

**FICHA IDENTIFICATIVA****Datos de la Asignatura**

Código	44655
Nombre	Analítica web
Ciclo	Máster
Créditos ECTS	3.0
Curso académico	2020 - 2021

Titulación(es)

Titulación	Centro	Curso	Periodo
2221 - M.U. en Ciencia de Datos	Escuela Técnica Superior de Ingeniería	1	Segundo cuatrimestre

Materias

Titulación	Materia	Caracter
2221 - M.U. en Ciencia de Datos	7 - Analítica web	Obligatoria

Coordinación

Nombre	Departamento
MAGDALENA BENEDICTO, JOSE RAFAEL	242 - Ingeniería Electrónica
SORIA OLIVAS, EMILIO	242 - Ingeniería Electrónica
VILA FRANCES, JOAN	242 - Ingeniería Electrónica

RESUMEN

El objetivo de la asignatura es introducir al alumno en el campo de la Analítica Web y desarrollar algunas de las principales herramientas que utilizan los científicos de datos en el desarrollo de sus funciones.

Después de realizar una introducción a la analítica web, se establecerán las bases de este campo de conocimiento y se definirán los KPI (Key Performance Indicators). A continuación se trabajara con una herramienta de analítica web gratuita (Google Analytics) para aplicar los conceptos aprendidos y su aplicación a la gestión de páginas web.

Algunas herramientas de analítica web realizan el proceso conocido como ingeniería de datos (Data Engineering) de forma automática, pero no siempre, se va a estar interesado en estos datos estándar, por lo que se precisaran de algunas herramientas para la captura y recogida de información, la limpieza de datos y la transformación de los mismos a través de procesos automatizados.



El científico de datos tiene, en muchas ocasiones, que diseñar experimentos para poder evaluar correctamente el impacto de los cambios que se van a introducir en páginas web de su organización, por lo que será necesario conocer que tipos de experimentos y en que condiciones se deben realizar.

Otra cuestión de relevancia dentro de la analítica web son los sistemas de recomendación que se utilizan para incrementar el uso o la venta de productos. Se estudiarán y aplicarán los sistemas de filtrado colaborativo y el algoritmo kNN para recomendadores.

Se finalizará la asignatura con la minería de texto y el análisis de sentimientos, con los que se pretende conocer la opinión de los usuarios de un producto o servicio a través de la información disponible en las redes sociales y de internet en general.

Para el desarrollo de la asignatura se utilizarán diferentes herramientas informáticas, entre las que cabe destacar R.

CONOCIMIENTOS PREVIOS

Relación con otras asignaturas de la misma titulación

No se han especificado restricciones de matrícula con otras asignaturas del plan de estudios.

Otros tipos de requisitos

COMPETENCIAS

2221 - M.U. en Ciencia de Datos

- Que los/las 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.
- Que los/las estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo
- 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.
- Ser capaces de valorar la necesidad de completar su formación técnica, científica, en lenguas, en informática, en literatura, en ética, social y humana en general, y de organizar su propio autoaprendizaje con un alto grado de autonomía
- Capacidad de análisis y síntesis, en la elaboración de informes, en la exposición, comunicación y defensa de ideas.
- Ser capaces de asumir la responsabilidad de su propio desarrollo profesional y de su especialización en uno o más campos de estudio, aplicando los conocimientos adquiridos en la identificación de salidas profesionales y yacimientos de empleo.



- Extraer conocimiento de conjuntos de datos en diferentes formatos.
- Analizar datos procedentes de la Web extrayendo conocimiento útil de ellos mediante la aplicación de técnicas de análisis de datos.

RESULTADOS DE APRENDIZAJE

Conocer las bases técnicas de la analítica web y sus principales herramientas de aplicación
Proporcionar al alumno el conocimiento de las principales técnicas para la ingeniería de datos (captura, limpieza y transformación de la información).
Proporcionar al alumno una base para que pueda desarrollar experimentos de analítica web sin necesidad de las herramientas estándar de analítica web.
Conocer y poner en práctica los principales sistema de recomendación, pudiendo aparatarlos a las necesidades de la empresa u organización.
Proporcionar al alumno el conocimiento de la minería de texto para su utilización dentro de la analítica web
Procesar texto para realizar un posterior análisis de los sentimientos del usuario.

DESCRIPCIÓN DE CONTENIDOS

1. Introducción a la Analítica Web y captura de datos

Introducción a la analítica web, definición y utilización de diferentes herramientas para la captura de datos.

2. Obtención de datos a través de APIs. Web scraping

Utilización de algunas APIs (Application Programming Interfaces) de empresas e instituciones para la captura de datos. Descripción de las técnicas más extendidas de web scraping

3. Herramientas de análisis web y contraste A/B

Uso de una herramienta de analítica web gratuita (p. ej. Google Analytics). Definición y utilización de los KPI (Key Performance Indicators) y los contrastes A/B.

4. Sistemas de Recomendación

Exposición de los diferentes sistemas de recomendación y sus aplicaciones actuales



5. Minería de texto y análisis de sentimiento

Minería de texto y sus aplicación a redes sociales. Análisis de sentimientos a través de internet.

VOLUMEN DE TRABAJO

ACTIVIDAD	Horas	% Presencial
Clases teórico-prácticas	30,00	100
Elaboración de trabajos individuales	6,00	0
Estudio y trabajo autónomo	4,00	0
Lecturas de material complementario	3,00	0
Preparación de actividades de evaluación	6,00	0
Preparación de clases de teoría	8,00	0
Preparación de clases prácticas y de problemas	9,00	0
Resolución de casos prácticos	9,00	0
TOTAL	75,00	

METODOLOGÍA DOCENTE

Las clases combinarán la parte teórica con la práctica sin distinción entre sesiones dedicadas a teoría y práctica. Todas las sesiones se impartirán en aula de informática.

En las sesiones teóricas se realizará un desarrollo expositivo de la materia con la participación del estudiante en la resolución de cuestiones puntuales. Las sesiones prácticas irán sincronizadas con las teóricas, y en ellas el aprendizaje se realizará mediante la resolución de problemas, ejercicios y casos de estudio a través de los cuales se adquieren competencias sobre los diferentes aspectos de la materia.

EVALUACIÓN

La evaluación del aprendizaje de los conocimientos y competencias conseguidas por los estudiantes se hará de forma continuada a lo largo del curso, y constará de los siguientes bloques de evaluación:

1. Ejercicios y trabajos entregados durante el curso y/o exámenes parciales: 80% de la nota final.
2. Examen final: 20% de la nota final.

Las calificaciones obtenidas en el apartado 1 se conservarán en las dos convocatorias del curso académico en que hayan sido realizadas, dado que su evaluación sólo es posible en el periodo de docencia.



REFERENCIAS

Básicas

- Aggarwal, C. (2016). Recommender Systems. The Textbook. Ed. Springer, New York
- Chan, Y., Talburt, J., Talley, T.M. (2010). Data Engineering. Mining, Information and Intelligence. Ed. Springer, New York
- Munzert, S. (2015) Automated Data Collection with R: A Practical Guide to Web Scraping and Text Mining. Ed. Wiley & Sons, New York.
- Zhai, C. y Massung, S. (2016) Text Data Management and Analysis: A Practical Introduction to Information Retrieval and Text Mining. Ed. ACM Books, Virginia.
- Practical Web Scraping for Data Science: Best Practices and Examples with Python (2018). Seppe vanden Broucke, Bart Baesens, Apress.
- Website Scraping with Python Using BeautifulSoup and Scrapy (2018) Hajba, Gábor, Apress.

Complementarias

- Croll, A. y Yoskovitz, B. (2014). Lean Analytics. Como utilizar los datos para crear más rápido una startup mayor. Universidad Internacional de La Rioja, S.A., Logroño
- Provost, F. y Fawcett, T. (2013). Data Science for Business: What you need to know about data mining and data-analytic thinking. Ed. O'Reilly Media
- Beasley, M. (2013) Practical Web Analytics for User Experience: How Analytics Can Help You Understand Your Users. Ed. Elsevier, New York

ADENDA COVID-19

Esta adenda solo se activará si la situación sanitaria lo requiere y previo acuerdo del Consejo de Gobierno

1. Contenidos.

Se mantendrán todos los contenidos inicialmente programados en la guía docente para las sesiones teóricas y se transforma en Tareas del Aula virtual determinados ejercicios prácticos a desarrollar por los alumnos.



2. Volumen de trabajo y planificación temporal de la docencia.

La guía preveía 30 horas de clases teórico prácticas de las que restaban el 100% en el momento de la docencia no presencial.

Se han reducido 7,5 horas de clases presenciales para adaptarlas a las video conferencias realizadas con Blackboard Collaborate a través del Aula virtual. Traslado de estas 7,5 horas a la realización de tareas por parte del alumno, bien individualmente, o bien en grupo.

Mantenimiento de la planificación temporal docente tanto en días como en horario. Las clases por videoconferencia se realizan en el horario habitual de clases.

3. Metodología docente.

Sustitución de la clase presencial por la videoconferencia síncrona mediante creación de tareas “Videoconferencia” en el Aula Virtual y ejecución de estas por Blackboard Collaborate el día y a la hora de la clase presencial. Subida al aula virtual de los materiales para estas sesiones (transparencias y apuntes).

Mismos materiales previstos en la guía original para la docencia presencial. Utilización del foro del aula virtual para atender las dudas al acabar la videoconferencia. Suministro de problemas resueltos junto a problemas propuestos a entregar mediante a opción de “Tarea” del aula virtual con resolución de dudas por el sistema de tutorías establecido y presentación de la solución correcta mediante videoconferencia

Sistema de tutorías. Se mantiene el programa de tutorías virtuales (atención en 48 horas laborables máximo por correo electrónico) y en el horario de tutorías presenciales disponibilidad del profesor a través de teléfono o mediante videoconferencia por Skype.

4. Evaluación.

Se mantiene el peso de la evaluación continua que es del 80%, y el examen final 20%.

La prueba de evaluación final será un examen tipo test en el Aula Virtual. El examen constará de 20 preguntas con 4 opciones, y una sola respuesta correcta. Las respuestas correctas se valorarán con 0,5 puntos, las erróneas con -0,15 puntos, y las que no se contesten no se valorarán. Simultáneamente, se abrirá una videoconferencia a través del Aula virtual para que los alumnos puedan realizar consultas con el profesor o transmitir cualquier problema.

Se ha hablado con los alumnos y todos disponen de los medios para acceder desde su lugar de confinamiento al Aula Virtual, a las videoconferencias y a las pruebas o tareas que se carguen en la misma.



5. Bibliografía.

La bibliografía recomendada se mantiene pues es accesible a través de la biblioteca de la Universitat de València o se facilita con enlaces o ficheros a través del Aula Virtual.

