Course guide

270959 - ASM - Advanced Statistical Modelling

Unit in charge: Barcelona School of Informatics
Teaching unit: 715 - EIO - Department of Statistics and Operations Research.
Degree: MASTER'S DEGREE IN DATA SCIENCE (Syllabus 2021). (Optional subject).
Academic year: 2022  ECTS Credits: 6.0  Languages: English

LECTURER

Coordinating lecturer: XAVIER PUIG ORIOL - JOSE ANTONIO SÀNCHEZ ESPIGARES
Others: Primer quadrimestre:
        XAVIER PUIG ORIOL - 10
        JOSE ANTONIO SÀNCHEZ ESPIGARES - 10

PRIOR SKILLS

Not specified

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
CE10. Identify machine learning and statistical modeling methods to use and apply them rigorously in order to solve a specific data science problem
CE12. Apply data science in multidisciplinary projects to solve problems in new or poorly explored domains from a data science perspective that are economically viable, socially acceptable, and in accordance with current legislation
CE5. Model, design, and implement complex data systems, including data visualization
CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
CE9. Apply appropriate methods for the analysis of non-traditional data formats, such as processes and graphs, within the scope of data science

Generical:
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats

Transversal:
CT4. INFORMATION LITERACY: Capacity for managing the acquisition, the structuring, analysis and visualization of data and information in the field of specialisation, and for critically assessing the results of this management.
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

Basic:
CB10. Possess and understand knowledge that provides a basis or opportunity to be original in the development and/or application of ideas, often in a research context.
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
CB7. Ability to integrate knowledges and handle the complexity of making judgments based on information which, being incomplete or limited, includes considerations on social and ethical responsibilities linked to the application of their knowledge and judgments.
TEACHING METHODOLOGY

There is a weekly 3 hours session. The first two hours are devoted to the exposition of the theoretical subjects by the teacher. The last hour is dedicated to implement these contents: Each student has his laptop in class and he or she performs the tasks proposed by the teacher. Each session ends with an assignment to students who must be delivered the following session.

LEARNING OBJECTIVES OF THE SUBJECT

1. Time Series
2. Bayesian Statistics

STUDY LOAD

<table>
<thead>
<tr>
<th>Type</th>
<th>Hours</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self study</td>
<td>96,0</td>
<td>64.00</td>
</tr>
<tr>
<td>Hours large group</td>
<td>54,0</td>
<td>36.00</td>
</tr>
</tbody>
</table>

Total learning time: 150 h

CONTENTS

**Time Series**

**Description:**
1. Box-Jenkins methodology (ARIMA models) for prediction
2. Extensions: outliers treatment, calendar effects and intervention analysis
3. Kalman State Space and Filter Models. Applications

**Bayesian Data Analysis**

**Description:**
1. Bayesian Model. The statistical model. The Likelihood function. The Bayesian model
2. Bayesian Inference. Point and Interval estimation. Hypothesis Test
4. Hierarchical Models
5. Checking and defining the model
ACTIVITIES

Presentation of Theme 1 (Time Series) in class

Description:
Presentation of Theme 1 (Time Series) in class

Specific objectives:
1

Related competencies:
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
CE10. Identify machine learning and statistical modeling methods to use and apply them rigorously in order to solve a specific
data science problem
CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make
decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
CE12. Apply data science in multidisciplinary projects to solve problems in new or poorly explored domains from a data science
perspective that are economically viable, socially acceptable, and in accordance with current legislation
CE5. Model, design, and implement complex data systems, including data visualization
CE9. Apply appropriate methods for the analysis of non-traditional data formats, such as processes and graphs, within the scope
of data science
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that
meets the needs of the profession and the labour market.

CT4. INFORMATION LITERACY: Capacity for managing the acquisition, the structuring, analysis and visualization of data and
information in the field of specialisation, and for critically assessing the results of this management.
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader
(or multidisciplinary) contexts related to their area of study.
CB10. Possess and understand knowledge that provides a basis or opportunity to be original in the development and/or
application of ideas, often in a research context.
CB7. Ability to integrate knowledges and handle the complexity of making judgments based on information which, being
incomplete or limited, includes considerations on social and ethical responsibilities linked to the application of their knowledge and
judgments.

Full-or-part-time: 75h
Theory classes: 22h 30m
Self study: 52h 30m
Presentation of theme 2 (Bayesian models) in class

Description:
Presentation of theme 2 (Bayesian Models) in class

Specific objectives:
3

Related competencies :
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
CE10. Identify machine learning and statistical modeling methods to use and apply them rigorously in order to solve a specific data science problem
CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
CE12. Apply data science in multidisciplinary projects to solve problems in new or poorly explored domains from a data science perspective that are economically viable, socially acceptable, and in accordance with current legislation
CE5. Model, design, and implement complex data systems, including data visualization
CE9. Apply appropriate methods for the analysis of non-traditional data formats, such as processes and graphs, within the scope of data science
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CT4. INFORMATION LITERACY: Capacity for managing the acquisition, the structuring, analysis and visualization of data and information in the field of specialisation, and for critically assessing the results of this management.
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
CB10. Possess and understand knowledge that provides a basis or opportunity to be original in the development and/or application of ideas, often in a research context.
CB7. Ability to integrate knowledges and handle the complexity of making judgments based on information which, being incomplete or limited, includes considerations on social and ethical responsibilities linked to the application of their knowledge and judgments.

Full-or-part-time: 75h
Theory classes: 22h 30m
Self study: 52h 30m

GRADING SYSTEM
Homeworks will be assigned during the course. Homework grades will be worth 50% of your course grade.

There will be an exam for the first part of the course (first theme), during the partial exams week, and another one for the second part (second theme), each one with a weight of 25%.

Course Grade = 0.5 * Hwk Grade + 0.25 * 1st part Exam Grade + 0.25 * 2nd part Exam Grade
BIBLIOGRAPHY

Basic: