Introduction to Unsupervised Machine Learning: Takeaways

Concepts

- Unsupervised machine learning is the type of machine learning that does not aim to predict values but to segment data.
- Common unsupervised machine learning types are:
    - `Clustering`
    - `Association`
    - `Anomaly detection`
- The K-means algorithm is an iterative algorithm designed to split the data into K clusters, where K is set by the user.
- The algorithm initializes random centers for each cluster (called centroids), then uses the mean distance from each point in a cluster to the cluster's centroids to calculate new centroids and tries to find a convergence point.
- The convergence point is found when the centroids are located in the mean point of their clusters.
- K-means uses Euclidean Distance to calculate the distance between two points. This is the formula for the Euclidean Distance in two-dimensional space as we worked in this lesson:

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Resources

- K-means algorithm
- Euclidean Distance