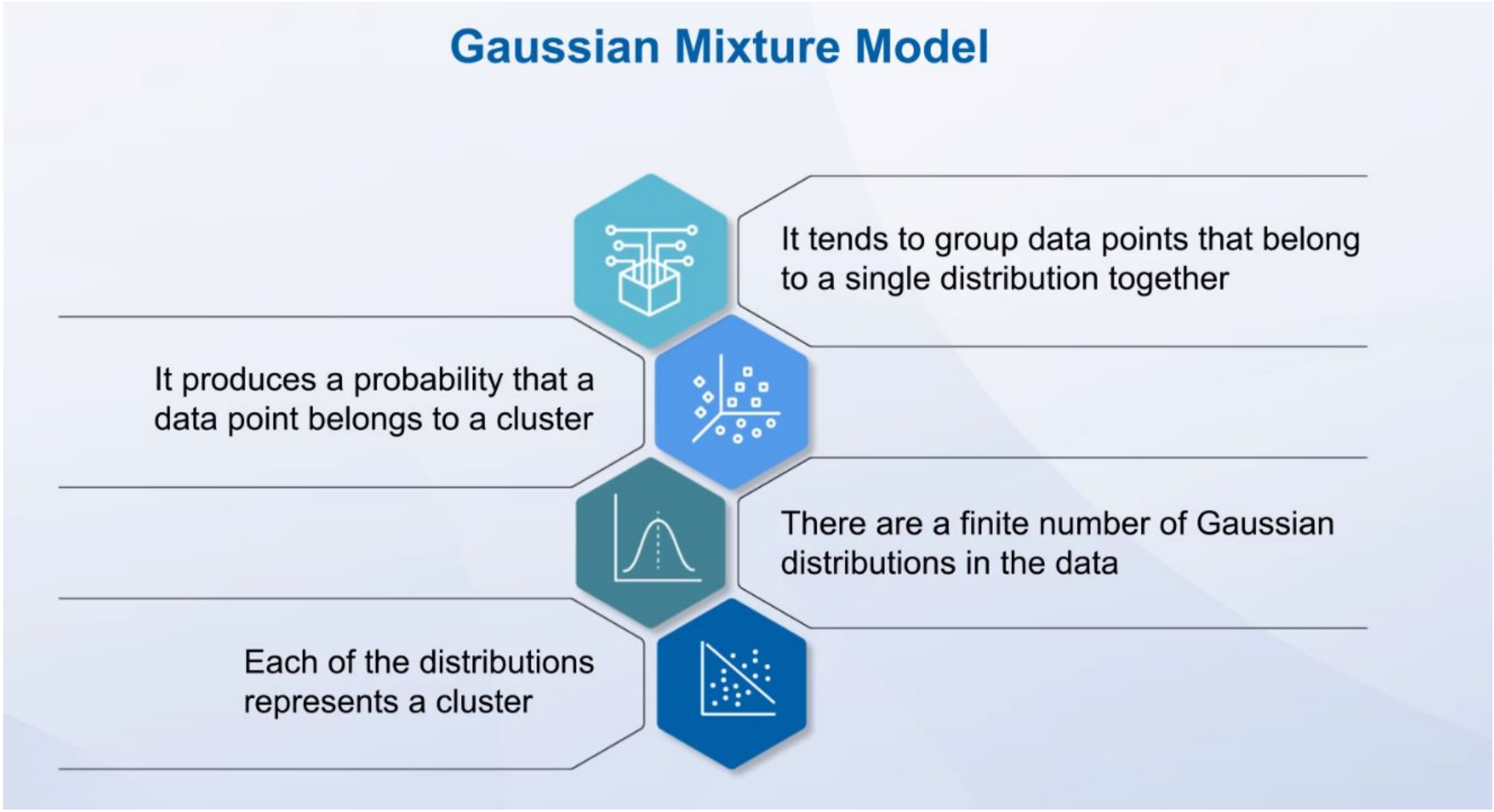
While more data generally yields more accurate results, it can also impact the performance of machine learning algorithms (e.g. overfitting) and it can also make it difficult to visualize datasets. Dimensionality reduction is a technique used when the number of features, or dimensions, in a given dataset is too high. It reduces the number of data inputs to a manageable size while also preserving the integrity of the dataset as much as possible. It is commonly used in the preprocessing data stage, and there are a few different dimensionality reduction methods that can be used

Principal Component Analysis
Singular Value Decomposition

K-Means

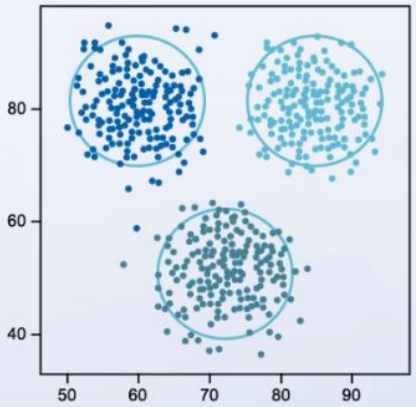

K-Means Clustering
Feature Y
Feature X
Cluster 0
Cluster 1
Centroids
• K-means clustering creates different bins of data or groups based on the similarity of characteristics.
• Groups having closest point to a particular data point are grouped into a bin.

Gaussian Mixture Model

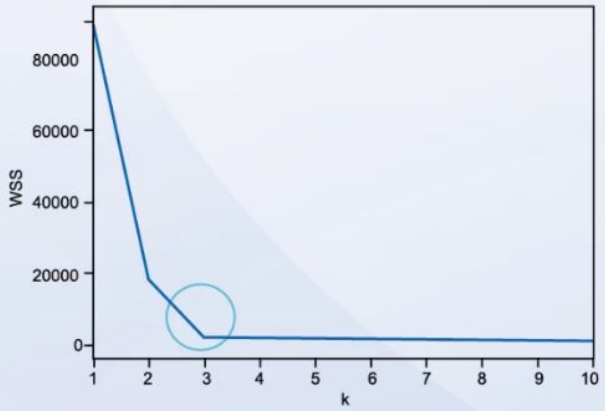


Example

Example: Chateau Winery



It uses the total monthly purchases of customers and uses cluster analysis to understand customer behaviour.

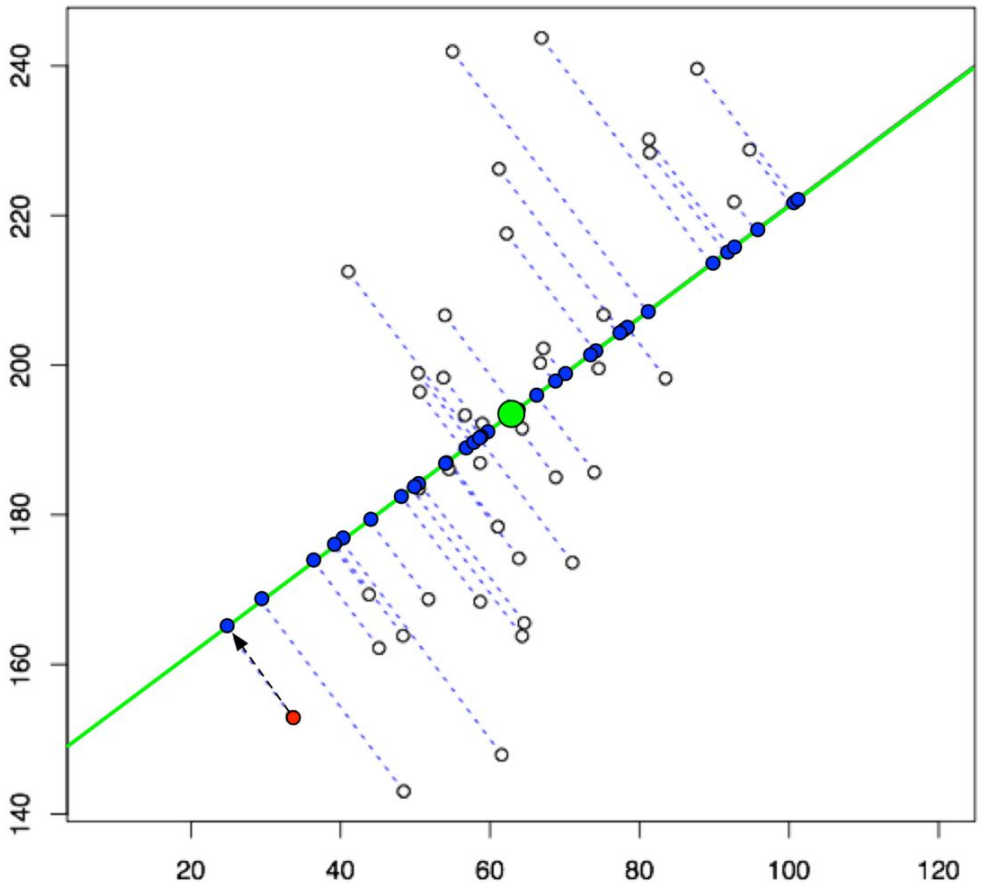


It figures the number of clusters, using Elbow Method,by minimising the within cluster sum of squares or silhouette distance.

Principal Component Analysis

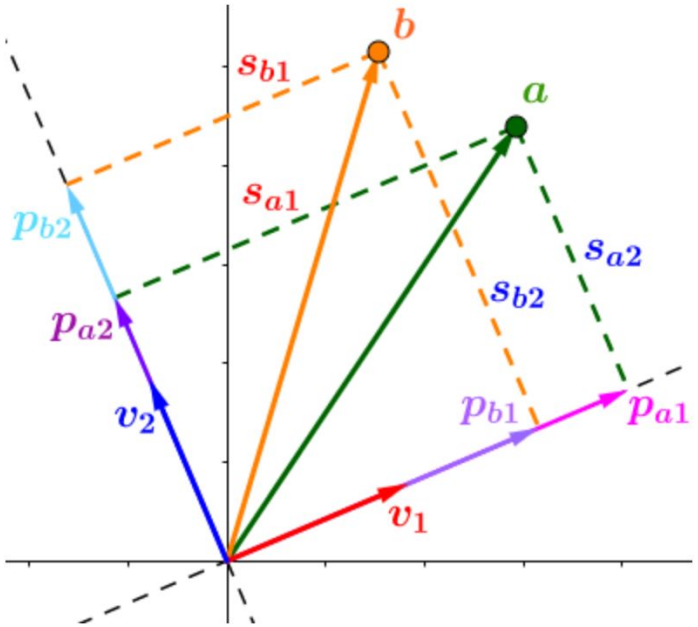
PCA is a very flexible tool and allows analysis of datasets that may contain, for example, multicollinearity, missing values, categorical data, and imprecise measurements. The goal is to extract the important information from the data and to express this information as a set of summary indices called **principal components (orthogonal to each other)**

Ex in picture:
The green line has been constructed using mathematical optimization to maximize the variance between the data points as much as possible along that line. We call this line our first principal component. Naturally, the points on the line are still closer to each other than in the original 2D space because you are losing a dimension to distinguish them. But in many cases, the simplification achieved by dimensionality reduction outweighs the loss in information, and the loss can be partly or fully reconstructed.



Singular Value Decomposition

Singular value decomposition (SVD) is another dimensionality reduction approach which factorizes a matrix, A, into three, low-rank matrices. SVD is denoted by the formula, A = USVT, where U and V are orthogonal matrices. S is a diagonal matrix, and S values are considered singular values of matrix A.



DISCUSSION

Discuss with you team what could be potential applications of **Unsupervised Learning**

Present the key points of your discussion back to the class

Unsupervised Learning Applications

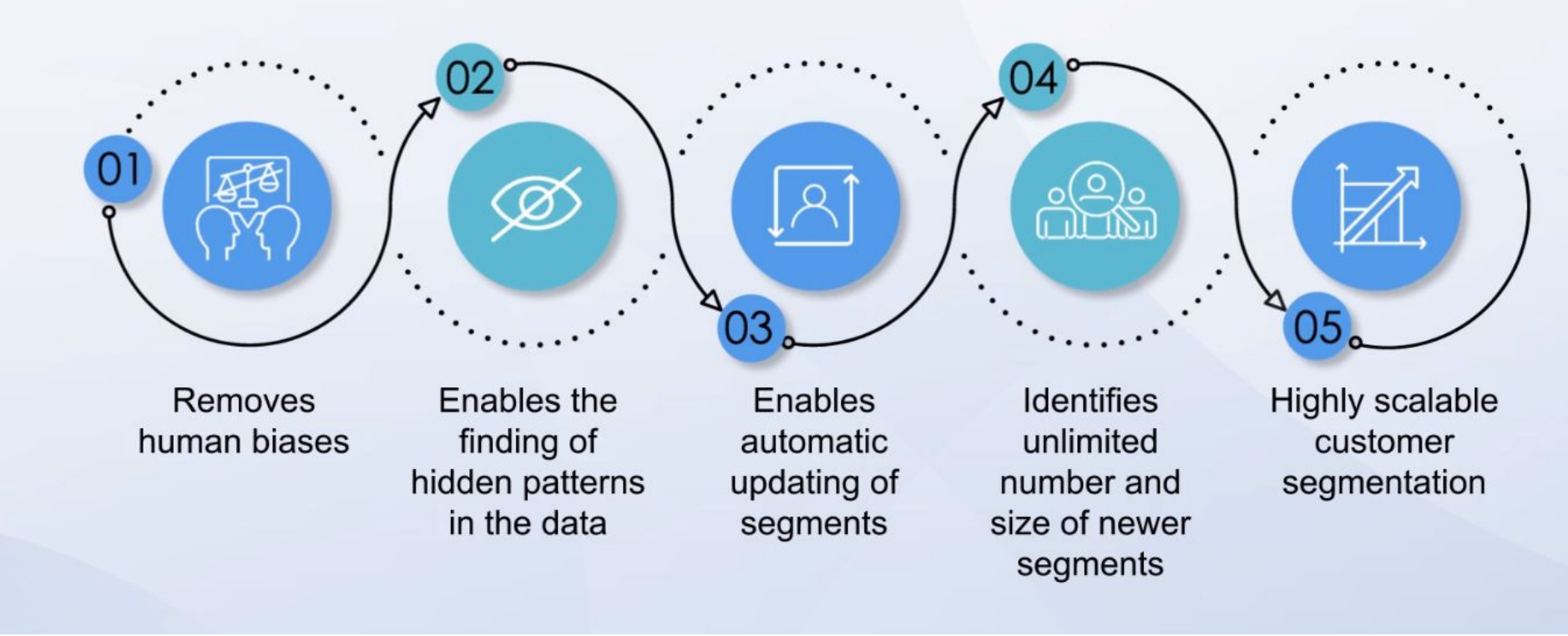
Clustering

- CUSTOMER and MARKET SEGMENTATION
- ANOMALY DETECTION (Fraud, maintenance)
- HEALTHCARE (diagnose cancer at earlier stage)

Association

- PRODUCT RECOMMENDATIONS
- TARGETED MARKETING

Applications: Segmentation



QUESTIONS?