Data Science with ML & MLOps Course in Chennai

Learn Data Science, Machine Learning and MLOps in Chennai with Real Projects and Placement Support.

A 33-day industry-focused course that takes you from data to production models. Learn SQL, Python, NumPy and Pandas, process big data with PySpark, build cloud data pipelines on AWS and Snowflake, train machine learning models with Scikit-learn, and run them in production with MLOps on Amazon SageMaker, Step Functions and CloudWatch.

Your Trainers

Gowri Shankar — AWS Data Engineering, MLOps & AI Engineer
8+ Years of Industry Experience

Course Curriculum

Phase 1: SQL Foundations 7 Days

Query, combine and optimise data with SQL.

1. SQL & DBMS Intro Day 1

- SQL introduction, DBMS vs RDBMS
- DDL, DML, DQL, TCL and DCL commands

2. Constraints & Operations Day 2

- Data types and constraints: PK, FK, UNIQUE
- Create/drop tables and import data

3. Filtering & Aggregations Day 3

- WHERE clause and logical operators
- ORDER BY, LIMIT and aggregate functions

4. Joins & Set Operations Day 4

- Inner, left, right and full joins
- UNION, INTERSECT and EXCEPT

5. Functions & Subqueries Day 5

- Subqueries in clauses
- Text, date/time and window functions

6. Advanced Objects Day 6

- CTEs, views and indexing strategies
- Stored procedure basics

7. SQL Optimization Day 7

- Sample interview questions
- Query optimization tips

Phase 2: Python, NumPy & Pandas 5 Days

Program in Python and manipulate data with NumPy and Pandas.

1. Python Basics & Syntax Day 8

- Setup, data types, variables and operators
- Simple programs and I/O

2. Data Structures Day 9

- Strings, lists, tuples, sets and dictionaries
- Hands-on exercises

3. Functions & Files Day 10

- User-defined and anonymous functions
- File I/O and regex basics

4. NumPy & Pandas Day 11

- Arrays, indexing and broadcasting
- Pandas Series and DataFrames

5. Advanced Analytics Day 12

- Missing-data and duplicate handling
- Visualization, merging, joining and pivot tables

Phase 3: Big Data & Cloud Data Engineering 5 Days

Process big data with PySpark and build data pipelines on AWS and Snowflake.

1. PySpark Architecture & Core Mechanics Day 13

- Introduction to Big Data and Spark evolution
- RDD concepts: immutability, partitioning and lazy evaluation
- DAG and execution hierarchy: jobs, stages and tasks

2. PySpark Data Ops & Performance Day 14

- SparkSession and data ingestion
- Transformations, filtering, SQL functions and window analytics
- Caching, persistence, broadcast joins and AQE
- Hands-on data processing lab

3. Storage & IAM Day 15

- IAM security and S3 storage

- Snowflake modern warehouse architecture

4. Compute & Monitoring Day 16

- AWS Lambda serverless functions
- CloudWatch Logs and EventBridge automation

5. ETL & Warehousing Day 17

- AWS Glue: crawlers, jobs and pipelines
- Snowflake data loading and warehouse concepts

Phase 4: Machine Learning Engineering 7 Days

Build, evaluate and tune machine learning models with Scikit-learn.

1. ML Fundamentals Day 18

- Supervised and unsupervised learning
- Training, validation and test data; evaluation concepts

2. Regression Basics Day 19

- Supervised learning and linear regression
- Metrics and evaluation

3. Decision Trees Day 20

- Trees and ensembles
- Random forests

4. Classification & Ensemble Learning Day 21

- Classification concepts and evaluation
- Ensemble model comparison

5. Clustering Day 22

- K-Means and DBSCAN techniques
- Clustering use cases

6. Feature Engineering Day 23

- Feature preparation and selection
- PCA and dimensionality reduction

7. ML Pipelines Day 24

- Hyperparameter tuning
- Scikit-learn pipelines and model selection

Phase 5: MLOps on AWS 7 Days

Train, version, deploy, monitor and retrain models in production with Amazon SageMaker.

1. MLOps Fundamentals Day 25

- ML lifecycle and reproducibility
- Model and dataset versioning

- Training, validation and monitoring workflow

2. SageMaker Training Day 26

- SageMaker training jobs
- Training data, configurations, evaluation and artifacts

3. SageMaker Pipelines Day 27

- Pipeline steps and parameters
- Data processing, training and evaluation workflow
- Automated pipeline execution

4. Model Registry & Versioning Day 28

- Model Registry
- Model versions, metadata and approval workflow
- Model lifecycle management

5. Monitoring & Data Drift Day 29

- Data quality and model monitoring
- Data drift and performance monitoring
- CloudWatch metrics, logs and alerts

6. MLOps Automation Day 30

- AWS Step Functions
- EventBridge-driven ML workflows
- Retries, failure handling and workflow monitoring

7. Production ML Operations Day 31

- Model retraining strategy
- Training-data and model validation
- Auditability and operational best practices

Phase 6: Projects 2 Days

Bring the whole course together in an end-to-end MLOps project and a capstone.

1. End-to-End MLOps Project Day 32

- Data validation → feature engineering → training → evaluation → registry → monitoring → retraining
- AWS S3, SageMaker, CloudWatch and workflow orchestration

2. Capstone Project Day 33

- End-to-end project integrating SQL, Python, Pandas, PySpark, AWS/Snowflake, machine learning and MLOps

Projects & Hands-On Labs

Big Data Processing with PySpark
Ingest a dataset with SparkSession, apply transformations, SQL functions and window analytics, and

tune the job with caching, broadcast joins and AQE.

Cloud ETL Pipeline on AWS & Snowflake
Store raw data in S3, catalogue and transform it with AWS Glue crawlers and jobs, automate with Lambda and EventBridge, and load the result into Snowflake.

End-to-End MLOps Pipeline
Build the full lifecycle: data validation, feature engineering, training, evaluation, Model Registry, monitoring for data drift and automated retraining on AWS.

Data Science & MLOps Capstone
An end-to-end project integrating SQL, Python, Pandas, PySpark, AWS/Snowflake, machine learning and MLOps into one production-style solution.

Enrol & Contact

Book a free demo class or a 1-on-1 counselling session.
Call / WhatsApp: +91 89399 15577
Email: contact@greenstechnologys.com
Website: www.greenstechnologys.com
25 centres across Chennai, Bengaluru, Coimbatore, Trichy and Salem, plus live online batches. Curriculum may be updated by the trainer to reflect current industry practice.