Universidad de Navarra

Técnicas de Recogida de Datos (Grupo A) (MBDS)
Guía docente 2026-27

PRESENTACIÓN

Breve descripción: Esta asignatura tiene como objetivo proporcionar a los estudiantes una comprensión técnica y práctica de los métodos y herramientas utilizados para la extracción y recopilación de datos de diversas fuentes. Estas fuentes incluyen redes sociales (como Facebook o X), páginas web mediante técnicas de web scraping, bases de datos gubernamentales (INE, UE), y dispositivos como smartwatches. El curso aborda la importancia de la recogida de datos desde una perspectiva técnica, pero siempre enmarcada en un contexto de negocio, donde la correcta extracción y tratamiento de los datos se convierte en un factor clave para la toma de decisiones estratégicas.

El lenguaje de programación principal que se utilizará será Python, empleando bibliotecas específicas como BeautifulSoup, Scrapy, Pandas, Requests y otras herramientas relevantes en el campo de la recogida de datos. Además, se discutirá cómo la calidad, validez y relevancia de los datos extraídos influyen directamente en los modelos predictivos y las decisiones de negocio.

- **Titulación**: Máster Universitario en Ciencia de Datos Masivos / Big Data Science
- **Módulo:** Programación y computación
- **Materia:** Preparación y recogida de datos
- **ECTS**: 1,5
- **Curso, semestre**: curso único del Máster, primer semestre
- **Carácter**: obligatoria
- **Profesor responsable**: Sergio Ardanza-Trevijano Moras
- **Profesorado**: Pablo Urruchi
- **Idioma**: castellano
- **Aula, Horario**: ver cronograma

RESULTADOS DE APRENDIZAJE (Competencias)

Competencias Básicas y Generales:

CB6 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación

CB7 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio

CB8 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

CB9 - Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades

CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.

CG2 - Explorar y explotar tanto tecnológica como estratégicamente de los datos un valor clave para diferentes empresas y organizaciones.

CG4 - Analizar los datos masivos y aportar medidas originales y novedosas para contribuir a la mejora de la competitividad de las empresas e instituciones públicas.

CG5 - Analizar los datos que se generan en el día a día, extrayendo conocimiento de los mismos, realizando predicciones y transformándolos en productos y servicios utilizando las herramientas estadísticas de Data Science.

CG6 - Trabajar en equipos de trabajo o grupos de investigación interdisciplinares de forma eficaz y colaborativa.

CG7 - Conocer y entender las herramientas habituales que se utilizan hoy día en el tratamiento de datos masivos.

CG8 - Saber aplicar los principios éticos relativos a la recogida, almacenamiento, y análisis de datos teniendo en cuenta las posibles discriminaciones directas o indirectas derivadas de la toma de decisiones.

Competencias Específicas:

CE11 - Aplicar técnicas y medidas de protección y control de la privacidad, en especial la evaluación del impacto de la protección de datos, la disociación de datos y la anonimización.

CE4.2 - (Programación) Programar con técnicas de programación con Python o lenguaje similar. Almacenamiento y manipulación con Python. Análisis de datos y programación de algoritmos.

PROGRAMA

Programa de la asignatura:

1. Introducción al web scraping y aplicaciones, consideraciones éticas y legales. Scraping de webs estáticas y dinámicas.

2. Extracción de datos estructurados y en tiempo real a partir de APIs.

3. Extracción de registros de firewall y análisis de patrones de seguridad en la red.

4. Extracción de datos de dispositivos IoT.

5. Técnicas de scraping de imágenes.

ACTIVIDADES FORMATIVAS

Clases presenciales teóricas	0
Prácticas con ordenador	5
Trabajos dirigidos	20
Tutorías personalizadas	5
Estudio y trabajo personal	2,5

Pruebas presenciales de evaluación	5

EVALUACIÓN

La nota final será un promedio con las siguientes ponderaciones.

EN LA CONVOCATORIA ORDINARIA

- Asistencia a clase 10%
- Trabajo final 40%
- Entregables 40%
- Presentación de negocio 10%

EN LA CONVOCATORIA EXTRAORDINARIA

- Asistencia a clase 10%
- Trabajo final 40%
- Entregables 40%
- Presentación de negocio 10%

HORARIOS DE ATENCIÓN

Contactar por correo electrónico (*se confirmará*)

- <purruchi>@external.unav.es

BIBLIOGRAFÍA

Selenium with Python https://selenium-python.readthedocs.io/

Chakrabarti, S. (2003) Mining the Web: Discovering Knowledge from Hypertext Data. The Morgan Kaufmann Series in Data Management Systems. https://doi.org/10.1016/B978-1-55860-754-5.X5000-9

László, G. (2018) Website Scraping with Python. Using Beautiful Soup and Scrapy. Springer. https://doi.org/10.1007/978-1-4842-3925-4