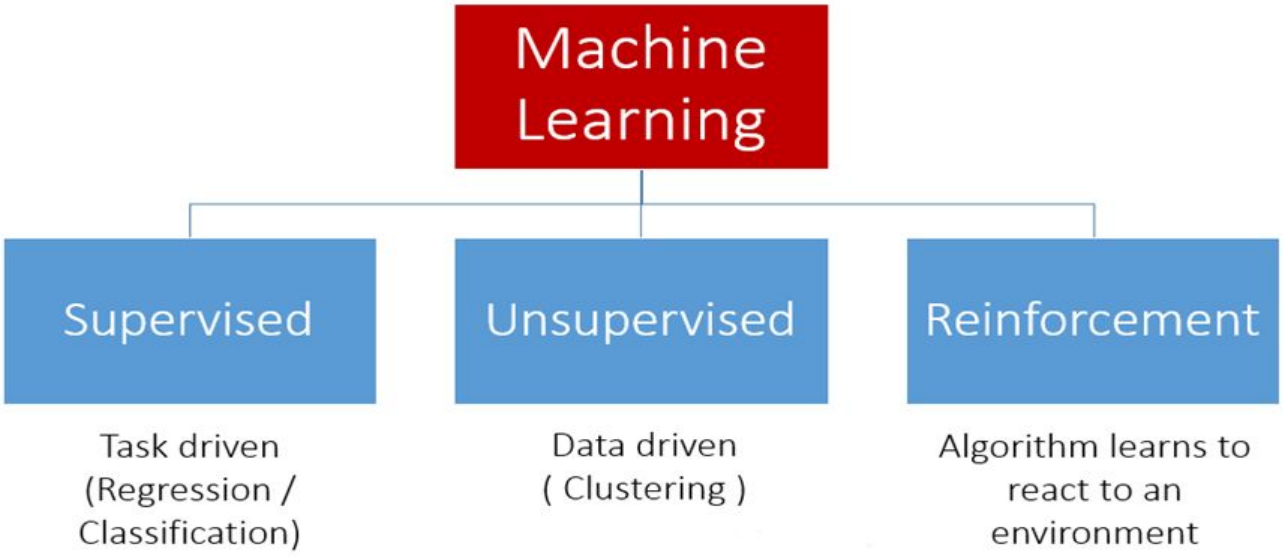
Types of Machine Learning

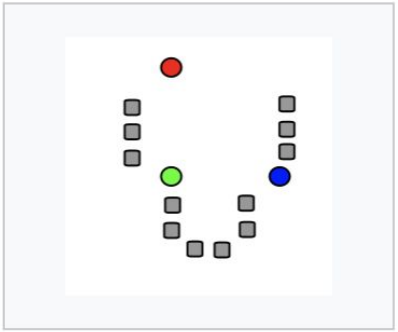


Unsupervised Learning

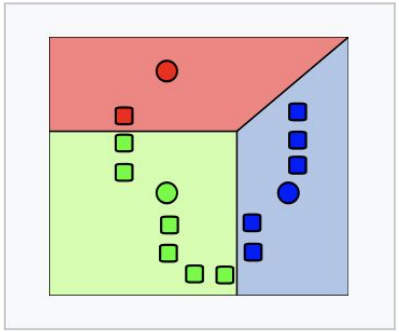
- Unsupervised learning is a type of machine learning algorithm used to draw inferences from datasets consisting of input data without labeled responses.
    - Examples:
        - Facial Recognition
        - Robotics (Roomba)
        - Natural Language Processing (NLP)

kMeans-Clustering

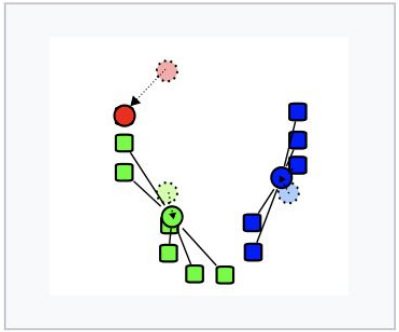
- Most common way of doing Unsupervised learning



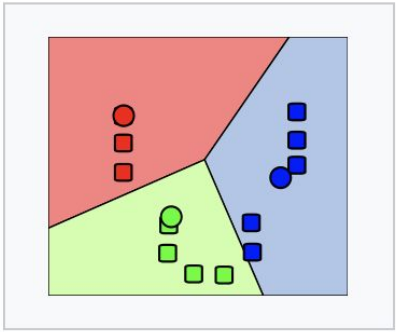
1. *k* initial "means" (in this case *k*=3) are randomly generated within the data domain (shown in color).



2. *k* clusters are created by associating every observation with the nearest mean. The partitions here represent the Voronoi diagram generated by the means.



3. The centroid of each of the *k* clusters becomes the new mean.



4. Steps 2 and 3 are repeated until convergence has been reached.

Pros and Cons

- Pros:
  - Simple, easy to implement.
  - Flexible, can easily adjust to changes in data.
  - Works well with large datasets, can efficient handle large datasets
  - Efficient, only requires simple arithmetics.
- Cons:
  - Difficult to determine the ideal number of clusters
  - Heavily influenced by the initial seeding of each point.

Use Cases:

- Document classification
- Delivery optimization
  - finding the optimal number of launch locations (clusters)
- Identifying crime localities
- Customer Segmentation
  - Classifying customer based on their purchase habits
- Insurance Fraud Detection
  - Flag activity outside of cluster regions.