



## Ciencia de datos en ambientes productivos (MLOps)

2024

---

**Profesores:** Federico Pousa y Agustin Mosteiro

**E-mail:** [fedepousa@gmail.com](mailto:fedepousa@gmail.com)

**Horario:** Mié de 19:15hs a 22:15hs

**Aula:**

---

*El programa de esta asignatura puede sufrir modificaciones a lo largo del curso a criterio de los profesores. Los criterios de evaluación son inalterables.*

**Descripción:** En este curso se introducen a las prácticas más comunes para disponibilizar soluciones de ciencia de datos en un ambiente productivo. El enfoque del curso es mayormente práctico con la idea de poder tener la base conceptual general de las diferentes herramientas que se utilizan en la industria y adicionalmente entendiendo y creando ejemplos específicos en dichas herramientas. Se espera que los alumnos participen proactivamente creando sus propias soluciones y logrando disponibilizarlas.

### Objetivos generales:

1. Introducir a los conceptos generales de MLOps y a las diferentes cuestiones cotidianas necesarias para disponibilizar soluciones de ciencia de datos en ambientes productivos.
2. Conocer ejemplos concretos de herramientas que se utilicen masivamente para la disponibilización de soluciones, entiendo las diferencias y similitudes que cada herramienta particular puede tener con sus alternativas.
3. Realizar diversos ejemplos para utilizar de manera real y concreta las herramientas introducidas.

### Programa:

Conceptos generales para el desarrollo de software: sistemas operativos, manejo de terminal, Secure Shell, Git, Docker, deploys enlatados como HuggingFace o Streamlit Cloud.

Conceptos sobre servidores y protocolos web: Arquitectura cliente-servidor. Web Frameworks. Web Servers.

Automatización de tareas en producción. Conceptos generales de orquestación de workflows y herramientas concretas como Airflow.

Desarrollo cloud: introducción al desarrollo en la nube. Interacción con AWS. Servicios específicos de desarrollo en la nube: EC2, RDS, ECS, S3.

Presentación de herramientas de disponibilización de modelos más avanzadas y casos reales de aplicación.

**Evaluación:** Un trabajo práctico final en grupos de 3 estudiantes. Para **aprobar** la materia, el alumno debe obtener **nota mayor a 50pts**.

La escala de conversión de notas numéricas a notas letra es la siguiente:

- [95.00 ; 100] A
- [85.00 ; 95.00) A-
- [75.00 ; 85.00) B+
- [65.00 ; 75.00) B
- [60.00 ; 65.00) B-
- [55.00 ; 60.00) C+
- [50.00 ; 55.00) C
- [00.00 ; 50.00) D

**Sobre la dinámica de trabajo en grupo:** Se espera que todos los integrantes del grupo se comprometan y participen de manera equitativa en la resolución de todos los TPs. En caso de existir problemas con el grupo de TPs de la materia, el alumno afectado deberá comunicarlo al docente previo a la entrega de los mismos. Una vez entregados, no se admitirán objeciones sobre la participación de los integrantes del grupo.

- Material de clase y diferentes papers introducidos oportunamente.
- Practical MLOps. Noah Gift, Alfredo Deza. O'reilly, 2021.
- Designing Machine Learning Systems. Chip Huyen. O'reilly, 2022.
- Introducing MLOps. Mark Treveil et al. O'reilly, 2020.
- Data Pipelines with Apache Airflow. Bas P. Harenslak, Julian Rutger de Ruiter. Manning. 2021.

- Data Science on AWS. Chris Fregly, Antje Barth. O'reilly, 2021.