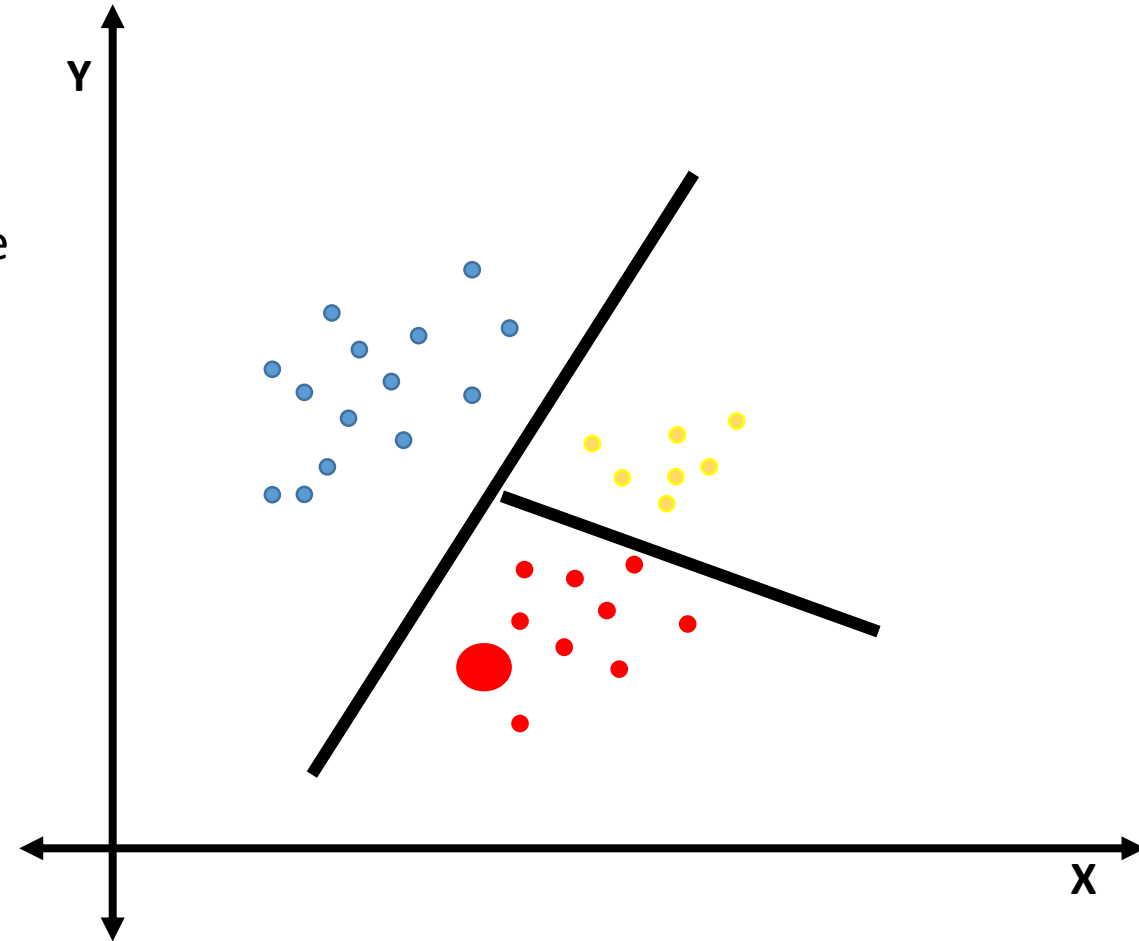


Classification

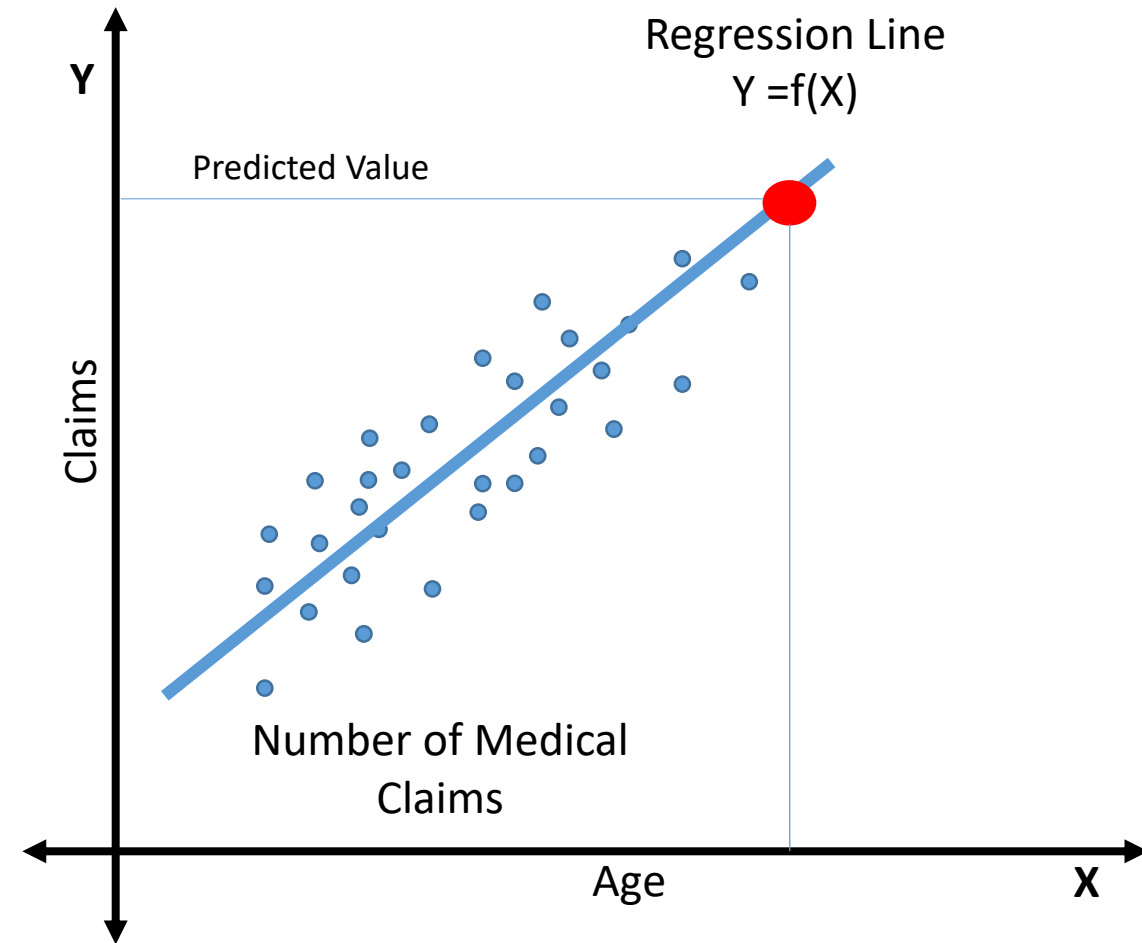
- Identification of category of data
- Binary/Two-Class Classification – Either/Or, Yes or No type
- Multi-Class Classification – One of the many alternatives
- Examples
 - Assigning a given email into "spam" or "non-spam" classes Or Primary, Social or Promotional emails
 - Will this customer default on loan repayment?
 - Will this customer buy my product?



Predicting the value for categorical variable.

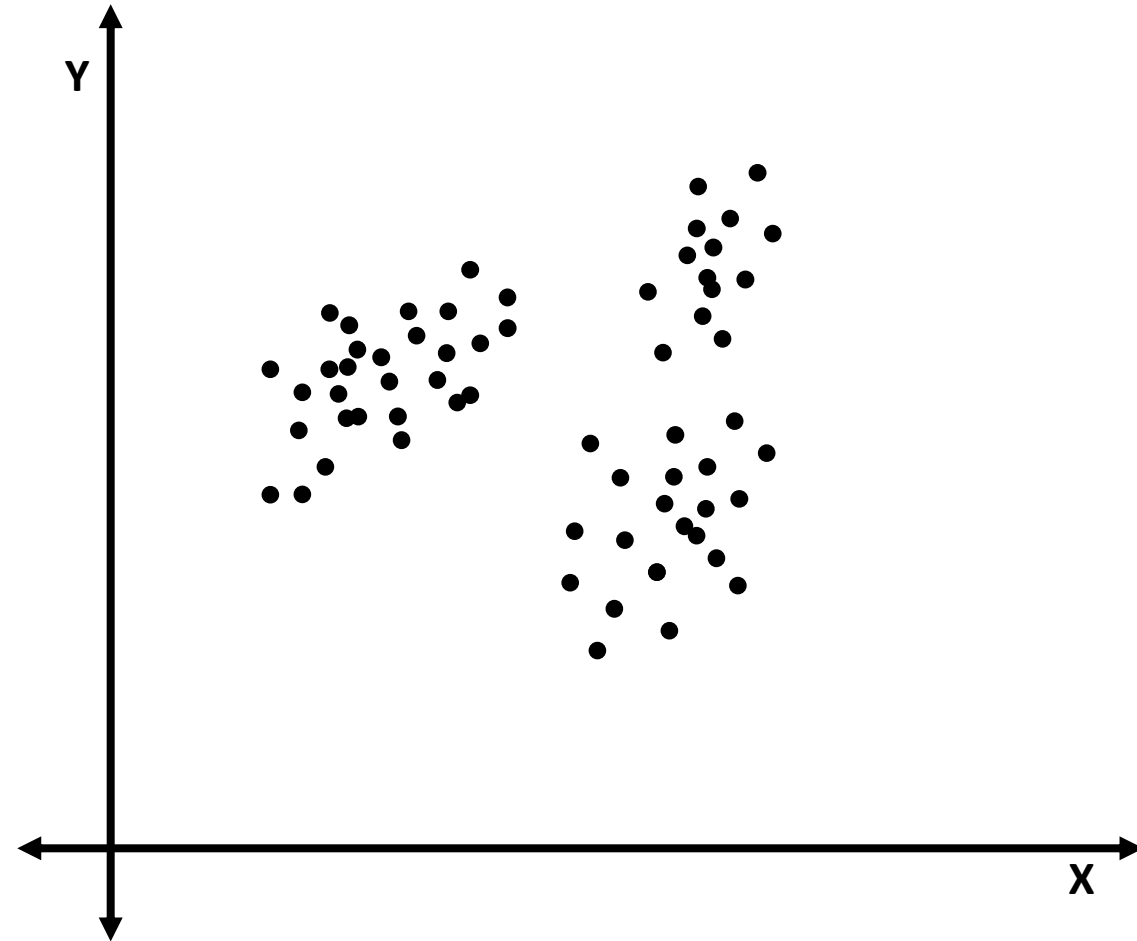
Regression Analysis

- Estimating the relationships among variables
- Predictor is a continuous variable
- Examples
 - Predicting the future sale of products
 - Computing fair price of the product or service
- One of the most common methods used in Machine Learning
- Infer causal relationships between dependent and independent variables.



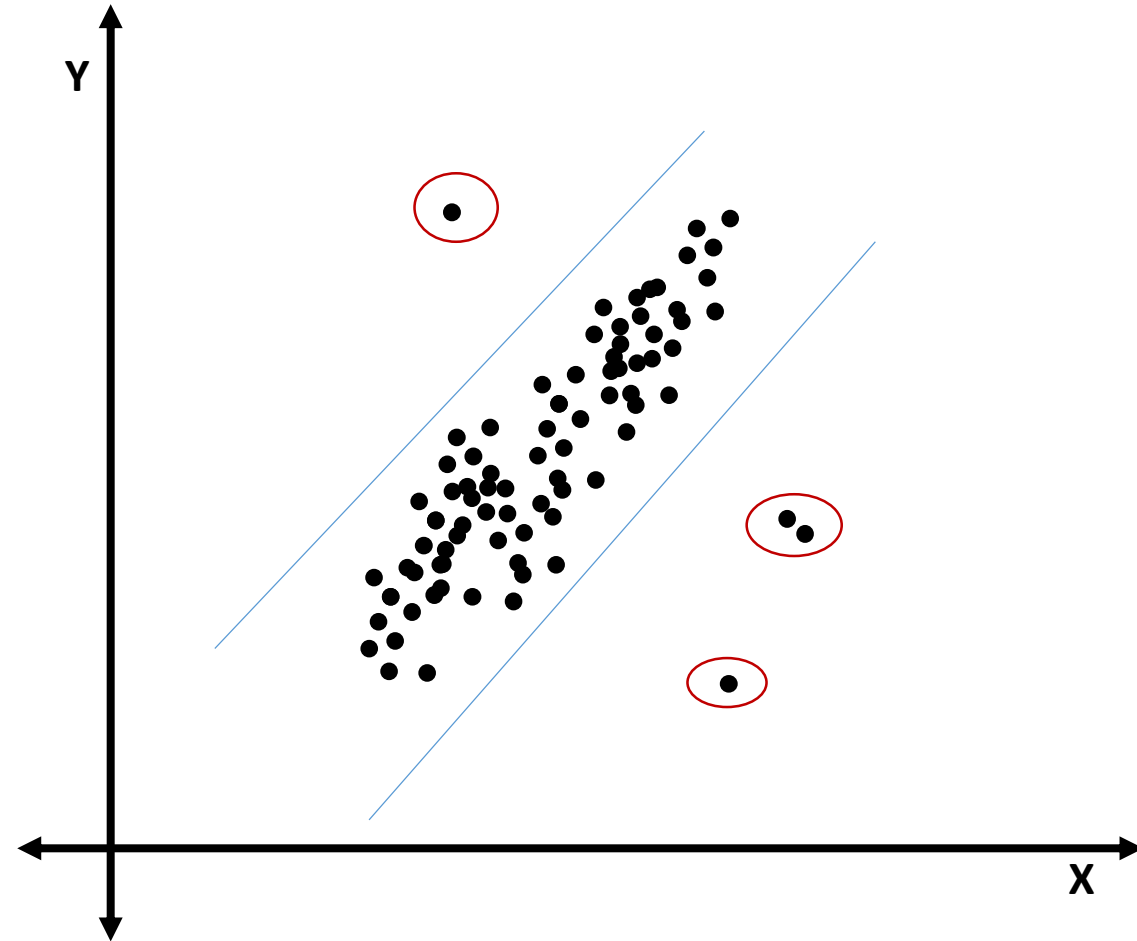
Clustering or Cluster Analysis

- Clustering is the task of grouping a set of objects in such a way that
 - objects in the same group (called a cluster) are
 - more similar (in some sense or another)
 - to each other than
 - to those in other groups (clusters)
- Unsupervised Learning model
- Customers who make lot of long-distance calls and don't have a job. Who are they?



Anomaly Detection

- Anomaly detection (also outlier detection) is the
- Identification of items, events or observations which
- Do not conform to an expected pattern or other items in a dataset.
- Typically the anomalous items will translate to some kind of problem such as
 - Bank fraud
 - Credit Card Fraud
 - Structural defect
 - Medical problems
- Anomalies are also referred to as outliers, novelties, noise, deviations and exceptions.



Thank You and Have a Great Time!