Foundations of AI and ML

Introduction to Data Science and AI & ML

- Data Science, AI & ML
- Use Cases in Business and Scope
- Scientific Method
- Modelling Concepts
- CRISP-DM Method

Python

- Install python
- jupyter Notebook
- http://jupyter.readthedocs.io/en/latest/install.html
- Spyder
- https://pypi.python.org/pypi/spyder
- Introduction to Python
- Basic Python Syntax
- Data types
- https://docs.python.org/3/library/datatypes.html

Python Overview

Getting Started with Python

- About Interpreted Languages
- Advantages/Disadvantages of Python pydoc
- Starting Python
- Interpreter PATH
- Using the Interpreter
- Running a Python Script
- Python Scripts on UNIX/Windows, Editors and IDEs
- Using Variables
- Keywords
- Built-in Functions
- StringsDifferent Literals
- Math Operators and Expressions
- Writing to the Screen
- String Formatting
- Command Line Parameters and Flow Control

Sequences and File Operations

- Lists
- Tuples
- Indexing and Slicing
- Iterating through a Sequence
- Functions for all Sequences
- Using Enumerate()
- Operators and Keywords for Sequences
- The xrange() function
- List Comprehensions
- Generator Expressions
- Dictionaries and Sets

Deep Dive – Functions Sorting Errors and Exception Handling

- Functions
- Function Parameters
- Global Variables
- Variable Scope and Returning Values. Sorting
- Alternate Keys
- Lambda Functions
- Sorting Collections of Collections, Dictionaries and Lists in Place
- Errors and Exception Handling
- Handling Multiple Exceptions
- The Standard Exception Hierarchy
- Using Modules
- The Import Statement
- Module Search Path
- Package Installation Ways
- Packages
- Machine Learning Python libraries
- NumPy
- SciPy
- Pandas etc.

Regular Expressionist's Packages and Object – Oriented Programming in Python

- The Sys Module
- Interpreter Information
- STDIO
- Launching External Programs
- path directories and Filenames
- Walking Directory Trees
- Math Function
- Random Numbers
- Dates and Times
- Zipped Archives
- Introduction to Python Classes
- Defining Classes
- Initializers
- Instance Methods
- Properties
- Class Methods and Data Static Methods
- Private Methods and Inheritance
- Module Aliases and Regular Expressions

Debugging, Databases and Project Skeletons

- Debugging
- Dealing with Errors
- Using Unit Tests
- Project Skeleton
- Required Packages
- Creating the Skeleton
- Project Directory
- Final Directory Structure
- Testing your Setup

- Using the Skeleton
- Creating a Database with SQLite 3
- CRUD Operations
- Creating a Database Object.

Statistics

What is Statistics?

- Descriptive Statistics
- Central Tendency Measures
- The Story of Average
- Dispersion Measures
- Data Distributions
- Central Limit Theorem
- What is Sampling
- Why Sampling
- Sampling Methods
- Inferential Statistics
- What is Hypothesis testing
- Confidence Level
- Degrees of freedom
- what is pValue
- Chi-Square test
- What is ANOVA
- Correlation vs Regression
- Uses of Correlation and Regression

Machine Learning

Machine Learning Introduction

- ML Fundamentals
- ML Common Use Cases
- Understanding Supervised and Unsupervised Learning Techniques
- Clustering
- Similarity Metrics
- Distance Measure Types: Euclidean, Cosine Measures
- Creating predictive models
- Understanding K-Means Clustering
- Understanding TF-IDF, Cosine Similarity and their application to Vector Space Model
- Case study
- Implementing Association rule mining
- Case study
- Understanding Process flow of Supervised Learning Techniques
- Decision Tree Classifier
- How to build Decision trees
- Case study
- Random Forest Classifier
- What are Random Forests
- Features of Random Forest
- Out of Box Error Estimate and Variable Importance
- Case study
- Naive Bayes Classifier
- Case study
- Project Discussion

- Problem Statement and Analysis
- Various approaches to solving a Data Science Problem
- Pros and Cons of different approaches and algorithms
- Linear Regression
- Case study
- Logistic Regression
- Case study
- Text Mining
- Case study
- Sentimental Analysis
- Case study

Machine Learning Using Python

- Introduction to Machine Learning
- Areas of Implementation of Machine Learning
- Why Python
- Major Classes of Learning Algorithms
- Supervised vs Unsupervised Learning
- Learning NumPy
- Learning Scipy
- Basic plotting using Matplotlib
- Machine Learning application

Supervised and Unsupervised learning

- Classification Problem
- Classifying with k-Nearest Neighbours (kNN)

Algorithm

- General Approach to kNN
- Building the Classifier from Scratch
- Testing the Classifier
- Measuring the Performance of the Classifier
- Clustering Problem
- What is K-Means Clustering
- Clustering with k-Means in Python and an

Application Example

- Introduction to Pandas
- Creating Data Frames
- GroupingSorting
- Plotting Data
- Creating Functions
- Converting Different Formats
- Combining Data from Various Formats
- Slicing/Dicing Operations.