

# Machine Learning

## Session 1

10th Feb 2018

Rama Krishna Bhupathi  
ramakris@gmail.com

# Agenda

- **Introduction to ML**
- **Installation and Setup**
- **Introduction to Python Libraries and hands on sessions.**
- **Linear Regression**

# What is Machine Learning?

*“A field of study that gives computers the ability to learn without being explicitly programmed” - Arthur Samuel*

*Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.*

# Top uses cases in *Machine Learning*

## Top 10 Use Cases for Data Science & Machine Learning



**HEALTHCARE:**  
Patient Diagnosis



**FINANCE:**  
Fraud Detection



**MANUFACTURING:**  
Anomaly Detection



**RETAIL:**  
Inventory Optimization



**GOVERNMENT:**  
Smarter Services



**TRANSPORTATION:**  
Demand Forecasting



**NETWORKS:**  
Intrusion Detection



**E-COMMERCE:**  
Recommender Systems



**MEDIA:**  
Interaction & Speed



**EDUCATION:**  
Research Insight

# Machine Learning ...

## Some key innovations/disruptions in the last decade

**Exponential Growth of Big Data**

**High Computational Power**

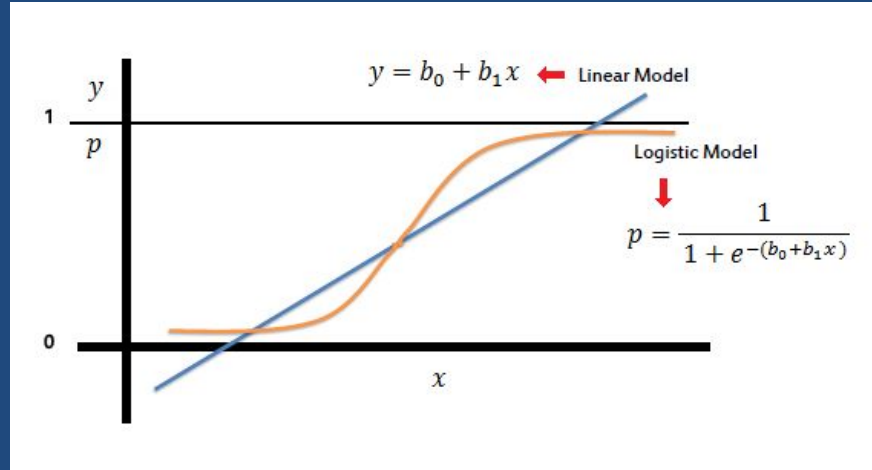
**Cloud Systems and Elasticity**

**Reduced Computational Costs**

**Availability of GPU's (Graphical Processing Units)  
available from nVidia and Google (TPU)**

# Machine Learning involves ....

Statistics  
High Level Algebra  
Probability  
Calculus



All these subject domains have been in vogue for a long time...but why is ML gaining traction now?

# Key ML Algorithms

**Decision Trees:** Useful for Credit Card Application processing.

**Naive Bayes:** Probabilistic Classifier

**Linear Regression:** Predict Sales volume next year

**Logistic Regression:** Binomial Outcome. Credit Scoring.

**Support Vector Machines :** Binary Classification Algorithm

**Clustering:** Grouping a set of Objects.

# Key ML Libraries

**SciPy:** SciPy is an ecosystem of Python libraries for mathematics, science and engineering. It is an add-on to Python that you will need for machine learning.

**NumPy :** Numerical Python Libraries with efficient implementation of multi-dimensional Array. Fast and efficient.

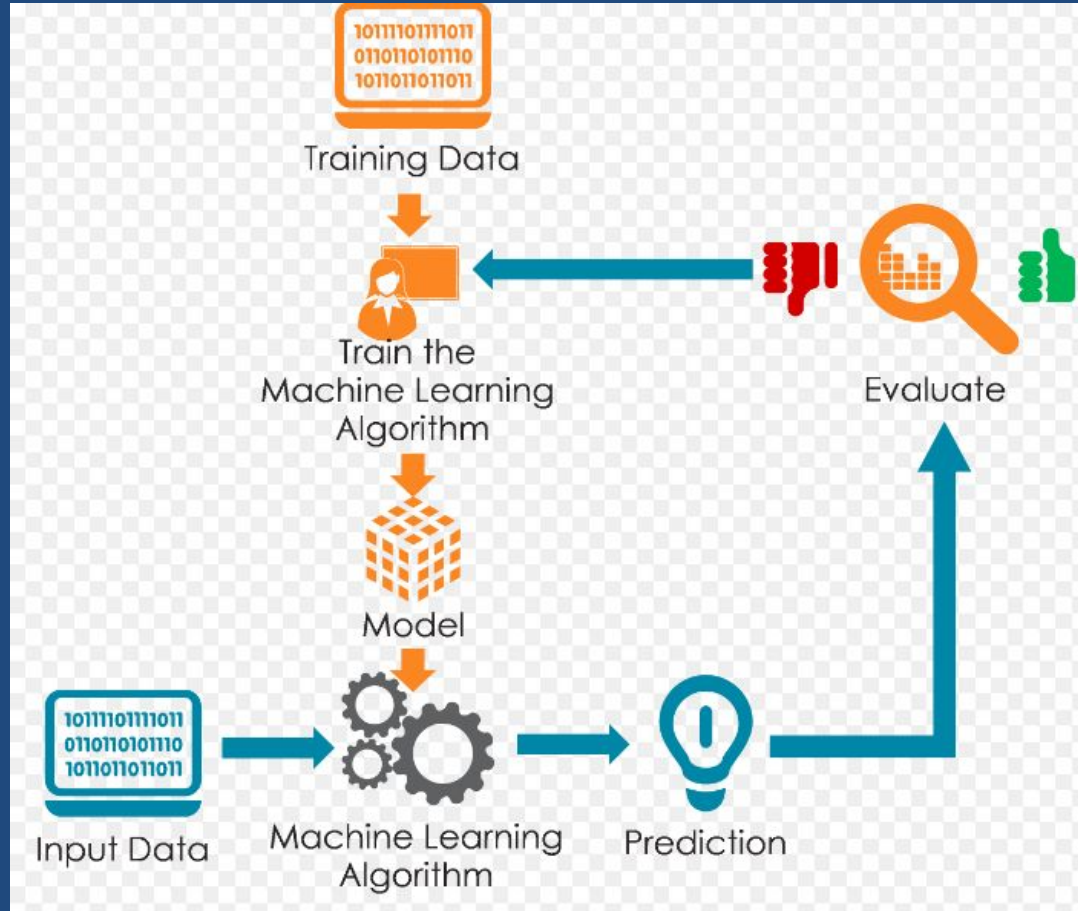
**MathPlotLib:** Plotting data in Graphs

**Pandas :** Efficient manipulating and processing of Data

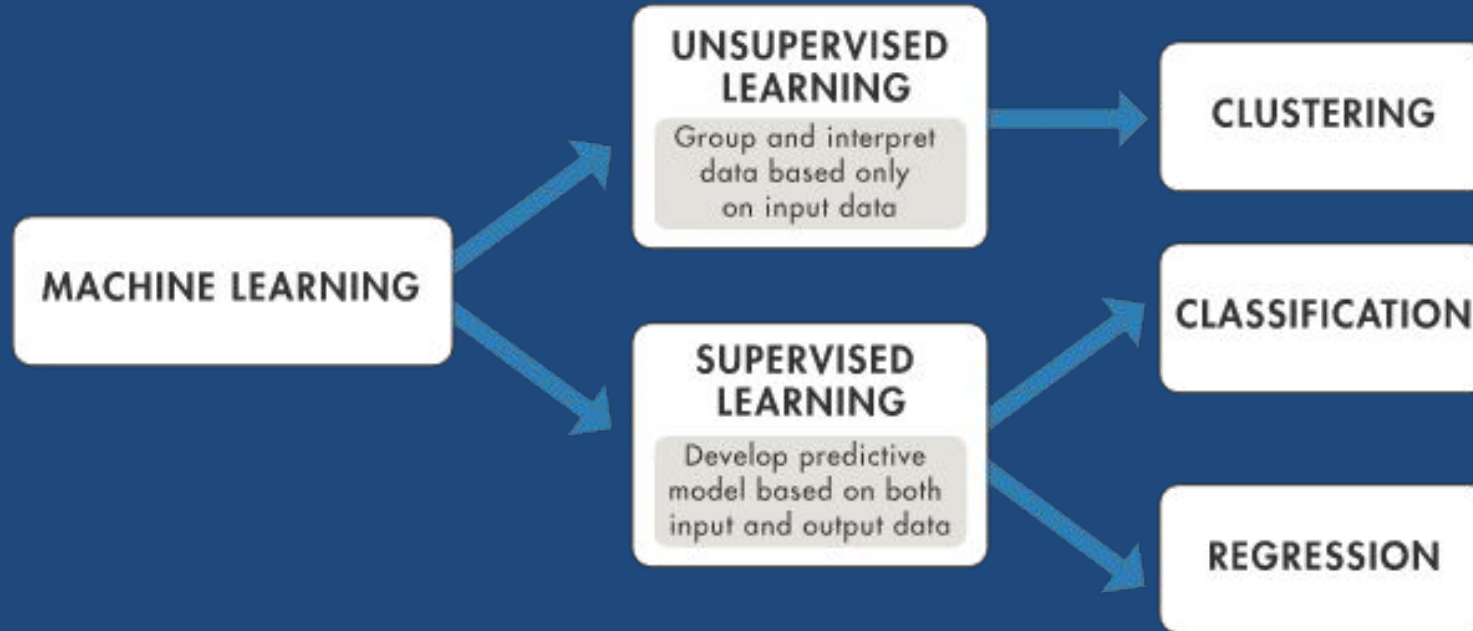
**SciKit-Learn:** Algorithm implementations , pre and post processing of data.



# So how do you do ML?



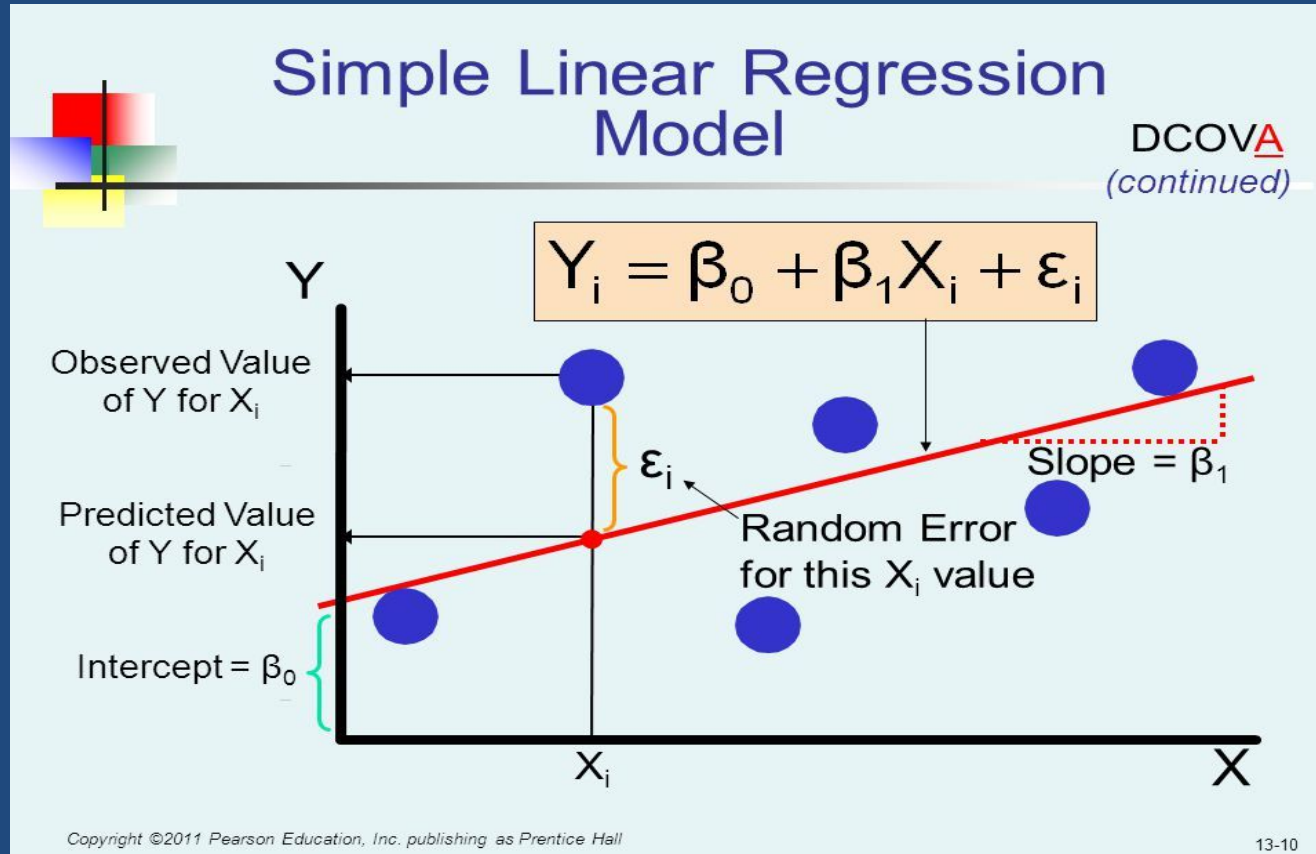
# Techniques in ML



# Simple Linear Regression

**Simple linear regression** is a statistical /machine learning method that allows us to summarize and study relationships between two continuous (quantitative) variables. Linear regression attempts to model the relationship between two variables by fitting a linear equation to observed data. One variable is considered to be an explanatory variable, and the other is considered to be a dependent variable.

# Simple Linear Regression



# Multiple Linear Regression

## Multiple Regression Model

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{i,p-1} + \varepsilon_i$$

- $Y_i$  is the value of the response variable for the  $i^{\text{th}}$  case
- $\beta_0$  is the intercept
- $\beta_1, \beta_2, \dots, \beta_{p-1}$  are the regression coefficients for the explanatory variables