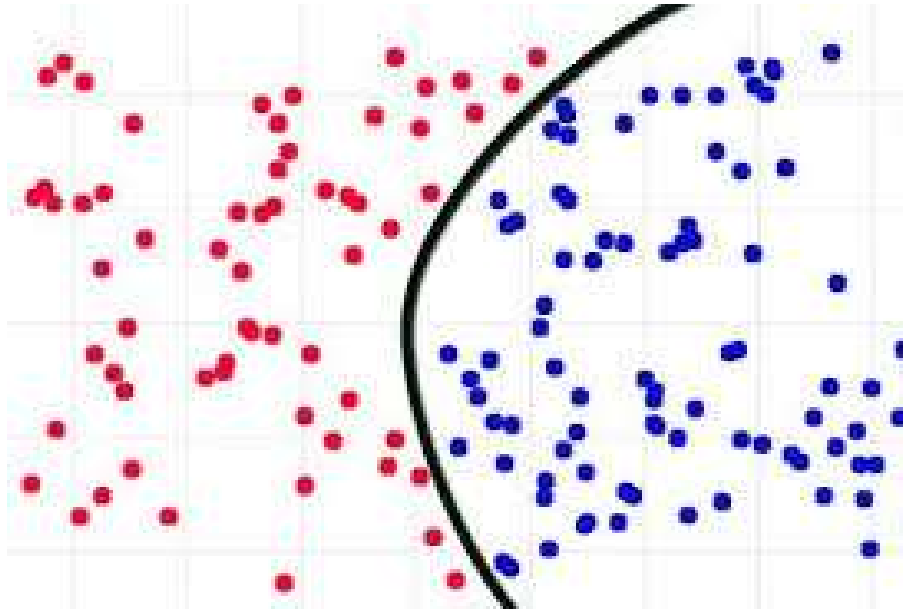
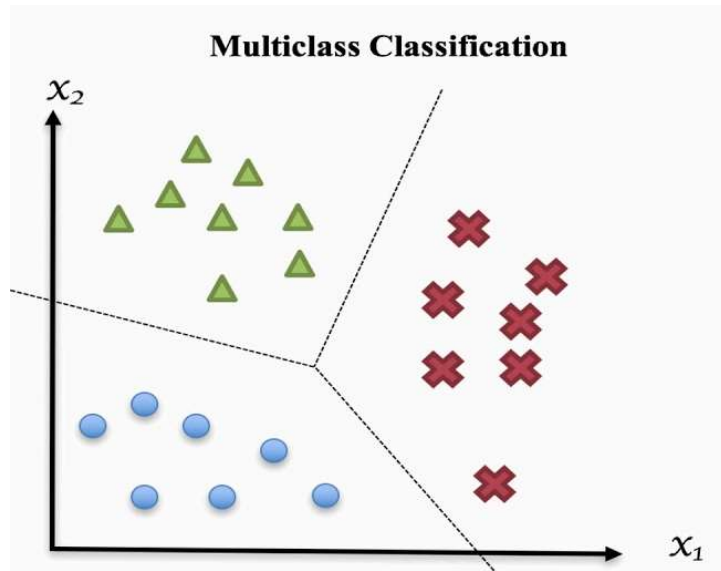


Classification, Regression and Clustering

Introduction

- When a machine learning task is presented, the first thing is to know whether the learning task is Classification or regression or clustering problem so that next you can pick the algorithm.
- These are simple concepts to understand.
- Regression and classification are supervised learning approach that maps an input to an output
- Clustering is a unsupervised learning approach.

Classification



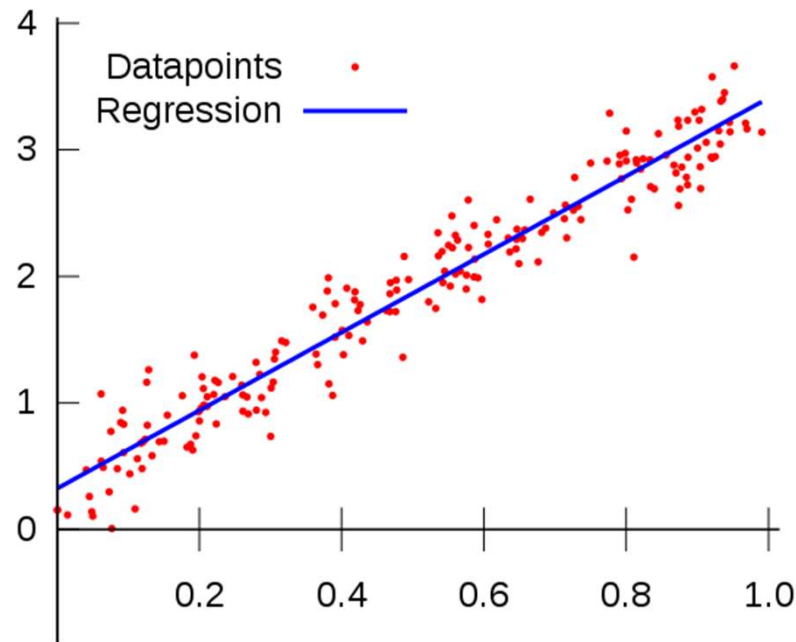
- Predicts discrete number of values.
- Binary classification — when there is only two classes to predict, usually 1 or 0 values.
- Multi-Class Classification — When there are more than two class labels to predict

Application of Classification Algorithms

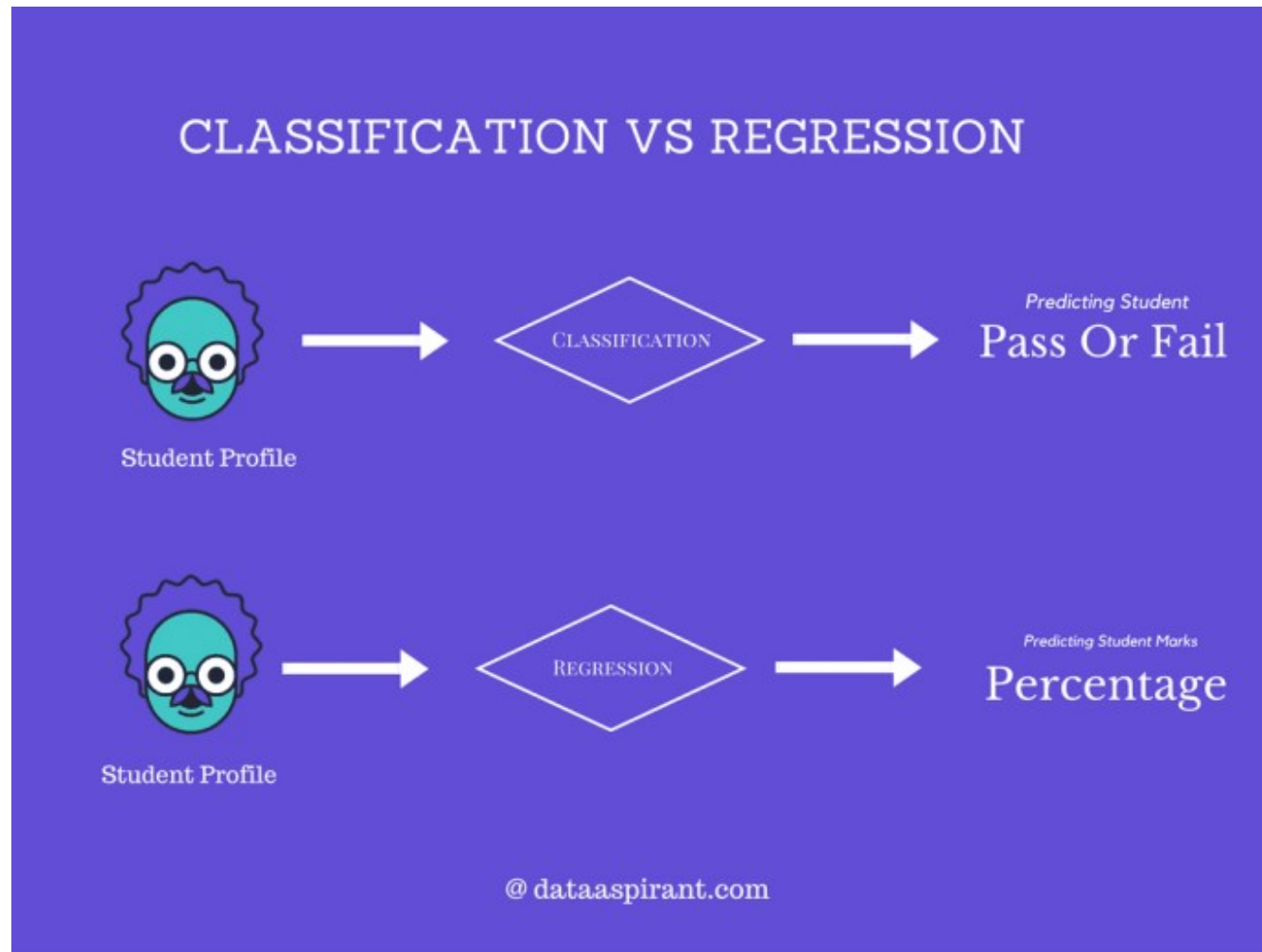
- Email spam classification
- Bank customers loan pay bank willingness prediction.
- Cancer tumor cells identification.
- Sentiment analysis.
- Drugs classification
- Facial key points detection
- Pedestrians detection in an automotive car driving.

Regression Problems

- In regression problems we try to predict continuous valued output.
- Example: Given a size of the house predict the price (real value).
- Example: Predicting a person's income from their age, education



Classification vs Regression

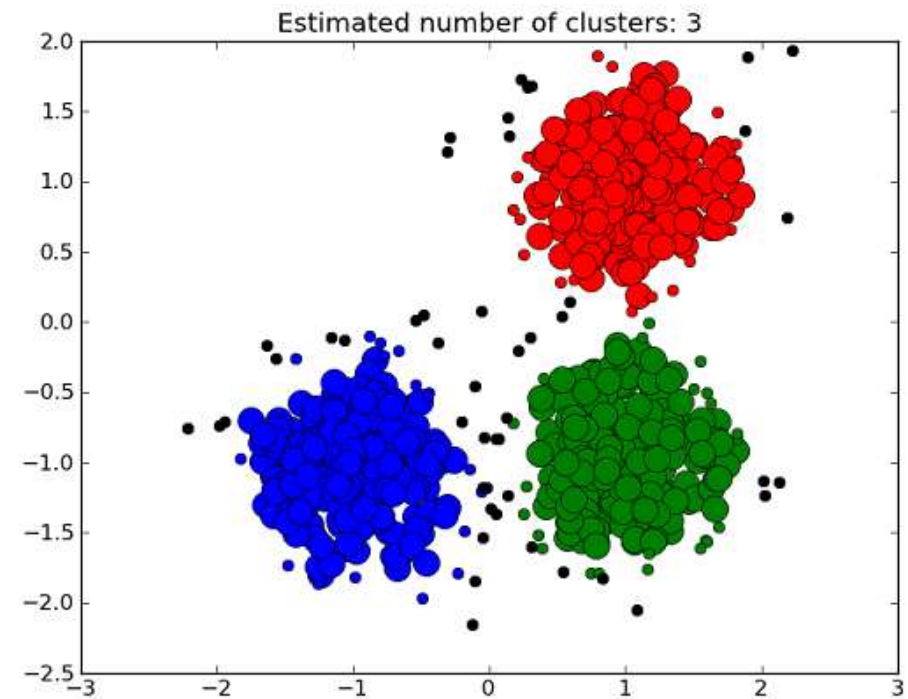
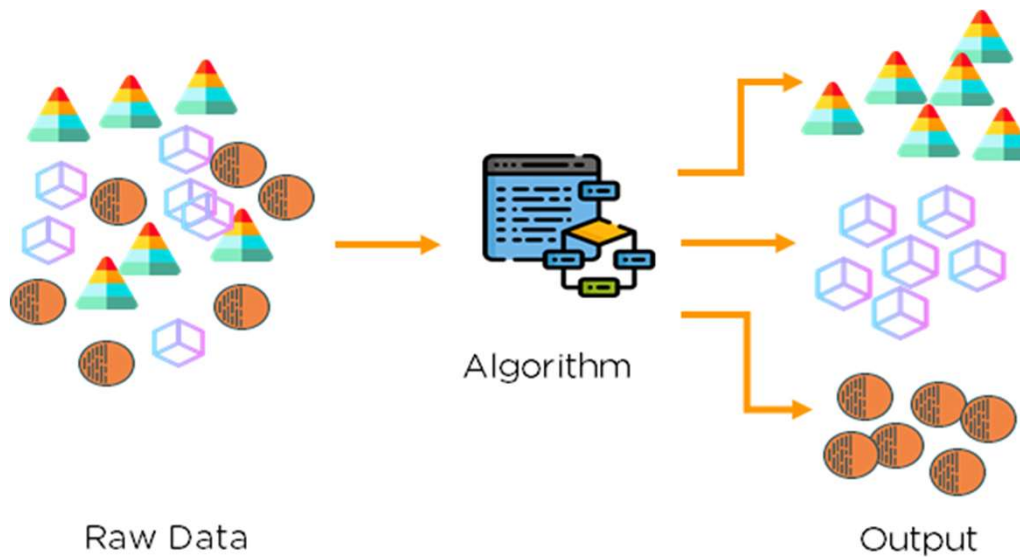


Clustering

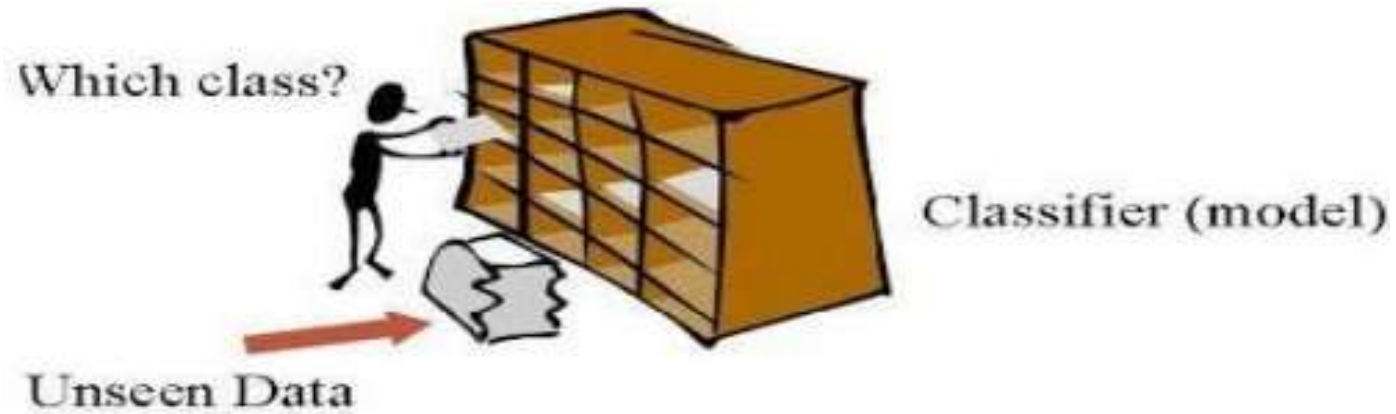
- Clustering is the task of partitioning the dataset into groups, called clusters.
- The goal is to split up the data in such a way that points within each cluster are very similar.
- It determines grouping among unlabeled data.
- Its grouping a set of new data based on similarities amongst them.
- Given a set of data points, we can use a clustering algorithm to classify each data point into a specific group.

Algorithms for Clustering

- K-Means Clustering
- Hierarchical clustering



Clustering vs Classification



Examples of Clustering

