



Plan de estudios

Materia

Data Science I

Curso: 1

Créditos: 9 ECTS

Idioma: Castellano

Competencias

Competencias básicas:

CB1. Que los estudiantes hayan demostrado poseer y comprender conocimientos en un área de estudio que parte de la base de la educación secundaria general, y se suele encontrar a un nivel que, si bien se apoya en libros de texto avanzados, incluye también algunos aspectos que implican conocimientos procedentes de la vanguardia de su campo de estudio

CB2. Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.

CB5. Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía.

Competencias Generales / Personales:

CG4. Pensamiento flexible y con mirada amplia

Competencias Específicas / Profesionales:

CE1. Entiende, conoce y aplica las teorías, herramientas y procesos para la captura fuentes de datos de diferente naturaleza y almacenamiento de datos

CE4. Es capaz de formular una problemática relativa al mundo de los datos en términos de modelización

CE7. Planifica, diseña y ejecuta proyectos vinculados con el mundo de los datos, siguiendo el ciclo de vida del dato

Resultados de aprendizaje

RA4. Capacidad para pensar e idear soluciones y respuestas más allá de lo común y estipulado por norma. Capacidad para dar respuesta a inesperadas circunstancias del momento y posibles del futuro, mediante una mirada más allá de lo inmediato (“out of the box”), lo que le permite entender el contexto en el que se encuentra.

RA7. Entender, conocer y aplicar las teorías, herramientas y procesos para la captura fuentes de datos de diferente naturaleza.

RA8. Entender, conocer y aplicar las teorías, herramientas y procesos para el almacenamiento de datos de diferente naturaleza.

RA11. Ser capaz de formular una problemática relativa al mundo de los datos en términos de modelización.

RA15. Planifica, diseña y ejecutar proyectos siguiendo el ciclo de vida del dato.

Contenidos

Introducción a la ciencia de datos

- Excel avanzado
- ¿Qué es la ciencia de datos?
- Fases del ciclo de vida del dato
- Metodología de trabajo en un proyecto de Data Science: CRISP-DM
- El proceso de data discovering

Data cleaning en R

- Proceso de data cleaning: detectando sesgos y errores en los datos
- Detección y tratamiento de valores outliers
- Detección, tipologías y tratamiento de valores missings

Fuentes de datos & preprocesamiento de datos

- Tipos de datos: Estructurado y no estructurados, open source y propietarios, Web Scrapping
- Formato de datos: CSV, TXT, ficheros, Json, bbdd relacionales, APIs, etc.
- Normalización y discretización: Chi-Merge, MDLP y CAIM
- Anonimización y pseudoanonimización: funciones de hasheo, datos sintéticos y privacidad diferencial
- Variables dummies

Diseño de una API

- ¿Qué es una API? Protocolo REST
- Consultas a APIs
- Uso de plumber en R

Repositorios: Git y GitHub

- Introducción a SCV: qué es Git
- Comandos básicos en Git: git init, git config, git add, git status, git commit.
- Trabajo básico en GitHub: clona repositorios, git push, git pull

Gobierno de datos

- Roles en la ciencia y en el gobierno de datos
- Gobernanza del dato

Experiencias de Data Science (visitas de empresa)

Actividades formativas

Las acciones formativas planificadas para este módulo son las siguientes:

- Aprendizaje basado en retos (3 ECTS)
- Trabajo en equipo (2 ECTS)
- Talleres (1 ECTS)
- Recursos online (0,5 ECTS)
- Reflexión (0,5 ECTS)
- Trabajo individual (1 ECTS)

- Realización de proyectos con empresas reales (1 ECTS)

Sistema de evaluación

La evaluación será mediante el sistema de evaluación continua a través de la cual se proporciona una información constante, tanto a los/las profesores como a los/las estudiantes, del proceso de aprendizaje a lo largo del período académico:

- Las actividades formativas de presentación de conocimientos y estudio individual se podrán evaluar con pruebas orales y/o escritas que corresponderán como máximo al 60% de la nota final.
- Las acciones formativas dirigidas a la adquisición de las competencias prácticas de las asignaturas se evaluarán a través de la realización de diferentes actividades (trabajos, casos, retos, etc.) correspondiendo como mínimo a un 40% de la nota final.

Los detalles de evaluación y calificación se harán explícitos en la planificación docente anual de las materias en función de los profesores responsables y de los condicionantes de cada curso.

Bibliografía

- Cady, F (2017). The Data Science Handbook. USA. Wiley