DATA SCIENCE WITH PYTHON

COURSE CURRICULUM

ENROLL NOW

CALL US
+91 630 171 9270

Course Name	Data Science with Python
Trainer Name	Udaykumar Gajavalli

Course Description:

In an era where data drives decision-making across every industry, Data Science has emerged as a vital discipline for extracting actionable insights from complex datasets. This course offers a comprehensive introduction to Data Science using Python, one of the most powerful and widely used programming languages in the field today.

Designed for both beginners and those with some programming background, the course covers core topics such as data wrangling, exploratory data analysis, statistical modeling, and machine learning. Students will gain hands-on experience with industry-standard Python libraries like Pandas, NumPy, Scikit-learn, and Matplotlib, learning how to turn raw data into meaningful insights and predictive models.

Beyond technical skills, the course emphasizes practical applications and ethical considerations in real-world scenarios—from data-driven decision-making in business to responsible AI in healthcare and finance. As the field continues to evolve, students will also explore emerging trends such as AutoML, responsible AI practices, and the growing role of generative models in data science workflows.

By the end of the course, students will have built a strong foundation in both the theory and application of Data Science with Python, equipping them to tackle data challenges across diverse industries and prepare for advanced study or career paths in this rapidly expanding field.

Pre-requisites:

- Basic Programming Knowledge: Familiarity with fundamental programming concepts such as variables, loops, functions, and data types. Prior experience with Python is recommended but not mandatory.

- Introductory Mathematics: Understanding of basic statistics and linear algebra concepts, including mean, median, standard deviation, probability, and matrices.

- Logical Thinking & Problem-Solving Skills: Ability to approach problems methodically and interpret data-driven questions.

Learning Outcomes:

- LO1: Develop a solid foundation in core data science concepts, including data preprocessing, exploratory data analysis, statistical inference, and machine learning, using Python and its key libraries (e.g., Pandas, NumPy, Scikit-learn, Matplotlib).

- LO2: Gain hands-on experience in designing and implementing end-to-end data science workflows—covering data collection, cleaning, modeling, and visualization—while applying best practices in reproducibility, performance evaluation, and interpretability.

- LO3: Analyze real-world datasets across diverse domains, apply predictive modeling techniques, and critically evaluate the ethical implications of data-driven decision-making, with a focus on fairness, bias mitigation, and responsible AI practices.

Course Duration: 180+ Hours Approx

Session	List of Topic Title
1	**Getting started** • Welcome to the course • Anaconda installation • Getting started with VS code
2	**Python programming language** • Getting started with vs code • Different ways of creating virtual environment • Python Basics-Syntax And Semantics • Variables In Python • Basics Data Types In Python • Operators In Python
3	**Python Control flow** • Conditional Statements (if, elif, else) • Loops In Python
4	**Inbuilt Data Structures in python** • List And List Comprehension In Python • Tuple In Python • Sets In Python • Dictionaries In Python • Real world Usecases Of List
5	**Functions in python** • Getting Started With Functions • Lambda Function In Python • Map Function In Python • Filter Function In Python

6	**Importing and creating a module in python** • Import Modules And Packages In Python • Standard Library Overview
7	**File Handling** • File Operation In Python • Working With File Paths
8	**Exception handling in python** • Exception Handling With try except else and finally blocks
9	**OOPS Concepts with Classes and Objects** • Classes And Objects In Python • Inheritance In OOPS • Polymorphism In OOPS • Encapsulation In OOPS • Abstraction In OOPS • Magic Methods In Python • Operator Overloading In Python • Custom Exception Handling
10	**Advance Python** • Iterators In Python • Generators With Practical Implementation • Function Copy, Closures And Decorators
11	**Data Analysis with Python** • Numpy In Python • Pandas- DataFrame And Series • Data Manipulation With Pandas And Numpy • Reading Data From Various Data Source Using Pandas • Data Visualization With Matplotlib • Data Visualization With Seaborn

12	**Getting starting with Statistics** • What is Statistics And its Application • Types Of Statistics • Population Vs Sample Data • Measure Of Central Tendency • Measure Of Dispersion • Why Sample Variance Is Divided By n-1? • Standard Deviation • What Are Variables? • What are Random Variables • Histograms- Descriptive Statistics • Percentile And Quartiles- Descriptive Statistics • 5 Number Summary-Descriptive Statistics • Correlation And Covariance
13	**Introduction to Probability** • Addition Rule (For Mutual And Non Mutual Exclusive Events) • Probability-Multiplication Rule(Independent And Dependent Events)
14	**Probability Distribution Function** • The Relationship Between PDF,PMF And CDF • Types Of Probability Distribution • Bernoulli Distribution • Binomial Distribution • Poisson Distribution • Normal/ Gaussian Distribution • Standard Normal Distribution And Z Score • Uniform Distribution • Log Normal Distribution • Power Law Distribution8 • Pareto Distribution • Central Limit Theorem • Estimates

15	**Inferential Statistics** • Hypothesis Testing And Mechanism • What is P value? • Z Test- Hypothesis Testing • Student t Distribution • T Stats With T test Hypothesis Testing • Z test Vs T test • Type 1 And Type 2 Error • Bayes Theorem • Confidence Interval And MArgin Of Error • What is Chi Square Test • ChiSquare Goodness OF Fit • Annova Test • Assumptions Of Annova • Types Of Annova • Partioning Of Variance In Annova
16	**EDA and Feature Engineering** • Red Wine Dataset EDA • EDA And Feature Engineering Flight Price Dataset • Data Cleaning With Google Playstore Dataset • Part 2- EDA Google Play Store Cleaned Dataset
17	**Introduction to machine learning** • Introduction • Types of Machine learning • Equation of Line, 3d, and Hyperplane • Distance of a point from a plane • Instance based Vs Model based learning

18	**Linear Regression** • Introduction to Simple Linear Regression • Understanding Simple Linear regression Equations • Cost Function • Convergence Algorithm • Convergence Algorithm Part02 • Multiple Linear regression • Performance Metrics • MSE, MAE, RMSE • Overfitting and Underfitting • Linear Regression with OLS • Simple Linear Regression Practical • Multiple Linear regression • Polynomial Regression Intuition • Polynomial Regression Implementation • Pipeline in Polynomial • Handling Missing Values • Handling Imbalanced Dataset
19	**Regularization Algorithms** • Ridge Regression • Lasso G ElasticNet • Types Of cross Validation • Cleaning the Dataset • EDA and Feature Engineering • Feature Selection • Model Training • Hyperparameter tuning
20	**Implementation of project with Life cycle of ML** • Basic Simple Linear Regression Project • Multiple Linear Regression Projects With Assumptions • Basic Regression Project From Scratch-EDA And Feature Engineering • Model Training With Cross Validation Using Lasso Regression • Model Training With Ridge and Elastic net With Cross Validation • Model Pickling In ML Project • End To End ML Project Implementation

21	**Logistic Regression** • Can Linear Regression Solve Classifier Problem. • Logistic Regression Indepth Math Intuition • Performance Metrics • Logistic Regression OVR • Logistic Regression Implementation • Grid Search Hyper Parameter • Randomised Search CV • Logistic OVR • Logistic Imbalanced Dataset • Logistic Regression ROC
22	**Support Vector Machine** • Introduction to support vector Machine • SoftMargin and Hard Margin • SVM Maths Intuition • SVC Cost function • Support Vector Regression • SVM Kernels • Support Vector Classifiers • SVM Kernels implementation • Support Vector Regression Implementation
23	**Naïve Bayes Algorithm** • Understanding Baye's Theorem • Variants Of Naive Baye's • Naive Baye's Practical Implementation

24	**KNN** • KNN Classification And Regression Indepth Intuition • Optimization Of KNN- KDtree And Ball Tree Indepth Intuition • KNN Classifier And Regressor Classification
25	**Decision Tree classifier and Regressor** • Introduction TO Decision Tree. • Entropy and Gini Impurity • Information Gain • Entropy vs Gini impurity • Decision Tree Split for Numerical Features • Post Pruning G Pre Pruning • Decision Tree Regression • Decision Tree Implementation • Decision tree Prepruning • Diabetes Prediction Using Decision Tree Regressor
26	**Random Forest** • Bagging G Boosting Ensemble Techniques. • Random Forest Regression • Problem Classification • Feature Engineering Part 01 • Feature Engineering Part 02 • Model Training Step • Random Forest Regression Project-Problem Statement • Feature Engineering • Model Training

27	**Adaboost** • Introduction to Adaboost ML algorithm • Creating Decision Tree Stump • Performance of Decision Tree Stump • Updating Weights. • Normalising Weights and Assigning Bins • Selecting New Datapoints for Next tree • Final Prediction for Adaboost • Adaboost Model Training • Adaboost Regressor Model Training
28	**Gradient Boosting** • Gradient Boosting Regression • Gradient Boost Classifier Training • Gradient Boost Regression Model Training
29	**Xgboost** • Xgboost Classification Indepth Intuit • Xgboost Regressor • Model Training Xgboost. • Xgboost Regressor Training
30	**Unsupervised Learning** • Introduction To Unsupervised Machine Learning
31	**PCA** • Curse Of Dimensionality • Feature Selection and Extraction • PCA Geometric Intuition • PCA Maths Intuition 01 • Eigen Decomposition on Covariance Matrix • PCA Implementation

32	**KMeans** • Kmeans Clustering Geometric Intuition • How to Find K Values? • Random Initialisation Trap(Kmeans++) • K means Clustering Implementation
33	**Hierarchical Clustering** • Hierarichal Clustering • Agglomerative Clustering Implementation • Kmeans vs Hierarical Mean Clustering
34	**DBSCAN** • How DBSCAN Works? • Examples After Applying DBSCAN • Pros and Cons Of DBSCAN • DBSCAN Clustering Implementation
35	**Anomaly Detection** • Anomaly Detection Using Isolation Forest Indepth Intuition • DBSCAN Clustering Anomaly Detection • Local Outlier Factor Anomaly Detection • Zscore • IQR • Isolation Forest

36	**NLP** • Roadmap to Learn NLP for Machine Learning • Practical Use cases of NLP • Tokenisation and Basic Terminologies • Tokenisation Practicals • Text Preprocessing Stemming using NLTK • Text Preprocessing Lemmatization NLTK • Text Preprocessing Stopwords • Parts of Speech Tagging Using NLTK • Named Entity Recognition • What's Next? • One Hot Encoding Intuition • Advantages and Disadvantages of OHE • Bag of Words Intuition • Advantages and Disadvantages BOW • BOW implementation using NLTK • N Grams • Gram BOW Implementation Using NLTK • TF-IDF Instituion • Advantages and Disadvantages of TF-IDF • TFIDF Practical implementation Python • Word Embeddings • Word2Vec Intuition • Word2Vec Cbow Intuition • SkipGram Indepth Intuition • Advantages of Word2Vec • AvgWord2vec Indepth Intuition • Word2vec Practical Implementation Gensim • Spam ham Project using BOW • Spam And Ham Project Using TFidf • Best Practises For Solving ML Problems • Part 1-Text Classification With Word2vec And AvgWord2vec • ff.Part 2- Text Classification With Word2vec And AvgWord2vec • Part 1-Kindle Review Sentiment Analysis • Part 2- Kindle Review Sentiment Analysis

37	**Deep Learning** • Introduction • Why Deep Learning is getting Popular? • 3 - Perception Intuition • Advantages and Disadvantages of Perceptron • ANN Intuition and Learning • Back Propogation and Weight Updation • Chain Rule of Derivatives • Vanishing Gradient Problem and Sigmoid • Sigmoid Activation Function • Sigmoid Activation Function 2.0 • Tanh Activation Function • Relu activation Function • Leaky Relu and Parametric Relu • ELU Activation Function • Softmax For Multiclass Classification • Which Activation Function To Apply When? • Loss Function Vs Cost Function • Regression Cost Function • Loss Function Classification Problem • Which Loss Function To Use When? • Gradient Descent Optimisers • SGD • Mini Batch With SGD • SGD With Momentum • Adagard • RMSPROP • Adam Optimiser • Exploding Gradient Problem • Weight Initialisation Techniques • Dropout Layers • CNN Introduction ff.Human Brain Vs CNN • All you need to Know about Image • Convolution Operation In CN • Padding In CNN • Operation Of CNN Vs ANN • Max, Min and Average Pooling • Flattening and Fully Connected Layers • CNN example with RGB

38	**RNN Indepth Intuition** • Understanding RNN Architecture- RNN Vs ANN • Forward Propogation With Time In RNN Training • Backward Propogation With Time In RNN Training • Problems With RNN
39	**LSTM and GRU** • Why LSTM RNN? • LSTM RNN Architecture • Forget Gate In LSTM RNN • Input Gate And Candidate Memory In LSTM RNN • Output Gate In LSTM RNN • Training Process In LSTM RNN • Variants Of LSTM RNN • GRU RNN Complete Indepth Intuition
40	**Bidirectional RNN Architecture and Intuition**
41	**Encoder and Decoder** • Intuition Of Encoder And Decoder-Sequence to Sequence Architecture • Problems With Encoder And Decoder
42	**Attention Mechanism Indepth Architecture Explanation**
43	**Transformers** • Plan Of Action • What And Why To Use Transformers • Understanding the basic architecture of transformers • Self Attention Layer Working • Multi Head Attention • Feed Forward Neural Network With Multi Head Attention • Positional Encoding Indepth Intuition • Layer Normalization • Layer Normalization Examples • Complete Encoder transformer architecture • Decoder Transformer- Plan Of Action • Decoder Transformer- Masked Multi Head Attention Working • Encoder Decoder Multi Head Attention • Final Decoder Linear And Softmax Layer

44	**End to End Projects:** • Sentiment Analysis • Regression • Classification • Next Word prediction • Topic Modeling
45	**Time Series Forecasting** • What Is Time Series Data and Why It Matters • Key Components of Time Series (Trend, Seasonality, Noise, Cyclicity) • Understanding Stationarity and Its Importance • Time Series Decomposition Techniques • Lag Features and Windowing Concepts • Autocorrelation and Partial Autocorrelation (ACF & PACF) • Classical Forecasting Methods (Moving Average, Exponential Smoothing) • AR, MA, ARMA, and ARIMA Models Explained • SARIMA for Seasonal Time Series • Diagnostics and Residual Analysis • Feature Engineering for Time Series Forecasting • Introduction to Machine Learning for Time Series • Using Models like Random Forest and XGBoost for Forecasting • Introduction to Deep Learning for Time Series • LSTM Architecture and Why It Works for Sequential Data • Building an LSTM Model for Forecasting • Encoder-Decoder LSTM for Multi-Step Forecasting • Using CNNs for Time Series Forecasting • Evaluating Forecasting Models (MAE, RMSE, MAPE, etc.) • Forecasting with Facebook Prophet • Time Series Cross-Validation Techniques • Handling Anomalies and Outliers in Time Series • Deploying Time Series Models in Production

46	**Intro of Functional Areas** • Price Optimization • Marketing Mix Modeling (MMM) / Media Mix Modeling • Price Sensitivity Analysis • Price Elasticity Measurement • Cannibalization Analysis • Market Basket Analysis • Recommendation Engine Development • Customer Segmentation • Markdown Optimization • Anomaly Detection • Fraud Detection • Churn Analysis • Risk Analysis • Share of Wallet Estimation • Customer Lifetime Value (CLTV) Prediction • Propensity Modeling • Promotion Effectiveness Analysis • Predictive Modeling
47	**Resume Preparation & Prepare for Interview**

Thank you

INFOTEKH

Thank You for Going Through Data Science with Python Curriculum
We hope this guide has provided a clear and structured learning path to strengthen your skills in Data Science with Python.

📌 **NEXT STEPS**

- Start practicing with real-world use cases and hands-on exercises
- Build personal or client-based projects for your portfolio
- Keep exploring updates and best practices in the industry
- Join discussions and stay connected with the community

📞 Need Help or Guidance?

Feel free to contact our course support team:

Course Coordinator

GS Infotekh

📧 **contact@gsinfotekh.com**

🌐 **www.gsinfotekh.com**

📞 **+91 630 171 9270**