ADVANCED DATA BASE DEVELOPMENT & MANAGEMENT

BACHELOR IN MANAGEMENT INFORMATION SYSTEMS
Professor: CELIA GUTIERREZ COSIO
E-mail: cgutierrez@faculty.ie.edu

Academic year: 17-18
Degree course: SECOND
Semester: 2º
Category: COMPULSORY
Number of credits: 6.0
Language: English

PREREQUISITES
The student should have knowledgements of the following areas (which correspond to subjects of the BIS program):
Databases and Data Modelling
Information Systems Analysis and Design
Operating Systems
Programming, Data Structures and Algorithms

SUBJECT DESCRIPTION
The purpose of this course is to provide the student with knowledgements in unstructured data, how to model this information in NoSQL databases, focusing in HBASE database.
HBASE is the NoSQL database provided by Cloudera Hadoop
The course also gives a full description of Cloudera Hadoop as it is one of the most popular Big Data frameworks.
Big Data concepts, relevant examples, and tools are also referred to in this course, with the aim of bringing a full perspective of unstructured data and NoSQL databases.
Last but not least, some analytic operations are also taught in order to extract knowledge from the information stored in NoSQL. This issue is important for company managers, so that they can think on new strategies with this valuable information.

OBJECTIVES AND SKILLS
The aim of this course is to provide the student with new skills: personal and technical skills
Personal skills are:
Active attitude
Entrepreneurship
Team work
Autonomous work
Oral skills
Presentation skills
Research attitude
Hands on skills
Reflexive thinking
Study habits

Technical skills:
Knowledgements of Big Data information and Data Science process.
Knowledgements of Big Data tools and frameworks with special emphasis on Cloudera Hadoop and how to install them.
Choose among types and examples of NoSQL databases.
Model unstructured data in different types of NoSQL databases.
Model unstructured data in HBase; manipulate data in HBase and HIV.
Build a complete NoSQL project

NoSQL Project manager skills: the students should be able to choose the right tools to use from the universe of Big Data and NoSQL

<table>
<thead>
<tr>
<th>Teaching methodology</th>
<th>Weighting</th>
<th>Estimated time a student should dedicate to prepare for and participate in</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lectures</td>
<td>33.33 %</td>
<td>50 hours</td>
</tr>
<tr>
<td>Discussions</td>
<td>6.67 %</td>
<td>10 hours</td>
</tr>
<tr>
<td>Exercises</td>
<td>20.0 %</td>
<td>30 hours</td>
</tr>
<tr>
<td>Group work</td>
<td>20.0 %</td>
<td>30 hours</td>
</tr>
<tr>
<td>Other individual studying</td>
<td>20.0 %</td>
<td>30 hours</td>
</tr>
<tr>
<td>TOTAL</td>
<td>100.0 %</td>
<td>150 hours</td>
</tr>
</tbody>
</table>
PROGRAM

SESSIONS 1 - 2
Presentation of the course
Introduction to Big Data

SESSIONS 3 - 4
Data Science Methodology
Big Data Applications

SESSIONS 5 - 6
Big Data Applications
Platforms and tools for Big Data

SESSIONS 7 - 8
Cloudera Hadoop
In this double session, operating systems concepts and Linux commands and shell will also be taught.

SESSIONS 9 - 10
Cloudera Hadoop

SESSIONS 11 - 12
NoSQL

SESSIONS 13 - 14
NoSQL

SESSIONS 15 - 16
NoSQL
Visit of Savanamed startup, which deals with unstructured data in the Health domain

SESSIONS 17 - 18
HBase

SESSIONS 19 - 20
HBase
SESSIONS 21 - 22
Hive

SESSIONS 23 - 24
Hive
Further considerations on Big Data, NoSQL and Cloudera Hadoop

SESSIONS 25 - 26
Presentation of the final project by the professor.
Initiation of the project by the groups with the professor’s assistance.

SESSIONS 27 - 28
Continuation of the final project by the groups with the professor’s assistance.

SESSIONS 29 - 30
Presentation of the final project by the groups
BIBLIOGRAPHY

Buy your books here

Basic bibliography:
Cloudera components https://www.cloudera.com
Retrieved 1 August 2017

Complementary bibliography:
Big Data course. Big Data University. On line course: https://courses.cognitiveclass.ai/courses/course-v1:BigDataUniversity+BD0101EN+2016_T2/courseware/c6143d9ff5764057a91e53fa8a3a6dff/884a7a2954a84eda9d6f3bca49bf02a/ Retrieved 30 November 2017

EVALUATION CRITERIA

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Percentage</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate Tests</td>
<td>10 %</td>
<td></td>
</tr>
<tr>
<td>Class Participation</td>
<td>30 %</td>
<td></td>
</tr>
<tr>
<td>Individual Work with Presentation</td>
<td>30 %</td>
<td></td>
</tr>
<tr>
<td>Final Project in group with presentation</td>
<td>30 %</td>
<td></td>
</tr>
</tbody>
</table>

With the exception of the intermediate test, which is an individual quiz that will take place without previous announcement, the other assignments are scheduled.
Students must be present the date of that assignment, otherwise, the score of that assignment will be 0.
The final project will be made in groups and presented by all the members of the group in the date scheduled.

PROFESSOR BIO
Professor: **CElia Gutiérrez Cosio**

E-mail: cgutierrez@faculty.ie.edu

**Celia Gutiérrez Cosio**

Member of one of the leading research groups in agent technologies in Spain. Celia Gutiérrez’s career highlights for the variety, quantity and quality of accumulated results in Computer Science and Software Engineering, since she started the PhD thesis in 1995 in Genetic Algorithms. For example, she has experience in companies like Indra and Espelsa (Eurofighter project). She has also worked for public and private universities, where she has taught a wide range of subjects in all levels, supervised master thesis, and published articles in JCR indexed journals, books, and prestigious international conferences. Her most recent experience is related to research evaluation and management, for the Spanish Agency of Innovation Certification and the European Commission. In the last case, she has not only evaluated Horizon2020 proposals, but also developed revision tasks.

Web page: [http://grasia.fdi.ucm.es/celia](http://grasia.fdi.ucm.es/celia)

LinkedIn: [https://www.linkedin.com/in/celia-gutierrez-0889715b/?ppe=1](https://www.linkedin.com/in/celia-gutierrez-0889715b/?ppe=1)

Research track: [http://dblp.uni-trier.de/pers/hd/g/Gutierrez:Celia](http://dblp.uni-trier.de/pers/hd/g/Gutierrez:Celia)

**OTHER INFORMATION**

The best way to arrange a meeting with me, is by sending me an email to this address: cgutierrez@faculty.ie.edu

**CODE OF CONDUCT IN CLASS**

1. **Be on time:** Students arriving more than 5 minutes late will be marked as “Absent”.

   Only students that notify in advance in writing that they will be late for a specific session may be granted an exception (at the discretion of the professor).

2. **If applicable, bring your name card and strictly follow the seating chart.** It helps faculty members and fellow students learn your names.

3. **Do not leave the room during the lecture:** Students are not allowed to leave the room during lectures. If a student leaves the room during lectures, he/she will not be allowed to re-enter and, therefore, will be marked as “Absent”.

   Only students that notify that they have a special reason to leave the session early will be granted an exception (at the discretion of the professor).

4. **Do not engage in side conversation.** As a sign of respect toward the person presenting the lecture (the teacher as well as fellow students), side conversations are not allowed. If you have a question, raise your hand and ask it. If you do not want to ask it during the lecture, feel free to approach your teacher after class.

   If a student is disrupting the flow of the lecture, he/she will be asked to leave the classroom and, consequently, will be marked as “Absent”.

5. **Use your laptop for course-related purposes only.** The use of laptops during lectures must be authorized by the professor. The use of Social Media or accessing any type of content not related to the lecture is penalized. The student will be asked to leave the room and, consequently, will be marked as “Absent”.

6. **No cellular phones:** IE University implements a “Phone-free Classroom” policy and, therefore, the use of phones, tablets, etc. is forbidden inside the classroom. Failing to abide by this rule entails expulsion from the room and will be counted as one absence.
7. Escalation policy: 1/3/5. Items 4, 5, and 6 above entail expulsion from the classroom and the consequent marking of the student as “Absent.” IE University implements an “escalation policy”: The first time a student is asked to leave the room for disciplinary reasons (as per items 4, 5, and 6 above), the student will incur one absence, the second time it will count as three absences, and from the third time onward, any expulsion from the classroom due to disciplinary issues will entail 5 absences.