Artificial Intelligence & Data Science Training
Course Agenda

S. No	Module	Topic Details
1	Introduction to Data Science & Machine Learning (ML) & Deep Learning(DL) & Artificial Intelligence(AI)	**Introduction to Data Science & Machine Learning (ML)** • Introduction to Data Science, ML, DL & AI - why is it so important? • Applications of Data science across industries • Business problems – Analytics scenarios • Analytics Industry in India, Job Market & Top Skills • Data science – CRIS DM Approach and DIPP framework • Data Scientist Toolbox, Tool of choice- Python: what & why? • Data Scientist - Tasks and Capabilities
2	SQL & DWH Concepts	Introduction to SQL & Data Warehouse Concepts • Introduction to Data Warehouse • Dimensions & Facts • Normalization & Schemas • Modelling • ETL
3	R programming	**Introduction to R Programming** • Data Structures • Data Visualization • Statistics for Data Science -1 • Statistics for Data Science -2 • Regression Analysis • Classification • Clustering • Association **Practice Assignment -1**
4	Python programming	**Introduction to Python** • Installation of Python framework and packages: Anaconda and pip • Writing/Running python programs using Spyder, Command Prompt • Working with Jupyter Notebooks • Creating Python variables: Numeric, string and logical operations • Basic Data containers: Lists, Dictionaries, Tuples & sets **Practice Assignment-2** **Operations & Functions in Python** • Writing for loops in Python • List & Dictionary Comprehension • While loops and conditional blocks • List/Dictionary comprehensions with loops • Writing your own functions in Python • Writing your own classes and functions as class objects **Practice assignment – 2A** **Numerical Summary of Data** • Summarizing numeric data and categorical data in pandas • Group wise summary of mixed data **Practice assignment – 2B** **Data Visualization using Python** • Need for visual summary • Introduction to Seaborn • Visual summary of different data combinations **Practice assignment – 2C** **Data Handling using NumPy and Pandas** • Introduction to NumPy arrays, functions &properties • Introduction to pandas, Data frame functions and properties

		• Reading and writing external data • Manipulating Data Columns **Practice assignment – 2D** **Regular expressions** • Introduction Regular expression – Data Preparation
5	Statistics and Linear Algebra	**Basics of Statistics** • Introduction to Univariate Statistics, Shape • Central Tendency and variability • Outliers • Correlation **Linear Algebra** • Introduction to Linear Algebra • Mathematics for Machine Learning • Vectors and Matrices • Matrices Operations • Applications to Data Problems
6	Machine Learning Basics	**Basics of Machine Learning** • Business Problems to Data Problems • Broad Categories of Business Problems • Supervised and Unsupervised Machine Learning Algorithm • Drivers of ML algorithms • Cost Functions • Brief introduction to Gradient Descent • Importance of Model Validation • Methods of Model Validation • Introduction to Cross Validation and Average Error
7	Machine Learning – Algorithms (Supervised Learning)	**Generalized Linear Models (Linear/Lasso/Ridge/Logistic)** • Linear Regression • Limitation of simple linear models and need of regularization • Ridge and Lasso Regression (L1 & L2 Penalties) • Introduction to Classification with Logistic Regression • Methods of threshold determination • Performance measures for classification score models **Case Study 1 – Linear Regression, Ridge, Lasso and Logistic Regression** **Practice assignment – 3** **Decision Trees & Random Forests** • Introduction to decision trees • Tuning tree size with cross validation • Introduction to bagging algorithm • Random Forests • Grid search and randomized grid search • Extra Trees (Extremely Randomized Trees) **Case Study 2 – DT and RF** **Practice assignment – 4** **Boosting Machines in Python** • Concept of weak learners • Introduction to boosting algorithms • Adaptive Boosting • Extreme Gradient Boosting (XGBoost) Case Study 3 – Boosting Machines **Practice assignment – 5** **K Nearest Neighbors** • Introduction to idea of observation-based learning • • Distances and Similarities • K Nearest Neighbors (KNN) for classification and Regression Case Study 4 - KNN **Practice assignment – 6** **Support Vector Machines**

		• Introduction to SVM for classification • Case Study 5 - SVM **Practice assignment – 7** **Neural Networks** • Introduction to Neural Networks • Single layer neural network • Multiple layer Neural network • Back propagation Algorithm • Neural Networks implementation in Python **Case study 6 - NN**
8	Machine Learning – Algorithms (Unsupervised Learning)	**Dimensionality Reduction** • Need for dimensionality reduction • Introduction to Principal Component Analysis (PCA) • Difference between PCAs and Latent Factors • Introduction to Factor Analysis **Case study 7 – PCA** **Case Study 8 - FA** **Segmentation in Python** • Patterns in the data in absence of a target • Segmentation with Hierarchical Clustering and Kmeans • Measure of goodness of clusters • Limitations of K-means • Introduction to density-based clustering (DBSCAN) **Case study 9 – K-Means** **Case study 10 - DBSCAN**
9	Web scraping & API	**Data collection with web scraping & APIs** • Gathering text data using web scraping with urllib • Processing raw web data • Interacting with Google search using urllib with custom user agent • Collecting twitter data with Twitter API • **Case study 11 – web scrapping** • **Case study 12 – API to extract Data**
10	Natural Language Processing	**Natural Language Processing (Text Mining)** • Quick Recap of string data functions and Introduction to Text Mining • Feature Engineering for text Data • Feature creation with TFIDF for text data **Case Study 13 – Text Data to model data** **Sentiment Analysis (NLP Supervised Learning) - Naïve Bayes/RF** • Introduction to Naive Bayes • Case Study 14 – Naïve Bayes Classifier using Text Data (SPAM/Not SPAM) **Case Study 15 – Sentiment analysis using Reviews Data** **Topic Modeling using LDA (NLP Unsupervised Learning)** • Introduction to Topic Modeling • Topic to word matrix and Document to topic matrix **Case Study 16 – LDA**
11	Ensemble Methods	**Ensemble Methods & Bokeh** • Making use of multiple ML models taken together • Simple Majority vote and weighted majority vote • Blending & Stacking **Case study 17 – ensemble method**
12	Big Data Analytics	**Big Data Analytics** • Big Data Hadoop Architecture, MapReduce • Apache Spark, PySpark, MLLib and Spark Tools • PySpark Integration with Jupyter Notebook
13	Version Control & Data Product	**Version control with Git & Interactive Data Product – prototyping solutions as Data Product** • Need and Importance of Version Control • Setting up git and github accounts on local machine

		• Creating and uploading GitHub Repos • Push and pull requests with GitHub App • Merging and forking projects • Pipeline and Pickle • Examples of static and interactive data products
14	AI & Deep Learning	**Deep Learning & Artificial Intelligence** • Installation of Tensor flow • Basics of Tensor flow with real-time project • Image classification with Tensor flow real-time project • Speech to Text with Tensor flow real-time project • OCR with OpenCv with real-time project • Object detection with Tensor flow real-time project • Deep Learning Concepts like Neural Networks, AI • image captioning with Tensor flow real-time project • Deep Learning: Searching for Images • Searching for images: A case study in deep learning • Learning very non-linear features with neural networks • Application of deep learning to computer vision • Deep learning performance • Demo of deep learning model on Image Net data • Deep learning ML block diagram • Deploying Tensorflow deep learning models in production • DNN/CNN/RNN • Building A chat bot with NLP **Case Study 18: Ecommerce product recommendation** **Case Study 19 :Chat Bot Implementation in Tensor Flow** **Case Study 20 :Chat Bot Implementation in Pytorch**
15	Business Analytics with Adv.Excel	**Introduction to Business Analytics** • Introduction • Introduction to business analytics • Formatting conditional formatting and logical functions • Analyzing data with pivot tables • Dashboarding • Business analytics with Adv.excel • Data analysis using statistics
16	Advanced Predictive Analysis	**Introduction to Advanced Predictive Analysis** • Data Understanding • Data Preparation • Data Transformation • Item sets and Associations
17	Story Telling Using BI	**Story Telling Using Business Intelligence (BI) Tool – Power BI/Tableau** • Introduction to BI and BI Tool • Exploratory Data Analysis (EDA) using BI Tool • Creating Dashboard using predictive model results **Case Study 21 – EDA on Sales Data**
18	Model Deployments Production in Cloud	• Model Building and Deployment in AZURE /AWS/GCP • Deep Learning Model Building and Deployment in AZURE/AWS/GCP

Other releated topics

Capstone Projects	Project -1	Project -2	Project -3
Interview Preparation	Resume Building	Mock Interviews	Interview Questions