BOARD INFINITY

COURSE 2 OF 3 · MANAGING AI SYSTEMS SPECIALIZATION

MLOps and LLMOps: Deploying and Scaling AI in Production

Full course syllabus - every module, video lesson, reading and assignment, from CI/CD pipelines to drift monitoring.

Board Infinity
Instructor · Offered on Coursera

4	36	16	2 weeks
Modules	Video lessons	Assignments	At 10 hours a week

INSIDE THE COURSE

01	Operationalizing AI Pipelines (CI/CD, CT/CD, Versioning)	5 hours to complete
02	LLMOps Fundamentals: Context, Prompts, Inference Optimization	5 hours to complete
03	Evaluation: From Vibes to Metrics	4 hours to complete
04	Observability & Tracing for Production AI	4 hours to complete

Enroll on Coursera

https://www.coursera.org/learn/mlops-and-llmops-deploying-and-scaling-ai-in-production

01 ABOUT THE COURSE

From AI prototype to production system.

MLOps and LLMOps is the bridge between an AI prototype and a production AI system - for ML engineers, data scientists and software engineers who have to keep models and LLM applications running reliably.

You work through reproducible pipelines, versioning and CI/CD, then inference optimization, then evaluation with LLM-as-a-Judge and quality gates, and finish with tracing and drift monitoring in MLflow, W&B and Arize Phoenix - 36 videos, 12 readings and 16 assignments in all.

WHO IT'S FOR

ML engineers
Production pipelines, deployment and reliability.

Data scientists
CI/CD from an ML-first view, no DevOps needed.

Software engineers
Serving, CI/CD and observability for AI apps.

AI / platform professionals
Operationalizing models and LLM apps at scale.

WHAT YOU'LL LEARN

01 **Configure CI/CD** Module 1
Configure CI/CD pipelines for ML and LLM systems using GitHub Actions and MLflow

02 **Optimize inference** Module 2
Optimize LLM inference pipelines for reduced latency, token cost, and improved reliability

03 **Evaluate** Module 3
Build automated evaluation frameworks using LLM-as-a-Judge and quality gates

04 **Observe** Module 4
Instrument production AI systems with tracing, drift detection, and observability dashboards

02 SYLLABUS

Four modules, pipelines to production.

2 weeks to complete at 10 hours a week. Flexible schedule - learn at your own pace.

01 Operationalizing AI Pipelines (CI/CD, CT/CD, Versioning)

5 hours
9 videos

CI/CD for ML, validation, versioning and reproducible environments

BY THE END OF THIS MODULE YOU CAN

- Configure CI/CD for ML and LLM systems with GitHub Actions and MLflow
- Version models, prompts and datasets; use registries, rollbacks and lineage
- Keep environments reproducible and promote them to production

CONTENTS: 9 videos (92 min) · 3 readings (90 min) · 4 assignments (105 min)

PLUGIN: Quick Course Check-In (5 min)

TOPICS: Foundations of CI/CD for ML & LLM Systems · Model Versioning & Release Management · Environment & Dependency Management

You'll leave with: a reproducible, versioned CI/CD pipeline with safe rollbacks.

02 LLMOps Fundamentals: Context, Prompts, Inference Optimization

5 hours
9 videos

Context windows, prompts, latency, token cost and prompt tracking

BY THE END OF THIS MODULE YOU CAN

- Optimize LLM inference for latency, token cost and reliability
- Design context windows and structured prompts that reduce hallucination
- Log and scientifically evaluate prompt variants with W&B and MLflow

CONTENTS: 9 videos (78 min) · 3 readings (90 min) · 4 assignments (105 min)

TOPICS: Context Windows & Prompt Structure · Inference Optimization: Latency & Token Cost · Prompt Versioning & Experiment Tracking

You'll leave with: a faster, cheaper, more reliable LLM pipeline.

03 Evaluation: From Vibes to Metrics

4 hours
9 videos

Evaluation data, LLM-as-a-Judge, automated scoring and quality gates

BY THE END OF THIS MODULE YOU CAN

- Build automated evaluation with LLM-as-a-Judge and quality gates
- Construct realistic evaluation data, with edge cases and without bias
- Trigger evaluation during deployment and judge release readiness

CONTENTS: 9 videos (66 min) · 3 readings (90 min) · 4 assignments (75 min)

TOPICS: Designing LLM Evaluation Datasets · LLM-as-a-Judge Methodologies · Integrating Evaluation into CI/CD

You'll leave with: quality gates that decide what ships.

04 Observability & Tracing for Production AI

4 hours
9 videos

Logging, tracing with Arize Phoenix, drift detection and alerting

BY THE END OF THIS MODULE YOU CAN

- Instrument AI systems with tracing, drift detection and dashboards
- Trace chains and agents in Phoenix to find hallucination triggers
- Tell data drift from behavioral drift, and decide when to retrain

CONTENTS: 9 videos (53 min) · 3 readings (90 min) · 4 assignments (105 min)

TOPICS: Experiment Tracking & Telemetry · Tracing and Debugging with Arize Phoenix · Monitoring Drift & System Health

You'll leave with: a traced, monitored system and a rule for when to retrain.

03 TOOLS & SKILLS

The platforms you work with.

MLflow Tracking & registry

Model registries, CI/CD for ML and experiment logging.

In this course: Registries in module one, prompt logging in module two.

Weights & Biases Experiments

Experiment tracking and comparing versions.

In this course: Prompt variants and evaluation automation templates.

Arize Phoenix Tracing

Tracing execution steps through chains and agents.

In this course: Tracing and debugging in module four.

GitHub Actions CI/CD

Automated build, test and deploy pipelines.

In this course: Named in the official outcome for module one.

SKILLS YOU LEAVE WITH

Shipping it

- CI/CD
- Continuous Deployment
- Release Management
- Containerization

Running it

- Continuous Monitoring
- Site Reliability Engineering
- Scalability
- AI Security

Building with LLMs

- MLOps
- LLM Application
- Retrieval-Augmented Generation
- Model Evaluation

04 CERTIFICATE & CAREER

Finish with a shareable certificate.

Add it to LinkedIn

Add this credential to your LinkedIn profile, resume, or CV.

Share it anywhere

Share it on social media and in your performance review.

Earned, not attended

Awarded on a graded quiz in each of the four modules.

Course 2 of 3

Part of Managing AI Systems: Development, Deployment, and Governance.

WHERE THESE SKILLS APPLY

ML engineering · AI platform · Data engineering · AI evaluation · Reliability & observability · LLM optimization



Board Infinity

Instructor · Offered on Coursera · 287 courses

Intermediate level · Taught in English · Flexible schedule

Ready to take AI to production?

Enrollment, lessons, grading and the certificate all happen on Coursera.

Enroll on Coursera

Course page: coursera.org/learn/mlops-and-llmops-deploying-and-scaling-ai-in-production

Want a coach alongside the course?

Board Infinity 1:1 coaching is a separate, optional Board Infinity offering - it is not part of the Coursera course. Details: https://www.boardinfinity.com/career-coaching/

Board Infinity is an independent education provider. Coursera and the Coursera logo are trademarks of Coursera, Inc. Not affiliated with or endorsed by MLflow, Weights & Biases, Arize, GitHub, Docker or Kubernetes; not a vendor certification. Capability areas are not guaranteed outcomes.