

Programa del curso

Nombre del curso:	Ciencia de Datos Aplicada
Course Name:	Applied Data Science
Créditos:	4
Profesor:	<ul style="list-style-type: none"> • Juan Pablo Reyes (jp.reyes39@uniandes.edu.co) • Fabián Camilo Peña (fc.pena@uniandes.edu.co) • Nancy Marcela Cubillos (ncubillos@uniandes.edu.co)
Versión PDF	Click Aquí

Descripción

Este curso busca principalmente desarrollar habilidades en los estudiantes para la selección y uso de técnicas y herramientas computacionales y estadísticas adecuadas para la preparación y el análisis de datos, incluyendo la construcción de mecanismos efectivos para la entrega de resultados y que a su vez puedan ser usados por los stakeholders de la organización interesados así como por otros sistemas computacionales que los requieran.

Objetivos

- Entender los fundamentos de la ciencia de datos, el rol del científico de datos en las organizaciones y las metodologías actuales para el desarrollo de este tipo de proyectos.
- Desarrollar habilidades para la selección y el uso adecuado de técnicas y herramientas computacionales y estadísticas, durante las diferentes etapas de un proyecto de ciencia de datos.
- Preparar y analizar datos de pequeña y mediana escala, provenientes de diferentes fuentes, con el fin de extraer información útil y accionable para una organización.
- Comprender los elementos básicos para la comunicación efectiva durante la fase de entrega de resultados a los stakeholders de la organización, habilitando la toma de decisiones con base en los datos.
- Aplicar mecanismos básicos de despliegue u operacionalización de modelos de machine learning para que puedan ser integrados efectivamente con los demás procesos o sistemas de la organización.

Plan de temas

1. Fundamentos de la ciencia de datos, el rol del científico de datos, metodologías, creación de valor, productos de datos.

2. Análisis Exploratorio de Datos (EDA), visualización de datos, inferencia estadística.
3. Limpieza y preparación de datos.
4. Fundamentos del Machine Learning, evaluación de modelos, interpretabilidad y Explicabilidad.
5. Despliegue y monitoreo de modelos, pruebas en producción.

Metodología

El curso se desarrolla en torno a casos de estudio los cuales deben ser desarrollados utilizando como guía el material proporcionado durante las sesiones de clase así como la bibliografía sugerida. Se proponen sesiones de clases teóricas complementadas con talleres prácticos que permitan apropiar los diferentes conceptos, habilidades y uso de herramientas que son objeto de aprendizaje del curso. Los estudiantes deben preparar previamente cada sesión con el fin de realizar discusiones que permitan profundizar en los temas vistos así como avanzar en los talleres basados en los casos propuestos durante la clase.