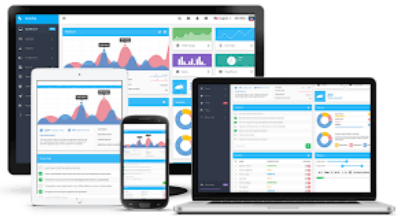
INTRO — DATA SCIENCE

Shah Ayub Quadri

Ayub.quadri89@gmail.com

Evolution of computations

Computation power

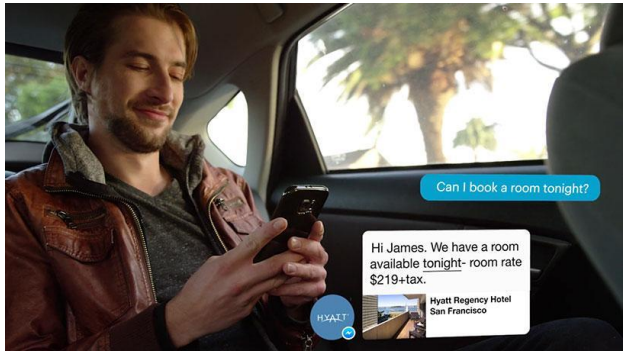


Time

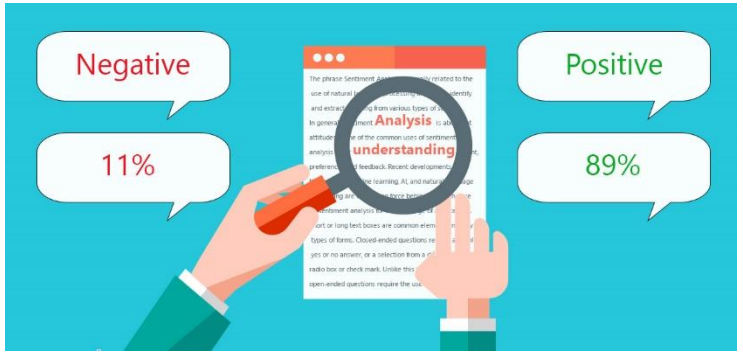
CURRENT TRENDS OF AI & ML

- Indian Languages Speech to text: https://liv.ai/
- Cosmo: https://www.anki.com/en-ca/cozmo

Future of AI & Machine learning



Chat Bots



Sentiment analysis

Sales Predictions

Self Driven Cars

Facial Exp recognition

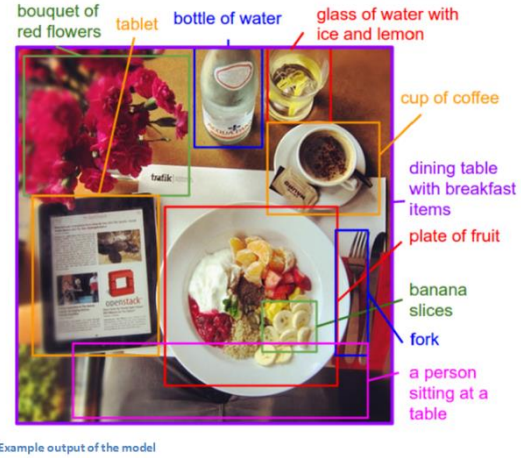
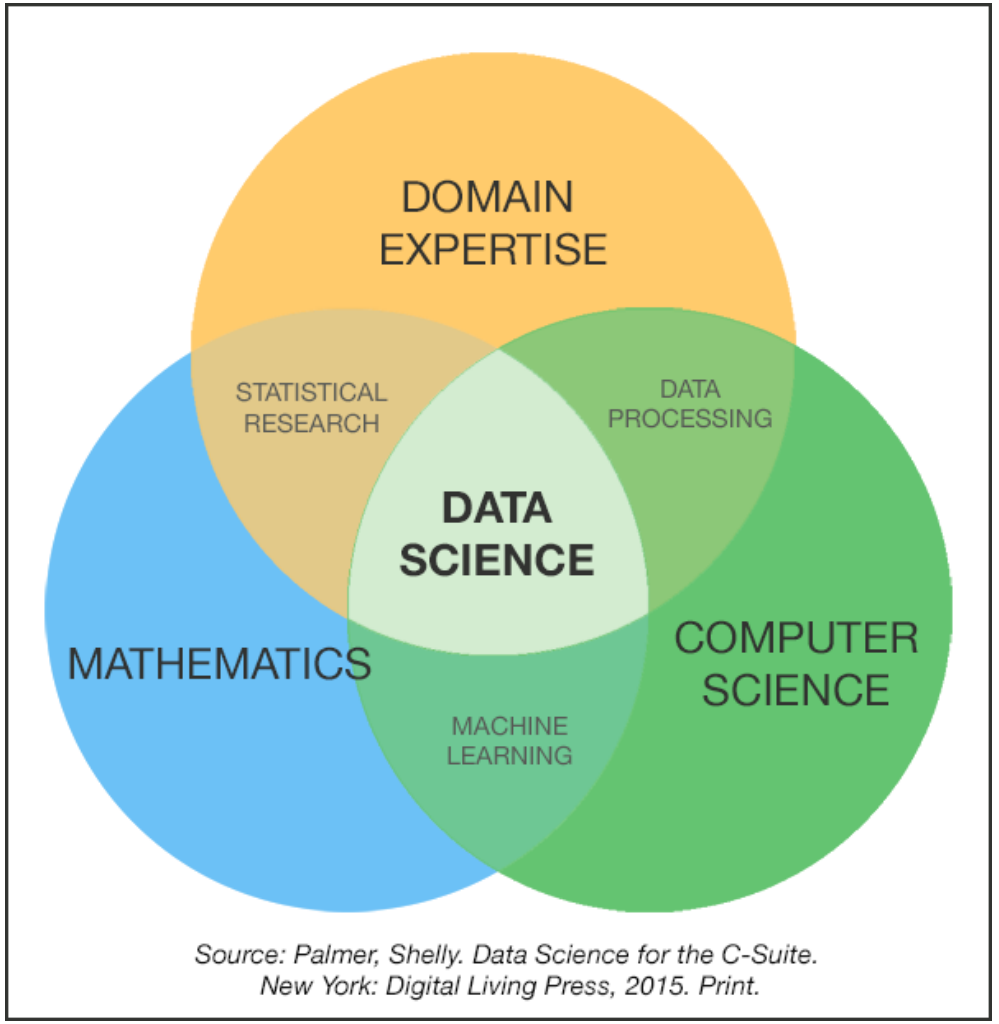


Image Tagging

Data Science

- Data science is the process of extracting meaning full insights from raw data which might be in structure or un-structured format.

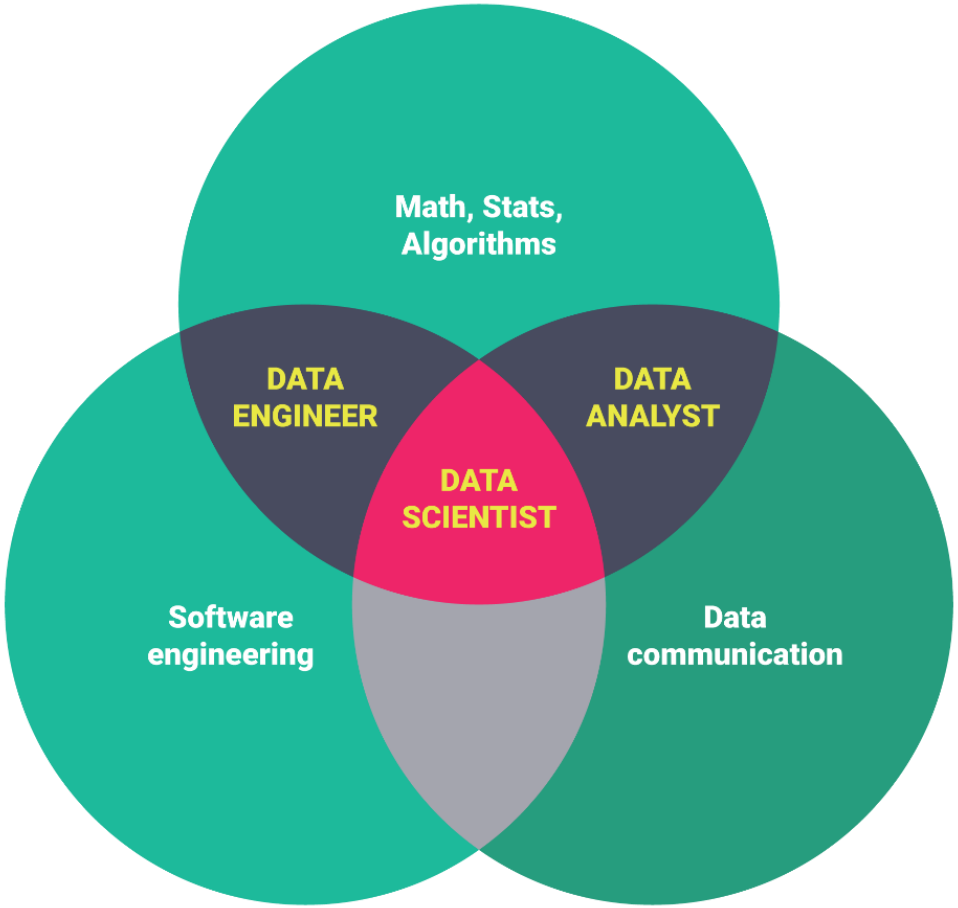
- What Data Science is not ?
  - No a piece of software – end to end App
  - Not about Visualization
  - Academic research – its more of providing business solution



Source: Palmer, Shelly. Data Science for the C-Suite. New York: Digital Living Press, 2015. Print.

Roles in Data Science

- Data Scientist
  - Uses Various Algorithms to solve complex business problems
- Data Engineer
  - Tries to handle huge data with proper constancy and availability
- Data Analysts
  - Analysts draws the insights from raw data and understands/study the data by visualizing and grouping it



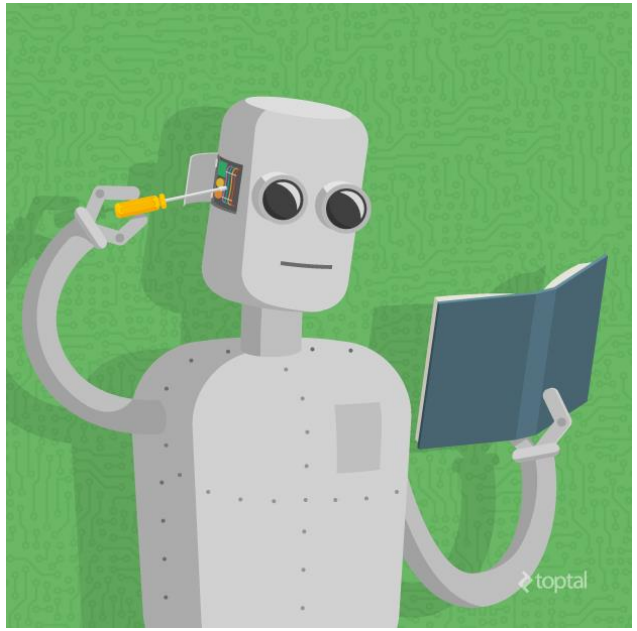
Different roles in DS: https://youtu.be/yR2wWQYiVKM

Machine Learning

- Filed of study which gives computer the ability to learn on its own or without explicitly programmed.

~ Arthur Samuel(1959)

- How to Teach Machines
  - Data
  - Machine learning Algorithms
  - Evaluate the model



Types of Learning

- Supervised
  - Model is trained on labeled data or outcome is known
- Unsupervised
  - Model learns it self with any labeled data or outcome
- Semi-supervised
  - Model is trained on both labeled and un labeled data
- Reinforcement Learning
  - Reinforcement learning is a self correcting way of learning, for every success its rewarded and for every failure it gets penalized.

Introduction & Basic Math's

- M1- Introduction to Data Science
  - Data Science & its roles
  - Types of Learning
    - Supervised, Unsupervised, Semi-Supervised & Reinforcement learning
  - Future of AI & Data Science
- M2- Basic Math's
  - Statistics
  - Probability
  - Conditional Probability (Bayes' Theorem)
  - Hypothesis testing
    - Z-test
    - T-test
    - Chi-square test
  - Introduction to R programming
- Module M1 & M2 Exam

Machine Learning Algorithms

- M3- Machine Learning Algorithms
  - Linear Regression
  - Multiple Regression
  - Logistic Regression
  - Time Series
  - Decision Trees
  - Clustering Techniques
  - SVM, PCA,
  - Neural Networks
  - Boosting Algorithms & Ensemble Techniques

Programming Languages

- R Studio
  - Majorly used for statistical analysis (regression Time series)
  - Good Community support
  - ggplot, ggplot2 packages for visualization
- Python
  - Majorly used for Text mining, sentiment analysis, neural nets, Deep Learning
  - Pretty good community with huge packages
  - Ntlk, TextBlob and other various packages for Text mining
  - Tensorflow and Keras for Deeplearning

Project submission & viva

- Self Learning
- Doubt Clearance
- Live Project

Take Away

- Data Science?
- Machine Learning?
- Types of Learning?
  - Supervised
  - Unsupervised
  - Semi Supervised
  - Reinforcement learning
- What is the future of AI & ML
- Programming Languages
  - R
  - Python