

2020

# VLR TRAINING

## Practical ML + DL + NLP with Python

By Dhamodhar

**WITH HANDS ON PROJECTS**

# Introduction

- Course Overview
- Installation of Anaconda
- Jupyter Notebook Basics
- DataSets

# Python Programming

- Operators
  - Arithmetic Operators
  - Comparison or Relational Operators
  - Logical or Boolean Operators
  - Bitwise Operators
  - Assignment Operators
  - Special Operators
- Math Library
- Variables
- Data Types
- Typecasting
- Booleans
- Strings
- Special Characters in a String
- Split and Strip a String
- Introduction to Lists
- Lists Slicing and Reverse Order
- Kinds of Lists
- Concatenate Strings Using join() method
- Add Lists

- Introduction to Dictionary
- Dictionary and It's Methods
- Nested Dictionary
- Create Dictionary Using zip() method
- Tuples
- Set
- If Condition
- While Loop
- Range() Method
- For Loop
- Reserve Keywords
- Built-In Functions
- User Defined Functions
- Anonymous or Lambda Functions
- File IO Operations

## Numpy

- Necessity of Numpy
- Creation & Metadata of Numpy Arrays
- Broadcasting
- Numpy Built-In Functions
- Data Types
- Typecasting
- Matrix Multiplication
- Change of Numpy Shape
- Numpy Slicing

- Boolean Indexing
- Filter Data
- Statistical Methods
- Sort, Min & Max of Numpy Arrays
- Stacking & Splitting
- Copy Vs. View

## Pandas

- Series
- DataFrame
- Metadata
- Rename Columns & Indices
- Transpose DataFrame
- Slice a DataFrame
- Boolean Indexing
- Missing Values
- Replace Values
- Search, Extract & Create NewColumns
- Set & Unset Index
- Built-In Customized Functions
- Value\_counts() Method
- Groupby() & Associated Methods
- Concat & Append
- Merge
- Reshape - Stack & Unstack
- Pivoting

- Melt
- Dummy Variables
- Crosstab() Method
- Upper, extract, replace & split Methods
- Regular Expressions
- Contains Method
- StartsWith Method
- Multiple String Method at a Time
- Manipulate Column Names
- Show Columns based on Keyword
- Read\_csv() method
- Tabbed File
- Fixed Width Files
- JSON Data
- HTML Data
- XML Data
- API
- Export DataFrame to CSV File
- Encoded Data Files
- Bad Data
- Select Columns Based on Datatype

## Time Series Analysis

- How to Convert Non-Timestamp To Timestamp
- Invalid Data
- Unix/Epoch Time
- Datetime Index
- Current Date Time

- Date\_range & bdate\_range Methods
- Pandas Slicing
- More components of Datetime
- Strftime() method
- Period Range
- Period
- Resample
- Handle TimeZone

## Matplotlib

- One Axis Plot
- Two Axis Plot
- Line Style & Color
- X and Y Limits
- Line Width
- Multiple Plots in One Chart
- Title, X & Y Labels
- Gridlines
- Annotations
- Ticks
- Spines
- Legend
- Subplots
- Line Plot
- Bar Graph
- Scatter Plot
- Area Plot
- Box Plot

- Histogram
- Pie Chart

## Seaborn

- Count Plot
- Box Plot
- Violin Plot
- Swarm Plot
- Overlaying Plot of Univariate Variables
- Facet Grid
- Lmplot & regplot
- Size & Shape of a Plot
- Pair Plot
- Join Plot
- Heat Map

## Statistics

- Types of Data
- Population Vs Sample
- Sampling Methods
- Branches of Statistics
- Distribution
- Variance Vs. Standard Deviation
- Z-Score
- Correlation
- Models
- Probability

## Machine Learning Basics

- Labelled Vs Unlabeled Data
- Types of ML Algorithms
- How ML Predict things
- Count Vectorizer
- Difference between fit and fit\_transform Methods
- Special & Numerical Characters
- Remove HTML Tags from Text Data
- Remove Stop words from Text
- Stemming
- Train Test Split
- Accuracy - MAE, MSE, RMSE & Variance Score

## Projects

- Simple Linear Regression
- Multiple Linear Regression
- Polynomial Regression
- Classification Algorithms
- Clustering Algorithms
- Univariate Timeseries Analysis



## Model Tuning

- Loss Functions
- Noise
- Penalty
- Dimensionality Reduction
- Principal Component Analysis
- $r^2$  score
- ROC & AUROC
- HyperParameters
- Project Skeletons

## Ensemble Algorithms

- Average
- Weighted
- Conditional
- Bagging
- Boosting

## GitHub Basics

## Principal Component Analysis (PCA)

- Data Scaling
- Covariance

- Eigen Values
- Eigen Vectors

## Natural Language Processing (NLP)

- Replace
- Spelling Correction
- Named Entity
- Parts of Speech (PoS)
- Text Cleaning
- N Grams
- Tokenization
- RegEx Stemmer
- Singular & Plural
- Translate
- TF-IDF

## NLP Projects

- Anomaly Detection
- Topic Modelling
- Sentiment Analysis
- Auto Tagging
- Spam Classification
- Text Generation - Deep Learning

## Amazon Web Services

- SageMaker

- Bucket Creation
- Regression - Deep Learning
- Classification – Deep Learning

(Near Jntu Hyderabad)

VLR Training

PlotNo 127/c,2nd floor,

Street Number 4,

Addagutta Society, Jal

Vayu Vihar, Kukatpally,

Hyderabad, Telangana

500072

Contact: +91 9059868766



<https://www.vlrtraining.in/>