

**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>	2019 - 2020

**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. Additionally, the subject addresses the field of Web Analytics so that the students can develop measuring plans and analyse data in websites, mobile apps, social networks, advertising campaigns, newsletters and other platforms. Finally, students are expected to achieve an advanced command of Google Analytics, which will allow them to get the Google Analytics Individual Qualification (GAIQ).



## 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.
- Knowing the basis of web analytics to define different strategies
- Develop and implementing a plan of online mediation
- Configuring Google Analytics and writing analysis reports of multichannel databases.
- Suggesting strategies for optimising development in websites and mobile apps, taking into consideration the previously conducted analysis reports of databases



## DESCRIPTION OF CONTENTS

### 1. Web mining

- Representation
- Clustering documents
- Classification of documents

### 2. Analysis of social data

- Analysis of tweets
- Analysis of feelings

### 3. Web Analytics

- Measurement methods, business objectives, macro and micro conversions
- Dimensions, metrics and KPI
- Defining a measurement plan

### 4. Google Analytics

- Configuring GA: accounts, property, visits, filters, objectives, conversion funnel, sections and labels
- Monitoring e-trade, events and multi-domains. Classifying channels, customizable dimensions and calculated metric.
- Audience and acquisition reports
- Behaviour and conversion reports
- Analysis of campaigns and analysis of data: advertising campaigns, organic traffic, social campaigns and e-mail campaigns.
- Google Analytics Individual Qualification (GAIQ)

**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.



- Analítica Web 2.0. Avinash KaushiK. Ed. Gestión 2000
- Introduction to Google Analytics: A Guide for Absolute Beginners. Todd Kelsey. Berkeley, CA : Apress : Imprint: Apress, 2017
- Google Analytics integrations. Daniel Waisberg. Indianapolis, Indiana: Wiley, 2015
- Advanced Web metrics with Google Analytics. Brian Clifton. Indianapolis, IN : Wiley, c2010

#### **Additional**

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

### **ADDENDUM COVID-19**

**This addendum will only be activated if the health situation requires so and with the prior agreement of the Governing Council**

#### Contents

The contents of the subject will not be modified and will be taught in its entirety.

#### Volume of work and teaching scheduling

Since the students can work autonomously and collaboratively through the tools provided, the dedication in hours by ECTS is not modified.

#### Teaching methodology

In this course, the following tools have been chosen for synchronous communication with students:

- Microsoft Teams
- Blackboard Collaborate
- Virtual Classroom Chat

As asynchronous tools, and all of them integrated into the Virtual Classroom community of the subject, the following will be used:

- Teaching / student forum, in order to visualize and consolidate for the whole group the specific actions for each session and the doubts raised by the students and the teachers' responses
- Glossary with more than 100 key topics of the subject, including detailed textual and graphic descriptions, along with additional references.
- Increase in the online bibliographic material specified in the teaching guide.
- Explanatory videos of each topic made with Blackboard Collaborate, in which all the content of the topic is fully explained.
- Videos explaining the practical exercises in which all the practical exercises are fully explained.
- Test exercises for each topic to implement a continuous evaluation system.
- Exercises of practical cases solved to model the realization of those proposed.
- Bank of 500 questions type preparatory test of the final test of 70 questions



The UV has a Virtual Classroom (Moodle) used by all the teachers of the master and in which all the learning of the subject is managed: the resources and the activities. Virtual Classroom operates with absolute normality during the second semester. Highlight the integrated multimedia server in which teachers can upload explanatory videos of the material they deem appropriate and make it available to the students of the subject.

The subject remains active online at the times set in the Teaching Organization Plan. During these hours, teachers are connected by videoconference and available to any student in the group to:

- Solve doubts
- Make indications and guidelines about the theoretical material housed in the Virtual Classroom and in Microsoft Teams
- Make indications and guidelines on practical exercises and explain practical cases resolved.

The objective is that in the hours of class of the subject almost the same tasks are carried out that would be carried out in the equivalent classroom.

The tutorials will be carried out by email and through the Teams tool that offers the possibility of using different types of chats: a chat for the subject to deal with general topics, chats in pairs to answer questions related to the practices and individual chats.

#### Assessment

The SE4 test of the original guide (face-to-face assessment) has been eliminated so that the assessment of the subject becomes focused on the activities of continuous assessment. The activities and their weight in the evaluation are indicated below:

- Continuous assessment questionnaires in Virtual Classroom (topic exercises): 20%
- Practical work delivered in Virtual Classroom: 60%
- Final theoretical-practical questionnaire type test and available online from virtual classroom: 20%

This criterion will be valid for both calls.

#### Bibliography

The recommended bibliography is available online so it is not modified.