Course guides

270224 - TAED1 - Advanced Topics in Data Engineering 1

Unit in charge: Barcelona School of Informatics
Teaching unit: 715 - EIO - Department of Statistics and Operations Research.
Degree: BACHELOR’S DEGREE IN DATA SCIENCE AND ENGINEERING (Syllabus 2017). (Compulsory subject).
Academic year: 2021 ECTS Credits: 6.0 Languages: Catalan, Spanish, English

LECTURER

Coordinating lecturer: PAU FONSECA CASAS - EVA MARIA VIDAL LOPEZ
Others: Segon quadrimestre:
PAU FONSECA CASAS - 11
JOAN GARCIA SUBIRANA - 11

PRIOR SKILLS

those obtained in the previous subjects

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Generical:

CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.

Transversal:

CT2. Sustainability and Social Commitment. To know and understand the complexity of economic and social phenomena typical of the welfare society; Be able to relate well-being to globalization and sustainability; Achieve skills to use in a balanced and compatible way the technique, the technology, the economy and the sustainability.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.

Basic:

CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.
CB5. That the students have developed those learning skills necessary to undertake later studies with a high degree of autonomy

TEACHING METHODOLOGY

Planned methodology for the part of data ethics:
Each week has an associated topic and an assigned group of students. The weekly topic is presented in the first hour to the whole class by an expert. The students, organized in small groups, are in charge of researching the topic in more depth and presenting the results to their classmates through activities that everyone can do to fully understand the ethic problems that the topic represents.
LEARNING OBJECTIVES OF THE SUBJECT

1. Know the basic concepts of modelling and simulation.
2. Recognize and understand the social and environmental impact of data science and engineering, and the ethical issues involved in their applications.

STUDY LOAD

<table>
<thead>
<tr>
<th>Type</th>
<th>Hours</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hours large group</td>
<td>30,0</td>
<td>20.00</td>
</tr>
<tr>
<td>Hours small group</td>
<td>30,0</td>
<td>20.00</td>
</tr>
<tr>
<td>Self study</td>
<td>90,0</td>
<td>60.00</td>
</tr>
</tbody>
</table>

Total learning time: 150 h

CONTENTS

**Modeling and simulation.**

Description:
System concept, model. In this point we can explain in more or less detail aspects related to how we interpret the reality and the phenomenon of experimentation and validation (positivism, empiricism, etc.). Stages in the process of Validation, Verification and Accreditation of a model: explaining the different techniques that we can apply, the hypotheses and their possible taxonomy. This brings us to the need to define conceptual models in order to properly develop the validation process. Main formal languages for defining conceptual models (SDL, DEVS, Petri Nets, SysML, BPMN, etc.). System Dynamics: Introducing Forrester’s Causal and Forrester diagrams and models related to social and environmental simulation, we present Jay Forrester’s World 03 model. Coding a discrete simulation model using a generic tool.

**Data ethics**

Description:
Development of topics related to science and data engineering based on values. Eg: welfare, privacy, freedom, equity, sustainability.
Case studies in different fields of application. Eg: advertising, teaching, work, justice, social networks
ACTIVITATS

**Aprenentatge sobre Modelat i Simulació**

**Specific objectives:**

1

**Related competencies:**

CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CB5. That the students have developed those learning skills necessary to undertake later studies with a high degree of autonomy
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.

**Full-or-part-time:** 72h
Theory classes: 21h
Laboratory classes: 8h
Self study: 43h

**Data ethics learning**

**Specific objectives:**

2

**Related competencies:**

CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CT2. Sustainability and Social Commitment. To know and understand the complexity of economic and social phenomena typical of the welfare society; Be able to relate well-being to globalization and sustainability; Achieve skills to use in a balanced and compatible way the technique, the technology, the economy and the sustainability.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.

**Full-or-part-time:** 71h
Theory classes: 14h
Laboratory classes: 14h
Self study: 43h
GRADING SYSTEM

The evaluation of the course will be carried out taking into account the blocks of knowledge of the syllabus: the B1 block of Modeling and simulation and the B2 block of ethics. The final grade will be calculated by a weighted average of the marks of each block:

Final grade = 0.5 B1 + 0.5 B2.

Grade B1 for the Modeling and Simulation part will be evaluated based on a practical work (written in English).

Grade B2 corresponding to the part of Data Ethics, will be evaluated taking into account:
Class participation - 25%
Development of a topic and presentation - 75%

The reassessment mark will be calculated as 0.5 B1 + 0.5 B2, where B1 and B2 will be the grades of the new submissions. Only students who have submitted and failed the regular assessment may be re-evaluated.

BIBLIOGRAPHY

Basic: