Universidad Zaragoza

Academic Year/course: 2021/22

60949 - Management of Large-Scale Data

Syllabus Information

Academic Year: 2021/22
Subject: 60949 - Manipulación y análisis de grandes volúmenes de datos
Faculty / School: 110 - Escuela de Ingeniería y Arquitectura
Degree: 533 - Master's Degree in Telecommunications Engineering
623 - Master's Degree in Telecommunications Engineering
ECTS: 6.0
Year: 2
Semester: Second semester
Subject Type: Optional
Module:

1. General information

2. Learning goals

3. Assessment (1st and 2nd call)

4. Methodology, learning tasks, syllabus and resources

4.1. Methodological overview

The methodology followed in this course is oriented towards achievement of the learning objectives. A wide range of teaching and learning tasks are implemented, such as

Activities with direct support from the teachers

- Theory sessions. Teacher's presentation of the course contents, including illustrative examples.
- Talks by experts. When possible, external experts to the university will explain some contents.
- Seminars. Sessions based on oral or written contributions by the students.
- Problem-based learning. Educative approach oriented towards a teaching and learning methodology where students tackle real problems in small groups under the supervision of a teacher.
- Practice sessions. Any practical or collaborative activity done with direct support from the teachers.
- Laboratory sessions. Activities developed with computers.
- Tutorials. Students can review or discuss with the teacher the materials and topics presented in the theory sessions.
- Assessment. A set of written/oral tests, lab assignments, projects, other assignments, etc.

Autonomous work

- Theoretical assignments. Preparing seminars, readings, research, assignments or write reports, etc., to be submitted or presented to the teacher in theory sessions.
- Practical assignments. Preparing activities to be submitted or presented to the teacher in practice sessions.
- Theoretical study. Study of contents related to the theory sessions: it includes any study activity not considered previously (study for exams, work in the library, complementary readings, solve problems and exercises, etc.).
- Practical study. Related to the practice sessions.

- Complementary activities. Formative activities related to the course, but not to the preparation of exams or included within the assessment activities: readings, seminars, videos, etc.

4.2. Learning tasks

The course (6 ECTS: 150 hours) includes the following learning tasks:

- Activities under direct supervision of the teachers of the course (60 hours). Theory sessions, expert talks, seminars, problem-solving and cases, and lab assignments.
- Practice and research assignments (60 hours).
- Tutorials (5 hours).
- Autonomous work and study (20 hours).
- Assessment (5 hours). Exam and defense of the course project.

4.3. Syllabus

The course will address the following topics:

1. Introduction and motivation to the problem of large volumes of data (Big Data).
2. Storage of large amounts of data
    1. Data warehouses. Star schema design.
    2. NoSQL databases.
3. Management of large amounts of data
    1. Data distribution.
    2. Information integration considering heterogeneous data sources.
    3. Use of knowledge representation techniques (ontologies) to represent data sources and their access and integration.
    4. Parallel processing techniques: MapReduce (Hadoop).
    5. Data Stream Management Systems.
    6. Other techniques: mobile agents.
4. Interaction with large amounts of data
    1. Visualization techniques.
    2. Design of appropriate user interfaces.
    3. Usability.
5. Analysis of large amounts of data
    1. Data mining.
    2. Text mining, sentiment analysis.
6. Use cases and applications
    1. Data provided by sensors.
    2. Unstructured data on the Web.
    3. Recommendation Systems.
    4. Analysis of blogs and social networks.
    5. Smart cities.
    6. Intelligent Transportation Systems.

Some of the topics will be addressed through autonomous work by the students and through the development of practical assignments, considering (to the extent possible) the preferences of the students.

4.4. Course planning and calendar

Further information concerning the timetable, office hours, assessment dates and other details regarding this course, will be provided on the first day of class and/or the faculty's website.

4.5. Bibliography and recommended resources

http://psfunizar10.unizar.es/br13/egAsignaturas.php?codigo=60949