

PLAN DE ESTUDIOS

Data Science





¿Quiénes somos?

Somos la academia digital que te acompaña en el desafío de aprender tecnología y te prepara para la nueva era del trabajo.

Ofrecemos programas ágiles, co-creados y certificados por referentes de la industria. Nuestro modelo educativo replica entornos profesionales de tecnología para que desarrolles, no sólo las habilidades técnicas, sino también las competencias digitales que necesitas para ser protagonista en un mundo cambiante, conectado y colaborativo. Data Science es sólo el comienzo de un mundo de posibilidades.

Objetivo del programa

¿Para qué estudiar Data Science?

Hoy en día, tenemos la capacidad de producir y almacenar infinitas cantidades de datos de distintas fuentes; ya sea en fotografías, reportes de ventas o códigos de ADN. El dilema es ¿cómo usamos esa inmensa cantidad de datos para crear impacto?

El perfil del data scientist hace uso de distintas herramientas y procesos para convertir datos en información y, de esta forma, tomar las decisiones que hoy están dándole forma al mundo. La ciencia de datos nos ayuda a hacernos preguntas sobre el pasado, el presente y el futuro, a analizar los datos pertinentes y trazar un nuevo camino hacia adelante. Es un campo aplicable, literalmente, a cualquier ámbito profesional. Lo que hace que su potencial sea infinito.



¿Cuál es la diferencia con Data Analytics?

Data Science y Data Analytics son dos disciplinas muy amigas entre sí. Son más las cosas en las que se parecen que aquellas en las que se diferencian, pero no por nada existen las dos. Ambas trabajan con datos, algo que ya nos los dice el nombre, pero la principal diferencia radica en lo que cada una hace con esos datos.

Por un lado, el perfil del Data Analyst se enfoca en un área específica del negocio o industria en la que trabaja. Se preocupa por acceder a datos y, a partir de ellos, extraer métricas y tendencias que sirvan al negocio para contar historias y tomar decisiones más acertadas. Es decir, como Data Analyst serás experto en relevar datos, entender cómo se relacionan entre sí y crear visualizaciones que permitan a la compañía utilizarlos para obtener el mejor resultado posible.

Por otro lado, el perfil del Data Scientist se ocupa de entender qué datos son necesarios para responder a una pregunta específica, de interpretar información extraída de esa inmensa cantidad de datos y, a través de Machine Learning, puede ir más allá y crear modelos para predecir ocurrencias futuras a partir de patrones pasados. Como Data Scientist no sólo podrás generar conclusiones basadas en el conocimiento de datos, sino que también podrás adelantarte a estas cuestiones creando sistemas autónomos predictivos.

Ambas habilidades se encuentran entre las más requeridas dentro del mundo de los negocios, independientemente de la industria o el rol.

¿A quién está dirigido este programa?

Perfil del ingresante

El programa intensivo de Data Science, está pensado para profesionales que buscan aprender a analizar datos de manera eficiente y aplicar herramientas modernas en su trabajo. También está orientada a personas que se interesan por saber de qué se trata esta área del conocimiento cercana al Machine Learning, el Deep Learning y la Inteligencia Artificial.

Este programa te desafía a potenciar tu creatividad, usando la lógica y la estadística para tomar decisiones que moldean el mundo.



Requisitos

Dedicación

La dedicación estimada para el programa es de 14 horas semanales que involucran 2 encuentros semanales online de 2.30 horas cada uno + la preparación individual asincrónica desde la plataforma de Acámica + el tiempo de trabajo individual en los sprint projects.

¿Necesito tener experiencia o conocimientos previos?

Programación: si bien no es obligatorio, contar con algún conocimiento previo de programación te será muy útil. En Acámica enseñamos desde cero, pero a buen ritmo.

Inglés: ten en cuenta que en este programa deberás leer textos en inglés, tal como lo hacen los Data Scientist en sus trayectorias profesionales. De esta forma, es importante que conozcas el idioma y puedas dominar su lectura.

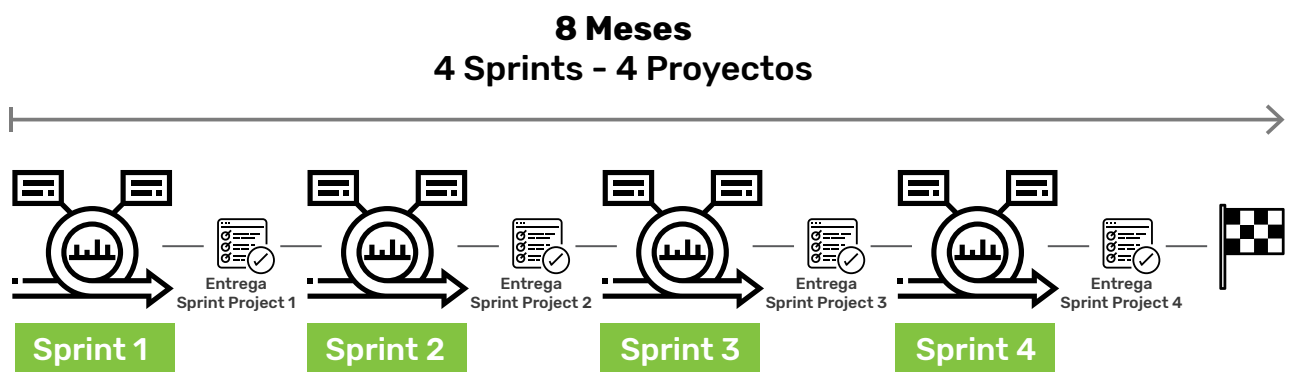
Matemática y Estadística: es importante que conozcas conceptos elementales tanto de matemática básica, -como funciones trigonométricas, el concepto de límite, las derivadas e integrales-, y ciertos conceptos fundamentales de estadística -como el promedio, la desviación estándar, la distribución normal y distribución binomial-.

Por último, durante el sprint 4, aprenderás a utilizar herramientas, usando la plataforma Amazon Web Services (AWS). Si bien es requisito contar con una tarjeta de crédito para suscribirse, las herramientas que utilizaremos en el marco de este programa no te traerán ningún costo adicional.

Hoja de ruta

¿Cómo está estructurado el programa?

Completa tu programa en 4 sprints de aprendizaje. Cada sprint tiene una duración de 9 semanas y finaliza con la entrega de un proyecto en el que pones en práctica los conceptos aprendidos mientras que comienzas a armar tu portafolio profesional.



Sprint 1

Introducción a Data Science

Sprint project: Mi primer modelo de Machine Learning.

Sprint 2

Herramientas Avanzadas

Sprint project: Ingeniería de Features, Modelos Avanzados e Interpretación de Modelos.

Sprint 3

Aplicaciones

Sprint project: Aplicaciones actuales.

Sprint 4

Predicción, inferencias e incertezas

Sprint project: Informe final.





SPRINT 1

1

INTRODUCCIÓN A DATA SCIENCE

Objetivo

En este primer sprint, empezarás a conocer las herramientas fundamentales de Data Science. Trabajarás particularmente en Análisis Exploratorio de Datos y verás una introducción a Machine Learning. Para eso, conocerás un lenguaje de programación (Python), junto con las librerías propias del análisis de datos. Abordarás problemas de Data Science concentrándote en estas preguntas:

- Dado un problema, ¿qué datos me pueden ayudar a abordarlo y resolverlo?
- ¿Dónde y cómo puedo conseguir esos datos?
- Dado un dataset, ¿qué problemas interesantes puedo trabajar con él?

Temas a abordar en el sprint

- Programación (clases y funciones)
- Python y librerías como Numpy, Pandas, Matplotlib, Seaborn
- Análisis exploratorio de datos (EDA)
- Introducción a Machine Learning: clasificación y regresión
- Tipos de gráficos
- Modelos básicos (árboles, KNN, regresión lineal, regresión logística)
- Métricas de evaluación
- MAE
- Train/test Split

Proyecto

Mi primer modelo de Machine Learning

En este proyecto, deberás seleccionar una problemática de entre varias opciones. El desafío consiste en indagar qué datos podrían ayudarte a abordarla, realizar un análisis exploratorio de datos y entrenar un modelo sencillo de Machine Learning para resolverla.



SPRINT 2

2

HERRAMIENTAS AVANZADAS

Objetivo

En este sprint, profundizarás en fundamentos estadísticos y conocerás modelos avanzados de Machine Learning. Además, podrás enfocarte en la interpretación de los modelos obtenidos: no sólo aprenderás un modelo que predice correctamente, sino que explorarás por qué predice lo que predice.

Temas a abordar en el sprint

- Machine Learning en profundidad
- Fundamentos estadísticos
- Interpretación de modelos
- Estadística: distribuciones, Teorema de Bayes
- Modelos avanzados: ensambles, SVM, Naive Bayes, redes neuronales
- Optimización de parámetros
- Interpretación de modelos
- Ingeniería de features: One Hot Encoding, LabelEncoding, Scaling, etc.

Proyecto

Ingeniería de Features, Modelos Avanzados e Interpretación de Modelos

En este proyecto, el desafío es aplicar ingeniería de features y modelos avanzados para desarrollar con mayor profundidad tu modelo de Machine Learning. ¿Qué puedes aprender del problema que estás abordando mediante el estudio de tu propio modelo?

SPRINT 3

3

APLICACIONES

Objetivo

En este sprint, utilizarás herramientas aprendidas en dominios específicos pero muy relevantes: procesamiento del lenguaje natural, sistemas de recomendación, análisis de imágenes, series de tiempo. Te impulsaremos a que puedas desempeñarte cada vez con mayor autonomía para resolver problemas.

Temas a abordar en el sprint

- Procesamiento del lenguaje natural
- Sistemas de recomendación
- Análisis de imágenes
- Series de tiempo
- Procesamiento del lenguaje natural: normalización (Tokenización, Stemming, Lemmatization, etc.) n-gramas, vectorización, Bag of Words, tfidf, embeddings, arquitecturas modernas.
- Sistemas de recomendación: matriz de utilidad, filtro colaborativo y basado en contenidos, descomposición UV. Ejemplos.
- Análisis de imágenes: tipos de imágenes, aplicaciones, redes neuronales, convolucionales, transfer learning
- Series de tiempo: tipos de series de tiempo, componentes (estacionalidad, tendencia, etc.). Problemáticas propias de series de tiempo: clasificación, regresión, predicción, detección de puntos de quiebre. One-step y Multi-step prediction.

Proyecto

Aplicaciones actuales

En este proyecto, el desafío consiste en resolver problemas de relevancia contemporánea aplicando el procesamiento del lenguaje natural, sistemas de recomendación, series de tiempo y análisis de imágenes.

SPRINT 4

4

PREDICCIONES, INFERENCIAS E INCERTEZAS

Objetivo

Cuando trabajamos con probabilidad y estadística, muchas veces suelen aparecer componentes que se nos presentan como anti intuitivos. En este último sprint, buscarás “bajar a tierra” estos cálculos. La pregunta que te guiará es: ¿qué incerteza puedo adjudicarle a esta predicción? También verás cómo las simulaciones computacionales nos pueden ayudar a evitar algunos obstáculos. Al finalizar, no sólo podrás utilizar modelos para hacer predicciones, sino también para estimar de manera más integral los errores e incertezas involucrados.

Temas a abordar en el sprint

- Estimación de parámetros y ajuste de funciones
- Test de hipótesis
- Métodos de muestreo
- Análisis discriminante lineal y cuadrático.
- Puesta en producción
- Repaso de probabilidad, probabilidad condicional, Teorema de Bayes, distribuciones.
- Ley de los grandes números, Teorema central del límite.
- Cuadrados mínimos, máxima verosimilitud.
- Significancia estadística, tamaño del efecto, P-valor. Relación con falsos positivos y falsos negativos. Corrección de Bonferroni
- Regresión logística en un marco probabilístico.

Proyecto

Informe final

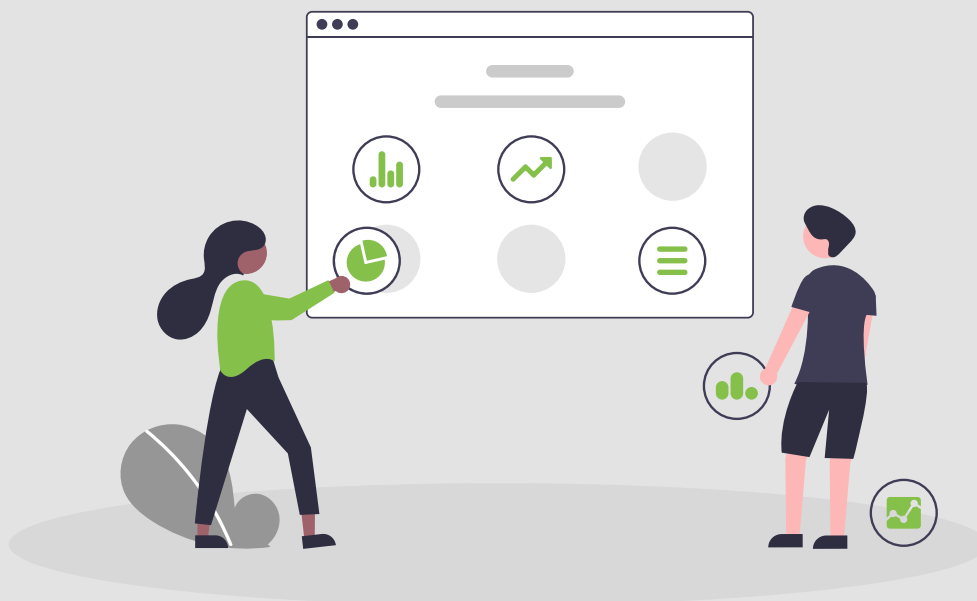
En este proyecto, el desafío es elaborar un informe final que describa el proceso de toma de decisiones. Deberás sustentar la razón por la cual usaste las librerías y aplicaste los métodos al dataset. Todos tus hallazgos deberán ser comunicados en función de la problemática y las preguntas que buscabas responder. No te limitarás únicamente a hacer predicciones, sino que además, podrás estimar la incerteza asociada a cada una.

¿Qué podrás hacer cuando termines?

Perfil del egresado y egresada

Al finalizar el programa, habrás adquirido los conocimientos y habilidades necesarias para tomar datos en estado crudo y aprender de ellos, ya sea visualizándolos de manera correcta y eficiente, o entrenando modelos (estadísticos y/o de Machine Learning) que aprendan de ellos y te ayuden a predecir. Llegarás a conclusiones que te ayudarán a tomar mejores decisiones para accionar, predecir y hasta moldear futuros comportamientos.

Además, la industria tecnológica se caracteriza por trabajar de manera ágil y cada vez más organizaciones incorporan este enfoque en sus equipos de trabajo. A lo largo del programa, no sólo entenderás de qué se tratan, sino que participarás activamente de ceremonias del mundo ágil que son llevadas a cabo por todos los equipos que forman parte de la industria digital. Nuestra metodología de aprendizaje fue diseñada para que, al tiempo que aprendes las habilidades técnicas, incorpores las competencias digitales que son fundamentales para la nueva era del trabajo. Es lo que marca la enorme diferencia entre enseñar habilidades tecnológicas y formar profesionales digitales.



ACÁMICA

CONTACTO

admisiones@acamica.com