

Data Science with Artificial Intelligence

Complete Professional Certification Program

Practical • Job-Oriented • Video-Based LMS Training • Industry-Ready

4.9 Rating | 18,500+ Students Enrolled | Updated May 2026

■ 30 Video Lessons	■ 65+ Hours Total	■ 60+ Hands-on Labs	■ 25+ Quizzes	■ 8+ Certifications	■ Lifetime Access
--------------------	-------------------	---------------------	---------------	---------------------	-------------------

■ WHAT YOU WILL LEARN

■ Master Python, Pandas, NumPy & SQL from scratch	■ Build & deploy real Machine Learning models
■ Perform complete EDA & data visualisation	■ Deep Learning with TensorFlow & PyTorch
■ NLP, Transformers & Hugging Face models	■ Generative AI, LLMs & LangChain applications
■ Build end-to-end AI projects for your portfolio	■ Deploy models with FastAPI, Docker & Cloud
■ Crack top certifications: Google, IBM, AWS, TensorFlow	■ Get hired with resume, interview & portfolio prep

■ THIS COURSE INCLUDES

■ 30 in-depth video lessons (HD, 1–3 hrs each)	■ 40+ downloadable PDF cheat sheets & notebooks
■ 60+ hands-on coding labs & mini-projects	■ 25+ quizzes after every section
■■ Access on Mobile, Tablet & Desktop	■■ Full lifetime access — watch anytime
■ ByteLearned Certificate of Completion	■■■ Live weekly doubt-clearing sessions
■ Complete source code for all 30 videos	■ Private student Discord community access

■ COMPLETE COURSE CONTENT — 30 Video Lessons • 65+ Hours • 25 Sections

■ PHASE 1 — FOUNDATIONS OF DATA SCIENCE

■ Video
1

Section 1: Introduction to Data Science & AI

■ 2 hrs

- What is Data Science, Machine Learning, and Artificial Intelligence — key differences explained
- The Data Science lifecycle — from business problem to production deployment
- Types of ML: Supervised, Unsupervised, Reinforcement Learning — with real examples
- Career paths in Data Science: Analyst, Data Scientist, ML Engineer, AI Engineer, MLOps
- Industry applications: Healthcare, Finance, E-commerce, Manufacturing, Education
- Setting up your complete Data Science environment — Anaconda, Jupyter, VS Code, Colab
- **Quiz — Data Science Fundamentals — 15 Questions**
- *Download: Data Science Career Roadmap 2026.pdf*
- *Download: DS Environment Setup Guide.pdf*

■ Video
2

Section 2: Mathematics & Statistics for Data Science

■ 3 hrs

- Linear Algebra essentials — Vectors, Matrices, Matrix multiplication, Transpose, Inverse
- Calculus for ML — Derivatives, Partial derivatives, Chain rule, Gradient concept
- Probability theory — Events, Conditional probability, Bayes theorem with ML applications
- Descriptive statistics — Mean, Median, Mode, Variance, Standard deviation, Skewness
- Probability distributions — Normal, Binomial, Poisson — when and why each is used
- Hypothesis testing — p-value, t-test, chi-square, ANOVA — practical significance testing
- Correlation vs Causation — the most important distinction in data science
- **Quiz — Math & Statistics — 20 Questions**
- *Download: Statistics Cheat Sheet.pdf*
- *Download: Probability Distributions Reference.pdf*

■ Video
3

Section 3: Python for Data Science — Part 1 (Core Language)

■ 3 hrs

- Python basics — variables, data types, operators, input/output, type conversion
- Control flow — if/elif/else, for loops, while loops, break, continue, pass
- Functions — defining, calling, default args, *args, **kwargs, lambda, map, filter
- Data structures — Lists, Tuples, Dictionaries, Sets — when to use each with examples
- Object-Oriented Python — Classes, Objects, Inheritance, Encapsulation, Polymorphism
- File handling — reading/writing CSV, JSON, TXT files with error handling
- Virtual environments, pip, requirements.txt — professional Python project setup
- **Quiz — Python Core — 20 Questions**
- *Download: Python Syntax Cheat Sheet.pdf*
- *Download: Python OOP Reference.pdf*

■ Video
4

Section 3: Python for Data Science — Part 2 (NumPy & Pandas)

■ 3 hrs

- NumPy — Arrays, array operations, broadcasting, indexing, slicing, reshaping
- NumPy — Mathematical operations, statistical functions, random module for ML
- Pandas — Series & DataFrames, creating, reading CSV/Excel/JSON data
- Pandas — Selection (loc, iloc), filtering, sorting, renaming, dropping columns
- Pandas — Handling missing data (dropna, fillna, interpolate) — real-world datasets
- Pandas — GroupBy, aggregation, pivot tables, merge, join, concatenation
- Lab: Clean and analyse a real messy dataset end-to-end using Pandas

■ **Quiz — NumPy & Pandas — 20 Questions**

■ *Download: Pandas Cheat Sheet.pdf*

■ *Download: NumPy Quick Reference.pdf*

■ Video
5

Section 4: SQL for Data Science

■ 2 hrs 30
min

- SQL fundamentals — SELECT, FROM, WHERE, ORDER BY, LIMIT with real examples
- Aggregations — COUNT, SUM, AVG, MIN, MAX with GROUP BY and HAVING clauses
- JOINS — INNER, LEFT, RIGHT, FULL OUTER — with practical multi-table queries
- Subqueries, CTEs (WITH clause), and temporary tables for complex analysis
- Window functions — ROW_NUMBER, RANK, DENSE_RANK, LAG, LEAD, running totals
- Lab: Answering 20 real business questions using SQL on an e-commerce database
- Connecting Python to SQL databases using SQLAlchemy and psycopg2

■ **Quiz — SQL for Data Science — 20 Questions**

■ *Download: SQL Cheat Sheet.pdf*

■ *Download: SQL Interview Questions Top 30.pdf*

■ PHASE 2 — DATA ANALYSIS & VISUALISATION

■ Video
6

Section 5: Data Cleaning & Preprocessing

■ 3 hrs

- Understanding dirty data — types of data quality issues in real-world datasets
- Handling missing values — when to drop, fill, interpolate, or flag missing data
- Outlier detection & treatment — IQR method, Z-score, visualisation approaches
- Fixing inconsistencies — standardising formats, units, categories, date parsing
- Data type conversions — ensuring correct dtypes for each column before modelling
- Feature engineering — creating new features from dates, text, and numerical data
- Encoding categorical variables — One-Hot, Label, Target, and Binary encoding
- Lab: Full cleaning pipeline on a real-world HR attrition dataset from Kaggle
- **Quiz — Data Cleaning & Preprocessing — 15 Questions**
- *Download: Data Cleaning Checklist.pdf*
- *Download: Feature Engineering Guide.pdf*

■ Video
7

Section 6: Exploratory Data Analysis (EDA)

■ 2 hrs 30
min

- What is EDA and why it is the most underrated skill in data science
- Univariate analysis — histograms, box plots, violin plots, KDE plots
- Bivariate analysis — scatter plots, bar charts, heatmaps, pair plots
- Correlation analysis — Pearson, Spearman, Kendall correlation matrices
- Statistical testing in EDA — t-test, chi-square test with real data
- Lab: Complete EDA on Titanic and E-commerce datasets — step by step
- Documenting EDA findings and presenting insights to non-technical stakeholders
- **Quiz — EDA — 15 Questions**
- *Download: EDA Template Notebook.ipynb*
- *Download: EDA Presentation Template.pdf*

- Matplotlib — figures, axes, subplots, styling, customisation, saving plots
- Seaborn — statistical plots, themes, colour palettes, FacetGrid
- Plotly — interactive charts, 3D plots, animations, geographic maps
- Storytelling with data — choosing the right chart for the right message
- Lab: Building an interactive dashboard with Streamlit from scratch
- Tableau basics — connecting data, building charts, creating dashboards
- Power BI basics — data modelling, DAX measures, building reports
- **Quiz — Data Visualisation — 15 Questions**
- *Download: Chart Type Selection Guide.pdf*
- *Download: Matplotlib & Seaborn Cheat Sheet.pdf*

■ PHASE 3 — MACHINE LEARNING

■ Video
9

Section 8: Supervised Learning — Regression

■ 2 hrs 30
min

- ML workflow — train/test split, cross-validation, overfitting vs underfitting
- Linear Regression — intuition, OLS, cost function, gradient descent from scratch
- Multiple Linear Regression — assumptions, multicollinearity, VIF analysis
- Polynomial Regression — when and how to handle non-linear relationships
- Regularisation — Ridge (L2), Lasso (L1), ElasticNet — controlling overfitting
- Evaluation metrics — MAE, MSE, RMSE, R^2 , Adjusted R^2
- Lab: Predicting house prices end-to-end — data to deployed model

■ **Quiz — Regression — 15 Questions**

■ *Download: Regression Algorithms Cheat Sheet.pdf*

■ *Download: House Price Project Notebook.ipynb*

■ Video
10

Section 9: Supervised Learning — Classification

■ 3 hrs

- Logistic Regression — sigmoid function, decision boundary, probability outputs
- Decision Trees — splitting criteria (Gini, Entropy), depth, pruning, visualisation
- Random Forest — bagging, feature importance, out-of-bag score, hyperparameters
- Support Vector Machines (SVM) — hyperplane, kernels (RBF, Polynomial), margin
- K-Nearest Neighbours (KNN) — distance metrics, choosing optimal K
- Naive Bayes — Gaussian, Multinomial — applications in spam and text classification
- Evaluation — Confusion matrix, Accuracy, Precision, Recall, F1, ROC-AUC, PR curve
- Lab: Customer churn prediction — compare all classifiers and select the best

■ **Quiz — Classification — 20 Questions**

■ *Download: Classification Algorithms Cheat Sheet.pdf*

■ *Download: Model Evaluation Reference.pdf*

■ Video
11

Section 10: Ensemble Methods & Boosting

■ 2 hrs 30
min

- Ensemble learning concepts — Bagging, Boosting, Stacking explained
- AdaBoost — adaptive boosting, weak learners, weight updates
- Gradient Boosting — gradient descent in function space, trees as learners
- XGBoost — regularisation, parallel processing, handling missing values
- LightGBM — leaf-wise growth, histogram-based splits, speed advantages
- CatBoost — handling categorical features natively without encoding
- Hyperparameter tuning — GridSearchCV, RandomSearchCV, Optuna
- Lab: Kaggle-style competition — achieve top accuracy on a real dataset

■ **Quiz — Ensemble Methods — 15 Questions**

■ *Download: XGBoost & LightGBM Cheat Sheet.pdf*

■ *Download: Hyperparameter Tuning Guide.pdf*

■ Video
12

Section 11: Unsupervised Learning

■ 2 hrs

- K-Means Clustering — algorithm, elbow method, silhouette score, limitations
- Hierarchical Clustering — dendrograms, linkage methods, cutting the tree
- DBSCAN — density-based clustering, eps and min_samples, handling outliers
- Principal Component Analysis (PCA) — dimensionality reduction, variance explained
- t-SNE & UMAP — visualising high-dimensional data in 2D/3D
- Anomaly Detection — Isolation Forest, One-Class SVM, Local Outlier Factor
- Lab: Customer segmentation for an e-commerce company using K-Means + PCA

■ **Quiz — Unsupervised Learning — 15 Questions**

■ *Download: Clustering Algorithms Reference.pdf*

■ *Download: Dimensionality Reduction Guide.pdf*

- Advanced feature engineering — interaction features, binning, log transforms
- Feature selection methods — Filter (correlation), Wrapper (RFE), Embedded (SHAP)
- Handling imbalanced datasets — SMOTE, ADASYN, class weights, threshold tuning
- Pipeline construction with Scikit-learn — ColumnTransformer, FunctionTransformer
- Model selection strategies — when to use which algorithm for which problem type
- AutoML overview — TPOT, Auto-Sklearn, H2O — automated model selection
- Lab: Build a complete reusable ML pipeline for a credit risk scoring project

■ **Quiz — Feature Engineering & Selection — 15 Questions**

■ *Download: Feature Engineering Techniques.pdf*

■ *Download: Scikit-learn Pipeline Template.ipynb*

■ PHASE 4 — DEEP LEARNING & NEURAL NETWORKS

■ Video
14

Section 13: Deep Learning Foundations with TensorFlow & PyTorch

■ 3 hrs

- Neural networks from scratch — neurons, layers, weights, biases, forward pass
- Activation functions — ReLU, Sigmoid, Tanh, Softmax — why and where to use each
- Loss functions — Binary CE, Categorical CE, MSE, MAE — for classification and regression
- Backpropagation — gradient computation, chain rule, weight update step by step
- Optimisers — SGD, Momentum, Adam, AdaGrad, RMSprop — comparison and choice
- Regularisation — Dropout, Batch Normalisation, L1/L2, Early Stopping
- TensorFlow + Keras — building, training, evaluating, saving models with callbacks
- PyTorch — tensors, autograd, nn.Module, training loop, DataLoader from scratch

■ Quiz — Deep Learning Foundations — 20 Questions

■ [Download: Neural Network Architecture Guide.pdf](#)

■ [Download: TensorFlow vs PyTorch Cheat Sheet.pdf](#)

■ Video
15

Section 14: Computer Vision with CNNs

■ 2 hrs 30
min

- Convolutional Neural Networks — convolution operation, filters, feature maps
- Pooling layers, stride, padding — building intuition with visualisations
- Classic CNN architectures — LeNet, AlexNet, VGG, ResNet, InceptionNet
- Transfer Learning — fine-tuning pre-trained models (ResNet50, EfficientNet, VGG16)
- Data augmentation — flipping, rotation, zoom, brightness for image training
- Object detection overview — YOLO, SSD, Faster R-CNN concepts
- Lab: Build an image classifier (cats vs dogs, medical X-ray) with 95%+ accuracy

■ Quiz — Computer Vision — 15 Questions

■ [Download: CNN Architecture Reference.pdf](#)

■ [Download: Transfer Learning Guide.pdf](#)

- Recurrent Neural Networks (RNN) — sequence modelling, vanishing gradient problem
- Long Short-Term Memory (LSTM) — cell state, forget gate, input gate, output gate
- Gated Recurrent Unit (GRU) — simplified LSTM, when to prefer GRU over LSTM
- Time series analysis — stationarity, ACF, PACF, ARIMA, SARIMA models
- Sequence-to-sequence models — encoder-decoder architecture for forecasting
- Lab: Stock price forecasting using LSTM — build, train, evaluate and visualise
- Lab: Sales forecasting for a retail company with seasonal decomposition

■ **Quiz — RNN & Time Series — 15 Questions**

- *Download: Time Series Cheat Sheet.pdf*
- *Download: LSTM Architecture Diagram.pdf*

■ PHASE 5 — NATURAL LANGUAGE PROCESSING & TRANSFORMERS

■ Video
17

Section 16: NLP Fundamentals

■ 2 hrs 30
min

- NLP pipeline — tokenisation, stemming, lemmatisation, stop word removal
- Text representation — Bag of Words, TF-IDF, N-grams for ML models
- Word embeddings — Word2Vec, GloVe, FastText — learning semantic relationships
- Sentiment analysis — VADER (rule-based), ML-based, and deep learning approaches
- Text classification — spam detection, topic classification, intent recognition
- Named Entity Recognition (NER) using spaCy and NLTK
- Lab: Build a movie review sentiment analyser and news topic classifier

■ **Quiz — NLP Fundamentals — 15 Questions**

■ *Download: NLP Pipeline Reference.pdf*

■ *Download: spaCy & NLTK Cheat Sheet.pdf*

■ Video
18

Section 17: Transformers & BERT — Modern NLP

■ 3 hrs

- Attention mechanism — self-attention, queries, keys, values explained visually
- Transformer architecture — encoder, decoder, multi-head attention, positional encoding
- BERT — Bidirectional Encoder, pre-training (MLM, NSP), fine-tuning for tasks
- Hugging Face Transformers library — complete workflow for classification, NER, QA
- Fine-tuning BERT for sentiment analysis, text classification, and Q&A;
- Other transformer models — RoBERTa, DistilBERT, ALBERT, XLNet comparison
- Lab: Fine-tune BERT on a custom dataset for multi-class text classification
- Lab: Question answering system using pre-trained transformer models

■ **Quiz — Transformers & BERT — 15 Questions**

■ *Download: Transformer Architecture Diagram.pdf*

■ *Download: Hugging Face Quick Start Guide.pdf*

■ PHASE 6 — GENERATIVE AI, LLMs & ADVANCED AI

■ Video
19

Section 18: Generative AI — GANs & Diffusion Models

■ 2 hrs

- Generative vs Discriminative models — the fundamental conceptual difference
- Generative Adversarial Networks (GANs) — generator, discriminator, training loop
- Variational Autoencoders (VAEs) — latent space, reconstruction, KL divergence
- Diffusion Models — denoising process, DDPM, Stable Diffusion concept
- Text-to-Image AI — DALL-E 3, Stable Diffusion, Midjourney — how they work
- Ethical considerations — deepfakes, bias in generative models, responsible use

■ **Quiz — Generative AI — 10 Questions**

■ *Download: Generative AI Landscape 2026.pdf*

■ Video
20

Section 19: Large Language Models (LLMs) & Prompt Engineering

■ 3 hrs

- What are LLMs — GPT architecture, parameters, pre-training, RLHF alignment
- Prompt engineering — zero-shot, few-shot, chain-of-thought, role prompting
- OpenAI API — calling GPT-4o, structured outputs, function calling, vision
- Google Gemini API and Anthropic Claude API — comparing capabilities
- Prompt chaining — building multi-step AI workflows with Python
- Fine-tuning LLMs — LoRA, QLoRA — adapting models to your domain data
- Lab: Build an AI-powered content generator and automated report writer
- Lab: AI code review assistant using GPT-4o API with function calling

■ **Quiz — LLMs & Prompt Engineering — 15 Questions**

■ *Download: Prompt Engineering Playbook.pdf*

■ *Download: OpenAI API Reference Guide.pdf*

- LangChain framework — chains, agents, tools, memory, callbacks architecture
 - Retrieval Augmented Generation (RAG) — why RAG beats fine-tuning for knowledge
 - Vector databases — embeddings, cosine similarity, FAISS, Pinecone, ChromaDB
 - Building a RAG pipeline — document loading, chunking, embedding, retrieval, generation
 - LangChain agents — ReAct agents, tool use, web search, code execution
 - LangGraph — stateful multi-agent workflows for complex AI pipelines
 - Lab: Build a full Document Q&A; chatbot over your own PDF files with Streamlit UI
 - Lab: Multi-tool AI agent that searches the web, queries SQL, and writes reports
 - **Quiz — LangChain & RAG — 15 Questions**
 - *Download: LangChain Architecture Diagram.pdf*
 - *Download: RAG System Design Template.pdf*
-

■ PHASE 7 — BIG DATA, CLOUD & MLOps

■ Video
22

Section 21: Big Data Processing with PySpark

■ 2 hrs 30
min

- Why Big Data — limitations of Pandas, when to use Spark vs Pandas
- Apache Spark architecture — driver, executors, DAG, transformations vs actions
- PySpark DataFrames — creating, reading, transforming large-scale data
- PySpark SQL — running SQL queries directly on Spark DataFrames at scale
- MLlib — Spark's distributed machine learning library for big data ML
- Databricks — cloud Spark environment, notebooks, Delta Lake, Unity Catalog
- Lab: Process a 10 million row dataset with PySpark in under 30 seconds

■ **Quiz — Big Data & Spark — 10 Questions**

■ *Download: PySpark Cheat Sheet.pdf*

■ *Download: Databricks Getting Started Guide.pdf*

■ Video
23

Section 22: Cloud ML Platforms — AWS, Azure & GCP

■ 2 hrs 30
min

- AWS for Data Science — S3, EC2, SageMaker, Glue, Athena, Lambda, Step Functions
- AWS SageMaker — training, tuning, deploying, monitoring models end-to-end
- Azure Machine Learning Studio — datasets, pipelines, compute, model registry
- Google Vertex AI — AutoML, custom training, model serving, Vertex Pipelines
- Choosing the right cloud platform — cost, features, and company context
- Lab: Train and deploy an ML model on AWS SageMaker with auto-scaling endpoint
- Cloud cost optimisation strategies for ML workloads

■ **Quiz — Cloud ML Platforms — 10 Questions**

■ *Download: AWS SageMaker Quick Start.pdf*

■ *Download: Cloud Platform Comparison Guide.pdf*

- What is MLOps — bridging the gap between ML and software engineering
- Model serialisation — Pickle, Joblib, ONNX, SavedModel formats
- Building REST APIs with FastAPI — endpoints, request validation, async handling
- Docker — containerising ML models, writing Dockerfiles, docker-compose
- CI/CD for ML — GitHub Actions pipeline to automatically test and deploy models
- MLflow — experiment tracking, model registry, model versioning, artifact storage
- Model monitoring — data drift detection with Evidently AI, performance dashboards
- Lab: Full MLOps pipeline — train → track → containerise → deploy → monitor
- **Quiz — MLOps & Deployment — 15 Questions**
- *Download: FastAPI ML Deployment Template.pdf*
- *Download: Docker Cheat Sheet.pdf*
- *Download: MLflow Guide.pdf*

■ PHASE 8 — SPECIALISATION, PROJECTS & CAREER

■ Video
25

Section 24: Data Engineering Fundamentals

■ 2 hrs

- Data engineering vs data science — roles, responsibilities and collaboration
- ETL & ELT pipelines — extraction, transformation, loading patterns explained
- Apache Airflow — DAGs, operators, scheduling, monitoring data pipelines
- dbt (data build tool) — transforming data in warehouses with SQL + version control
- Data warehouse design — star schema, snowflake schema, fact and dimension tables
- Real-time data streaming — Kafka concepts, producers, consumers, topics
- Lab: Build an automated daily ETL pipeline using Airflow + PostgreSQL + dbt

■ Quiz — Data Engineering — 10 Questions

- *Download: Data Pipeline Design Patterns.pdf*
- *Download: Airflow DAG Template.py*

■ Video
26

Section 25: Responsible AI — Ethics, Fairness & Governance

■ 1 hr 30
min

- AI bias — sources of bias in data, model outputs, and real-world consequences
- Fairness metrics — demographic parity, equalised odds, individual fairness
- Explainability — SHAP values, LIME, model cards, explainable AI (XAI) principles
- Data privacy — GDPR, India PDPB compliance, data anonymisation, differential privacy
- AI governance frameworks — EU AI Act, RBI guidelines for AI in banking
- Environmental impact of AI — carbon cost of training large models

- *Download: Responsible AI Checklist.pdf*
- *Download: SHAP & LIME Reference Guide.pdf*

■ Video
27

Section 26: Capstone Project 1 — End-to-End ML Project

■ 3 hrs

- Project brief: Customer Lifetime Value prediction for an e-commerce platform
- Step 1: Business problem framing, success metrics definition, data audit
- Step 2: Full data cleaning, EDA, and feature engineering pipeline
- Step 3: Model building — compare Regression, Random Forest, XGBoost, LightGBM
- Step 4: Model explainability with SHAP — feature importance and business insights
- Step 5: Deploy model as a FastAPI endpoint with Docker on AWS
- Step 6: Build a Streamlit dashboard showing predictions with business context
- Capstone 1 final submission, peer review, and instructor feedback
- *Download: Capstone 1 Project Guide.pdf*
- *Download: Submission Rubric.pdf*

■ Video
28

Section 27: Capstone Project 2 — Generative AI Application

■ 3 hrs

- Project brief: Build a smart AI assistant for business document analysis
- Step 1: RAG pipeline design — document ingestion, chunking strategy, embedding choice
- Step 2: Vector store setup with FAISS/ChromaDB, similarity search optimisation
- Step 3: LangChain integration — retriever, prompt template, LLM chain assembly
- Step 4: Multi-turn conversation memory with LangChain ConversationBufferMemory
- Step 5: Streamlit UI with chat interface, document upload, and answer citations
- Step 6: Deploy on Hugging Face Spaces or AWS — share publicly for portfolio
- Capstone 2 final submission and professional portfolio writeup guidance
- *Download: Capstone 2 Project Guide.pdf*
- *Download: RAG Architecture Template.pdf*

■ Video
29

Section 28: Certification Preparation

■ 2 hrs

- Google Data Analytics Professional Certificate — syllabus review and exam tips
- IBM Data Science Professional Certificate — 10-course structure overview
- TensorFlow Developer Certificate — practical exam format and key topics
- AWS Certified Machine Learning Specialty — service coverage and prep strategy
- DeepLearning.AI Deep Learning Specialisation — Andrew Ng's 5-course roadmap
- Practice test session — 80 mixed questions across all certification domains
- Exam day strategy — time management, elimination technique, common traps
- *Download: Certification Study Planner.pdf*
- *Download: Practice Test Answer Key.pdf*

- Building a data science resume that gets shortlisted — structure, keywords, ATS tips
- GitHub portfolio — how to present notebooks, projects, and README files professionally
- Kaggle profile — competitions, datasets, notebooks — building public credibility
- LinkedIn optimisation for data science — headline, about, featured section, skills
- Top 60 Data Science interview Q&A; — Statistics, ML, SQL, Python, Behavioural
- System design for ML — designing ML systems at scale in senior interviews
- Negotiating your first data science offer — salary benchmarks and scripts
- Freelancing in data science — Upwork, Toptal, direct consulting, rates to charge
- *Download: DS Resume Template.pdf*
- *Download: Top 60 Interview Q&A.pdf;*
- *Download: LinkedIn Optimisation Guide.pdf*
- *Download: Salary Negotiation Scripts.pdf*

■ JOB ROLES & SALARY GUIDE

Job Role	Specialisation	Experience	Avg Salary India
Data Analyst	Analytics	0–1 yr	Rs. 4 – 8 LPA
Junior Data Scientist	ML/AI	0–2 yrs	Rs. 5 – 10 LPA
Business Intelligence Analyst	Analytics	1–2 yrs	Rs. 5 – 9 LPA
Machine Learning Engineer	MLOps/ML	1–3 yrs	Rs. 8 – 18 LPA
Data Engineer	Data Infra	1–3 yrs	Rs. 7 – 16 LPA
NLP / Generative AI Engineer	NLP/GenAI	2–4 yrs	Rs. 12 – 25 LPA
Senior Data Scientist	ML/AI	3–5 yrs	Rs. 18 – 35 LPA
MLOps / Platform Engineer	MLOps	3–5 yrs	Rs. 20 – 38 LPA
Computer Vision Engineer	CV/DL	2–5 yrs	Rs. 15 – 32 LPA
AI Research Scientist	Research	4–6 yrs	Rs. 25 – 50 LPA
Lead / Principal Data Scientist	ML/AI	6–8 yrs	Rs. 35 – 60 LPA
Head of AI / Director	Leadership	8–10 yrs	Rs. 50 – 80 LPA+

■ 50+ TOOLS & TECHNOLOGIES YOU WILL MASTER

Python	R	SQL	Pandas	NumPy	Matplotlib
Seaborn	Plotly	Scikit-learn	XGBoost	LightGBM	TensorFlow
PyTorch	Keras	Hugging Face	LangChain	OpenAI API	FAISS
Pinecone	ChromaDB	Streamlit	FastAPI	Flask	Docker
Kubernetes	Git/GitHub	MLflow	W&B	Airflow	PySpark
dbt	Snowflake	BigQuery	AWS SageMaker	Azure ML	Vertex AI
Databricks	Kafka	Tableau	Power BI	PostgreSQL	MongoDB
Redis	Grafana	SHAP	LIME	Optuna	NLTK
spaCy	Jupyter/Colab				

ByteLearned

bytelearned.online

info@bytelearned.online

7396633051