Introduction

About Me

Gaurav Prakash

Associate Product Manager
HCL Technology
(Artificial Intelligence & Blockchain)

B.Tech. – Computer Science & Engineering

MBA – Marketing & Analytics

- **Product Management (AI & Blockchain), HCL Tech.**
- **Pre-Sales (Analytics & AI), HCL Tech.**
- **Business Intelligence, Infosys Ltd.**

GitHub https://github.com/gauravpks/ml-repo

LinkedIn https://www.linkedin.com/in/gauravpks/


4
Q W E R T Y U I O P
A S D F G H J K L
Z X C V B N M
123
space
return


2

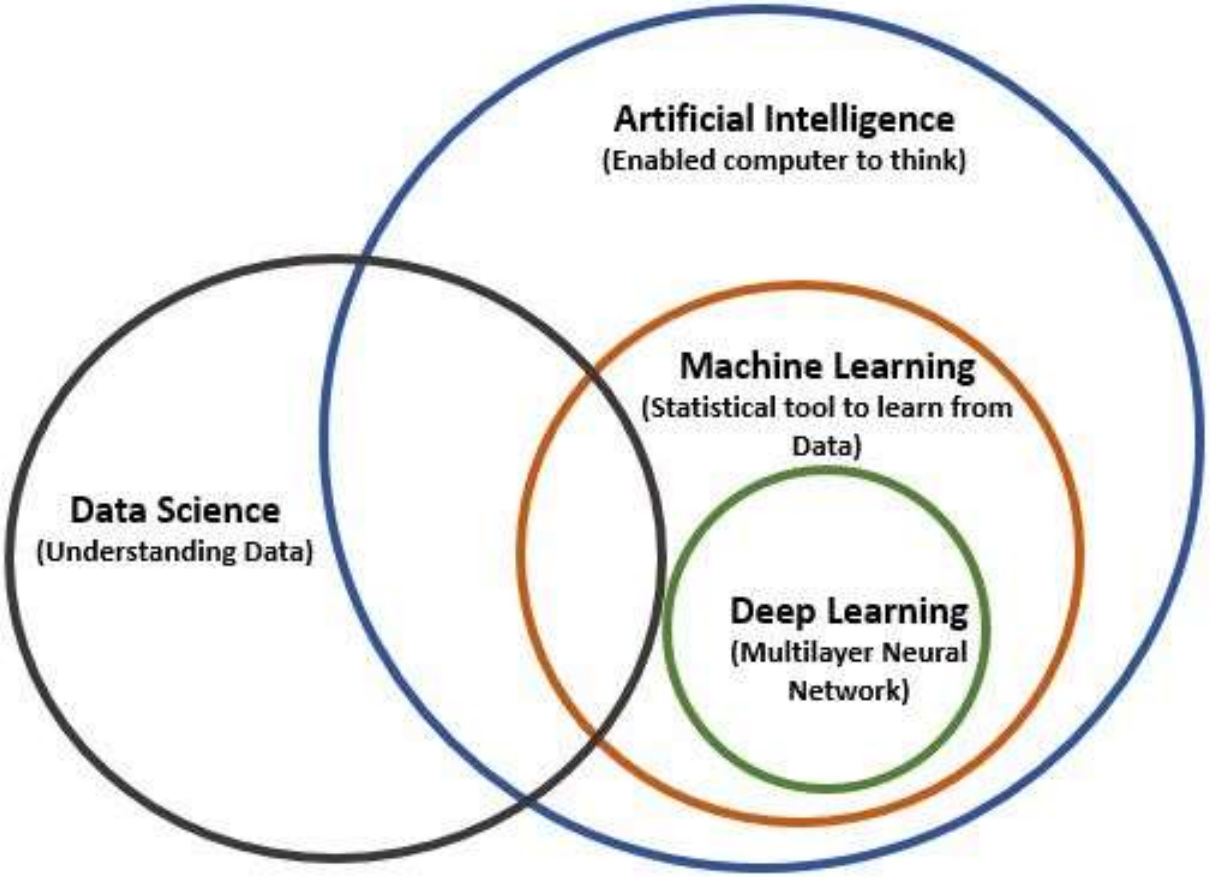

5


Advertising


amazon


NETFLIX

The Universe – Data Science & Artificial Intelligence



Artificial Intelligence… Enables computer to think

- **Artificial Intelligence** refers to the overall gamut which **enables computer to think**
- This has evolved from a **rule-based system** to modern application in defense, healthcare, automotive, retail, education, and more.
- The core objective of AI is to **impart human intelligence to machines** and enable them to **act like humans**.

Self Driven Car

Credit Decision

Face Recognition

Evolution from ML To DL

Machine Learning… Statistical tool to learn from data

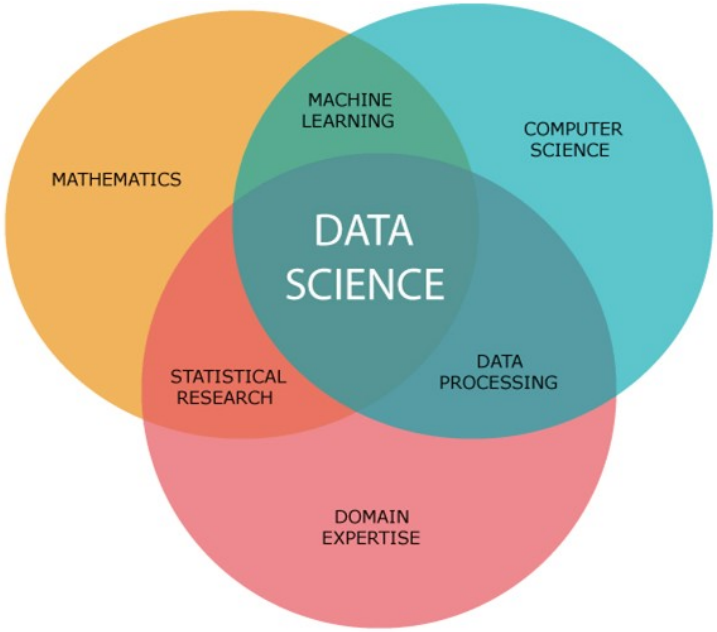
- **Machine Learning** is a **sub area of Artificial Intelligence** with a bunch of statistical tools to learn from data
- It enables the computer **to decide based on data** rather than **explicitly rule based programs** to perform a specific task.
- The logic and algorithms are **statistically designed in a certain way which learns** and improved over time and enables user to make better decision

Deep Learning… learning inspired by human brain

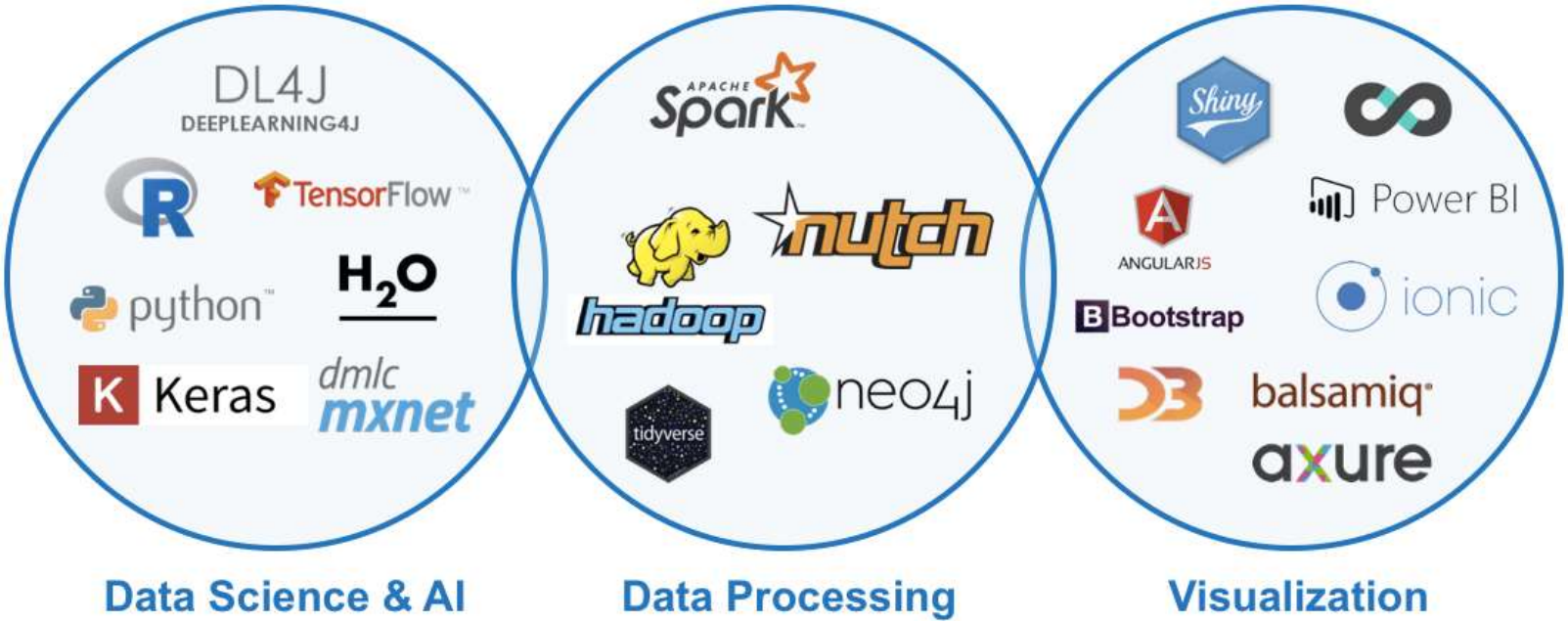
- **Deep Learning** is a recent area which has taken **shape since 2006 and has given** a new approach to Machine Learning.
- It uses **Multilayered Neural Network** (set of task-specific algorithms that uses neural networks inspired by human brain) for its advances.
- Some of the most important advances in last **6 to 8 years in Artificial Intelligence** has happened using Deep Learning

Data Science… Understanding the data

- **Data Science** is all about making sense of data.
- It overlaps with Machine Learning and Artificial Intelligence techniques, and not so much with Deep Learning
- It has its own areas like **visualization, statistics and more**. It enables us to analyze and manipulate large volumes of data to find meaning and appropriate information.



Tools & Techniques



Machine Learning

Data Exhaust

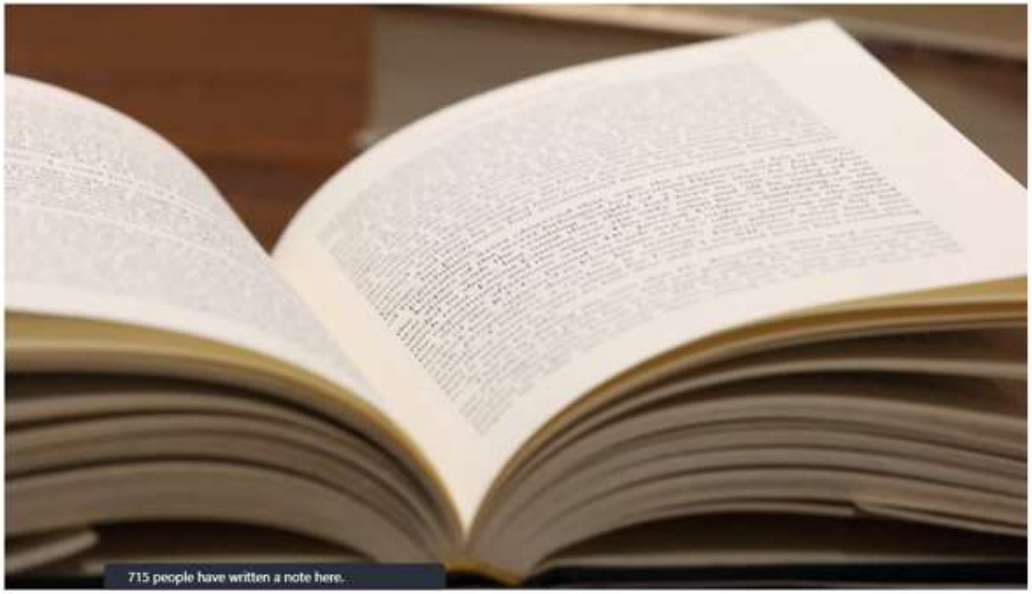
Since the Dawn of Time…
Until 2005…
Humans had created…
130 Exabytes of Data

A

Chapter I

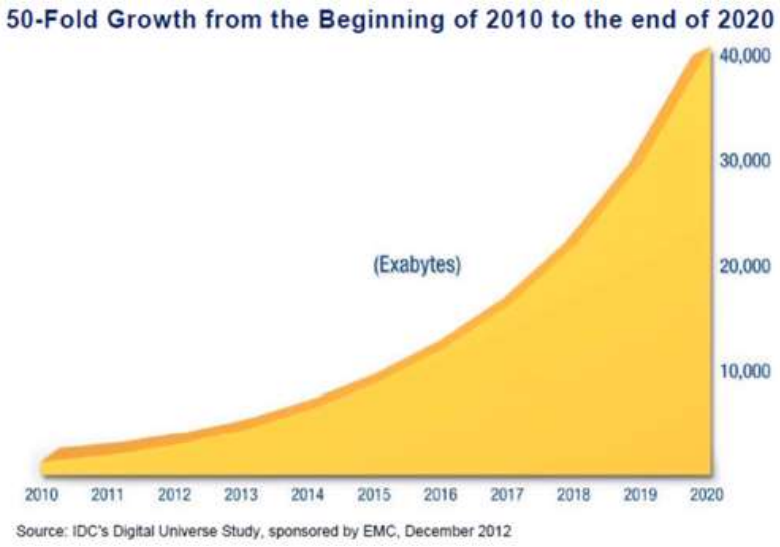
Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, 'and what is the use of a book,' thought Alice, 'without pictures or conversations?' So she was considering in her own mind, (as well as she could, for the hot day made her feel very sleepy and stupid,) whether the pleasure of making a daisy-chain was worth the trouble of getting up and picking the daisies, when a White rabbit with pink eyes ran close by her.

There was nothing very remarkable in that; nor did Alice think it so very much out of the way to hear the rabbit say to itself, 'Oh dear, oh dear! I shall be too late!' (when she thought it over afterwards, it occurred to her that she ought to have wondered at this, but at the time it all seemed quite natural); but when the rabbit actually took a watch out of its waistcoat-pocket, looked at it, and then hurried on, Alice started to her feet,

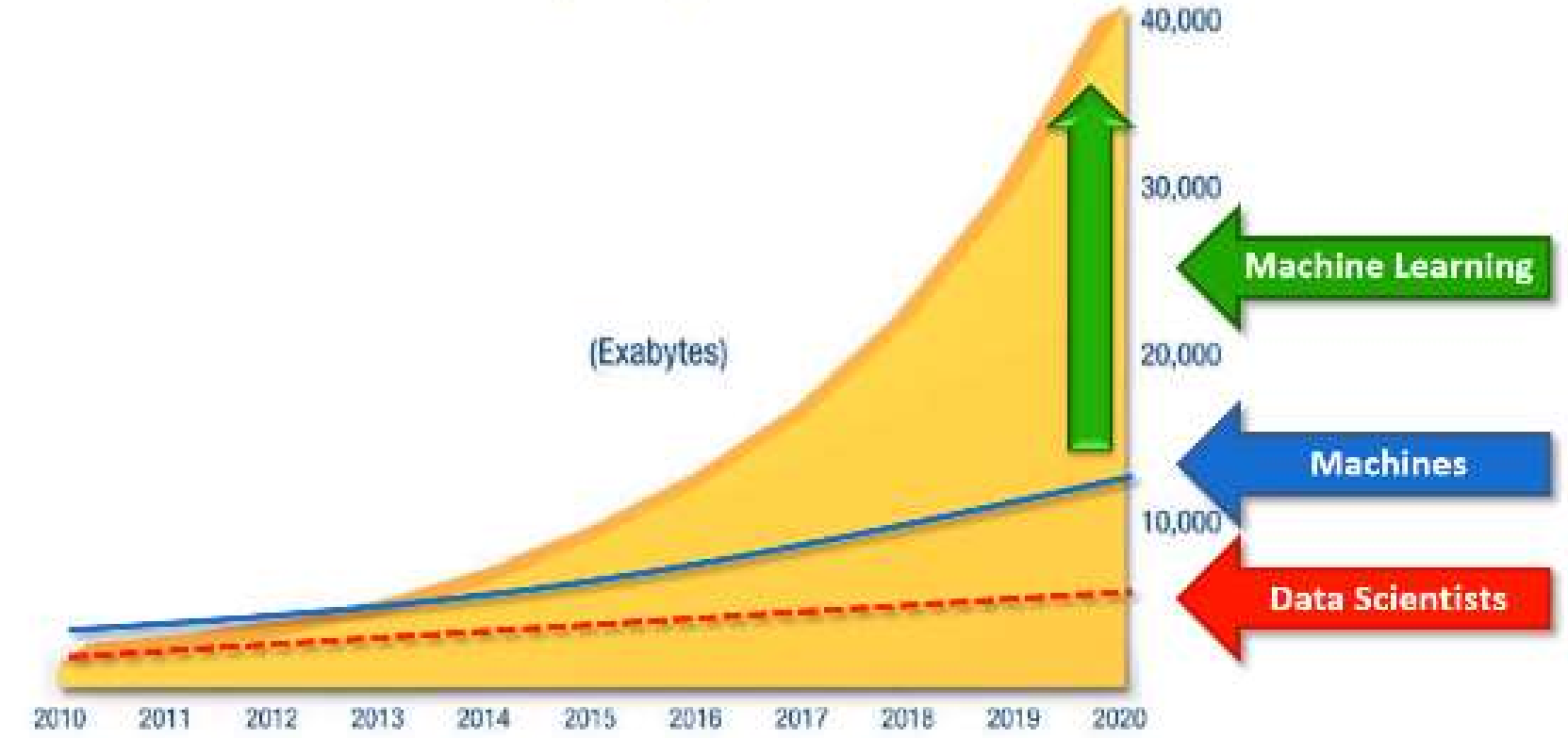

715 people have written a note here.


HD
SAMSUNG

- Until 2005…130 Exabytes of Data
- 2010… 1200 Exabytes
- 2015… 8000 Exabytes
- 2020…40000 Exabytes of Data



50-Fold Growth from the Beginning of 2010 to the end of 2020



Source: IDC's Digital Universe Study, sponsored by EMC, December 2012

Course Structure

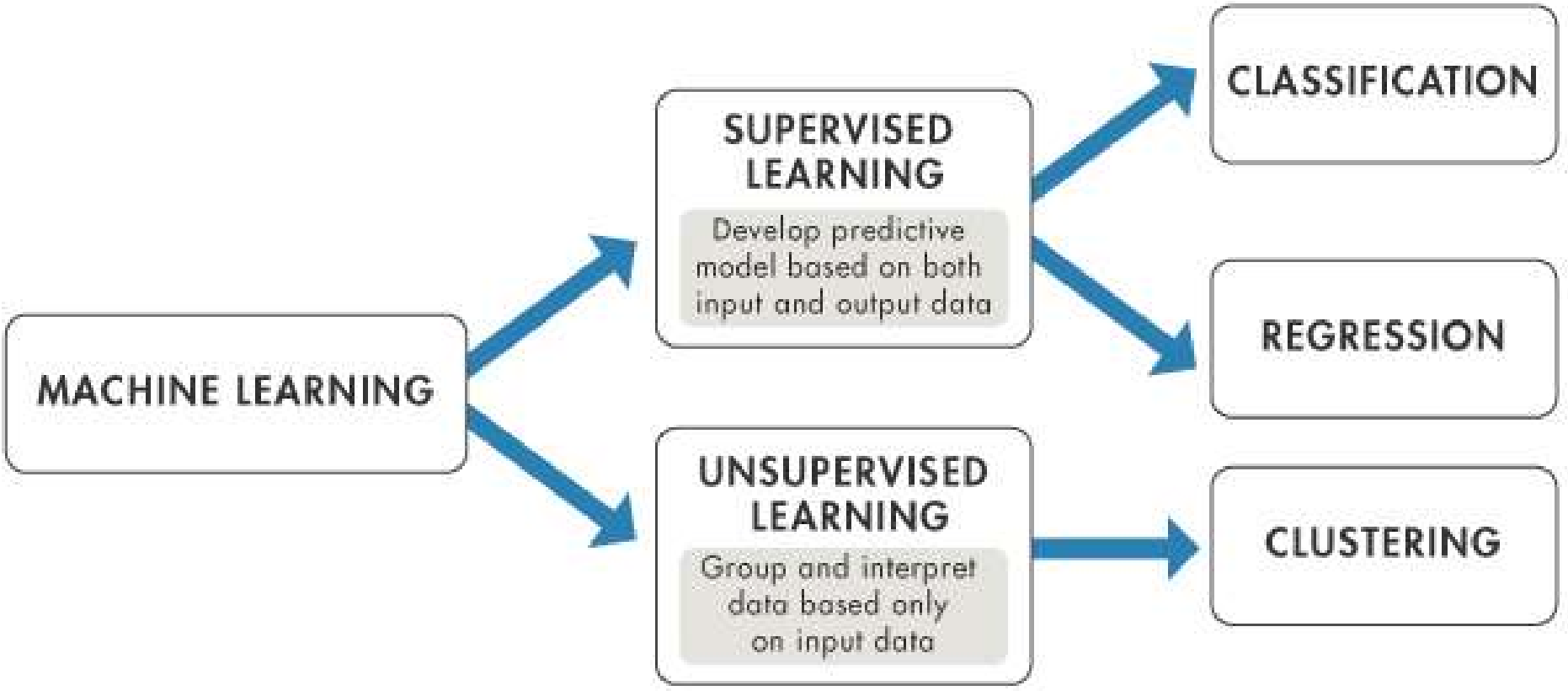
Machine Learning

- Part 0 - Welcome to Machine Learning
  - Welcome to Machine Learning
- Part 1 - Data Preprocessing
  - Data Preprocessing
- Part 2- Regression
  - Simple Linear Regression
  - Multiple Linear Regression
  - Polynomial Regression
  - Logistic Regression
- Part 3 – Classification
  - Support Vector Machine
  - Kernal Support Vector Machine
  - Naïve Bayes
  - Decision Tree
  - Random Forest
- Part 4 – Clustering
  - K-Means
  - Hierarchical
- Part 5 – Association Rule Learning
  - Apriori
  - Eclat
- Part 7- Natural Language Processing
  - Bag of Words Model
  - Sentiment Analysis
- Part 8 - Time series Forecasting
  - Time Series Components
  - Classification Technique
  - Techniques – Smoothening , ARIMA etc.
- Part 9 – Dimensionality Reduction
  - Principal Component Analysis
  - Linear Discriminant Analysis
  - Kernel PCA
- Part 10- Model Selection & Boosting
  - Cross validation
  - XG Boost
  - Ada Boost

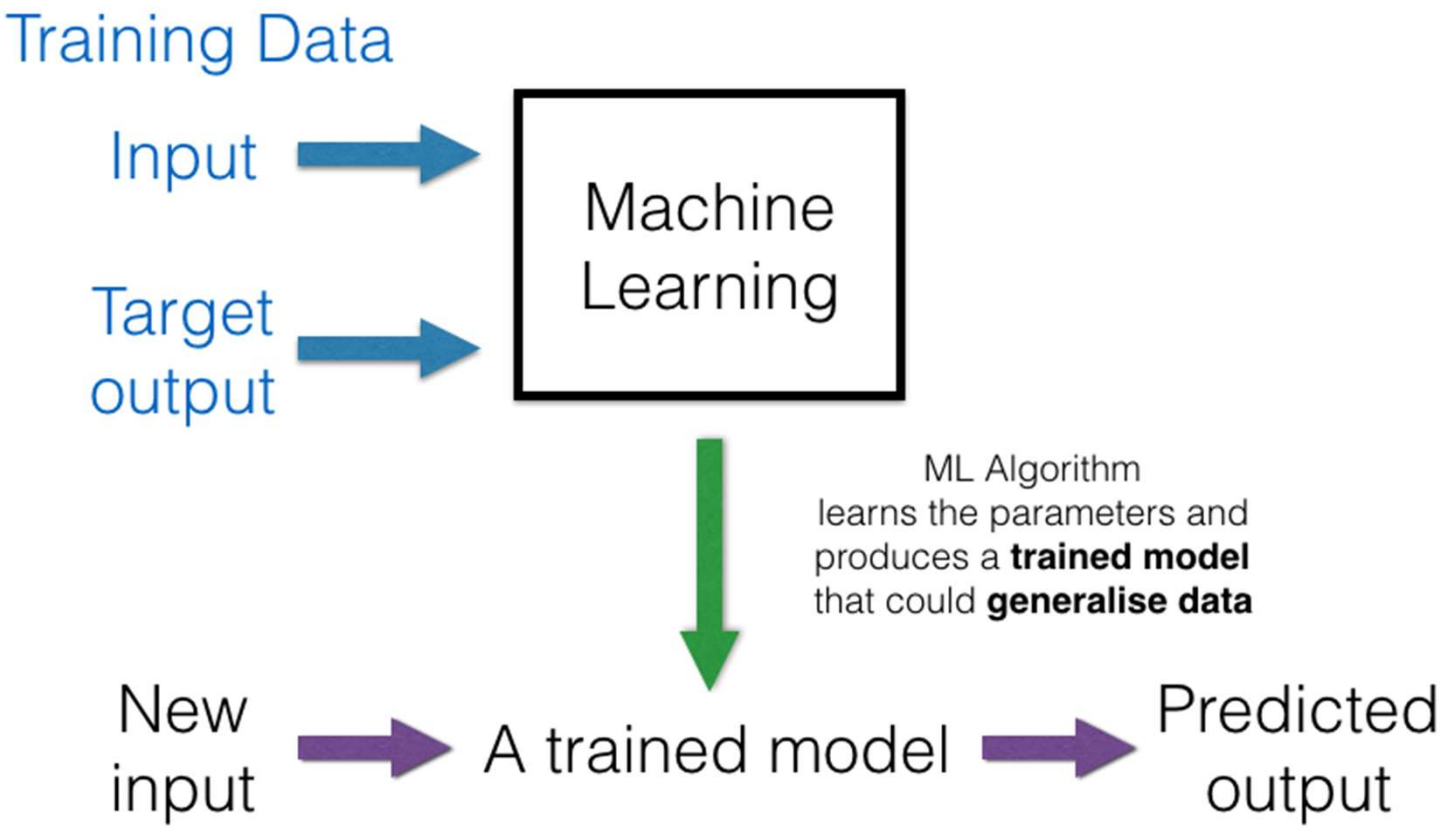
Python

NumPy Stack- NumPy, Matplotlib, Pandas

Machine Learning- Categorization



Machine Learning- Approach



Stages in Machine Learning