

How Data Used for discovery

General Methods (Supervised Learning, Unsupervised learning)

Why Estimate  $F$ ? (**Machine learning Idea which finds function for dependent variable**)

How do we estimate  $F$ ?

- **Statistics and Linear Algebra for Machine Learning**

- Sample mean, variance
- Normal Distribution (not mandatory just need some brief)
- Vectors and matrices
- Matrix operations (multiplication, dot product, Cross product, inverse, Transpose)
- Multi dimensional data

**>> Introduction to Python/R**

- What is machine learning
  - Supervised Learning
  - Unsupervised Learning
- Model Based Learning and Instance Based Learning
- Main challenge In machine learning (Testing and Validating)

- **Calculus for Machine Learning**

- Idea of Minimization and maximization of function
- First and second derivative test
- Chain Rule
- Partial Derivatives
- Gradients
- Gradient descent
- Gradient ascent
- Multivariate Gradient descent/ ascent
- Linear Algebra Basics (Matrix/ Dot Product)

- **Classification using K Nearest Neighbours Algorithm**

- Idea of similarity between samples
- Euclidean distance
- Hamming distance
- Cosine Distance
- Manhattan Distance
- Detailed algorithm

- Code using Package
- Hyper parameter tuning (Number of K)
- **Linear Regression**
  - Idea of Curve fitting
  - Least Square Regression
  - Correlation
  - R squared
  - Ridge Regression and Lasso Regression
  - Linear Regression code
  - Polynomial regression
- **Model validation**
  - Test train split
  - Cross validation /Leave one out validation
  - Bias variance tradeoff
  - Hyper parameter tuning
- **Basics of Statistical Inference**
  - Continuous statistical distribution
  - Normal distribution
  - Law of Large Numbers
  - Central Limit Theorem
  - Bootstrap
  - Maximum Likelihood Estimator (EM algorithm)
- **Classification Models**
  - Why not linear regression for Classification?
  - Logistic Regression
    - Estimating regression Coefficient
    - Making Prediction
    - Multiple Logistic Regression
    - Odds Ratio
    - Logistic regression for more than 2 classes
  - Naive Bayes Classifier
  - Linear Discriminant analysis
  - Quadratic Discriminant analysis
- **Tree Based Methods**
  - Entropy
  - Information Gain
  - Decision Tree
  - Decision Tree Use Case

- Random Forests
- **Unsupervised Learning**
  - Introduction to Clustering
  - K-means Clustering
  - Hierarchical Clustering
- **Principal Component Analysis**
  - Variance as a proxy for interest
  - Principal component analysis
  - Choosing D
  - Limitation in PCA
- **Time Series**
  - Serial Correlation in time series
  - White noise and Random Walks in Time Series
  - Trends and Seasonality in time series
  - Time series Forecasting
  - Autoregressive Moving Averages -1
  - Autoregressive Moving Averages -2
  - Autoregressive Moving Averages -3
  - Autoregressive Integrated Moving averages
  - Hidden Markov Models for Time Series

**(Following are some Dense Topics will take at least 15 to 20 hours )**

- **Random Forest and Support Vector Machines (SVM)**
- Gaussian Mixture Models
- Outlier Detection
- Handling Imbalance Data
- **Neural Networks**
  - Applications of neural networks
  - Perceptron
  - Feed forward Neural networks
  - Back propagation in basic form
  - Introduction to Keras