

## Diplomatura en Data Science

La Modalidad de Cursado es Virtual Sincrónica, con una clase virtual teórica/práctica los días martes 18:00 a 22:00 horas, grabadas y disponibles para su posterior visualización. Además, habrá laboratorios prácticos a desarrollar por el asistente, para lo cual se dispondrá de consultas a través de medios de comunicación digitales. Los módulos son 7 y se contará con un proyecto de ciencia de datos por módulo a desarrollar en grupos. Dichos proyectos se trabajarán en modalidad de sprints, como en la industria, y durante el sprint los docentes asumirán el rol de Product Owner.

## PROPUESTA CURRICULAR

### Objetivos generales:

Que los asistentes:

- Entiendan la importancia del análisis de datos y el valor que tienen los datos en la cuarta revolución industrial
- Comprendan el proceso de Ciencia de Datos, independientemente de las herramientas y tecnologías utilizadas.
- Conozcan los distintos aportes de valor que puede brindar la ciencia de datos a las diferentes industrias.
- Comprendan los distintos tipos de Aprendizaje Automático, incluyendo métodos supervisados, no supervisados y redes neuronales.
- Diseñen, implementen y gestionen efectivamente proyectos de Analítica Avanzada de Datos.
- Asimilen los fundamentos de Big Data y Cloud Computing.

### Objetivos específicos:

- Analizar, integrar y validar distintos orígenes de datos, así como la calidad de dichos datos.
- Aplicar técnicas de Exploración, Limpieza y Curado de Datos, reconociendo los fundamentos teóricos detrás de estas técnicas.
- Planificar, dirigir y ejecutar procesos completos de Ciencia de Datos.
- Reconocer las distintas alternativas a la gestión de proyectos de Ciencia de Datos y seleccionar la más adecuada al caso en cuestión.
- Identificar los puntos críticos del proceso de análisis de datos.
- Conocer las distintas herramientas disponibles y utilizarlas para mejorar la eficacia y eficiencia del proceso de Ciencia de Datos.

### CONDICIONES DE ADMISIÓN

- Profesionales universitarios de carreras tecnológicas
- Graduados de tecnicaturas superiores universitarias
- Egresados de Institutos Superiores en carreras vinculadas con las Tecnologías Informáticas.
- Egresados de Tecnicaturas Superiores en carreras vinculadas con las Tecnologías Informáticas.

- Profesionales ejerciendo tareas de Big Data, Data Science o afines.
- Estudiantes Avanzados, Graduados e Investigadores de cualquier carrera con conocimiento de grado en estadísticas y lógica computacional, a validar por la Comisión Académica. En caso de no poseer conocimientos de programación, se puede sugerir un módulo 0, de nivelación, a cargo del asistente.

## Temario Abreviado

### **Módulo 1: Ciencia de Datos y Análisis Exploratorio (30 horas)**

**Breve descripción:** En este módulo se abordarán los fundamentos de Ciencia de datos, los roles dentro de la disciplina, la estructuración de un proyecto y las metodologías de trabajo más frecuentes. Durante los laboratorios se llevará a cabo un análisis exploratorio sobre un dataset propuesto y se construirá un caso de uso de Churn.

**Caso de Uso:** Modelo de Churn.

### **Módulo 2: Fundamentos de Aprendizaje Automático (30 horas)**

**Breve descripción:** En este módulo se trabajará sobre los fundamentos de Machine Learning, profundizando en el aprendizaje supervisado. Se abordarán los métodos menos complejos (Regresiones y árboles de decisión) y el proceso de Ingeniería de Variables. En el Laboratorio se desarrollará un modelo predictivo.

**Caso de Uso:** Modelo de Predicción.

### **Módulo 3: Aprendizaje Automático no Supervisado - Clustering (30 horas)**

**Breve descripción:** Se profundizará sobre el aprendizaje automático no supervisado, abordando también el tratamiento de variables utilizado para segmentación. Se evaluarán distintas técnicas de Segmentación y de reducción de dimensiones. Finalmente se abordará el proceso de Descripción de los Clústers y las mejores formas para transmitir la información obtenida de esa segmentación. Durante los laboratorios se llevará a cabo un modelo de Segmentación de Clientes.

**Caso de Uso:** Modelo de Clustering.

### **Módulo 4: Redes Neuronales (30 horas)**

**Breve descripción:** En este módulo se desarrollarán los fundamentos de redes neuronales, trabajando sobre el concepto de neuronas, tipos de activación y capas dentro de una red. También se abordará el concepto de Cloud Computing así como sus ventajas y desventajas. Se revisarán algunos casos de uso de utilidad y se llevará adelante un laboratorio de predicción a través de redes neurales.

**Caso de Uso:** Modelo de Red Neuronal.

### **Módulo 5: Deep Learning y Computer Vision (30 horas)**

**Breve descripción:** Se brindarán los conceptos necesarios para entender el Aprendizaje Profundo, incluyendo los distintos tipos de casos de uso con los que se puede utilizar. Se trabajará realizando sobre el mecanismo de fine tuning para aprovechar el aprendizaje ya entrenado, realizando transferencia de conocimiento. Se abordarán conceptos clave de computer vision y las tecnologías utilizadas. Durante el laboratorio se realizará un proyecto de clasificación de objetos / personas en imágenes.

**Caso de Uso:** Clasificación de Objetos a través de Aprendizaje Profundo.

### **Módulo 6: Procesamiento de Lenguaje Natural (30 horas)**

**Breve Descripción:** Este módulo abordará los fundamentos del Procesamiento de Lenguaje Natural junto a los desafíos más frecuentes. Se abordarán los conceptos de Lenguaje y del valor semántico de las palabras. Se brindarán las herramientas necesarias para realizar el análisis de sentimientos y chatbots. En laboratorio se llevará adelante un proyecto de análisis de sentimientos en datos de redes sociales.

**Caso de Uso:** Procesamiento de Lenguaje Natural en Análisis de Sentimientos.

### **Módulo 7: NLP en Sistemas de Recomendación (30 horas)**

**Breve Descripción:** Este módulo es totalmente opcional y aquellos asistentes que decidan desarrollarlo lo llevarán a cabo de forma completamente virtual, contando con seguimiento del cuerpo docente. En este módulo se trabajará sobre qué es un sistema de recomendación y cuáles son los casos de uso más frecuentes, además de los fundamentos y herramientas necesarias para la construcción del mismo. Durante el laboratorio se construirá un sistema de recomendación de películas o música sencillo.

**Caso de Uso:** Construcción de un Sistema de Recomendación.