MockAttempt

TECH ACADEMY • OFFICIAL COURSE SYLLABUS BROCHURE

MockAttempt Academy Pvt. Ltd. • Live Architect Mentorship, AI Assessments & Client Production

Admissions & Support: support@mockattempt.com
Phone / WhatsApp: +91 98765 43210
Course Code: DSAI • **Updated:** 01 Nov 2026
https://mockattempt.com/courses/applied-data-science-genai-mlops-engineering

DSAI • DATA SCIENCE & AI | **BEGINNER TO ADVANCED** | **100% LIVE INDUSTRY EXPERT COHORT + INTERNSHIP LETTER**

Applied Data Science, GenAI & MLOps Engineering

Master Python data workflows, machine learning modeling, LLM RAG agents, and FastAPI model serving with verified topic assessments.

Duration: 18–22 Weeks • Live + Hands-On Projects (211 Hrs) • **Modules:** 14 • **Live Topics:** 42 • **Labs & Tasks:** 84 • **Mocks / AI Vivas:** 84 / 84 • **Projects:** 14 Module + 1 Live Client Prod

Fixed Non-Negotiable Tuition Fee: Base Fee **₹9,999** + 18% GST (₹1,800) = **Total Incl. GST: ₹11,799**

Live Batch Tracks: Weekday Live Online Batch (Mon – Fri) (Monday to Friday (5 Live Interactive Classes / Week) • 1.5 Hours / Day (60 Mins Live Architecture & Coding + 30 Mins Hands-On Lab) • 18–20 Weeks Fast-Track) | Weekend Intensive Live Online Batch (Sat – Sun) (Saturday & Sunday (Designed for Working Professionals & College Students) • 3.0 Hours / Day (2 Hrs Deep-Dive Live Coding + 1 Hr Guided Project Build) • 20–22 Weeks Weekend Track)

Core Technologies & Engineering Stack: Python • PyTorch • Scikit-Learn • FastAPI • LangChain • PostgreSQL pgvector

1. PROGRAM MASTERY OUTCOMES & ELIGIBILITY

End-to-End Engineering Outcomes:

1. Train, evaluate, and serve ML models behind FastAPI microservices
2. Build grounded RAG pipelines with hybrid vector search & evaluation metrics
3. Reproducible Python data pipelines (`NumPy`, `Pandas`, feature stores) with strict train/validation leakage prevention.
4. Classical ML & Deep Learning (`Scikit-Learn`, `XGBoost`, `PyTorch`) tuned via cross-validation, ROC-AUC, and SHAP explainability.
5. Production Generative AI: chunking strategies, embeddings, `pgvector` hybrid retrieval, RAG grounding, and hallucination evals.
6. Low-latency model & LLM serving via async FastAPI endpoints with token streaming, guardrails, and cost/latency telemetry.
7. Vectorized NumPy, Pandas Pipelines & Statistical Validation — Clean messy million-row datasets 50× faster and separate real business signals from random noise.
8. Leak-Free Feature Engineering & Classical ML (Scikit-Learn / Boosting) — Predict customer churn, credit risk, demand forecasts, and fraud with calibrated accuracy.

Eligibility & Prerequisites: No prior Applied Data Science, GenAI & MLOps Engineering industry experience required — Module 1 starts from environment setup, core mental models, and plain-English fundamentals. • Basic computer & internet familiarity on a Windows, macOS, or Linux laptop (8 GB+ RAM recommended for smooth live coding labs). • Ideal for B.Tech / BCA / MCA / B.Sc IT students, non-CS beginners, career switchers, and working developers upgrading to 2026 production standards. • Commitment to attend live Weekday or Weekend classes, complete daily hands-on lab checklists, and practice topic-wise Mock Tests & AI Interviews. • Dedication to write original code in your GitHub portfolio and complete the mentor-guided Live Client Production Sprint. • No prior Applied Data Science, GenAI & MLOps Engineering industry experience required — Module 1 starts from environment setup, core mental models, and plain-English fundamentals.

2. DETAILED MODULE & TOPIC-WISE SYLLABUS (14 MODULES • 42 LIVE TOPICS)

MODULE 1 OF 14 • 25 LIVE + LAB HOURS • 1 TOPICS • 1 CAPSTONE PROJECT (100 MARKS)

Python Data Structures, Vectorization & Feature Engineering

NumPy memory layouts, Pandas pipelines, and statistical validation.

Objectives: ✓ Eliminate Python row loops using contiguous NumPy vectorization and broadcasting • ✓ Build leak-free Scikit-Learn ColumnTransformer feature pipelines • ✓ Serve validated ML & RAG inference endpoints behind async FastAPI services

1.1 NumPy Broadcasting & High-Performance Data Pipelines [45 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply NumPy Broadcasting & High-Performance Data Pipelines in real Applied Data Science, GenAI & MLOps Engineering systems — specifically: Eliminating s...

Module 1 Hands-On Capstone Project (100 Marks): Real-Time Credit Risk Feature Pipeline & FastAPI Scorer — Engineer a vectorized feature pipeline and deploy a sub-25ms FastAPI credit-risk scoring endpoint with drift metrics.

MODULE 2 OF 14 • 16 LIVE + LAB HOURS • 4 TOPICS • 1 CAPSTONE PROJECT (100 MARKS)

Data Wrangling with Pandas, Polars & Analytical Window SQL

Production-grade Applied Data Science, GenAI & MLOps Engineering engineering module covering Data Wrangling with Pandas, Polars & Analytical Window SQL with hands-on labs and capstone build.

Objectives: ✓ Pandas DataFrame Indexing, GroupBy Aggregations & Memory Optimization — Purpose: Clean messy multi-million-row CSV/Parquet datasets fast without exhausting RAM. • ✓ Analytical SQL for Data Scientists: CTEs, Window Functions & Funnel Queries — Purpose: Extract ML-ready customer feature tables directly inside PostgreSQL/DuckDB. • ✓ High-Speed Lazy Dataframes with Polars & DuckDB Local Lakehouse — Purpose: Process 20GB+ datasets on a standard laptop using Rust/C++ vectorized query engines. • ✓ Automated Ingestion Pipelines, Pydantic Data Schemas & Idempotent Backfills — Purpose: Ingest partner APIs and webhook JSON logs cleanly into partitioned feature tables.

2.1 Pandas DataFrame Indexing, GroupBy Aggregations & Memory Optimization [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Pandas DataFrame Indexing, GroupBy Aggregations & Memory Optimization in real Applied Data Science, GenAI & MLOps Engineering systems — specifical...

2.2 Analytical SQL for Data Scientists: CTEs, Window Functions & Funnel Queries [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Analytical SQL for Data Scientists: CTEs, Window Functions & Funnel Queries in real Applied Data Science, GenAI & MLOps Engineering systems — spec...

2.3 High-Speed Lazy Dataframes with Polars & DuckDB Local Lakehouse [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply High-Speed Lazy Dataframes with Polars & DuckDB Local Lakehouse in real Applied Data Science, GenAI & MLOps Engineering systems — specifically: Pr...

2.4 Automated Ingestion Pipelines, Pydantic Data Schemas & Idempotent Backfills [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Automated Ingestion Pipelines, Pydantic Data Schemas & Idempotent Backfills in real Applied Data Science, GenAI & MLOps Engineering systems — spec...

Module 2 Hands-On Capstone Project (100 Marks): Automated Multi-Source E-Commerce Cohort & Retention Data Pipeline — Design, implement, test, and document 'Automated Multi-Source E-Commerce Cohort & Retention Data Pipeline' applying every production concept from Module 2 (Data Wrangling with Pandas, Polars...

MODULE 3 OF 14 • 16 LIVE + LAB HOURS • 4 TOPICS • 1 CAPSTONE PROJECT (100 MARKS)

Exploratory Data Analysis, Probability Distributions & A/B Hypothesis Testing

Production-grade Applied Data Science, GenAI & MLOps Engineering engineering module covering Exploratory Data Analysis, Probability Distributions & A/B Hypothesis Testing with hands-on labs and capstone build.

Objectives: ✓ Diagnostic EDA, Outlier Detection & Feature Correlation Forensics — Purpose: Uncover data skew, Simpson’s paradox, and collinearity before fitting a single model. • ✓ Probability Distributions, Bayes’ Theorem & Confidence Intervals — Purpose: Quantify uncertainty around churn rates, fraud probabilities, and demand forecasts. • ✓ Statistical A/B Testing: Sample Size Power, p-Values & Multiple-Testing Guardrails — Purpose: Design valid online product experiments that distinguish real lift from random noise. • ✓ Causal Uplift Thinking, Guardrail Metrics & Network Interference Checks — Purpose: Ensure product launches improve net retention without hurting latency or support tickets.

3.1 Diagnostic EDA, Outlier Detection & Feature Correlation Forensics [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Diagnostic EDA, Outlier Detection & Feature Correlation Forensics in real Applied Data Science, GenAI & MLOps Engineering systems — specifically:...

3.2 Probability Distributions, Bayes’ Theorem & Confidence Intervals [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Probability Distributions, Bayes’ Theorem & Confidence Intervals in real Applied Data Science, GenAI & MLOps Engineering systems — specifically: Q...

3.3 Statistical A/B Testing: Sample Size Power, p-Values & Multiple-Testing Guardrails [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Statistical A/B Testing: Sample Size Power, p-Values & Multiple-Testing Guardrails in real Applied Data Science, GenAI & MLOps Engineering systems...

3.4 Causal Uplift Thinking, Guardrail Metrics & Network Interference Checks [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Causal Uplift Thinking, Guardrail Metrics & Network Interference Checks in real Applied Data Science, GenAI & MLOps Engineering systems — specific...

Module 3 Hands-On Capstone Project (100 Marks): Product Experimentation & Statistical A/B Test Decision Engine — Design, implement, test, and document 'Product Experimentation & Statistical A/B Test Decision Engine' applying every production concept from Module 3 (Exploratory Data Analysis, Probability...

MODULE 4 OF 14 • 14 LIVE + LAB HOURS • 3 TOPICS • 1 CAPSTONE PROJECT (100 MARKS)

Leak-Free Feature Engineering, Scikit-Learn Pipelines & Imbalanced Classification

Production-grade Applied Data Science, GenAI & MLOps Engineering engineering module covering Leak-Free Feature Engineering, Scikit-Learn Pipelines & Imbalanced Classification with hands-on labs and capstone build.

Objectives: ✓ ColumnTransformer Pipelines, Target Encoding & Leakage Prevention — Purpose: Package preprocessing and modeling into a single serializable Scikit-Learn Pipeline. • ✓ Regularized Linear Models (Ridge, Lasso, ElasticNet) & Cross-Validation Strategies — Purpose: Build fast, interpretable baseline models and choose time/group-aware CV splits. • ✓ Imbalanced Learning: PR-AUC, Cost-Sensitive Thresholds & Probability Calibration — Purpose: Optimize fraud, medical, and churn classifiers where positive cases are < 2% of rows.

4.1 ColumnTransformer Pipelines, Target Encoding & Leakage Prevention [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply ColumnTransformer Pipelines, Target Encoding & Leakage Prevention in real Applied Data Science, GenAI & MLOps Engineering systems — specifically:...

4.2 Regularized Linear Models (Ridge, Lasso, ElasticNet) & Cross-Validation Strategies [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Regularized Linear Models (Ridge, Lasso, ElasticNet) & Cross-Validation Strategies in real Applied Data Science, GenAI & MLOps Engineering systems...

4.3 Imbalanced Learning: PR-AUC, Cost-Sensitive Thresholds & Probability Calibration [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Imbalanced Learning: PR-AUC, Cost-Sensitive Thresholds & Probability Calibration in real Applied Data Science, GenAI & MLOps Engineering systems —...

Module 4 Hands-On Capstone Project (100 Marks): End-to-End Credit Default & Churn Prediction Pipeline — Design, implement, test, and document 'End-to-End Credit Default & Churn Prediction Pipeline' applying every production concept from Module 4 (Leak-Free Feature Engineering, Scikit-Learn Pip...

MODULE 5 OF 14 • 14 LIVE + LAB HOURS • 3 TOPICS • 1 CAPSTONE PROJECT (100 MARKS)

Gradient Boosted Ensembles (XGBoost, LightGBM), Optuna & SHAP Explainability

Production-grade Applied Data Science, GenAI & MLOps Engineering engineering module covering Gradient Boosted Ensembles (XGBoost, LightGBM), Optuna & SHAP Explainability with hands-on labs and capstone build.

Objectives: ✓ Gradient Boosting Math, Histogram Split Finding & XGBoost/LightGBM Tuning — Purpose: Master the #1 tabular ML algorithm family used across FinTech, retail, and ad-tech. • ✓ Bayesian Hyperparameter Optimization with Optuna & Early Stopping — Purpose: Search hyperparameter spaces intelligently in a fraction of GridSearch compute cost. • ✓ Model Explainability with TreeSHAP, Permutation Importance & Fairness Audits — Purpose: Generate regulator-ready per-customer prediction explanations and detect subgroup bias.

5.1 Gradient Boosting Math, Histogram Split Finding & XGBoost/LightGBM Tuning [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Gradient Boosting Math, Histogram Split Finding & XGBoost/LightGBM Tuning in real Applied Data Science, GenAI & MLOps Engineering systems — specif...

5.2 Bayesian Hyperparameter Optimization with Optuna & Early Stopping [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Bayesian Hyperparameter Optimization with Optuna & Early Stopping in real Applied Data Science, GenAI & MLOps Engineering systems — specifically:...

5.3 Model Explainability with TreeSHAP, Permutation Importance & Fairness Audits [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Model Explainability with TreeSHAP, Permutation Importance & Fairness Audits in real Applied Data Science, GenAI & MLOps Engineering systems — spe...

Module 5 Hands-On Capstone Project (100 Marks): Explainable Real-Time Fraud Scoring Engine with SHAP Reason Codes — Design, implement, test, and document 'Explainable Real-Time Fraud Scoring Engine with SHAP Reason Codes' applying every production concept from Module 5 (Gradient Boosted Ensembles (XGBoost...

MODULE 6 OF 14 • 14 LIVE + LAB HOURS • 3 TOPICS • 1 CAPSTONE PROJECT (100 MARKS)

Unsupervised Learning, Dimensionality Reduction, Anomaly Detection & Forecasting

Production-grade Applied Data Science, GenAI & MLOps Engineering engineering module covering Unsupervised Learning, Dimensionality Reduction, Anomaly Detection & Forecasting with hands-on labs and capstone build.

Objectives: ✓ K-Means, HDBSCAN Density Clustering & Behavioral RFM Segmentation — Purpose: Group millions of unlabeled users or SKUs into actionable business segments. • ✓ PCA, UMAP Visualization & Unsupervised Anomaly Detection (Isolation Forest) — Purpose: Compress wide sensor/embedding spaces and catch novel zero-day anomalies without labels. • ✓ Time-Series Decomposition, Walk-Forward Backtesting & Multi-Step Forecasting — Purpose: Forecast daily inventory demand, cloud spend, and revenue across multiple horizons.

6.1 K-Means, HDBSCAN Density Clustering & Behavioral RFM Segmentation [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply K-Means, HDBSCAN Density Clustering & Behavioral RFM Segmentation in real Applied Data Science, GenAI & MLOps Engineering systems — specifically:...

6.2 PCA, UMAP Visualization & Unsupervised Anomaly Detection (Isolation Forest) [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply PCA, UMAP Visualization & Unsupervised Anomaly Detection (Isolation Forest) in real Applied Data Science, GenAI & MLOps Engineering systems — spec...

6.3 Time-Series Decomposition, Walk-Forward Backtesting & Multi-Step Forecasting [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Time-Series Decomposition, Walk-Forward Backtesting & Multi-Step Forecasting in real Applied Data Science, GenAI & MLOps Engineering systems — spe...

Module 6 Hands-On Capstone Project (100 Marks): RFM Customer Segmentation & Multi-Horizon Demand Forecaster — Design, implement, test, and document 'RFM Customer Segmentation & Multi-Horizon Demand Forecaster' applying every production concept from Module 6 (Unsupervised Learning, Dimensionality Red...

MODULE 7 OF 14 • 14 LIVE + LAB HOURS • 3 TOPICS • 1 CAPSTONE PROJECT (100 MARKS)

Deep Learning with PyTorch: Autograd, Neural Architectures & Fine-Tuning

Production-grade Applied Data Science, GenAI & MLOps Engineering engineering module covering Deep Learning with PyTorch: Autograd, Neural Architectures & Fine-Tuning with hands-on labs and capstone build.

Objectives: ✓ PyTorch Tensors, Dynamic Autograd Graphs & Custom Dataset/DataLoader — Purpose: Build GPU-accelerated neural training loops with clean data batching. • ✓ Optimizers (AdamW), Cosine LR Schedules, Mixed Precision (AMP) & Checkpointing — Purpose: Train neural networks 2x faster in BF16/FP16 with reproducible model checkpoints. • ✓ Transfer Learning, Vision Backbones & ONNX Runtime Quantized Export — Purpose: Fine-tune pretrained backbones on small client datasets and serve them on low-cost CPUs.

7.1 PyTorch Tensors, Dynamic Autograd Graphs & Custom Dataset/DataLoader [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply PyTorch Tensors, Dynamic Autograd Graphs & Custom Dataset/DataLoader in real Applied Data Science, GenAI & MLOps Engineering systems — specificall...

7.2 Optimizers (AdamW), Cosine LR Schedules, Mixed Precision (AMP) & Checkpointing [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Optimizers (AdamW), Cosine LR Schedules, Mixed Precision (AMP) & Checkpointing in real Applied Data Science, GenAI & MLOps Engineering systems — s...

7.3 Transfer Learning, Vision Backbones & ONNX Runtime Quantized Export [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Transfer Learning, Vision Backbones & ONNX Runtime Quantized Export in real Applied Data Science, GenAI & MLOps Engineering systems — specifically...

Module 7 Hands-On Capstone Project (100 Marks): PyTorch Multi-Modal Defect & Document Classifier with Mixed Precision — Design, implement, test, and document 'PyTorch Multi-Modal Defect & Document Classifier with Mixed Precision' applying every production concept from Module 7 (Deep Learning with PyTorch: Aut...

MODULE 8 OF 14 • 14 LIVE + LAB HOURS • 3 TOPICS • 1 CAPSTONE PROJECT (100 MARKS)

Transformers, Tokenization, Dense Embeddings & Parameter-Efficient Fine-Tuning (LoRA)

Production-grade Applied Data Science, GenAI & MLOps Engineering engineering module covering Transformers, Tokenization, Dense Embeddings & Parameter-Efficient Fine-Tuning (LoRA) with hands-on labs and capstone build.

Objectives: ✓ Self-Attention Mechanics, Subword Tokenization & Encoder vs Decoder LLMs — Purpose: Understand exactly how modern Transformers process context windows, tokens, and KV caches. • ✓ Dense Bi-Encoder Embeddings, Cosine Similarity & Cross-Encoder Reranking — Purpose: Map sentences, resumes, and products into semantic vector spaces. • ✓ Prompt Engineering Systems, Structured JSON Outputs & LoRA / QLoRA Fine-Tuning — Purpose: Control LLM outputs deterministically and know when prompting beats fine-tuning.

8.1 Self-Attention Mechanics, Subword Tokenization & Encoder vs Decoder LLMs [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Self-Attention Mechanics, Subword Tokenization & Encoder vs Decoder LLMs in real Applied Data Science, GenAI & MLOps Engineering systems — specifi...

8.2 Dense Bi-Encoder Embeddings, Cosine Similarity & Cross-Encoder Reranking [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Dense Bi-Encoder Embeddings, Cosine Similarity & Cross-Encoder Reranking in real Applied Data Science, GenAI & MLOps Engineering systems — specifi...

8.3 Prompt Engineering Systems, Structured JSON Outputs & LoRA / QLoRA Fine-Tuning [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Prompt Engineering Systems, Structured JSON Outputs & LoRA / QLoRA Fine-Tuning in real Applied Data Science, GenAI & MLOps Engineering systems — s...

Module 8 Hands-On Capstone Project (100 Marks): Domain-Adapted Support Ticket Triage & Semantic Similarity Model — Design, implement, test, and document 'Domain-Adapted Support Ticket Triage & Semantic Similarity Model' applying every production concept from Module 8 (Transformers, Tokenization, Dense Em...

MODULE 9 OF 14 • 14 LIVE + LAB HOURS • 3 TOPICS • 1 CAPSTONE PROJECT (100 MARKS)

Vector Databases (PostgreSQL pgvector), HNSW Indexing & Hybrid Search

Production-grade Applied Data Science, GenAI & MLOps Engineering engineering module covering Vector Databases (PostgreSQL pgvector), HNSW Indexing & Hybrid Search with hands-on labs and capstone build.

Objectives: ✓ Document Parsing, Structure-Aware Chunking & Metadata Enrichment — Purpose: Prepare PDFs, policies, and codebases into high-signal chunks that retrieve cleanly. • ✓ PostgreSQL pgvector, HNSW vs IVFFlat Indexes & Pre-Filtered Tenant Search — Purpose: Run vector search right inside ACID PostgreSQL alongside relational RBAC permissions. • ✓ Hybrid Search: Combining Lexical Full-Text BM25 with Reciprocal Rank Fusion (RRF) — Purpose: Handle exact SKU codes, error IDs, and semantic natural-language queries simultaneously.

9.1 Document Parsing, Structure-Aware Chunking & Metadata Enrichment [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Document Parsing, Structure-Aware Chunking & Metadata Enrichment in real Applied Data Science, GenAI & MLOps Engineering systems — specifically: P...

9.2 PostgreSQL pgvector, HNSW vs IVFFlat Indexes & Pre-Filtered Tenant Search [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply PostgreSQL pgvector, HNSW vs IVFFlat Indexes & Pre-Filtered Tenant Search in real Applied Data Science, GenAI & MLOps Engineering systems — specif...

9.3 Hybrid Search: Combining Lexical Full-Text BM25 with Reciprocal Rank Fusion (RRF) [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Hybrid Search: Combining Lexical Full-Text BM25 with Reciprocal Rank Fusion (RRF) in real Applied Data Science, GenAI & MLOps Engineering systems...

Module 9 Hands-On Capstone Project (100 Marks): Multi-Tenant Hybrid Semantic + BM25 Search Engine on pgvector — Design, implement, test, and document 'Multi-Tenant Hybrid Semantic + BM25 Search Engine on pgvector' applying every production concept from Module 9 (Vector Databases (PostgreSQL pgvector),...

MODULE 10 OF 14 • 14 LIVE + LAB HOURS • 3 TOPICS • 1 CAPSTONE PROJECT (100 MARKS)

Production Enterprise RAG Systems, Grounded Citations & Safety Guardrails

Production-grade Applied Data Science, GenAI & MLOps Engineering engineering module covering Production Enterprise RAG Systems, Grounded Citations & Safety Guardrails with hands-on labs and capstone build.

Objectives: ✓ Query Rewriting, Multi-Query Expansion & Conversational Memory Compression — Purpose: Resolve follow-up pronouns (\"what about its refund fee?\") into standalone retrieval queries. • ✓ Strict Citation Grounding, Abstention Logic & Faithfulness Verification — Purpose: Ensure enterprise copilots cite exact source paragraphs and abstain when evidence is absent. • ✓ Prompt Injection Defense, PII Redaction & Semantic Response Caching — Purpose: Protect GenAI endpoints against jailbreaks, data exfiltration, and runaway API bills.

10.1 Query Rewriting, Multi-Query Expansion & Conversational Memory Compression [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Query Rewriting, Multi-Query Expansion & Conversational Memory Compression in real Applied Data Science, GenAI & MLOps Engineering systems — speci...

10.2 Strict Citation Grounding, Abstention Logic & Faithfulness Verification [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Strict Citation Grounding, Abstention Logic & Faithfulness Verification in real Applied Data Science, GenAI & MLOps Engineering systems — specific...

10.3 Prompt Injection Defense, PII Redaction & Semantic Response Caching [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Prompt Injection Defense, PII Redaction & Semantic Response Caching in real Applied Data Science, GenAI & MLOps Engineering systems — specifically...

Module 10 Hands-On Capstone Project (100 Marks): Hallucination-Resistant Enterprise Policy & Compliance RAG Copilot — Design, implement, test, and document 'Hallucination-Resistant Enterprise Policy & Compliance RAG Copilot' applying every production concept from Module 10 (Production Enterprise RAG Systems...

MODULE 11 OF 14 • 14 LIVE + LAB HOURS • 3 TOPICS • 1 CAPSTONE PROJECT (100 MARKS)

Agentic AI Workflows, Function/Tool Calling & Automated LLM Evaluation

Production-grade Applied Data Science, GenAI & MLOps Engineering engineering module covering Agentic AI Workflows, Function/Tool Calling & Automated LLM Evaluation with hands-on labs and capstone build.

Objectives: ✓ Deterministic Function Calling, Tool Schemas & Human-in-the-Loop Approvals — Purpose: Connect Gemini/LLMs safely to internal SQL, CRM, and ticketing APIs. • ✓ Multi-Step Graph Workflows (Router -> Retriever -> Verifier) vs Free-Form Agents — Purpose: Design predictable state-machine AI workflows that hit 99% production reliability. • ✓ Automated Golden-Dataset Evals (RAGAS / LLM-as-Judge) in GitHub Actions CI — Purpose: Prevent prompt edits or model upgrades from silently breaking accuracy.

11.1 Deterministic Function Calling, Tool Schemas & Human-in-the-Loop Approvals [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Deterministic Function Calling, Tool Schemas & Human-in-the-Loop Approvals in real Applied Data Science, GenAI & MLOps Engineering systems — speci...

11.2 Multi-Step Graph Workflows (Router -> Retriever -> Verifier) vs Free-Form Agents [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Multi-Step Graph Workflows (Router -> Retriever -> Verifier) vs Free-Form Agents in real Applied Data Science, GenAI & MLOps Engineering systems —...

11.3 Automated Golden-Dataset Evals (RAGAS / LLM-as-Judge) in GitHub Actions CI [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Automated Golden-Dataset Evals (RAGAS / LLM-as-Judge) in GitHub Actions CI in real Applied Data Science, GenAI & MLOps Engineering systems — speci...

Module 11 Hands-On Capstone Project (100 Marks): Autonomous Customer Operations Triage Agent with Golden-Set CI Evals — Design, implement, test, and document 'Autonomous Customer Operations Triage Agent with Golden-Set CI Evals' applying every production concept from Module 11 (Agentic AI Workflows, Function/...

MODULE 12 OF 14 • 14 LIVE + LAB HOURS • 3 TOPICS • 1 CAPSTONE PROJECT (100 MARKS)

Production MLOps: Async FastAPI Serving, SSE Token Streaming & Drift Monitoring

Production-grade Applied Data Science, GenAI & MLOps Engineering engineering module covering Production MLOps: Async FastAPI Serving, SSE Token Streaming & Drift Monitoring with hands-on labs and capstone build.

Objectives: ✓ Async FastAPI Model Serving, Lifespan Warmup & SSE Token Streaming — Purpose: Expose ML and GenAI pipelines as low-latency, high-concurrency production APIs. • ✓ Experiment Tracking, Model Registry (MLflow) & Reproducible Docker Packaging — Purpose: Version datasets, hyperparameters, prompts, and container artifacts with full lineage. • ✓ Data/Concept Drift Detection (PSI), Token Cost Governance & Retraining Triggers — Purpose: Monitor live models after launch so silent distribution shifts never erode business ROI.

12.1 Async FastAPI Model Serving, Lifespan Warmup & SSE Token Streaming [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Async FastAPI Model Serving, Lifespan Warmup & SSE Token Streaming in real Applied Data Science, GenAI & MLOps Engineering systems — specifically:...

12.2 Experiment Tracking, Model Registry (MLflow) & Reproducible Docker Packaging [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Experiment Tracking, Model Registry (MLflow) & Reproducible Docker Packaging in real Applied Data Science, GenAI & MLOps Engineering systems — spe...

12.3 Data/Concept Drift Detection (PSI), Token Cost Governance & Retraining Triggers [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Data/Concept Drift Detection (PSI), Token Cost Governance & Retraining Triggers in real Applied Data Science, GenAI & MLOps Engineering systems —...

Module 12 Hands-On Capstone Project (100 Marks): Containerized ML + GenAI Inference Service with Drift & Cost Telemetry — Design, implement, test, and document 'Containerized ML + GenAI Inference Service with Drift & Cost Telemetry' applying every production concept from Module 12 (Production MLOps: Async FastA...

MODULE 13 OF 14 • 14 LIVE + LAB HOURS • 3 TOPICS • 1 CAPSTONE PROJECT (100 MARKS)

RAG Evaluation Harnesses, Hallucination Guardrails & Hybrid Search

Self-explanatory list-and-bullet production mastery module covering RAG Evaluation Harnesses, Hallucination Guardrails & Hybrid Search in Applied Data Science, GenAI & MLOps Engineering, paired with hands-on labs, capstone repository, Topic Mock Tests, and Spoken Gemini AI Viva.

Objectives: ✓ Chunking Strategies, Metadata Filtering & Hybrid Keyword + Vector Retrieval — Purpose: Retrieve exact enterprise policy clauses with high recall instead of noisy semantic-only matches. • ✓ Cross-Encoder Re-Ranking, Citation Enforcement & Prompt-Injection Guards — Purpose: Block jailbreak attempts and force LLM responses to cite verifiable source passages. • ✓ Automated RAG Evaluation (Faithfulness, Context Precision & Latency Budgets) — Purpose: Catch prompt regressions and hallucination spikes automatically before shipping model updates. • ✓ Build, verify, and document 'Grounded Hybrid BM25 + pgvector RAG Pipeline with Automated Faithfulness Scoring' with clean Git commits and edge-case protection. • ✓ Clear 100% of Topic Timed Mock Tests and Spoken Voice + Google Gemini AI Technical Interviews.

13.1 Chunking Strategies, Metadata Filtering & Hybrid Keyword + Vector Retrieval [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Chunking Strategies, Metadata Filtering & Hybrid Keyword + Vector Retrieval in real Applied Data Science, GenAI & MLOps Engineering systems — spec...

13.2 Cross-Encoder Re-Ranking, Citation Enforcement & Prompt-Injection Guards [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Cross-Encoder Re-Ranking, Citation Enforcement & Prompt-Injection Guards in real Applied Data Science, GenAI & MLOps Engineering systems — specifi...

13.3 Automated RAG Evaluation (Faithfulness, Context Precision & Latency Budgets) [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Automated RAG Evaluation (Faithfulness, Context Precision & Latency Budgets) in real Applied Data Science, GenAI & MLOps Engineering systems — spe...

Module 13 Hands-On Capstone Project (100 Marks): Grounded Hybrid BM25 + pgvector RAG Pipeline with Automated Faithfulness Scoring — Architect, implement, test, and document 'Grounded Hybrid BM25 + pgvector RAG Pipeline with Automated Faithfulness Scoring' applying every bullet-wise engineering concept from Module 13 (RAG...

MODULE 14 OF 14 • 14 LIVE + LAB HOURS • 3 TOPICS • 1 CAPSTONE PROJECT (100 MARKS)

Production MLOps, Model Drift Monitoring & AI System Design

Self-explanatory list-and-bullet production mastery module covering Production MLOps, Model Drift Monitoring & AI System Design in Applied Data Science, GenAI & MLOps Engineering, paired with hands-on labs, capstone repository, Topic Mock Tests, and Spoken Gemini AI Viva.

Objectives: ✓ Feature Skew Detection, Data Drift Alerts (PSI/KS) & Shadow Deployments — Purpose: Detect when live customer distributions diverge from training data before prediction accuracy degrades. • ✓ Async Batching, KV-Cache / Semantic Caching & Token Cost Governance — Purpose: Cut LLM and ML inference latency and cloud GPU/API bills by up to 60% in production. • ✓ End-to-End ML & GenAI System Design Case Studies for Product Interviews — Purpose: Design recommendation, fraud-scoring, and enterprise copilot architectures step-by-step. • ✓ Build, verify, and document 'Canary Model Serving Gateway, Drift Alerting & Token-Cost Telemetry Service' with clean Git commits and edge-case protection. • ✓ Clear 100% of Topic Timed Mock Tests and Spoken Voice + Google Gemini AI Technical Interviews.

14.1 Feature Skew Detection, Data Drift Alerts (PSI/KS) & Shadow Deployments [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Feature Skew Detection, Data Drift Alerts (PSI/KS) & Shadow Deployments in real Applied Data Science, GenAI & MLOps Engineering systems — specific...

14.2 Async Batching, KV-Cache / Semantic Caching & Token Cost Governance [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply Async Batching, KV-Cache / Semantic Caching & Token Cost Governance in real Applied Data Science, GenAI & MLOps Engineering systems — specifically...

14.3 End-to-End ML & GenAI System Design Case Studies for Product Interviews [65 mins • Assignment (100M) • Hands-On Lab (50M) • 2 Mock Sets • 2 AI Vivas]

Purpose: To understand and apply End-to-End ML & GenAI System Design Case Studies for Product Interviews in real Applied Data Science, GenAI & MLOps Engineering systems — specific...

Module 14 Hands-On Capstone Project (100 Marks): Canary Model Serving Gateway, Drift Alerting & Token-Cost Telemetry Service — Architect, implement, test, and document 'Canary Model Serving Gateway, Drift Alerting & Token-Cost Telemetry Service' applying every bullet-wise engineering concept from Module 14 (Producti...

3. LIVE CLIENT PRODUCTION PROJECT, GLOBAL CERTIFICATIONS & CAREER SCOPE

End-of-Course Live Client Production Project + Official Internship Confirmation Letter

Project: DocuSense Enterprise GenAI RAG Knowledge Copilot & Predictive Scoring API • **Confirmed Role:** AI, Data Science & GenAI MLOps Intern

Domain & Production Infra: Healthcare & Enterprise Knowledge Operations Client (Python + FastAPI + PostgreSQL pgvector + Cloud

GPU/CPU Inference Endpoint)

- Built automated Python feature-engineering pipelines and calibrated predictive ML models with Scikit-Learn & PyTorch.
- Engineered a grounded Retrieval-Augmented Generation (RAG) service over PostgreSQL pgvector with citation verification.
- Exposed low-latency async FastAPI inference endpoints with token streaming, guardrails, and structured evaluation metrics.
- Deployed model-serving containers to live client production with drift monitoring and automated regression tests.

Aligned Global Vendor Certifications: **AIF-C01** (AWS Certified AI Practitioner (Generative AI & Foundation Models)) • **NCA-GENL** (NVIDIA-Certified Associate: Generative AI & LLMs) • **PCAP-31-03** (PCAP™ – Certified Associate in Python Programming)

Target Job Roles: Applied AI & GenAI Engineer, Data Scientist / Machine Learning Engineer, LLM RAG & MLOps Engineer, Predictive Analytics & AI Developer

Realistic Annual CTC Scope: **Fresher / Associate Data Scientist & AI Engineer (0–1 Yr):** ₹6.5 LPA – ₹13.5 LPA **Mid-Level ML / GenAI Engineer (1–3 Yrs):** ₹14.0 LPA – ₹28.0 LPA **Senior AI Architect / Principal ML Engineer (3–6+ Yrs):** ₹28.0 LPA – ₹58.0+ LPA

MockAttempt Academy Pvt. Ltd. • Email: support@mockattempt.com • Phone/WhatsApp: +91 98765 43210 • Web: https://mockattempt.com/courses/applied-data-science-genai-mlops-engineering

Try 2 Free Topic-Wise 50-MCQ Mock Tests & Spoken Gemini AI Interviews daily on every topic above before enrolling.