

**COURSE DATA****Data Subject**

<b>Code</b>	44833
<b>Name</b>	Analysis of Web and Social Data
<b>Cycle</b>	Master's degree
<b>ECTS Credits</b>	4.0
<b>Academic year</b>	2017 - 2018

**Study (s)**

<b>Degree</b>	<b>Center</b>	<b>Acad. year</b>	<b>Period</b>
2234 - M.D. in Web Technology, Cloud Computing and Mobile Apps	School of Engineering	1	Second term

**Subject-matter**

<b>Degree</b>	<b>Subject-matter</b>	<b>Character</b>
2234 - M.D. in Web Technology, Cloud Computing and Mobile Apps	4 - Information and Content Management and Processing	Obligatory

**Coordination**

<b>Name</b>	<b>Department</b>
LOZANO IBAÑEZ, MIGUEL	240 - Computer Science

**SUMMARY**

In this course are presented the computational models aimed at the extraction of knowledge, potentially useful and previously unknown, from the World Wide Web. Within this field, known as Web Mining, the student is expected to know the computational techniques used for web data analysis, organized mainly in four groups: structure, content, use and social data/relations analysis. Each of these groups will be shown, as well as the characteristic problems of each one of them. The goal is to provide a broad view of the data analysis techniques that are applied to the web environment so that the student is able to handle them and solve analysis problems in this context.



## PREVIOUS KNOWLEDGE

### Relationship to other subjects of the same degree

There are no specified enrollment restrictions with other subjects of the curriculum.

### Other requirements

Basic statistics, Basic programming

## OUTCOMES

### 2234 - M.D. in Web Technology, Cloud Computing and Mobile Apps

- Students should apply acquired knowledge to solve problems in unfamiliar contexts within their field of study, including multidisciplinary scenarios.
- Students should be able to integrate knowledge and address the complexity of making informed judgments based on incomplete or limited information, including reflections on the social and ethical responsibilities associated with the application of their knowledge and judgments.
- Students should communicate conclusions and underlying knowledge clearly and unambiguously to both specialized and non-specialized audiences.
- Students should demonstrate self-directed learning skills for continued academic growth.
- Students should possess and understand foundational knowledge that enables original thinking and research in the field.
- Ability to apply acquired knowledge and solve problems in new or little-known environments within broader and multidisciplinary contexts, being able to integrate this knowledge.
- To foster, in academic and professional contexts, technological, social or cultural advancement within a society based on In knowledge and respect for: a) fundamental rights and equal opportunities between men and women; b) principles of equal opportunities and universal accessibility of persons with disabilities; and, c) the values of a culture of peace and democratic values.
- Ability to model, design, define the architecture, implement, manage, operate, and maintain applications, systems, services, networks and content in the field of Web technologies, cloud computing and mobile applications.
- Ability to understand and apply the operation and organization of component models, intermediary software and services.
- Ability to design and evaluate servers, applications and systems based on distributed computing.



- Ability to collect data and apply machine learning methods to model, design and develop applications and services.

## LEARNING OUTCOMES

- Specify and complete computer tasks that are complex, incompletely defined or unfamiliar
- Describe and explain techniques and methods applicable to their particular area of study and identify their limitations
- Organize your own work independently, demonstrating initiative and exercising personal responsibility
- Perform bibliographic searches and reviews using databases and other information sources.
- Learning and improving personal performance as the basis for lifelong learning and professional development
- Communicate effectively both verbally and through other media to a variety of audiences and preferably in a second language
- Obtain and analyze research data and use appropriate analytical tools to address unfamiliar problems, such as those with uncertain or incomplete data or specifications, through innovation, use or adaptation of analytical methods.
- Apply the main techniques used to predict user behavior when interacting with the Web
- Manage the link graph representing the Web, as well as the main structural models of the Web
- Solve problems arising from the extraction of knowledge from the Web and social data.
- Discover and identify relationships and influences between different datasets.
- Use technologies and tools for the analysis of social data.
- Know applications of social data analysis.
- Know the structure and operation of a search engine, how the searches are performed and the documents are classified (crawler, indexer).
- Apply the main linguistic and machine learning techniques used in the mining of the content (text) of the Web.
- Implement Web services for applications based on text mining.

## DESCRIPTION OF CONTENTS



### 1. Web structure analysis

### 2. Web use analysis

### 3. Social data analysis

### 4. Web content mining: automatic learning as a service

## WORKLOAD

ACTIVITY	Hours	% To be attended
Theoretical and practical classes	40,00	100
Development of group work	6,00	0
Study and independent work	35,00	0
Preparation of practical classes and problem	16,00	0
Resolution of online questionnaires	3,00	0
<b>TOTAL</b>	<b>100,00</b>	

## TEACHING METHODOLOGY

- Lectures
- Problem resolution
- Project-oriented learning

## EVALUATION

The assesment modalities used in this subject are:

SE1: Online assessment and/or degree of participation

SE2: Assessment of problems, works, reports and/or memories



SE4: Exam or face-to-face assessment

SE6: Assessment of laboratory

- First call:

In the first call the note will be obtained as follows:

$$SE1*0.1+SE2*0.3+SE6*0.3+SE4*0.3$$

SE6 is not recoverable for the second call.

- Second call:

$$SE6*0.3+SE4*0.7$$

The grading system is specified at the following link:

<http://www.uv.es/uvweb/universidad/es/estudios-postgrado/informacion-administrativa-postgrado/permanencia-calificaciones/calificaciones-1285897761928.html>

The applicable regulations can be found at the following link:

<http://www.uv.es/uvweb/universidad/es/estudios-grado/informacion-academica-administrativa/normativas/normativas-universidad-valencia-1285850677111.html>

## REFERENCES



### Basic

- Natural Language Processing with Python. Steven Bird, Ewan Klein, and Edward Loper .  
<http://www.nltk.org/book/>
- Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Ipython. Wes McKinney, W (2012). Ed. O'Reilly Media
- Building an Intelligent Web: Theory and Practice . Pawan Lingras , Rajendra Akerkar. Ed. Jonnes & Bartlett Learning
- Social big data mining. Ishikawa, H. CRC Press.
- Semantic mining of social networks. Synthesis Lectures on the Semantic Web: Theory and Technology, 5(2), 1-205. Tang, J., & Li, J.
- Networks, crowds, and markets: Reasoning about a highly connected world. Easley, D., & Kleinberg, J. Cambridge University Press.

### Additional

- <http://scikit-learn.org/stable/>