

Curso: Data Science (Introducción)
Sección A: Lunes y Viernes 13:00 – 14:20 horas
Catedrático: Franz Haidacher - franzhaidacher@ufm.edu (5051-0103)
Auxiliar: TBA

Estructura del curso y metodología:

Para el desarrollo de este curso se aplicarán diversas metodologías de aprendizaje:

- Clase magistral
- Lectura e investigación de temas clave
- Aprendizaje basado en proyectos (Project-Based Learning)

El curso contará con algunas clases teóricas que apoyen la metodología PBL. El curso estará sujeto a exámenes según los horarios planificados por la UFM.

Se realizarán algunos ejercicios en clase para la inmersión de los alumnos en la metodología. Los alumnos serán asignados a grupos de trabajo de una forma aleatoria que el maestro considere conveniente. Se buscará que los estudiantes trabajen con diversos compañeros como un ejercicio de cara a la vida profesional.

Materiales del curso:

La metodología a seguir en este curso, permite que los alumnos aprendan de acuerdo a su propio estilo de aprendizaje. Los alumnos tienen la libertad de utilizar el material de referencia que consideren conveniente.

Este curso tiene una gran diversidad de contenido (fundamentos de estadística, recopilación de datos, análisis de información, presentación de datos, etc.), razón por la cual, no existe un único libro que sirva como guía única del curso. Sin embargo, hay libros, materiales y recursos que se utilizarán como referencia para apoyar el proceso de aprendizaje. Los mismos se harán disponibles en el portal de MiU.

Contenido del Curso:

El contenido del curso está basado en varias fuentes, incluyendo el libro de texto “Foundations of Data Science” y diversos recursos en línea.

Hay 33 sesiones disponibles para clase magistral y trabajo en equipo.

A continuación se describen los temas que serán cubiertos en clase.

Temas:

- Introducción
- Análisis exploratorio de datos (EDA)
- Transformaciones y tidy data
- Conceptos esenciales de visualización
- Presentaciones efectivas
- Aplicaciones y herramientas de Data Science
- Data Science Roadmap
- Tipos de problemas en Data Science (Scikit-Learn)
- Clasificación
- Regresión
- Clustering
- Reducción de dimensionalidad
- Feature engineering
- Deep learning

Distribución de la calificación:

Quizzes, discusiones de lectura, ejercicios en clase y presentaciones	40 puntos
Primer parcial	15 puntos
Segundo parcial	15 puntos
Examen final	30 puntos

Asistencia a clase:

La asistencia a clase no recibirá calificación; sin embargo, los alumnos deberán haber asistido al menos al 75% de las clases a efecto de tener derecho a la nota final del curso. Si el alumno debe ausentarse por motivos justificados (e.g. por viaje) deberá notificarse con anticipación para que la ausencia no impacte el porcentaje de asistencia. Ausencias por enfermedades comunes cuentan contra el porcentaje permitido de inasistencia.

Bibliografía recomendada:

- Lantz, Brett. Machine Learning with R. 2015 Packt Publishing ISBN 978-1-78439-390-8
- Gollapudi, Sunila. Practical Machine Learning. 2016 Packt Publishing ISBN 978-1-78439-968-9
- Garren, Witten, Hastie & Tibshirani. An Introduction to Statistical Learning 2017 Springer ISBN 978-1-4614-7137-0