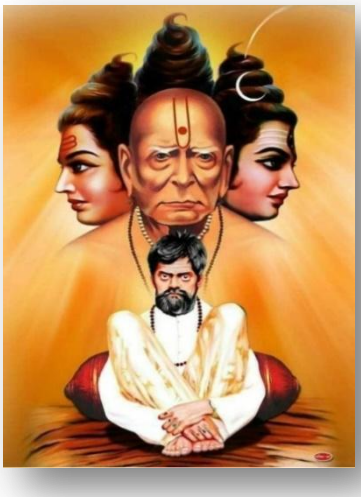


|| Jai Shankar||

PerfectLearn Software Testing Institute

----------Get UR Dream Job with our small support----------

Data Science and Machine Learning

---------- By Projects only----------

WHO CAN LEARN?

- Aspirants looking for permanent jobs in big IT corporations
- Students/scholars willing to start their professional career as a “Test Engineer” with any of the IT giants
- Proficient IT professionals who are willing to switch their field of work towards Testing field
- Aspirants with additional IT skills and educational background who are willing to gain practical freelancing experience in the related field.

PRE-REQUISITES?

- Any Graduate can enter in IT field like:

  Any engineering degree,B.Com,M.com,B.A.,BCA,MCA,BCS,MCS,MBA,BBA etc.
- No need of any programming language knowledge, only basic computer knowledge is essential.

TRAINING OUTCOME?

- Confidence on Automation software testing tool and manual software testing to crack interview and to work efficiently in company as QA.
- No need to view other videos or topics to take extra knowledge we take care of all.
- Programming language i.e. Core Java for selenium, student also perfect in that.
- Any student can work in automation testing at any level in company.
- In manual software Testing tool –You are perfect in to analyzing documents, to prepare test scenario, to prepare test cases and to execute test cases on given build. Also bug logging and bug tracking process.
- In Automation software testing tool- You are perfect in Selenium, maven, Jenkins, GIT, TestNG, core java etc. Also you are work effectively in automation framework as like experienced QA.

DATA SCIENCE AND MACHINE LEARNING

Context of the Study……

Data science is an area of math and science that is still growing. We use mobile phones, laptops, different digital platforms, credit cards, debit cards, social media, banks, transactions in different shops, our opinions and comments on social media platforms, hospitals, insurance, our purchases, and a lot more. These devices and platforms can keep track of your information if you give them permission to do so. Today, we have better ways to store data and can save it so we can look at it again later. This led to the field of data science, and now many businesses are interested in data and how to analyze it.

- This is a field that has been around for a long time, but computers have changed it. In data science, there are many jobs, such as Data Analyst, Data Scientist, Expert in Machine Learning, Database Management (using SQL), and Expert in Data Visualization with Tableau, Power BI, Business Analyst, and many more.
- This field has many uses, such as Real-time chat bot agents, Decision support, Customer recommendation engines, Customer churn modeling, Dynamic pricing tactics, Market research and customer segmentation, Fraud detection, Natural Language Processing (NLP) etc.
- This is also useful for people those who are doing their higher studies. Lots of researches are going on in this field. For Doctorial work data science and regression techniques are really very useful.
- There are many interesting jobs for recent college graduates and people who have worked in the same field before. You will definitely have a successful career in data science if you work out and learn a lot.
- MNCs and small IT companies are recruiting continuously for the above said profiles. Some companies are also hiring for paid internships through them.

Who Can Learn?

- Aspirant looking for permanent jobs in Big IT corporations
- Students/Scholars willing to start their professional career as a Data Analyst, Data Scientist, Expert in Machine Learning, Expert in Data Visualization with Tableau, Power BI, Business Analyst etc.
- Proficient IT professionals who are willing to switch their field of work towards Data Science Field

- Aspirants with additional IT skills and educational background who are willing to gain practical free lancing experience in the related field
- PhD Research Scholars /M. Tech. / B. Tech. students can learn for academic purpose and for their own experimental data analysis
- Anyone who has good knowledge of sensors, IoT, networking, image processing etc.
- Prerequisites:
- Enthusiasm towards learning new techniques
- Basic Math, probability, very basic knowledge of algorithms

Preferred skills:

- These are not mandatory skills. If you know it then it will be an advantage for learning.
- Python programming
- Any programming language knowledge(for logic developmemts)

Training Outcome

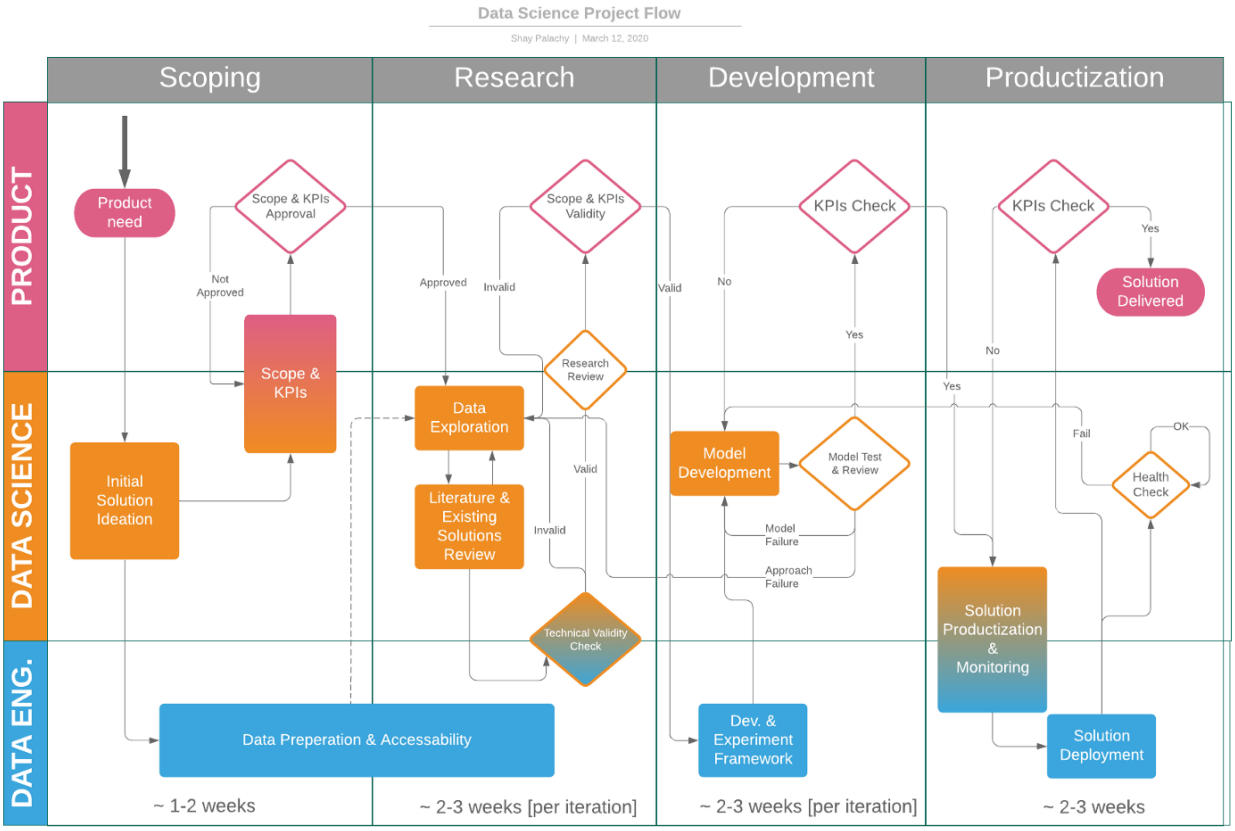
After completion of course learner will be able to:

- understand and solve exact business /research related data science problem
- acquire sufficient data science skills relevant to industry
- fetch data from different websites using web scrapping and apply end to end case studies related to data science and machine learning projects using software tools
- work efficiently and confidently in data driven projects
- deploy machine learning models in production

Data Science Syllabus

Data Science and Machine Learning project implementation in industry

Following is the general idea of how the project flows in industry. In this the project is divided in verticals and horizontals.



Syllabus is framed by considering above project idea.

Part 1 Python for data Science

Almost all industry asks for python programming knowledge for DSML. Following points will be covered in this module.

- Basics of Python
- Why anaconda and python? Different IDEs, their usage, Jupyter over other IDEs, Superiority of Python over other languages, Alternatives for Python
- Arithmetic operations in python
- All arithmetic operations, array, matrix creation, Set, Tuples, Lists their operation and usages
- Types of variables in python
- Data types in python, conversion on one to other data type, variable declaration, local and global variables
- Loops, if-else statements
- If-else, elif, while, for loops, try except statements
- Functions
- Types of functions, Creation and calling a function,
- Libraries in Python: Numpy, Pandas
- 1 D, 2D, 3D and n dimensional array creation and operations on it using Numpy, Data Manipulation using Pandas, Data Frame creation and operations related to dataframe
- Basic Data Visualization in python using Seaborn, Matplotlib and plotly
- All types of graphs, Univariate and bivariate analysis

Pedagogy

All above topics are prepared and practiced through assignments, quizzes and Project using python Jupyter as a tool

Part 2: Decision Based Statistics

Sometimes data science is also known as advanced statistics. Most of the business decisions are taken on the basis of data statistics. Prediction techniques and model evaluation is done on the basis of concepts of statistics only. So it becomes important to study decision based statistics. Following points will be covered in the second module.

- Terminologies of Statistics
- Probability and distributions
- Measures of Centers, Measures of Spread Probability
- Normal Distribution
- Binary Distribution
- Hypothesis Testing
- Chi-Square Test ,P-test, Z-test
- ANOVA

Pedagogy

All above topics are prepared and practiced through assignments, quizzes and Project using python Jupyter as a tool

Part 3: Exploratory Data Analysis (EDA)

It is one of the important primary stages before going for the Machine learning model fitting on a given data set. It takes almost 80% of the whole cycle because it includes data cleaning and data interpretation. Following topics will be covered in this module

- In depth Data Visualization
- Data Preprocessing and cleaning
- Imputation of null values
- Mean, Mode, Median
- Understanding data skewness
- Label encoding, One hot encoding, Dummy variables
- Heat map, correlation plot
- Pandas Profiling

Pedagogy

- All above topics are prepared and practiced through assignments, quizzes and Project using python Jupyter as a tool

Part 4: Core Machine Learning and Data Mining techniques

Main task of machine learning is to find hidden patterns and to predict. Following techniques namely supervised and unsupervised learning is core of ML. These are the basic algorithms of ML

Unsupervised Learning

- PCA (Principal Component Analysis)
- Clustering Techniques (K-Nearest Neighbor, Dendograms)

Supervised Learning

A. Classification

- Decision Tree
- Random forest
- Neural Networks
- Support Vector Machine
- Linear Discriminate Analysis
- Naïve Bays Model
- Ensemble Techniques
- Boosting and bagging
- XG and ADABOOST
- Deep Learning
- Feature selection

B. Regression

- Linear Regression, Lasso, Ridge and ElasticNet
- Linear Discriminate Analysis
- Logistic Regression
- Other Regressors

Bias Variance Trade off

- For all above models check classification matrix and best model selection

Pedagogy

All above topics are prepared and practiced through assignments, quizzes and Project using python Jupyter as a tool.

Part 5: Kaggle competitions

There are many competitions and Hackathon continuously going on the website known as www.kaggle.com.It is one of the favorite data science website. This gives a plat form to participate all over the globe. It is also famous for free data sets and codes. In this regards competition and ranking ideas will be given to create the profile on Kaggle. Along with Git Hub tutorials will also be provided.

Part 6: Model Deployment

Selecting Best model after end to end case study is then deployed in actual practice known as production deployment. It is done on many platforms. **Heroku and flask** are two of those platforms. These platforms will be covered in last module

Other Content if needed on demand

Additional Contents for fast learners

- Hadoop
- SQL
- Tableau
- KNIME
- Web Scrapping
- Model Deployment
- PySpark, Pytorch
- Open CV
- and many more