

CIENCIA DE DATOS APLICADA PARA TODOS

JUSTIFICACIÓN

La ciencia de datos o Analítica de datos está siendo aplicada cada vez más en todas las áreas del conocimiento a nivel profesional. Esta tendencia exige que el profesional conozca las herramientas básicas y potenciales aplicaciones de la ciencia de datos en su contexto. Por otro lado, a nivel personal y empresarial el volumen y variedad de datos crece a un ritmo acelerado, para lo que se requiere profesionales con las herramientas para procesar, analizar, predecir y tomar decisiones.

En el curso de ciencia de datos aplicada para todos se ofrecen las herramientas para análisis de datos y entrenamiento de modelos predictivos, empezando desde un nivel básico hasta un nivel intermedio garantizando que profesionales con poca o ninguna experiencia en análisis de datos, adquiera los conocimientos necesarios para aplicar análisis descriptivos, exploratorios y predictivos. El enfoque aplicado del programa busca que a través de la práctica y entrenamiento en casos de estudio y proyectos de clase aplicados los estudiantes adquieran confianza y experiencia para implementarlo a nivel profesional.

La ciencia de datos ya no representa el futuro sino el presente de las herramientas que se aplican para resolver los grandes problemas a nivel público y privado. Este curso busca complementar la formación profesional del estudiante con técnicas de vanguardia para la solución de problemas.

RESULTADOS DE APRENDIZAJE

- Identificar los tipos de análisis y problemas que se pueden resolver con la ciencia de datos para detectar oportunidades de aplicación en el entorno profesional
- Desarrollar análisis exploratorios de datos en contextos multidisciplinarios para identificar relaciones entre variables y plantear hipótesis de trabajo
- Visualizar datos de forma interactiva y creativa para la identificación de patrones e información clave que faciliten la toma de decisiones
- Entrenar y validar modelos supervisados de clasificación y regresión para predecir e identificar variables claves en problemas multidisciplinarios
- Desarrollar modelos no supervisados como técnicas de clustering para la identificación de patrones en problemas multidisciplinarios
- Desarrollar procesamiento de datos, construcción de modelos y análisis a través de Python (Pandas, seaborn, plotly, statsmodels, scikit-learn) para conocer las herramientas de programación más utilizadas a nivel mundial
- Desarrollar proyectos de análisis de datos desde la recolección de datos hasta la modelación predictiva para la toma de datos en el sector empresarial o a nivel personal.
- Aprender tendencias de la ciencia de datos e inteligencia artificial para hacer uso de procesamiento en la nube o algoritmos de deep learning cuando el contexto lo requiera.

METODOLOGÍA

Todas las sesiones tienen un alto contenido práctico con acompañamiento del instructor. Las metodologías de las sesiones incluyen principalmente

- Casos de estudio con datos reales en problemas multidisciplinarios
- Trabajo participativo y colaborativo durante las sesiones para desarrollar casos asignados
- Aprendizaje por proyectos con acompañamiento para formular y desarrollar proyecto aplicado

DIRIGIDO A

El curso está diseñado para profesionales de ingeniería, negocios, y áreas afines que tengan el deseo de aprender de analítica de datos y aprendizaje de máquina para aplicarlo en un contexto empresarial. El enfoque multidisciplinario permite que profesionales de otras áreas con interés de aprender de estos conceptos y herramientas realicen el programa. No se requiere experiencia previa en programación.

El programa también está dirigido a empresas que desean capacitar a sus profesionales en áreas de planeación de operación para que aprendan el marco conceptual y herramientas propias de la ciencia de datos, que puedan ser aplicadas en la organización.

CONTENIDO

MÓDULO I. Introducción a la ciencia de datos (12 horas)

- Metodología de análisis de datos
- Exploración de problemas de aprendizaje de máquina
- Introducción al análisis de datos con Python
- Análisis exploratorio de datos (EDA)
- Visualización de datos

MÓDULO II. Modelos predictivos supervisados (20 horas)

- Problemas de regresión y sus aplicaciones
- Problemas de clasificación binaria y multi-clase y aplicaciones en empresas
- Análisis de redes sociales (Twitter)

MÓDULO III Modelos no supervisados (10 horas)

- Clustering para la segmentación de clientes y grupos
- Problemas de recomendación y asociación
- Introducción a aprendizaje por refuerzo y aplicaciones

MÓDULO IV Modelos y aplicaciones avanzadas (6 horas)

- Introducción a redes neuronales y Deep learning
- Introducción a servicios de computo en la nube (AWS, Azure, etc)
- Proyecto aplicado

EXPERTO FACILITADOR

DANIEL ROMERO RODRIGUEZ

M.Sc y PhD en Ingeniería industrial de University of South Florida. Docente de la universidad del Norte en cursos de análisis de datos 1 y 2. Investigador en proyecto internacional para la redistribución voluntaria de migrantes con herramientas de machine learning en colaboración con el Immigration Policy Lab de la universidad de Stanford. Director del Analytics Research Lab UN, semillero para la divulgación y aplicación de la ciencia de datos.

Profesional en ingeniería con experiencia en ciencia de datos con Python para realizar acompañamiento a los estudiantes durante las sesiones de clase. Maestría en curso o finalizada es deseada.

DURACIÓN DEL PROGRAMA

48 HORAS

MODALIDAD

REMOTO



CECUninorte

Mayores informes
Teléfonos: (57-5) 3509222
cec@uninorte.edu.co

www.uninorte.edu.co/web/educacion-continuada