



UNIVERSIDAD
POLITÉCNICA
DE MADRID

PROCESO DE
COORDINACIÓN DE LAS
ENSEÑANZAS PR/CL/001



E.T.S. de Ingenieros
Informáticos

ANX-PR/CL/001-01

GUÍA DE APRENDIZAJE

ASIGNATURA

103000866 - Massively Parallel Machine Learning

PLAN DE ESTUDIOS

10AZ - Master Universitario En Innovación Digital

CURSO ACADÉMICO Y SEMESTRE

2025/26 - Primer semestre

Índice

Guía de Aprendizaje

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

1. Datos descriptivos

1.1. Datos de la asignatura

Nombre de la asignatura	103000866 - Massively Parallel Machine Learning
No de créditos	4.5 ECTS
Carácter	Optativa
Curso	Segundo curso
Semestre	Tercer semestre
Período de impartición	Septiembre-Enero
Idioma de impartición	Inglés/Castellano
Titulación	10AZ - Master Universitario en Innovación Digital
Centro responsable de la titulación	10 - E.T.S. De Ingenieros Informáticos
Curso académico	2025-26

2. Profesorado

2.1. Profesorado implicado en la docencia

Nombre	Despacho	Correo electrónico	Horario de tutorías *
Bonifacio Alberto Mozo Velasco (Coordinador/a)	4307	a.mozo@upm.es	X - 16:00 - 17:30 J - 16:00 - 17:30

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

3. Competencias y resultados de aprendizaje

3.1. 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

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.

3.2. Resultados del aprendizaje

RA47 - Design massively parallel versions of supervised machine learning algorithms

RA44 - Understand the basic mechanisms for designing parallel applications in Big Data regimes

RA46 - Apply parallelization strategies to machine learning algorithms

4. Descripción de la asignatura y temario

4.1. Descripción de la asignatura

This course introduces to the parallelization of Machine Learning algorithms (supervised and unsupervised) on Big Data distributed computing platforms.

In particular, MapReduce and RDD paradigms are explained jointly with the corresponding distributed platforms supporting them (Hadoop and Apache Spark).

Using as examples several well-known machine learning algorithms, the student will learn how to parallelize machine learning algorithms using the Apache Spark API.

4.2. Temario de la asignatura

1. Introduction to Big Data
2. Machine Learning (supervised and unsupervised techniques)
3. MapReduce and Hadoop
4. RDDs and Spark
5. Parallelization of machine learning algorithms

5. Cronograma

5.1. Cronograma de la asignatura *

Sem	Actividad tipo 1	Actividad tipo 2	Tele-enseñanza	Actividades de evaluación
1	Lecture on Unit 1 Duración: 02:00 LM: Actividad del tipo Lección Magistral			
2	Lecture on Unit 2 Duración: 02:00 LM: Actividad del tipo Lección Magistral			
3	Practical classes Duración: 03:00 PR: Actividad del tipo Clase de Problemas			
4	Practical classes Duración: 03:00 PR: Actividad del tipo Clase de Problemas			
5	Practical classes Duración: 03:00 PR: Actividad del tipo Clase de Problemas			
6	Lecture on Unit 3 Duración: 02:00 LM: Actividad del tipo Lección Magistral			
7	Lecture on Unit 4 Duración: 02:00 LM: Actividad del tipo Lección Magistral			
8	Lecture on Unit 4 Duración: 02:00 LM: Actividad del tipo Lección Magistral			
9				Individual Exam EX: Técnica del tipo Examen Escrito Evaluación Progresiva y Global Presencial Duración: 02:00
10	Explanation of Project P1 Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio	Project Development Duración: 03:00 AC: Actividad del tipo Acciones Cooperativas		
11	Explanation of Project P1 Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio	Project Development Duración: 03:00 AC: Actividad del tipo Acciones Cooperativas		
12	Explanation of Project P2 Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio	Project Development Duración: 03:00 AC: Actividad del tipo Acciones Cooperativas		

13	Explanation of Project P2 Duración: 02:00 PL: Actividad del tipo Prácticas de Laboratorio	Project Development Duración: 03:00 AC: Actividad del tipo Acciones Cooperativas		
14				Student Project Presentations P1 PG: Técnica del tipo Presentación en Grupo Evaluación Progresiva y Global Presencial Duración: 02:00
15				Student Project Presentations P2 PG: Técnica del tipo Presentación en Grupo Evaluación Progresiva y Global Presencial Duración: 02:00
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.

6. Actividades y criterios de evaluación

6.1. Actividades de evaluación de la asignatura

6.1.1. Evaluación (progresiva)

Sem.	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
9	Individual Exam	EX: Técnica del tipo Examen Escrito	Presencial	02:00	40%	5 / 10	CB06 CB07 CG03 CG07
14	Student Project Presentations P1	PG: Técnica del tipo Presentación en Grupo	Presencial	02:00	30%	5 / 10	CB06 CG03 CG07
15	Student Project Presentations P2	PG: Técnica del tipo Presentación en Grupo	Presencial	02:00	30%	5 / 10	CB06 CG03 CG07

6.1.2. Prueba evaluación global

Sem	Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
9	Individual Exam	EX: Técnica del tipo Examen Escrito	Presencial	02:00	40%	5 / 10	CB06 CB07 CG03 CG07
14	Student Project Presentations P1	PG: Técnica del tipo Presentación en Grupo	Presencial	02:00	30%	5 / 10	CB06 CG03 CG07
15	Student Project Presentations P2	PG: Técnica del tipo Presentación en Grupo	Presencial	02:00	30%	5 / 10	CB06 CG03 CG07

6.1.3. Evaluación convocatoria extraordinaria

Descripción	Modalidad	Tipo	Duración	Peso en la nota	Nota mínima	Competencias evaluadas
The student will take an individual written exam that covers all units of the course	EX: Técnica del tipo Examen Escrito	Presencial	02:00	40%	5 / 10	CB06 CB07 CG03 CG07
The student will hand practical assignments P1 and P2 up to the day of the written exam.	EP: Técnica del tipo Examen de Prácticas	Presencial	00:00	60%	5 / 10	CB06 CG03 CG07

6.2. Criterios de evaluación

Assessment Criteria

Continous assessment

The following partial grades will be obtained (grades from 0 to 10):

N1: Grade for the individual exam.

P1: Grade for the student project P1

P2: Grade for the student project P2

The final grade is obtained as follows: $N = 0.40 \cdot N1 + 0.30 \cdot P1 + 0.30 \cdot P2$

As a general criterion, to pass the course (in the first or second call) it is required:

- $N1 \geq 5$ and $P1$ and $P2 \geq 5$.

Final exam

The final grade is obtained as follows: $N = 0.40 \cdot N1 + 0.30 \cdot P1 + 0.30 \cdot P2$

N is the grade obtained in the final exam

P1: Grade for the student project P1

P2: Grade for the student project P2

As a general criterion, to pass the course in the final exam it is required:

$N1 \geq 5$ and $P1$ and $P2 \geq 5$.

7. Recursos didácticos

7.1. Recursos didácticos de la asignatura

Nombre	Tipo	Observaciones
Resilient Distributed Datasets	Bibliografía	Zaharia, Matei, et al. "Resilient distributed datasets: A fault-tolerant abstraction for in-memory cluster computing." Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation. USENIX Association, 2012.
MapReduce	Bibliografía	Dean, Jeffrey, and Sanjay Ghemawat. "MapReduce: simplified data processing on large clusters." Communications of the ACM 51.1 (2008): 107-113.
UPM Moodle Web Site	Recursos web	The web site UPM-Moodle (http://moodle.upm.es) will be used to provide course material to students. Students will use this web site to submit for evaluation the results of their assignments and individual works.

Machine Learning. A probabilistic approach.	Bibliografía	Murphy, K. P. (2012). Machine learning: a probabilistic perspective. MIT press.
---	--------------	---

8. Otra información

8.1. Otra información sobre la asignatura

Sustainable Development Goals:

This subject is aligned with Goal 4 Quality Education that aims to "Guarantee inclusive, equitable and quality education and promote lifelong learning opportunities for all" that defined within the Sustainable Development Goals (SDGs) established by the United Nations Organization (UN).