

# MLOPS & AI DEPLOYMENT PROFESSIONAL PROGRAM- 2025 EDITION

"From models to market—deploy AI with confidence"

# **Market Demand Note**

MLOps is one of the fastest-growing skills in AI engineering. Companies want ML/AI solutions that are not just developed but deployed, monitored, and updated seamlessly in production environments.

Duration: 12 Weeks | Mode: Online/Offline

# MLOps & Al Deployment Professional Program— 2025 Edition

# **Key Tools & Technologies:**

Python, Docker, Kubernetes, Git, GitHub Actions, MLflow, DVC, TensorFlow, PyTorch, FastAPI, AWS/GCP/Azure, Kubeflow.

# **Learning Objectives**

By the end of this program, learners will be able to:

- 1. Design reproducible ML workflows with version control for data and models.
- 2. Deploy AI models into scalable production environments.
- 3. Automate CI/CD pipelines for ML projects.
- 4. Monitor model performance and retrain pipelines.
- 5. Optimize AI systems for cost, performance, and scalability.

#### **Table of Contents**

Week	Topic	Key Concepts
1	MLOps Foundations	MLOps lifecycle, vs. DevOps, industry trends
2	Data & Model Versioning	DVC, MLflow, Git, reproducibility
3	Containerization & Environments	Docker, Conda, Pipenv, reproducible environments
4	Orchestration & Workflow Management	Airflow, Kubeflow, scalable pipelines
5	API Deployment for ML Models	FastAPI, REST endpoints, model serving
6	Cloud Deployment Strategies	AWS SageMaker, GCP Vertex Al, Azure ML, Kubernetes
7	CI/CD for Machine Learning	GitHub Actions, continuous training/deployment, automation
8	Model Monitoring & Logging	Performance tracking, drift detection, alerting (Prometheus, Grafana)
9	Security & Governance	API/auth security, data governance, compliance (GDPR, HIPAA)
10	Advanced Deployment Patterns	Multi-model serving, A/B testing, shadow deployments
11	Optimization & Cost Management	GPU/TPU optimization, serverless, cloud cost tracking
12	Capstone & Final Assessment	End-to-end MLOps pipeline, monitoring, scaling, portfolio project

#### **Detailed Content**

Week 1: MLOps Foundations

© 2025 Udaan Tech Academy. All Rights Reserved.

- MLOps Defined: End-to-end lifecycle of machine learning in production—data, training, deployment, monitoring, retraining.
- MLOps vs. DevOps: ML-specific challenges (data drift, model decay, experiment tracking).
- ML Lifecycle: From problem definition to model retirement; roles (Data Scientist, ML Engineer, DevOps).
- Industry Trends: Shift-left testing, GitOps for ML, model registries, and reproducibility.
- Hands-on: Map a real ML project's lifecycle; compare with a software project.

# Week 2: Data & Model Versioning

- Data Version Control (DVC): Track datasets, metrics, and artifacts alongside code.
- Model Versioning (MLflow): Log experiments, register models, compare runs.
- Git Integration: Version control for code, data, and models.
- Reproducibility: Recreate any model/dataset state at any time.
- Hands-on: Set up DVC+Git for a sample project; log experiments with MLflow.

#### Week 3: Containerization & Environments

- Docker Basics: Images, containers, Dockerfiles, multi-stage builds for ML.
- Environment Management: Conda environments, Pipenv, Poetry.
- Reproducible Builds: Pin libraries, avoid "it works on my machine."
- Hands-on: Containerize a Python ML app; build and share Docker images.

# Week 4: Orchestration & Workflow Management

- Airflow: DAGs, operators, scheduling, monitoring.
- Kubeflow: Kubernetes-native ML pipelines, component reuse.
- Scalable Pipelines: Design for parallel training, validation, deployment.
- Hands-on: Build an Airflow DAG for ML training; deploy a Kubeflow pipeline.

# Week 5: API Deployment for ML Models

- FastAPI: Async, typed, OpenAPI/Swagger docs, high performance.
- REST Endpoints: Predict, health check, versioning.
- Model Packaging: ONNX, joblib, pickle, TensorFlow Serving, Triton.
- Hands-on: Deploy a Scikit-learn or TensorFlow model as a REST API; test with Postman.

# Week 6: Cloud Deployment Strategies

- AWS SageMaker: Managed training, hosting, AutoML, endpoints.
- GCP Vertex AI: Pipelines, custom training, model registry.
- Azure ML: Workspaces, experiments, deployment.
- Kubernetes: Deploy anywhere, autoscaling, GPU/TPU scheduling.
- Hands-on: Train and deploy a model on AWS/GCP/Azure; scale with Kubernetes.

# Week 7: CI/CD for Machine Learning

- GitHub Actions: Trigger on code/data/model changes, run tests, deploy.
- Continuous Training: Retrain on new data, validate, promote.
- Continuous Deployment: Roll out new models, blue-green, canary.
- Hands-on: Set up GitHub Actions CI/CD for an ML project.

# Week 8: Model Monitoring & Logging

- Performance Tracking: Latency, throughput, error rates.
- Data/Model Drift: Detect drift, trigger retraining.
- Alerting: Prometheus, Grafana, Slack/email alerts.
- Hands-on: Instrument a live model with logging and drift detection.

#### Week 9: Security & Governance

## MLOps & Al Deployment Professional Program @ https://udaan.x-fuzion.com/

- API Security: Auth (OAuth, API keys), rate limiting, input validation.
- Data Governance: Row-level security, masking, audit logs.
- Compliance: GDPR, HIPAA, AI ethics, bias/fairness checks.
- Hands-on: Secure an ML API; implement basic compliance checks.

## Week 10: Advanced Deployment Patterns

- Multi-Model Serving: Deploy ensembles, A/B test variants.
- Shadow Deployment: Run new model alongside old, compare outputs.
- Canary Releases: Gradual rollout, monitor, rollback if needed.
- Hands-on: Deploy two model versions; A/B test with traffic splitting.

# Week 11: Optimization & Cost Management

- GPU/TPU Optimization: Mixed precision, quantization, pruning.
- Serverless: AWS Lambda, GCP Cloud Functions, Azure Functions.
- Cost Tracking: Cloud billing dashboards, spot instances, autoscaling.
- Hands-on: Optimize a model for inference speed; deploy serverless.

#### Week 12: Capstone & Final Assessment

- Project Scope: Implement a full MLOps pipeline—data → train → deploy → monitor
   → retrain.
- Monitoring: Track performance, drift, and costs.
- Scaling: Handle production traffic, autoscaling, logging.
- Presentation: Demo your pipeline, explain architecture and tradeoffs.
- Portfolio: GitHub repo with code, docs, and a live demo.

# Capstone Project Ideas

• Fraud Detection Pipeline: Ingest transaction data, train model, deploy as API, monitor for drift, retrain on new fraud patterns.

- Al-Powered Recommendation System: Continuous retraining on user behavior, A/B test new algorithms, monitor engagement.
- Image Classification on Kubernetes: Train CNN, deploy with autoscaling, monitor accuracy and latency.

# Current Tools & Industry Trends (2025)

- Open Source: MLflow, DVC, Airflow, Kubeflow, FastAPI, Prometheus, Grafana.
- Cloud: AWS SageMaker, GCP Vertex AI, Azure ML, Kubernetes (EKS, GKE, AKS).
- CI/CD: GitHub Actions, GitLab CI, Jenkins.
- Monitoring: Evidently, Arize, WhyLabs (for drift), OpenTelemetry.
- Security: OAuth, JWT, Vault, GDPR tools, bias detection libraries.
- Optimization: ONNX Runtime, TensorRT, TVM, mixed-precision training.
- Trends: GitOps for ML, model registries, multi-model serving, responsible AI, edge deployment.

# Why This Structure?

- End-to-end: Covers data, code, model, deployment, monitoring, and governance—exactly what enterprises need.
- Hands-on: Every week has real, implementable labs with open-source and cloud tools.
- Portfolio-ready: Capstone projects are exactly what employers look for in interviews.
- Job relevance: Skills in MLOps, cloud, and automation are in highest demand globally.