

ANX-PR/CL/001-01

LEARNING GUIDE

SUBJECT

103000839 - Complex Data In Health

DEGREE PROGRAMME

10AX - Master Universitario Innovación Digital Ciencia de Datos Itinerario Health

ACADEMIC YEAR & SEMESTER

2020/21 - Semester 1

Index

Learning guide

1. Description.....	1
2. Faculty.....	1
3. Prior knowledge recommended to take the subject.....	2
4. Skills and learning outcomes	3
5. Brief description of the subject and syllabus.....	5
6. Schedule.....	7
7. Activities and assessment criteria.....	9
8. Teaching resources.....	11

1. Description

1.1. Subject details

Name of the subject	103000839 - Complex Data In Health
No of credits	4.5 ECTS
Type	Optional
Academic year of the programme	Second year
Semester of tuition	Semester 3
Tuition period	September-January
Tuition languages	English
Degree programme	10AX - Master Universitario Innovación Digital Ciencia de Datos Itinerario Health
Centre	10 - Escuela Tecnica Superior de Ingenieros Informaticos
Academic year	2020-21

2. Faculty

2.1. Faculty members with subject teaching role

Name and surname	Office/Room	Email	Tutoring hours *
Antonio Jesus Diaz Honrubia	4302	antoniojesus.diaz@upm.es	Sin horario. contact the professor
Ernestina Menasalvas Ruiz	4303	ernestina.menasalvas@upm.es	Sin horario. contact the professor

Alejandro Rodriguez Gonzalez (Subject coordinator)	4302	alejandro.rg@upm.es	Sin horario. contact the professor
--	------	---------------------	--

* The tutoring schedule is indicative and subject to possible changes. Please check tutoring times with the faculty member in charge.

2.3. External faculty

Name and surname	Email	Institution
Lucia Prieto Santamaría	lucia.prieto.santamaria@upm.es	Centro de Tecnología Biomédica
Massimiliano Zanin	massimiliano.zanin@ctb.upm.es	Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB)
Javier Rodriguez Vidal	javier.rodriguez.vidal@upm.es	Centro de Tecnología Biomédica

3. Prior knowledge recommended to take the subject

3.1. Recommended (passed) subjects

The subject - recommended (passed), are not defined.

3.2. Other recommended learning outcomes

- relational data model
- Databases
- SQL
- programming skills
- graphs

4. Skills and learning outcomes *

4.1. Skills to be learned

CB06 - Poseer y comprender conocimientos que aporten una base u oportunidad de ser originales en el desarrollo y/o aplicación de ideas, a menudo en un contexto de investigación

CB07 - Que los estudiantes sepan aplicar los conocimientos adquiridos y su capacidad de resolución de problemas en entornos nuevos o poco conocidos dentro de contextos más amplios (o multidisciplinares) relacionados con su área de estudio

CB08 - Que los estudiantes sean capaces de integrar conocimientos y enfrentarse a la complejidad de formular juicios a partir de una información que, siendo incompleta o limitada, incluya reflexiones sobre las responsabilidades sociales y éticas vinculadas a la aplicación de sus conocimientos y juicios

CE-HMDA01 - Capacidad para seleccionar las soluciones de almacenamiento para datos estructurados y no estructurados adecuadas en función del problema a resolver

CE-HMDA02 - Capacidad para aplicar técnicas para la generación de visualizaciones adecuadas para el análisis y la exploración de datos en un contexto médico, y para la correcta comunicación de los resultados del análisis

CE-HMDA03 - Capacidad para seleccionar las técnicas y herramientas para visualización de grandes cantidades de datos más adecuadas para resolver un determinado problema en el campo de la salud

CE-HMDA05 - Capacidad para usar herramientas de procesamiento de big data tanto en online como en modo batch

CE-HMDA07 - Capacidad para diseñar y gestionar proyectos de salud y datos médicos

CE-HMDA09 - Capacidad para solventar problemas reales en el área clínica seleccionando la mejor de las técnicas posibles

CG01 - Que los estudiantes sean capaces de predecir y controlar la evolución de situaciones complejas mediante el desarrollo de nuevas e innovadoras metodologías de trabajo adaptadas al ámbito científico/investigador, tecnológico o profesional concreto, en general multidisciplinar, en el que se desarrolle su actividad.

CG02 - Que los estudiantes desarrollen la autonomía suficiente para participar en proyectos de investigación y colaboraciones científicas o tecnológicas dentro su ámbito temático explorando y generando nuevas ideas sistemáticamente, en contextos interdisciplinares y, en su caso, con una alta componente de transferencia del

conocimiento.

CG03 - La capacidad de usar la lengua inglesa de manera competente, es decir, con capacitación para tareas complejas de trabajo y estudio.

CG07 - Capacidad de trabajar y comunicarse también en contextos internacionales.

4.2. Learning outcomes

RA17 - Apply methods for knowledge acquisition to create knowledge bases using other sources of information

RA18 - Use computer languages or software tools for knowledge representation and reasoning for building intelligent systems

RA2 - Conocer técnicas de visualización y procesos de análisis de datos, y de programación, diseño y depuración de algoritmos, para computación de altas prestaciones

RA20 - Manage bibliographic sources in the domain, including manuals, online documentation and scientific papers

RA8 - To be able to deal with unstructured sources as device raw data.

RA3 - Conocer cómo se aplican las técnicas de computación científica en algún campo específico de ciencia o ingeniería

RA5 - Develop projects of data science

RA4 - Ser capaz de procesar datos masivos

RA29 - Deal with unstructured health data

RA11 - To know and apply the main techniques to explore, describe and analyse multivariate data.

RA30 - Extract and manipulate data from public sources

RA31 - Extract and manipulate data from social media

* The Learning Guides should reflect the Skills and Learning Outcomes in the same way as indicated in the Degree Verification Memory. For this reason, they have not been translated into English and appear in Spanish.

5. Brief description of the subject and syllabus

5.1. Brief description of the subject

The aim of the subject is to identify various types of data that can co-exist in the context of health data and how to approach its manipulation, processing, and analysis.

More specifically, the subject will try to provide the students with a set of skills and competencies that will allow them to work in multidisciplinary environments, mainly related to the use of health information.

The subject will be taught by providing a set of specific classes providing theoretical background about the subject, but with special emphasis in the work that needs to be done by the student itself.

In this context, most of the time to be spent in this subject must be done by the students working by themselves and applying the theoretical work.

5.2. Syllabus

1. Introduction

1.1. Types of biomedical information

1.2. Sources of biomedical information

2. Extraction of information: APIs, frameworks and tools

2.1. Crawlers

2.2. Natural Language Processing

3. Complex networks

3.1. Basics of complex networks. Cytoscape for plotting

3.2. Physical networks. For instance, connectome, calculate basic topological features, and network randomisation

3.3. Functional networks. Reconstructing brain networks with correlation. Difference with causality

3.4. Data mining and networks. Optimisation of networks. Using networks in classification tasks

3.5. Other topics. MST. Link filtering and prediction. Multi-layer and time-evolving

4. Disease networks

- 4.1. Human disease networks
- 4.2. Human symptom disease networks
- 4.3. Disease understanding
- 4.4. Approaches and utilities of disease networks
- 5. Health data in social media
- 6. Textual health data

6. Schedule

6.1. Subject schedule*

Week	Face-to-face classroom activities	Face-to-face laboratory activities	Distant / On-line	Assessment activities
1	Subject presentation Duration: 03:00		Subject presentation Duration: 03:00	
2	Introduction Duration: 03:00		Introduction Duration: 03:00	
3	Extraction of information: APIs, frameworks and tools Duration: 03:00		Extraction of information: APIs, frameworks and tools Duration: 03:00	
4	Complex networks Duration: 03:00		Complex networks Duration: 03:00	
5	Complex networks Duration: 03:00		Complex networks Duration: 03:00	
6	Complex networks Duration: 03:00		Complex networks Duration: 03:00	
7	Disease networks Duration: 03:00		Disease networks Duration: 03:00	
8	Disease networks (Individual practical work) Duration: 03:00		Disease networks (Individual practical work) Duration: 03:00	
9	Disease networks (Individual practical work) Duration: 03:00		Disease networks (Individual practical work) Duration: 03:00	
10	Health data in social media Duration: 03:00		Health data in social media Duration: 03:00	
11	Health data in social media (Individual practical work) Duration: 03:00		Health data in social media (Individual practical work) Duration: 03:00	
12	Textual health data Duration: 03:00		Textual health data Duration: 03:00	

13	Textual health data (Individual practical work) Duration: 03:00		Textual health data (Individual practical work) Duration: 03:00	
14	Textual health data (Individual practical work) Duration: 03:00		Textual health data (Individual practical work) Duration: 03:00	
15	Assignment work Duration: 03:00		Assignment work Duration: 03:00	
16	Assignment presentations Duration: 03:00		Assignment presentations Duration: 03:00	Assignment delivery and presentation Continuous assessment Presential Duration: 01:00 Final exam Final examination Presential Duration: 01:00
17				

Depending on the programme study plan, total values will be calculated according to the ECTS credit unit as 26/27 hours of student face-to-face contact and independent study time.

* The schedule is based on an a priori planning of the subject; it might be modified during the academic year, especially considering the COVID19 evolution.

7. Activities and assessment criteria

7.1. Assessment activities

7.1.1. Continuous assessment

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
16	Assignment delivery and presentation		Face-to-face	01:00	100%	5 / 10	CE-HMDA01 CE-HMDA02 CE-HMDA03 CE-HMDA05 CE-HMDA07 CE-HMDA09 CB06 CB07 CB08 CG01 CG02 CG03 CG07

7.1.2. Final examination

Week	Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
16	Final exam		Face-to-face	01:00	100%	5 / 10	CE-HMDA01 CE-HMDA02 CE-HMDA03 CE-HMDA05 CE-HMDA07 CE-HMDA09 CB06 CB07 CB08 CG01 CG02 CG03 CG07

7.1.3. Referred (re-sit) examination

Description	Modality	Type	Duration	Weight	Minimum grade	Evaluated skills
Final exam		Face-to-face	01:00	100%	5 / 10	CE-HMDA01 CE-HMDA02 CE-HMDA03 CE-HMDA05 CE-HMDA07 CE-HMDA09 CB06 CB07 CB08 CG01 CG02 CG03 CG07

7.2. Assessment criteria

Continuous evaluation: The course will be evaluated by:

- 3 practical assignments (includes presentation)

Attendance is mandatory.

Practical assignments will be done in groups or individually, depending on the number of students enrolled, among those enrolled in the course at the beginning of the academic year (nature and the number of components will be established at the beginning of the course, depending on the number of students enrolled). In order to pass the course in the fall semester, the requirement is to obtain a minimum of 50 points out of 100 in the added evaluation.

The final score will be calculated as follows:

- 30% for each practical assignment;
- 10% based on attendance and the professors' opinion.

Final exam evaluation: Those students whose extraordinary circumstances cannot perform the continuous evaluation, and having done the final exam evaluation written petition during the first 15 days of the course, will perform the final exam evaluation without having the opportunity to do the continuous evaluation;

In these premises, the final exam evaluation will consist of an exam as stated by the head of studies. A minimum of 5 points over 10 will be required to pass this exam.

Measures against copies and fraud Rights and duties of college students are gathered on the status of the Universidad Politécnica de Madrid (BOCM de 15 de noviembre de 2010) and in the statutes of the college student (RD 1791/2010 de 30 de diciembre). Article 124 a) of EUPM fixes the duty of the student... "to follow with responsibility and taking advantage of the learning process, knowledge acquisition correspondent to its condition of college student"... and the article 13 of the statutes of the college student in its point d) also specifies as duty of the college student "abstain from the use or cooperation in fraudulent procedures in the evaluation assessments, in the assignments developed or in the official documents of the university". In the case that in the development of the evaluation assessments it is appreciated a breach in the duties as college student, the subject coordinator may communicate the headmaster as established in the article 74 (n) of EUPM to have the competences to "propose the initiation of a disciplinary procedure to any College member, by its own initiative or as instance from the "Comisión de Gobierno"" to the Rector, pursuant to the statutes and rules of application.

8. Teaching resources

8.1. Teaching resources for the subject

Name	Type	Notes
IBM Watson - How it works	Web resource	http://www.youtube.com/watch?v=_Xcmh1LQB9I
Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Introduction to Data Mining, Pearson Addison Wesley (May, 2005).	Bibliography	

Ian Witten, Eibe Frank, Mark Hall, Data Mining: Practical Machine Learning Tools and Techniques, 3rd Edition, Morgan Kaufmann, ISBN 978-0-12-374856-0, 2011.	Bibliography	
Moodle	Web resource	Moodle platform (https://moodle.upm.es/)
Disease networks and their contribution to disease understanding: A review of their evolution, techniques and data sources.	Bibliography	https://www.ncbi.nlm.nih.gov/pubmed/31077818
The human disease network	Bibliography	https://www.pnas.org/content/104/21/8685
Human symptoms?disease network	Bibliography	https://www.nature.com/articles/ncomms5212
Cytoscape	Web resource	https://cytoscape.org/
Information Retrieval: A Health and Biomedical Perspective	Bibliography	Hersh WR, Information Retrieval: A Health and Biomedical Perspective, Third Edition, 2009
Introduction to information retrieva	Bibliography	Introduction to information retrieval / Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze
information retrieval in practice	Bibliography	Croft, W. Bruce, Search engines : information retrieval in practice / W. Bruce Croft, Donald Metzler, Trevor Strohman