

Aprender Online UTN

Diplomatura en IA Generativa para Data Science

FACULTAD REGIONAL RECONQUISTA

APRENDERONLINE.FRRQ.UTN.EDU.AR

Acerca del curso

Programa integral enfocado en formar profesionales capaces de diseñar, construir y orquestar arquitecturas de datos modernas y escalables. Cubre desde la ingesta, automatización de pipelines y procesamiento masivo en entornos de Big Data, hasta la gobernanza y disponibilidad de datos para Business Intelligence y Machine Learning.

¿A quien va dirigido?

Está dirigido a desarrolladores, analistas de datos, administradores de bases de datos, profesionales de IT y estudiantes de carreras técnicas o afines que cuenten con conocimientos básicos de lógica de programación y manejo de datos, y deseen especializarse en el diseño y la gestión de infraestructura de datos de alto rendimiento.

Perfil Profesional

Al completar la diplomatura, la persona estará capacitada para:

- **Diseño de repositorios**: Modelar y administrar bases de datos SQL y NoSQL (BigQuery, MongoDB), así como Data Warehouses y Data Lakes.
- **Construcción de pipelines**: Diseñar, automatizar y orquestar flujos de datos (ETL/ELT) robustos utilizando Shell scripting, Apache Kafka y Apache Airflow.
- **Procesamiento de Big Data**: Procesar grandes volúmenes de información distribuida mediante motores como Apache Spark y Hadoop.
- **Gobernanza y seguridad**: Implementar estándares de calidad, seguridad, compliance y buenas prácticas a lo largo del ciclo de vida del dato.
- **Visualización de datos**: Crear paneles interactivos y comunicar hallazgos estratégicos utilizando Tableau y Looker.
- **Preparación para Machine Learning**: Estructurar y disponibilizar conjuntos de datos optimizados para el entrenamiento y despliegue de modelos predictivos.

Características generales

Modalidad

Online • On-demand • Asíncrona

Duración

9 semanas

Nivel de dificultad

Medio

Dedicación

Moderada

Pre-requisitos

Conocimientos previos

Es recomendable contar con nociones básicas de lógica de programación (en cualquier lenguaje), familiaridad con el uso general de sistemas operativos y conceptos elementales de tablas o bases de datos (como Excel o SQL básico). Si bien el programa introduce las herramientas paso a paso, no está pensado para usuarios sin experiencia informática básica, sino para personas con curiosidad técnica y pensamiento analítico interesadas en adentrarse en la infraestructura de datos.

Requerimientos técnicos

- Computadora: Laptop o PC con Windows (64 bits), macOS o Linux (mínimo 8 GB de RAM).
- Permisos: Acceso de administrador para instalar programas y herramientas.
- Conectividad: Conexión a Internet estable.
- Cuentas: Navegador web actualizado y cuenta de correo electrónico activa.

Modulo 1: Fundamentos de IA Generativa aplicada a DS

- Introducción: el Data Scientist frente a una revolución conceptual
- ¿Qué es realmente la IA generativa y por qué importa ahora?
- La tríada fundamental: interfaz, instrucción y modelo
- La IA como copiloto cognitivo del Data Scientist
- El impacto transversal: más producción al mismo costo
- Privacidad, riesgos y la necesidad de gobernanza
- Del Data Science tradicional al Data Science moderno
- Nuevas arquitecturas: automatización, agentes, RAG y MCP
- Requisitos fundamentales para trabajar profesionalmente con IA
- Conclusión: la IA como extensión intelectual del Data Scientist

Modulo 2: Insights de datos no estructurados

- Introducción: cuando el Data Science deja de ser solo números
- El surgimiento del Data Science aplicado a lenguaje humano
- El potencial oculto de los datos no estructurados
- Un caso concreto: extraer valor real de una reunión
- Qué diferencia a Data Science de Machine Learning en este nuevo contexto
- Incorporación de IA al ciclo clásico del Data Science
- Herramientas tecnológicas que habilitan el proceso
- La ingeniería de prompts como disciplina analítica
- Conversar con los datos: un nuevo paradigma
- La necesidad de supervisión humana: límites y riesgos
- Conclusión: el Data Scientist como intérprete del lenguaje organizacional

Modulo 3: Newsletter con IA

- Introducción: cuando las conversaciones se convierten en datos
- Por qué ahora: el salto tecnológico que habilita un nuevo tipo de análisis
- El caso práctico: la reunión semanal como fuente de conocimiento estratégico
- El rol del Data Scientist en esta nueva modalidad: analizar lenguaje, no solo números
- El flujo completo: de conversación informal a newsletter profesional
- Automatización total o asistencia híbrida: dos caminos para implementar el sistema
- El arte de construir prompts que realmente producen análisis
- Responsabilidades diferenciadas: IA, Data Scientist y Community Manager
- Riesgos, limitaciones y necesidad de supervisión humana
- Conclusión: un nuevo tipo de inteligencia organizacional

Modulo 4: Introducción a sistemas multi-agente (MCP-Básico)

- Introducción: el paso de usar IA a construir sistemas inteligentes
- ¿Qué es un sistema multiagente? Una nueva forma de inteligencia distribuida
- MCP como cerebro del sistema: orquestación, estrategia y control
- RAG como memoria externa: cómo las máquinas recuerdan
- La sinergia MCP + RAG: pensar, buscar y responder
- Cliente y servidor: el sistema distribuido que sostiene a los agentes
- Herramientas: los “superpoderes” del MCP
- ¿Cómo razona realmente un MCP?
- Aplicaciones reales de sistemas multiagente en empresas
- La práctica: de Relevance a Python, y de Python a n8n•
- Conclusión: el Data Scientist como arquitecto de inteligencia distribuida

Modulo 5: Analisis de datos con GenIA

- Introducción: cuando la IA generativa acelera todo el ciclo de Data Science
- El desafío del negocio: construir una unidad de detección automática de fugas
- El rol de IA en la búsqueda de datasets: investigación automática en minutos
- El desafío técnico detrás de los sensores IoT: entender la deriva, el ruido y los umbrales
- Definir correctamente el problema de Machine Learning
- EDA y limpieza de datos asistidos por IA generativa
- Generación del modelo: cuando la IA crea código y propone enfoques predictivos
- Reglas de negocio y umbrales: el puente entre ingeniería y ciencia de datos
- Integración con n8n: llevar el modelo a producción y activar alertas reales
- Conclusión: IA generativa como acelerador, no como reemplazo

Modulo 6: Buenas prácticas de para la gestión de datos

- Introducción: cuando la IA se vuelve parte del sistema operativo de la empresa
- La IA como activo económico, no como juguete tecnológico
- El principio rector: costo–beneficio como marco permanente
- Diligencia técnica: la base invisible de la confianza
- Valor agregado: cuando la automatización realmente importa
- Feedback humano: el mecanismo que evita la degradación del sistema
- Monitoreo de costos: visibilidad como condición de control
- Documentación automática: construir memoria organizacional
- Evaluar, deprecar y priorizar: la vida útil de los agentes
- La IA como sistema gobernado, no como entidad autónoma

Proyecto Final

- Proyecto Final



Aprender Online UTN

UNIVERSIDAD TECNOLÓGICA NACIONAL
FACULTAD REGIONAL RECONQUISTA

aprenderonline.frrq.utn.edu.ar

Aprender Online se reserva el derecho de modificar este programa, o cualquiera de los componentes del curso, cuando lo considere necesario a fin de generar mejoras o actualizaciones.