

Course guide

205231 - IBD - Introduction to Big Data

Last modified: 02/04/2024

Unit in charge: Terrassa School of Industrial, Aerospace and Audiovisual Engineering
Teaching unit: 709 - DEE - Department of Electrical Engineering.

Degree: BACHELOR'S DEGREE IN AUDIOVISUAL SYSTEMS ENGINEERING (Syllabus 2009). (Optional subject).
BACHELOR'S DEGREE IN CHEMICAL ENGINEERING (Syllabus 2009). (Optional subject).
BACHELOR'S DEGREE IN ELECTRICAL ENGINEERING (Syllabus 2009). (Optional subject).
BACHELOR'S DEGREE IN INDUSTRIAL ELECTRONICS AND AUTOMATIC CONTROL ENGINEERING (Syllabus 2009). (Optional subject).
BACHELOR'S DEGREE IN MECHANICAL ENGINEERING (Syllabus 2009). (Optional subject).
BACHELOR'S DEGREE IN TEXTILE TECHNOLOGY AND DESIGN ENGINEERING (Syllabus 2009). (Optional subject).
BACHELOR'S DEGREE IN AEROSPACE TECHNOLOGY ENGINEERING (Syllabus 2010). (Optional subject).
BACHELOR'S DEGREE IN AEROSPACE VEHICLE ENGINEERING (Syllabus 2010). (Optional subject).
BACHELOR'S DEGREE IN INDUSTRIAL DESIGN AND PRODUCT DEVELOPMENT ENGINEERING (Syllabus 2010). (Optional subject).
BACHELOR'S DEGREE IN INDUSTRIAL TECHNOLOGY ENGINEERING (Syllabus 2010). (Optional subject).

Academic year: 2024 **ECTS Credits:** 3.0 **Languages:** English

LECTURER

Coordinating lecturer: Coordinador: Alvaro Luna Alloza

Others: Altres: Stoyan Danov

TEACHING METHODOLOGY

La metodologia docent es divideix en tres parts:

☐ Sessions presencials d'exposició - participació dels continguts i realització d'exercicis.

☐ Sessions presencials de treball de laboratori.

☐ Treball autònom d'estudi i realització d'exercicis i activitats.

Aquestes sessions presencials es poden realitzar per videoconferència, utilitzant plataformes como Google Meet o per streaming en Youtube, en cas de que les circumstàncies ho requereixin.

En les sessions d'exposició -participació dels continguts, el professorat introduirà les bases teòriques de la matèria, conceptes, mètodes i resultats il·lustrant-los amb exemples convenients i sol·licitant, si escau, la realització d'exercicis per facilitar-ne la seva comprensió.

En les sessions de treball de laboratori, el professorat guiarà l'estudiantat en l'aplicació dels conceptes teòrics per a la resolució de muntatges experimentals, fonamentant en tot moment el raonament crític. Es proposaran activitats que l'estudiantat resolgui a l'aula i fora de l'aula, per tal d'afavorir el contacte i utilització de les eines bàsiques necessàries per a la realització d'un sistema d'instrumentació.

L'estudiantat, de forma autònoma, ha de treballar el material proporcionat pel professorat i el resultat de les sessions de treball-problemes per tal d'assimilar i fixar els conceptes. El professorat proporcionarà un pla d'estudi i de seguiment d'activitats (ATENEA).

En quant a la evaluació es realitzarà en base a treballs, i les presentacions dels treballs es faràn en forma de video, on els i les estudiant hauran de fer una defensa curta però efectiva.

LEARNING OBJECTIVES OF THE SUBJECT

En acabar l'assignatura l'estudiant o estudianta haurà vist el potencial de les aplicacions Big Data en sistemes d'energia i més concretament a les xarxes intel·ligents. Convertir-se en un expert/experta en Big Data no s'aconsegueix cursant una assignatura, però aquesta en concret té l'objectiu d'aportar uns fonaments de coneixement sòlid per a que els/les estudiants puguin introduir-se a un sector clarament a l'alça. Amb aquest objectiu en ment l'assignatura presentarà les eines principals amb les que es treballa en aplicacions big data, així com les tècniques d'anàlisi i de disseny que s'implementen. Tanmateix, els i les estudiants accediran a un coneixement valuós sobre la dimensió, les plataformes, sistemes de programació, els temps de càlcul i el ventall d'aplicacions que ofereix aquesta tecnologia.

En l'àmbit de l'aplicació es centrarà l'atenció en els sistemes dissenyats per a xarxes elèctriques, els quals permeten millorar l'eficiència del sistema energètic i poder assolir un major nivell de control. Malgrat que l'aplicació tingui una vessant elèctrica els coneixements adquirits són prou transversals, essent anàlegs en molts casos al necessaris en anàlisis de mobilitat, data mining o sistemes econòmics.

Aquesta assignatura compta amb un altre de tres crèdits com a continuació natural; 205232 - Big Data Tools and Applications. La combinació de les dues permet consolidar les competències.

Aquest assignatura es desenvolupa plenament en anglès, desde les classes, els materials i les entregues per part dels i de les estudiants.

STUDY LOAD

Type	Hours	Percentage
Hours large group	15,0	20.00
Hours medium group	15,0	20.00
Self study	45,0	60.00

Total learning time: 75 h

CONTENTS

Mòdul 1: Introducció al Big Data

Description:

- 1.1.- Qué es Big Data?
- 1.2.- Introducció de conceptes i terminologia en Big Data
- 1.3.- Arquitectures informàtiques necessàries
- 1.4.- Estructures i tipus de processat de dades

Full-or-part-time: 16h

Theory classes: 4h

Laboratory classes: 2h

Self study : 10h

Mòdul 2: Eines de desenvolupament en Big Data

Description:

- 2.1.- Introducció a Apache Hadoop i el seu ecosistema de software
- 2.2.- Introducció a Hadoop Distributed File System
- 2.3.- Introducció a Apache Hive
- 2.4.- Introducció a Apache HBase

Full-or-part-time: 22h

Theory classes: 8h

Laboratory classes: 4h

Self study : 10h

Mòdul 3: Introducció al MapReduce

Description:

- 3.1.- Introducció al concepte MapReduce
- 3.2.- Introducció a les aplicacions de MapReduce per al processament de grans volums de dades en paral·lel

Full-or-part-time: 11h

Theory classes: 4h

Laboratory classes: 2h

Self study : 5h

GRADING SYSTEM

- 1er Treball: 25 %
- 2on Treball: 25 %
- Exposicions Orals: 25 %
- Treball de Laboratori: 25 %

BIBLIOGRAPHY

Basic:

- Stimmel, Carol L. Big data analytics strategies for the smart grid. Boca Raton, FL: CRC Press, Taylor & Francis Group, 2015. ISBN 9781482218282.
- Grolemond, Garrett; Grolemond, Hadley. Hands-on programming with R. Sebastopol: O'Reilly Media, 2014. ISBN 9781449359010.

Complementary:

- Arghandeh, Reza; Zhou, Yuxun. Big data application in power systems. Amsterdam, Netherlands,: Elsevier, 2018. ISBN 9780128119686.
- Grolemond, Garrett; Wickham, Hadley. R for data science : import, tidy, transform, visualize and model data. O'Reilly, 2016. ISBN 9781491910399.
- White, Tom. Hadoop : the definitive guide. 4a ed. Sebastopol, Calif: O'Reilly, 2015. ISBN 9781491901632.
- Walkowiak, Simon. Big data analytics with R. Birmingham: Packt Publishing, 2016. ISBN 9781786466457.
- James, Gareth [et al.]. An introduction to statistical learning. New York: Springer, 2013. ISBN 9781461471370.

RESOURCES

Hyperlink:

- Miner, Donald; Shook, Adam. MapReduce design patterns, O'Reilly, 2012, ISBN 9781449327170. Resource. https://discovery.upc.edu/discovery/fulldisplay?docid=alma991005093379206711&context=L&vid=34CSUC_UPC:VU1&lang=ca

Other resources:

Donald Miner and Adam Shook, MapReduce Design Patterns, O'Reilly, 2012, ISBN: 978-1-449-32717-0