



UNIVERSIDAD
POLITÉCNICA
DE MADRID

PROCESO DE
COORDINACIÓN DE LAS
ENSEÑANZAS PR/CL/001



E.T.S. de Ingenieros
Industriales

ANX-PR/CL/001-01

GUÍA DE APRENDIZAJE

ASIGNATURA

53001521 - Tools For Big Data Analytics

PLAN DE ESTUDIOS

05BD - Master Universitario En Ingenieria De La Organizacion

CURSO ACADÉMICO Y SEMESTRE

2025/26 - Primer semestre

Índice

Guía de Aprendizaje

1. Datos descriptivos.....	1
2. Profesorado.....	1
3. Conocimientos previos recomendados.....	2
4. Competencias y resultados de aprendizaje.....	2
5. Descripción de la asignatura y temario.....	4
6. Cronograma.....	7
7. Actividades y criterios de evaluación.....	9
8. Recursos didácticos.....	12
9. Otra información.....	13

1. Datos descriptivos

1.1. Datos de la asignatura

Nombre de la asignatura	53001521 - Tools For Big Data Analytics
No de créditos	3 ECTS
Carácter	Optativa
Curso	Segundo curso
Semestre	Tercer semestre
Período de impartición	Septiembre-Enero
Idioma de impartición	Inglés/Castellano
Titulación	05BD - Master Universitario en Ingeniería de la Organización
Centro responsable de la titulación	05 - E.T.S. De Ingenieros Industriales
Curso académico	2025-26

2. Profesorado

2.1. Profesorado implicado en la docencia

Nombre	Despacho	Correo electrónico	Horario de tutorías *
Joaquin Bienvenido Ordieres Mere (Coordinador/a)	UD Proyectos	j.ordieres@upm.es	M - 11:00 - 13:00 X - 11:00 - 14:00 V - 08:30 - 09:30

* Las horas de tutoría son orientativas y pueden sufrir modificaciones. Se deberá confirmar los horarios de tutorías con el profesorado.

3. Conocimientos previos recomendados

3.1. Asignaturas previas que se recomienda haber cursado

- Análisis De Datos Para La Gestión
- Métodos Cuantitativos Avanzados Para La Gestión

3.2. Otros conocimientos previos recomendados para cursar la asignatura

- Programming Python
- Knowledge about Operating Systems such as Unix or Linux at user level

4. Competencias y resultados de aprendizaje

4.1. Competencias

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

CB10 - Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo

CE01 - Conocer y aplicar técnicas y herramientas para el manejo y análisis de grandes masas de datos

CG02 - Analizar situaciones estructuradas y poco estructuradas de empresas y otras organizaciones, estableciendo diagnósticos apropiados, en particular, de carácter estratégico

CG03 - Concebir soluciones para afrontar problemas previamente diagnosticados, y evaluarlas desde diferentes criterios correspondientes a los distintos actores concernidos

CT01 - Aplica. Habilidad para aplicar conocimientos científicos, matemáticos y tecnológicos en sistemas relacionados con la práctica de la ingeniería

CT02 - Experimenta. Habilidad para diseñar y realizar experimentos así como analizar e interpretar datos

CT03 - Diseña. Habilidad para diseñar un sistema, componente o proceso que alcance los requisitos deseados teniendo en cuenta restricciones realistas tales como las económicas, medioambientales, sociales, políticas, éticas, de salud y seguridad, de fabricación y de sostenibilidad

CT05 - Resuelve. Habilidad para identificar, formular y resolver problemas de ingeniería

CT08 - Entiende los impactos. Educación amplia necesaria para entender el impacto de las soluciones ingenieriles en un contexto social global

CT09 - Se actualiza. Reconocimiento de la necesidad y la habilidad para comprometerse al aprendizaje continuo

CT10 - Conoce. Conocimiento de los temas contemporáneos

CT11 - Usa herramientas. Habilidad para usar las técnicas, destrezas y herramientas ingenieriles modernas necesarias para la práctica de la ingeniería

CT12 - Es bilingüe. Capacidad de trabajar en un entorno bilingüe (inglés/español)

4.2. Resultados del aprendizaje

RA71 - RA1 - Elegir y aplicar técnicas de análisis exploratorios de grandes bases de datos

RA12 - Comprender y evaluar los efectos de la incorporación de las Tecnologías de la Información y las Comunicaciones (TIC) en todos los niveles de la organización

RA13 - Identificar y analizar el impacto que las tecnologías digitales, de la comunicación y los datos y de la inteligencia artificial tienen en la aparición de nuevos modelos de negocio disruptivos

RA74 - RA47 - Enumerar, seleccionar y aplicar herramientas y metodologías para la explotación de datos mediante modelos

RA46 - Distinguir el significado y alcance de las soluciones BigData y los casos de uso más importantes de aplicación de las mismas

5. Descripción de la asignatura y temario

5.1. Descripción de la asignatura

This course is taught in the third semester of the master in Organizational Engineering, in the fall season. As a general idea it aims to,

- Understand the real meaning of Big Data problems and solutions, through different use cases, to compare with classical IT solutions regularly found
- Be aware and able to apply ETL solutions for those environments.
- Be aware of the CI/CD contexts and microservice oriented platforms.
- Be able to create models and visualize the forecasts connecting them with decision making processes

The course is mainly oriented by case enrollment.

In previous editions this course took advantage of RosettaHub middleware to interface with AWS, making it possible to have a context for distributing computing. However, the UPM authorities did not continue the agreement with RosettaHub, therefore this solution is no longer available.

Therefore, students are kindly requested to install anaconda on their own computers and practical knowledge regarding Hadoop/Spark will be provided by the instructor but no hardware will be granted for students to play with.

Application cases will be gathered from the real life, including datasets and other referential characteristics. The motivation for the analysis to be carry out will emphasize their contribution to the SDGs, highlighting how these technologies can help organizations to do better regarding such aspects.

Deeply, the SDGs to be worked with, are:

SDG8: Good Jobs and Economic Growth. Because of students will realize how value creation can be linked to the successful contribution to solve the raised cases due to their effects

SDG9: Industry, Innovation and Infrastructure. Most of the examples will be linked to the industry examples, as cases allow to observe effects from the shopfloor to the organizational level as well.

SDG3: Good Health and Well being. Because of workers are key aspects in the value stream of the organizations and cases will care about sensing them.

5.2. Temario de la asignatura

1. INTRODUCTION AND BUSINESS PERSPECTIVE

- 1.1. Why do we need big data solutions? Who are using such approaches?
- 1.2. What does IoT and IIoT mean ? What is the foreseen impact in regular business?
- 1.3. How to store data in BD? How does it work? A Technology perspective

2. BIG-DATA ARCHITECTURE

- 2.1. Clusters
- 2.2. Data bases in BigData

3. BIGDATA ANALYTICS

- 3.1. Doing operations with data living in BD environments
- 3.2. Spark solution and other related stuff around

4. COMMERTIAL vs HOME MADE SOLUTIONS

- 4.1. The big players
- 4.2. OpenSource solutions

5. DEEP LEARNING

- 5.1. What does DL stand for?
- 5.2. Is it the same that Machine Learning?
- 5.3. Creating ML models in python
- 5.4. Creating DL models in python with Keras and TF

6. PRACTICAL ASSIGNMENT

- 6.1. CRISP/DM Framework

6.2. Dealing with the data and the business case

6.3. Growing models

6.4. Analysis of results

6.5. Value proposal

6. Cronograma

6.1. Cronograma de la asignatura *

Sem	Actividad tipo 1	Actividad tipo 2	Tele-enseñanza	Actividades de evaluación
1	Introduction to the course. Lecture 1 Duración: 02:00 LM: Actividad del tipo Lección Magistral			
2	Lecture 1 Duración: 02:00 LM: Actividad del tipo Lección Magistral			Gamification activity to assess the knowledge gathered. ET: Técnica del tipo Prueba Telemática Evaluación Progresiva Presencial Duración: 00:05
3	Architecture(s) analysis and discussion Duración: 01:00 LM: Actividad del tipo Lección Magistral	Case selection to be developed during the course. Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio		Gamification activity to assess the knowledge gathered. ET: Técnica del tipo Prueba Telemática Evaluación Progresiva Presencial Duración: 00:05
4	Big Analytics analysis and discussion Duración: 01:00 LM: Actividad del tipo Lección Magistral	Case development. Practical activity following the CRISP/DM methodology. Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio		Gamification activity to assess the knowledge gathered. ET: Técnica del tipo Prueba Telemática Evaluación Progresiva Presencial Duración: 00:05
5	Implementation strategies: Isolated, Distributed, CD/CI Duración: 01:00 LM: Actividad del tipo Lección Magistral	Case development. Practical activity following the CRISP/DM methodology. Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio		Gamification activity to assess the knowledge gathered. ET: Técnica del tipo Prueba Telemática Evaluación Progresiva Presencial Duración: 00:05
6	Refinement of the Group/Individual assignment Duración: 01:00 AC: Actividad del tipo Acciones Cooperativas	Case development. Practical activity following the CRISP/DM methodology. Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio		Gamification activity to assess the knowledge gathered. ET: Técnica del tipo Prueba Telemática Evaluación Progresiva Presencial Duración: 00:05
7		Case development. Practical activity following the CRISP/DM methodology. Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio	Effective development of the individual/group assignments Duración: 01:00 OT: Otras actividades formativas / Evaluación	Gamification activity to assess the knowledge gathered. ET: Técnica del tipo Prueba Telemática Evaluación Progresiva Presencial Duración: 00:05
8		Case development. Practical activity following the CRISP/DM methodology. Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio	Effective development of the individual/group assignments Duración: 01:00 OT: Otras actividades formativas / Evaluación	Gamification activity to assess the knowledge gathered. ET: Técnica del tipo Prueba Telemática Evaluación Progresiva Presencial Duración: 00:05

9		<p>Case development. Practical activity following the CRISP/DM methodology. Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio</p>	<p>Effective development of the individual/group assignments Duración: 01:00 OT: Otras actividades formativas / Evaluación</p>	
10		<p>Case development. Practical activity following the CRISP/DM methodology. Duración: 01:00 PL: Actividad del tipo Prácticas de Laboratorio</p>	<p>Effective development of the individual/group assignments Duración: 01:00 OT: Otras actividades formativas / Evaluación</p>	
11		<p>Case development. Practical activity following the CRISP/DM methodology. Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio</p>		
12		<p>Case development. Practical activity following the CRISP/DM methodology. Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio</p>		
13		<p>Case development. Practical activity following the CRISP/DM methodology. Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio</p>		
14		<p>Case development. Practical activity following the CRISP/DM methodology. Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio</p>		<p>Individual delivery of the practical assignment. TI: Técnica del tipo Trabajo Individual Evaluación Global No presencial Duración: 00:30</p>
15				<p>Presentation of the achievements PG: Técnica del tipo Presentación en Grupo Evaluación Progresiva Presencial Duración: 02:00</p> <p>Delivery of the video presenting the strengths and weaknesses of the work carried out. PI: Técnica del tipo Presentación Individual Evaluación Global No presencial Duración: 00:30</p>
16				
17				

Para el cálculo de los valores totales, se estima que por cada crédito ECTS el alumno dedicará dependiendo del plan de estudios, entre 26 y 27 horas de trabajo presencial y no presencial.

7. Actividades y criterios de evaluación

7.1. Actividades de evaluación de la asignatura

7.1.1. Evaluación (progresiva)

Sem.	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
2	Gamification activity to assess the knowledge gathered.	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	3%	3 / 10	
3	Gamification activity to assess the knowledge gathered.	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	4%	3 / 10	
4	Gamification activity to assess the knowledge gathered.	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	3%	3 / 10	
5	Gamification activity to assess the knowledge gathered.	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	4%	3 / 10	
6	Gamification activity to assess the knowledge gathered.	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	3%	3 / 10	
7	Gamification activity to assess the knowledge gathered.	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	4%	3 / 10	
8	Gamification activity to assess the knowledge gathered.	ET: Técnica del tipo Prueba Telemática	Presencial	00:05	4%	3 / 10	CB06 CT12 CE01
15	Presentation of the achievements	PG: Técnica del tipo Presentación en Grupo	Presencial	02:00	75%	5 / 10	CB06 CT03 CE01

7.1.2. Prueba evaluación global

Sem	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
14	Individual delivery of the practical assignment.	TI: Técnica del tipo Trabajo Individual	No Presencial	00:30	75%	5 / 10	CB06 CT03 CE01
15	Delivery of the video presenting the strengths and weaknesses of the work carried out.	PI: Técnica del tipo Presentación Individual	No Presencial	00:30	25%	5 / 10	CB06 CT03 CT12

7.1.3. Evaluación convocatoria extraordinaria

No se ha definido la evaluación extraordinaria.

7.2. Criterios de evaluación

This course has a clear practical orientation with just a few seminar to be aware of the existing technologies for massive data storage and data process

Consistently, it will be assessed by following the next rules

It is highly recommended to undertake a group-oriented practical assignment and present it publicly. This recommendation is based on the connection between the assignment and the process of improving the Sustainable Development Goals (SDGs), as well as the recognition that addressing the various commitments requires more effort than what individuals can provide.

The score obtained from this activity will be taken into consideration.

The choice of the group assignment topic can be suggested by either the instructor or the students. However, if students propose a topic, it is essential to obtain the instructor's approval.

Promoting diversity in topics and ensuring data availability, whether from public sources or provided by stakeholders, will be encouraged.

In certain situations, the instructor may share a scraper tool if it is deemed appropriate.

The assessment will always encompass a 360° perspective, and the development of the assignment will follow the

Agile methodology as much as possible.

Individual Assessment will be in place for special cases, although it is not recommended.

Scoring Rules

During the first week of the course, students have the option to choose between two main approaches:

1. Continuous Assessment:

- This method involves short and gamified individual tests that assess knowledge based on the course topics, as well as practical questions. In this approach, there are a total of eight short gamified tests, and the score for this component (25% of the final score) will be based on the best six test results.
- The score for the group assignment will be determined when it is submitted. Each team member will evaluate and contribute to the scoring of other group members. The remaining part of the final score (75%) will be determined by the practical assessment, where the opinions of other team members will impact 20% of the score.

2. Alternatively, for those who opt for the Final Test (**Traditional style**), the final score is determined by the following components:

- Score from the individual test, with a minimum passing threshold of 4 points: This component carries a weight of 25% towards the final score.
- Individual Practical Assignment: This assignment contributes 50% towards the final score.
- Self-explanatory video showcasing a critical assessment of the findings: This component holds a weight of 25% towards the final score.

If students are unable to successfully complete the Continuous Assessment method, they are kindly requested to participate in the Final test, but only if the individual test can resolve the issue.

If the failure is due to practical teamwork, it cannot be rectified since there are specific competences that must be demonstrated through group work (CG02, CG03, CT01, CT03, etc.).

However, students who fail to pass the course on their first attempt can always reorganize their efforts for the next examination period. They have the option to enhance their practical work or switch to the Individual mode (Traditional Style). They must clearly communicate their decision by sending an email to the course advisor no later than two weeks after the start of the next semester.

Please note that partial components of the score are not carried over beyond the specific academic year.

8. Recursos didácticos

8.1. Recursos didácticos de la asignatura

Nombre	Tipo	Observaciones
Data-Intensive Text Processing with MapReduce by Jimmy Lin and Chris Dyer Morgan & Claypool Publishers, 2010.	Bibliografía	
Hadoop Real World Solutions Cookbook by Jonathan R. Owens, Brian Femiano, and Jon Lentz Publication Date: February 7, 2013 ISBN-10: 1849519129 ISBN-13: 978-1849519120	Bibliografía	
BlueMix	Equipamiento	Plataforma de uso de IBM
Amazon	Equipamiento	Plataforma de desarrollo
GCP	Equipamiento	Plataforma de desarrollo
UPM Virtual Desktops	Recursos web	The UPM provides with virtual desktops on remote infrastructure, ready to be used by students from their laptops. https://escritorio.upm.es/uds/page/services

python	Recursos web	
keras	Recursos web	
jupyter	Recursos web	

9. Otra información

9.1. Otra información sobre la asignatura

As the course is handled in a practical way, the usage of computing resources is enforced. Therefore, it is expected the students bring their own computers.

The course is related to the following SDGs:

- **SDG8: Good Jobs and Economic Growth.** Because of students will realize how value creation can be linked to the successful contribution to solve the raised cases due to their effects
- **SDG9: Industry, Innovation and Infrastructure.** Most of the examples will be linked to the industry examples, as cases allow to observe effects from the shopfloor to the organizational level as well.
- **SDG3: Good Health and Well being.** Because of workers are key aspects in the value stream of the organizations and cases will care about sensing them.

This course plays as an EELISA initiative, therefore, its practical application looks to have specific impact out of the academic environment.