



TECHNICAL SHEET OF THE SUBJECT

| Data of the subject | |
|---------------------|--|
| Subject name | Social and Web Analytics |
| Subject code | DTC-BA-523 |
| Main program | Grado en Análisis de Negocios/Business Analytics |
| Involved programs | Grado en Ingeniería en Tecnologías de Telecom. y Grado en Análisis de Negocios/Business Analytics [Fifth year] |
| Level | Reglada Grado Europeo |
| Quarter | Semestral |
| Credits | 6,0 ECTS |
| Type | Obligatoria (Grado) |
| Department | Department of Telematics and Computer Sciences |
| Coordinator | Francisco Javier Ruiz de Ojeda Fernández |
| Office hours | Send an email to arrange an appointment. |

| Teacher Information | |
|---------------------|--|
| Teacher | |
| Name | Francisco Javier Ruiz de Ojeda Fernández |
| Department | Department of Telematics and Computer Sciences |
| E-Mail | fjruijosed@icai.comillas.edu |

SPECIFIC DATA OF THE SUBJECT

| Contextualization of the subject |
|--|
| Contribution to the professional profile of the degree |
| Social networks have had a significant impact on technology, the economy, and society itself. Being able to exploit and analyze the data generated on social networks allows for the improvement of citizens' quality of life. It enhances the efficiency of processes related to such data. |
| Prerequisites |
| Basic knowledge of programming (Python). |

| Competencies - Objectives | |
|---------------------------|---|
| Competences | |
| GENERALES | |
| CG01 | Capacidad de organización y planificación en la identificación de problemas en el contexto de datos masivos |



| | | |
|-------------|---|---|
| | RA1 | Identificar y organizar temporalmente las tareas necesarias para la realización sus actividades de aprendizaje, cumpliendo con los plazos establecidos de entrega de actividades. |
| | RA2 | Ser capaz de integrarse y participar en el desarrollo organizado de un trabajo en grupo. |
| CG02 | Capacidad de análisis de datos masivos procedentes de diversas fuentes: texto, audio, numérica e imagen | |
| | RA1 | Describir, relacionar e interpretar situaciones y planteamientos sencillos |
| | RA2 | Seleccionar las fuentes de información digital más significativas |
| | RA3 | Identificar las fuentes de información más relevantes para cada problema |
| | RA4 | Utilizar las herramientas digitales idóneas para la adquisición y gestión de la información y es capaz de integrar esa información de una manera racional |
| CG03 | Resolución de problemas y toma de decisiones en un entorno de datos masivos tanto cuantitativos como cualitativos | |
| | RA1 | Resolver problemas y realizar trabajos aplicando una diversidad metodológica de análisis |
| | RA2 | Identificar correctamente los conocimientos aplicables a cada situación |
| | RA3 | Identificar la metodología más apropiada para la resolución de los problemas planteados, conociendo las herramientas software más relevantes y aportando una visión crítica. |
| CG07 | Capacidad de liderazgo y trabajo en equipo, en la sociedad de la información | |
| | RA1 | Participar de forma activa en el trabajo de grupo compartiendo información, conocimientos y experiencias |
| | RA2 | Orientarse a la consecución de acuerdos y objetivos comunes |
| | RA3 | Ser capaz de organizar y/o coordinar equipos de trabajo |
| CG08 | Capacidad crítica y autocrítica en la sociedad de la información | |
| | RA1 | Identificar los supuestos y las limitaciones de los modelos y metodologías a aplicar en cada situación |
| | RA2 | Ejercer una valoración crítica de los modelos y metodologías a aplicar en cada situación |
| CG09 | Compromiso ético en la sociedad de la información | |
| | RA1 | Ser honesto en el desarrollo de las actividades académicas y en otros aspectos de la vida y no ser pasivo ante la deshonestidad de otros. |
| | RA2 | Conocer y desarrollar las bases de la ética profesional |
| CG11 | Capacidad para aprender y trabajar autónomamente en la sociedad de la información | |
| | RA1 | Organizar los materiales de la asignatura con el fin de articular su aprendizaje |



| | | |
|--------------------|---|--|
| | RA2 | Poseer criterio para decidir la utilización de los medios humanos y otros recursos puestos a su disposición para el aprendizaje |
| | RA3 | Realizar sus trabajos y su actividad necesitando sólo unas indicaciones iniciales y un seguimiento básico |
| ESPECÍFICAS | | |
| CE23 | Conocer y utilizar las bases de datos relacionales | |
| | RA1 | Conocer el concepto de base de datos y su importancia para el almacenamiento ordenado de la información |
| | RA2 | Comprender y aplicar el concepto de entidad relación y es capaz de diseñar bases de datos relacionales |
| | RA3 | Conocer la sintaxis básica de los sistemas de recuperación de información más utilizados |
| CE24 | Conocer y aplicar técnicas más efectivas de visualización de datos y su importancia para la extracción de conocimiento | |
| | RA1 | Conocer los tipos de representaciones más importantes y su idoneidad en función del tipo de información. |
| | RA2 | Identificar los parámetros más relevantes en un conjunto de datos mediante técnicas elementales de análisis exploratorio |
| | RA3 | Ser capaz de producir representaciones gráficas que sintetizan conjuntos de datos voluminosos |
| CE25 | Comprender los conceptos más importantes asociados al Big Data y sus elementos característicos | |
| | RA1 | Conocer el concepto de paralelismo y la importancia de la distribución de datos voluminosos |
| | RA2 | Identificar las necesidades específicas asociadas al manejo, adquisición y almacenamiento de volúmenes masivos de datos. |
| | RA3 | Comprender y conocer las tecnologías de comunicaciones (redes) y computación precisas para la distribución y recuperación de información, así como de su análisis. |
| CEO33 | Conocer y usar técnicas de análisis de información para extraer el conocimiento para la toma de decisiones de negocio. | |
| | RA01 | Conoce y comprende las herramientas de apoyo a la decisión en el ámbito de la empresa |
| | RA02 | Es capaz de implementar una solución computerizada de apoyo a la decisión. |
| | RA03 | Comprende y aplica herramientas de adquisición, representación e inferencia, y aquellas técnicas de aprendizaje automático y basadas en sistemas inteligentes |
| CEO34 | Entender la computación y las bases de datos distribuidas, con énfasis en la arquitectura Hadoop y las herramientas actuales más extendidas en el contexto de Big Data. | |



| | | |
|--------------|---|--|
| | RA01 | Conoce y comprende la problemática del Big Data y las técnicas más avanzadas de computación y almacenamiento usadas en la actualidad. |
| | RA02 | Comprende y es capaz de administrar una arquitectura Hadoop y otras arquitecturas NoSQL. |
| | RA03 | Comprende y aplica las técnicas de MapReduce y otras técnicas avanzadas de computación específicas para el análisis de datos voluminosos. |
| CEO35 | Comprender y ser capaz de desarrollar herramientas de visualización avanzadas. | |
| | RA01 | Conoce y aplica las metodologías y herramientas de desarrollo de aplicaciones orientadas a la visualización. |
| | RA02 | Identifica las necesidades de síntesis de información y la representación visual automatizada. |
| CEO36 | Conocer y comprender las herramientas más extendidas de gestión de datos. | |
| | RA01 | Conoce y comprende el modelo de extracción, transformación y carga de datos orientados a la gestión de la empresa. |
| | RA02 | Conoce las tecnologías OLAP y su uso en el contexto de los datawarehouses corporativos. |
| | RA03 | Conoce y es capaz de generar informes basados en la gestión de datos para la toma de decisiones. |
| CEO37 | Conocer y aplicar las técnicas de extracción de información de redes sociales y portales web. | |
| | RA01 | Conoce y comprende las características de las distintas redes sociales y la importancia de su análisis para la extracción de información relevante para la empresa |
| | RA02 | Conoce y es capaz de administrar un sistema de recopilación, agregación y evaluación del <i>¿sentimiento¿</i> asociado a la imagen corporativa en redes sociales |
| | RA03 | Es capaz de analizar métricas para el análisis del tráfico web de un portal así como la cuantificación de la experiencia de usuario de un servicio web. |

THEMATIC BLOCKS AND CONTENTS

Contents - Thematic Blocks

1. Data Exploitation Process
 - Introduction to Social and Web Data
 - Data Exploitation Process
 - Work Environment
2. Data Acquisition and Storage
 - APIs, Authentication, and Data Parsing
 - Data Cleaning
 - Data Storage

- 3. Data Processing
 - Data Processing
- 4. Sentiment Analysis
 - Sentiment Analysis
- 5. User Experience Metrics
 - User Experience Metrics
 - Markov Chains

TEACHING METHODOLOGY

General methodological aspects of the subject

In-class Methodology: Activities

Expository Lesson: The teacher develops the topic that the students have previously read, explaining it on the whiteboard and on the computer. Once the theoretical concepts have been developed, they are applied to a real case through a small program, in which the students contribute and suggest ideas. Collaborative tools are used for the development of this program. The resulting code from each class is made available to the students through the subject's website.

CE25, CEO36, CEO37

Collaborative work, exercise resolution, and tests: In these sessions, doubts will be resolved through collaborative coding of programs to facilitate the understanding of theoretical concepts and the completion of some short programming tests in class. Participation and attitude in class by the student will be valued.

CG01, CG02, CG03,
CG07, CG08, CG09,
CG11, CE23, CE24,
CEO36, CEO37

Laboratory practices. Throughout the course, students will carry out individual and group practices for each of the topics covered. They will start working on the practice at home, and in the laboratory, they will solve any doubts they may have, concluding the practice.

CG03, CG07, CG08,
CG09, CG11, CEO37

Non-Presential Methodology: Activities

Theoretical study:

1. Study and preparation of the topics or concepts to be presented by the teacher.
2. Individual and personal study by the student of the concepts already presented in the lectures.

CG08, CG11, CEO37

Casos prácticos: Preparación y comienzo del desarrollo de las prácticas de laboratorio propuestas semanalmente por el profesor.

CG08, CG11, CEO36,
CEO37

Project: Development of the final course project that will be worked on incrementally throughout the course. Partial deliveries will be made to the teacher.

CG03, CG08, CE25,
CEO37

SUMMARY STUDENT WORKING HOURS

CLASSROOM HOURS

Lecciones de Carácter
expositivo

Ejercicios y resolución de casos y de problemas



| | | | |
|---|--|---|---|
| 40.00 | 20.00 | | |
| NON-PRESENTIAL HOURS | | | |
| Exposición pública de temas o trabajos | Estudios individual y/o en grupo, y lectura organizada | Ejercicios y resolución de casos y de problemas | Trabajos monográficos y de investigación, individuales o colectivos |
| 10.00 | 30.00 | 40.00 | 40.00 |
| ECTS CREDITS: 6,0 (180,00 hours) | | | |

EVALUATION AND CRITERIA

| Evaluation activities | Evaluation criteria | Weight |
|--|---|--------|
| <p>Objectives:</p> <ul style="list-style-type: none"> Understanding concepts through open tests, focused on practical cases Application of acquired concepts in the form of software development for web and social data collection and analysis Written presentation and communication of the proposed problems. | <p>Breakdown:</p> <ul style="list-style-type: none"> Intersemester Exam (15%) Final Exam (35%) | 50 % |
| <p>Objectives:</p> <ul style="list-style-type: none"> Understanding concepts Applying concepts to the development of cases (small thematic programs) based on the new concepts presented Applying all the knowledge acquired during the course | <p>Cases to be developed in the practical class. It also includes the final practice.</p> | 10 % |
| <p>Objectives:</p> <ul style="list-style-type: none"> Understanding concepts Applying concepts to the development of cases (small thematic programs) on the new concepts proposed Performing tests in the laboratory to verify skills in a real development environment. | <p>Group work</p> | 20 % |
| <p>Objectives:</p> <ul style="list-style-type: none"> Understanding of concepts Application of concepts to the development of cases (small thematic programs) on the new concepts presented Performing tests in the laboratory to verify proficiency in a real development | <p>It implies the completion of tasks and a participative attitude in the classroom, submission within deadlines and with accuracy...</p> | 20 % |



environment

Ratings

The grade in the regular call for the subject will be obtained as follows:

- 15% for the grade of the mid-semester exam.
- 35% for the grade of the final exam.
- 50% will be the grade for the final project.

The grade in the extraordinary call:

- The final project will account for 50%.
- Final exam: 50%.

To pass the subject, students must have at least 5 points out of 10 in the final exam, in the final practice of the subject, and in each of the intermediate practices required, both in the regular and extraordinary call.

BIBLIOGRAPHY AND RESOURCES

Basic Bibliography

- Siddhartha Chatterjee, Michal Krystyanczuk, *Python Social Media Analytics*, Packt Publishing 2017
- Krishna Raj P.M., Ankith Mohan, Srinivasa K.G., *Practical Social Network Analysis with Python*, Springer 2018
- Matthew A. Russell, Mikhail Klassen, *Mining the Social Web, 3rd Edition*, O'Reilly 2019

Complementary Bibliography

- Ryan Mitchell, *Web Scraping with Python, 2nd Edition*, O'Reilly 2018
- Gabor Szabo, Gungor Polatkan, P. Oscar Boykin, Antonios Chalkiopoulos, *Social Media Data Mining and Analytics*, Wiley 2018

In compliance with current regulations on the **protection of personal data**, we would like to inform you that you may consult the aspects related to privacy and data [that you have accepted on your registration form](#) by entering this website and clicking on "download"

<https://servicios.upcomillas.es/sedelectronica/inicio.aspx?csv=02E4557CAA66F4A81663AD10CED66792>