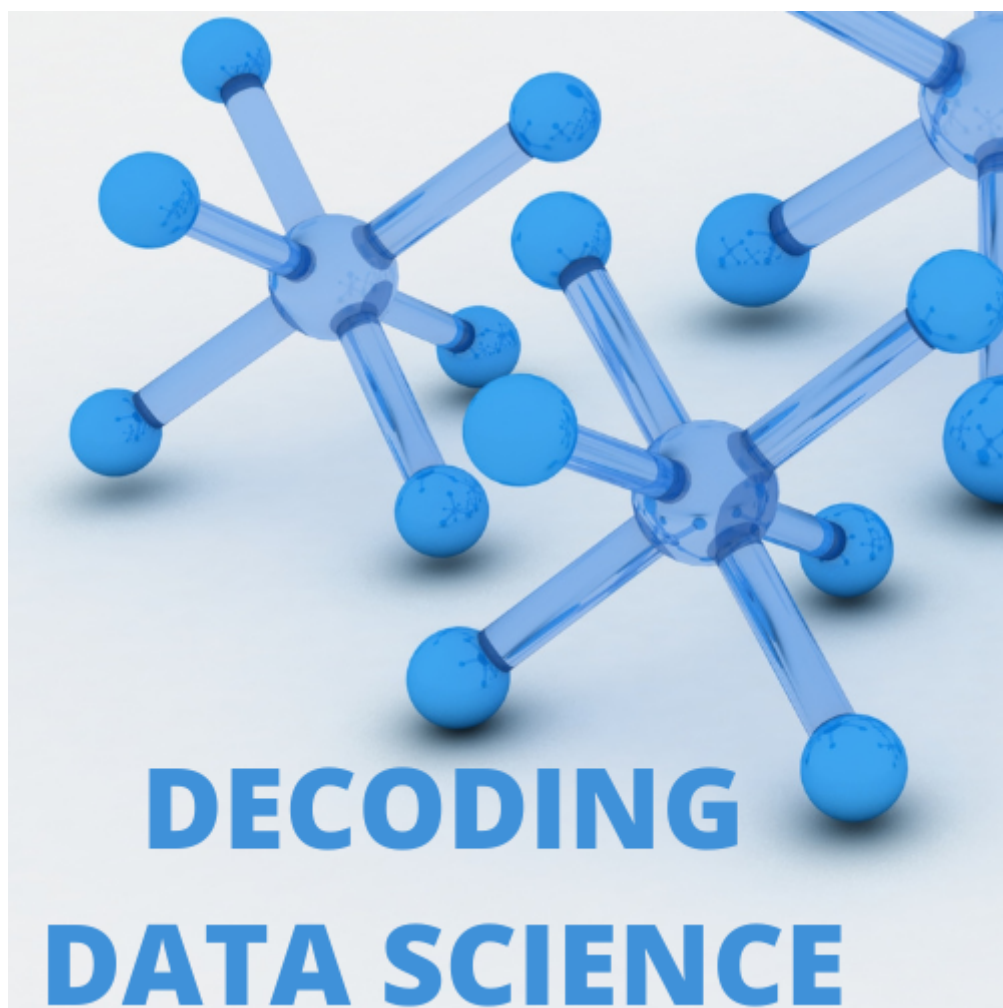


Data Science Road Map



Mathematics

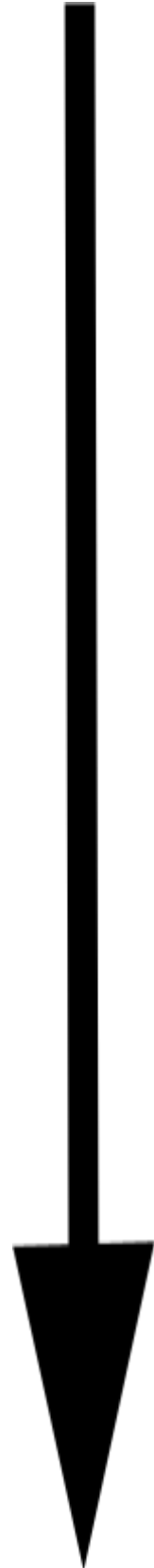
- Linear Algebra
- Analytics Geometry
- Matrix
- Vector Calculus
- Optimization
- Regression
- Dimensionality Reduction
- Density Estimation
- Classification

Probability

- Discrete Distribution
 - Binomial
 - Bernouli
 - Geometric etc
- Continuous Distribution
 - Uniform
 - Exponential
 - Gamma
- Normal Distribution
- Introduction to Probability
- 1D Random Variable
- Function of One Random Variable
- Joint Probability Distribution

Statistics

- Introduction to Statistics
- Data Description
- Random Samples
- Sampling Distribution
- Parameter Estimation
- Hypotheses Testing
- ANOVA
- Reliability Engineering
- Stochastic Process
- Computer Simulation
- Design of Experiments
- Simple Linear Regression
- Correlation
- Multiple Regression
- Nonparametric Statistics
 - Sign Test
 - The Wilcoxon Signed-Rank Test
 - The Wilcoxon Rank Sum Test
 - The Kruskal-Wallis Test
- Statistical Quality Control
- Basic of Graphs



Programming

Python

- **Python Basics**
 - List
 - Set
 - Tuples
 - Dictionary
 - Function, etc.
- NumPy
- Pandas
- Matplotlib/Seaborn, etc.

R

- **R Basic**
 - Vector
 - List
 - Data Frame
 - Matrix
 - Array, etc
- dplyr
- ggplot2
- TidyR
- Shiny, etc.

DataBase

- SQL
- MongoDB

Other

- **Data Structure**
 - Array, etc
- Web Scraping
- Linux
- Git

Machine Learning

Introduction

- How Model Works
- Basic Data Exploration
- First ML Model
- Model Validation
- Underfitting & Overfitting
- Random Forests
- scikit-learn

Intermediate

- Handling Missing Values
- Handling Categorical Variables
- Pipelines
- Cross-Validation
- XGBoost
- Data Leakage

Deep Learning

- Artificial Neural Network
- Convolutional Neural Network
- Recurrent Neural Network
- Keras
- PyTorch
- TensorFlow

- A Single Neuron
- Deep Neural Network
- Stochastic Gradient Descent
- Overfitting and Underfitting
- Dropout Batch Normalization
- Binary Classification

Feature Engineering

- Baseline Model
- Feature Generation
- Categorical Encodings
- Feature Selection

Natural language Processing

- Text Classification
- Word Vectors

Data Visualization Tools

- Excel VBA
- Bi (Business Intelligence)
 - Tableau
 - Power BI
 - Qlik View
 - Qlik Sense

Deployment

- Microsoft Azure
- Flask
- Heroku
- Django
- Google Cloud Platform

Other Points

- Domain Knowledge
- Case Studies
 - Data Science at Netflix
 - Data Science at Flipkart
 - Project on Credit Card Fraud Detection
 - Project on Movie Recommendation , etc.
- Communication Skill
- Reinforcement Learning

Keep Practicing

