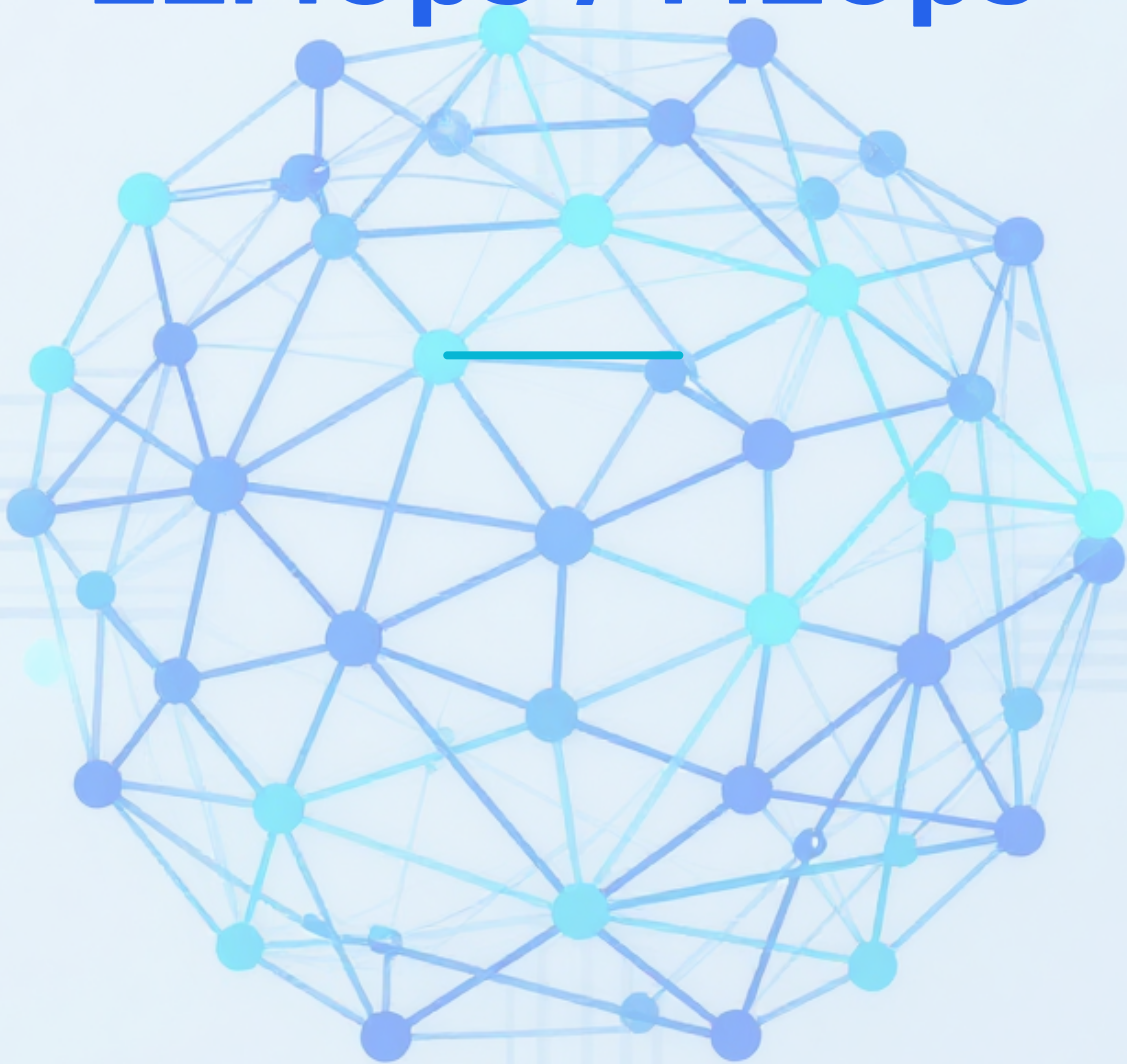


BOOTCAMP

LLM0ps / MLOps



Operaciones de IA en Producción

Objetivos del Curso

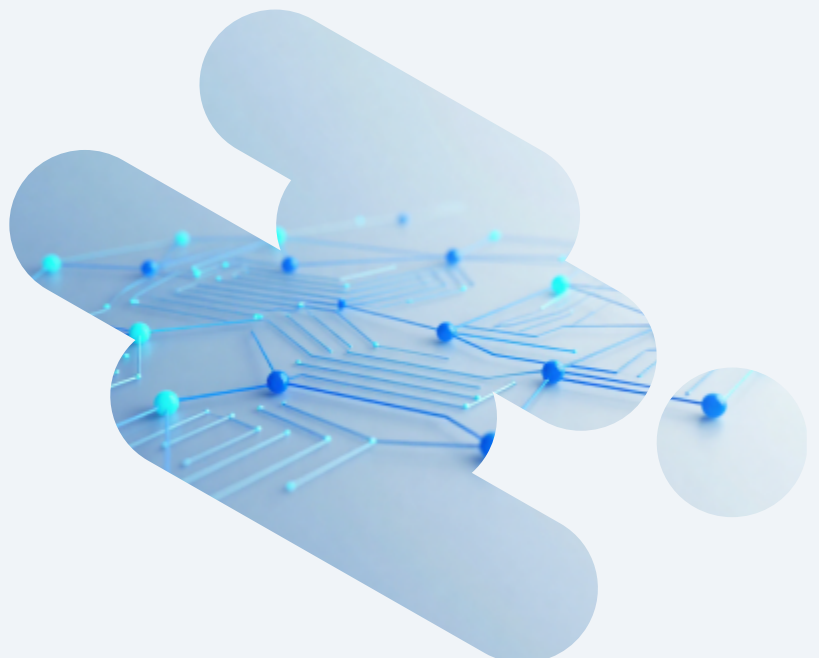
Comprender los fundamentos de MLOps y LLMOps, identificando el ciclo de vida de los modelos de Machine Learning y aplicaciones basadas en LLMs, desde el desarrollo hasta su operación en producción.

Gestionar el versionado de datos, modelos, prompts y experimentos, aplicando buenas prácticas de reproducibilidad, trazabilidad y control de cambios en proyectos de inteligencia artificial.

Automatizar pipelines de entrenamiento, validación y despliegue, utilizando herramientas de CI/CD, contenedores y flujos automatizados para llevar modelos y aplicaciones IA a entornos productivos.

Implementar monitoreo, observabilidad y evaluación de modelos IA, considerando métricas de rendimiento, calidad de respuestas, costos, consumo de tokens, drift y comportamiento del modelo en producción.

Diseñar y desplegar soluciones IA escalables, seguras y gobernadas, aplicando buenas prácticas de arquitectura, seguridad, retraining, cloud, control de costos y operación continua de modelos ML y sistemas basados en LLMs.



Horarios

INICIODE CLASES:

INICIO : 24 de Julio

Martes y Viernes | 7:30 pm - 10:00 pm (Hora Peru)



Temario

DURACIÓNTOTAL: CLASES EN VIVO (96 HORAS ACADÉMICAS)

Modalidad: Onlineenvivo 

Duración total : **3 MESES**

Nivel: Básico - Intermedio



Módulo 1: Fundamentos de MLOps y LLMOps

- Introducción a MLOps y LLMOps
- IA tradicional vs IA Generativa
- Ciclo de vida de modelos ML
- Ciclo de vida de aplicaciones con LLMs
- Roles dentro de un equipo MLOps
- Arquitectura de soluciones IA
- Buenas prácticas para proyectos de IA
- Herramientas del ecosistema MLOps
- Laboratorio: Diseño del ciclo de vida de un proyecto ML

M2 Módulo 2: Versionado y Gestión de Modelos

- Versionado de datasets
- Versionado de modelos
- Gestión de prompts
- Seguimiento de experimentos
- Gestión de artefactos
- Reproducibilidad de modelos
- Registro de modelos y Model Registry
- Gestión de versiones de aplicaciones IA
- Herramientas: MLflow, DVC, Git, Weights & Biases
- Laboratorio: Implementación de un flujo completo de versionado

M3 Módulo 3: Automatización y CI/CD para Machine Learning

- DevOps aplicado a IA
- GitFlow para proyectos ML
- Pipelines CI/CD
- Automatización de entrenamiento
- Automatización de pruebas y validación automática
- Integración continua y entrega continua
- Infraestructura como código para proyectos IA
- Herramientas: GitHub Actions, GitLab CI/CD, Jenkins, Docker
- Laboratorio: Pipeline automático para entrenamiento y despliegue



M4

Módulo 4: LLMOps, Observabilidad y Evaluación

- Arquitectura LLMOps
- Evaluación de respuestas y evaluación automática de prompts
- Prompt Versioning y Prompt Testing
- RAG Evaluation
- Observabilidad de aplicaciones IA
- Trazabilidad y métricas de calidad
- Métricas de costo y control de tokens
- Gestión de alucinaciones
- Herramientas: LangSmith, LangFuse, OpenTelemetry, Phoenix
- Laboratorio: Monitoreo completo de una aplicación basada en LLM

M5

Módulo 5: Despliegue, Escalabilidad y Seguridad

- Docker para aplicaciones IA
- Kubernetes para IA
- APIs para modelos con FastAPI
- Inferencia en producción
- Escalabilidad horizontal y balanceo de carga
- GPU y CPU
- Seguridad en modelos IA y gestión de secretos
- Protección de APIs y gobierno de modelos
- Cloud: AWS, Azure, Google Cloud
- Laboratorio: Despliegue completo de un servicio IA



M6 **Módulo 6: Monitoreo y Proyecto Final**

- Model Monitoring y Drift Detection (Data Drift, Concept Drift)
- Retraining automático
- Alertas, Dashboards y Logging
- Auditoría y Cost Optimization
- FinOps aplicado a IA
- Buenas prácticas empresariales
- Proyecto Final: Solución completa MLOps/LLMOps
 - Desarrollo de modelo o aplicación basada en LLM
 - Versionado de datasets, modelos y prompts
 - Automatización mediante CI/CD
 - Despliegue con Docker y FastAPI en la nube
 - Observabilidad, monitoreo, costos y seguridad
 - Presentación de arquitectura y demostración funcional

