



UNIVERSIDAD
POLITÉCNICA
DE MADRID

PROCESO DE
COORDINACIÓN DE LAS
ENSEÑANZAS PR/CL/001



E.T.S. de Ingenieros de
Telecomunicacion

ANX-PR/CL/001-01

GUÍA DE APRENDIZAJE

ASIGNATURA

93000944 - Analítica De Contenidos Multimedia A Gran Escala

PLAN DE ESTUDIOS

09AT - Master Universitario En Teoria De La Señal Y Comunicaciones

CURSO ACADÉMICO Y SEMESTRE

2024/25 - Segundo 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..... | 3 |
| 5. Cronograma..... | 6 |
| 6. Actividades y criterios de evaluación..... | 8 |
| 7. Recursos didácticos..... | 11 |
| 8. Otra información..... | 12 |

1. Datos descriptivos

1.1. Datos de la asignatura

| | |
|--|--|
| Nombre de la asignatura | 93000944 - Analítica de Contenidos Multimedia a Gran Escala |
| No de créditos | 4 ECTS |
| Carácter | Optativa |
| Curso | Primer curso |
| Semestre | Segundo semestre |
| Período de impartición | Febrero-Junio |
| Idioma de impartición | Inglés/Castellano |
| Titulación | 09AT - Master Universitario en Teoría de la Señal y Comunicaciones |
| Centro responsable de la titulación | 09 - Escuela Tecnica Superior De Ingenieros De Telecomunicacion |
| Curso académico | 2024-25 |

2. Profesorado

2.1. Profesorado implicado en la docencia

| Nombre | Despacho | Correo electrónico | Horario de tutorías * |
|---|-----------------|---------------------------|--|
| Alberto Belmonte Hernandez (Coordinador/a) | D-112.2 | alberto.belmonte@upm.es | Sin horario. Appointment arranged by email |
| Federico Alvarez Garcia | D-103 | federico.alvarez@upm.es | Sin horario. Appointment arranged by email |

| | | | |
|--------------------------------|-------|--------------------|--|
| Jose Manuel Menendez Garcia | C-300 | jm.menendez@upm.es | Sin horario. Appointment arranged by email |
|--------------------------------|-------|--------------------|--|

* 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

CB09 - Que los estudiantes sepan comunicar sus conclusiones y los conocimientos y razones últimas que las sustentan a públicos especializados y no especializados de un modo claro y sin ambigüedades

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 - Analizar y aplicar técnicas para el diseño y desarrollo avanzado de equipos y sistemas, basándose en la teoría de la señal y las comunicaciones, en un entorno internacional

CE02 - Evaluar y sintetizar los resultados de un trabajo en equipo en proyectos relacionados con la teoría de la señal y las comunicaciones, en un entorno internacional.

CT05 - Capacidad para gestionar la información, identificando las fuentes necesarias, los principales tipos de documentos técnicos y científicos, de una manera adecuada y eficiente

3.2. Resultados del aprendizaje

RA34 - Capability to develop and evaluate machine-learning techniques and to design big data learning systems

RA43 - Ability to develop basic applications in relevant current use cases in the media industry (media search, content recommendation, etc.)

RA42 - knowledge on Big Data technologies and their application to multimedia content

RA41 - Ability to select and apply adequate machine learning techniques to large-scale multimedia datasets and evaluate their performance

4. Descripción de la asignatura y temario

4.1. Descripción de la asignatura

Current data analysis applications require the management of extremely large collections of heterogeneous multimedia data. The extraction of knowledge from these huge datasets is a difficult problem with a broad scope. The course begins with a comprehensive exploration of data analysis techniques, focusing on Exploratory Data Analysis (EDA). Students will learn about analytic graphics and techniques essential for EDA, including data transformation, dimensionality reduction, clustering, and outlier detection. The curriculum also covers the analysis of algorithms, hyperparameters, and evaluation metrics, emphasizing the importance of Explainable AI (XAI) in understanding and interpreting model decisions. Practical application is reinforced through a lab session where students will apply EDA and XAI techniques to real datasets, solidifying their understanding of these crucial concepts.

Further, the course delves into advanced topics in image and audio feature extraction, integrating deep learning methodologies. Students will study computer vision and image descriptors, including histograms, texture, and

spatio-temporal features, alongside techniques for analyzing image and video movement. Deep learning's role in extracting features from images and audio signals is explored, complemented by a detailed examination of audio descriptors and deep learning for audio analysis. This section includes a lab session focused on the analysis and knowledge extraction from images, videos, and audio. Additionally, the course covers recommender systems, text analysis, and feature extraction, with practical sessions to develop and test various recommender algorithms and text analytics techniques. Students will engage with machine and deep learning applications for multimedia content, learning about advanced techniques and recent architectures. The course concludes with a project development, where students work in pairs or individually to apply their knowledge to real-world problems, culminating in a project presentation. During the project, each week students will present their progress to the rest of the class, followed by a brief discussion between professors and students to address any questions or highlight interesting parts of the projects.

4.2. Temario de la asignatura

1. Exploratory Data Analysis
 - 1.1. Analytic graphics and techniques for exploratory data analysis
 - 1.2. Data transformation and dimensionality reduction, clustering and outliers detection
 - 1.3. Algorithms analysis, hyperparameters, evaluation metrics
 - 1.4. Explainable AI
 - 1.5. Lab session: EDA, analysis and XAI applied to datasets
2. Image and Audio Feature Extraction techniques and deep learning
 - 2.1. Computer vision and image descriptors: Histograms, Texture, Spatio-Temporal
 - 2.2. Image/Video Movement
 - 2.3. Deep Learning for image feature extraction
 - 2.4. Audio signal and audio descriptors
 - 2.5. Deep Learning for audio analysis
 - 2.6. Lab session: Image, video, audio analysis and knowledge extraction
3. Text analysis and feature extraction techniques
 - 3.1. Text mining and cleaning
 - 3.2. Vector space models
 - 3.3. Keywords extraction and summarization
 - 3.4. Text clustering, classification and recommendation

- 3.5. Deep learning for text
- 3.6. Lab session: Text analytics
- 4. Recommender Systems
 - 4.1. Simple recommenders
 - 4.2. Content-based recommenders
 - 4.3. Collaborative filtering
 - 4.4. Hybrid recommenders
 - 4.5. Machine/Deep Learning for recommendation
 - 4.6. Lab session: Simple, content-based and collaborative filtering recommenders
- 5. Machine/Deep Learning Applications with multimedia content
 - 5.1. Advanced techniques for multimedia content analysis
 - 5.2. Recent deep learning architectures for multimedia content
- 6. Project development in pairs or individually
 - 6.1. Project guidance
 - 6.2. Project results and presentation

5. Cronograma

5.1. Cronograma de la asignatura *

| Sem | Actividad tipo 1 | Actividad tipo 2 | Tele-enseñanza | Actividades de evaluación |
|-----|---|---|----------------|--|
| 1 | Chapter 1 - EDA and data understanding, XAI Duración: 03:00 LM: Actividad del tipo Lección Magistral | | | |
| 2 | | Lab 1 - EDA and data understanding, XAI Duración: 03:00 PL: Actividad del tipo Prácticas de Laboratorio | | Lab session report and homework 1 EP: Técnica del tipo Examen de Prácticas Evaluación Progresiva Presencial Duración: 03:00 |
| 3 | Chapter 2 - Image feature extraction techniques, analysis, ML and DL. Duración: 03:00 LM: Actividad del tipo Lección Magistral | | | |
| 4 | Chapter 2 - Image/Video and audio feature extraction techniques and Deep Learning Duración: 03:00 LM: Actividad del tipo Lección Magistral | | | |
| 5 | | Lab 2 - Image/Video/Audio feature extraction Duración: 03:00 PL: Actividad del tipo Prácticas de Laboratorio | | Lab session report and homework 2 EP: Técnica del tipo Examen de Prácticas Evaluación Progresiva Presencial Duración: 03:00 |
| 6 | Chapter 3 - Text feature extraction techniques, analysis and Deep learning Duración: 03:00 PL: Actividad del tipo Prácticas de Laboratorio | | | |
| 7 | | Lab 3 - Text feature extraction techniques, analysis and Deep learning Duración: 03:00 PL: Actividad del tipo Prácticas de Laboratorio | | Lab session report and homework 3 EP: Técnica del tipo Examen de Prácticas Evaluación Progresiva Presencial Duración: 03:00 |
| 8 | Chapter 4 - Recommender systems Duración: 03:00 LM: Actividad del tipo Lección Magistral | | | |
| 9 | | Lab 4 - Recommender systems Duración: 03:00 PL: Actividad del tipo Prácticas de Laboratorio | | Lab session report and homework 4 EP: Técnica del tipo Examen de Prácticas Evaluación Progresiva Presencial Duración: 03:00 |

| | | | | |
|----|---|---|--|---|
| 10 | Chapter 5 - Advanced techniques and analysis in multimedia data Duración: 03:00 LM: Actividad del tipo Lección Magistral | | | |
| 11 | | Project session: Project proposal and dataset used Duración: 03:00 AR: Aprendizaje basado en retos | | |
| 12 | | Project session: Feature extraction techniques Duración: 03:00 AR: Aprendizaje basado en retos | | |
| 13 | | Project session: Preliminary results and deep analysis Duración: 03:00 AR: Aprendizaje basado en retos | | |
| 14 | | Project session: Final projects ideas and results Duración: 00:00 AR: Aprendizaje basado en retos | | |
| 15 | | | | Project development and presentation TG: Técnica del tipo Trabajo en Grupo Evaluación Progresiva Presencial Duración: 00:30 |
| 16 | | | | |
| 17 | | | | Exam: Test / Short Questions EX: Técnica del tipo Examen Escrito Evaluación Progresiva Presencial Duración: 02:00 Lab sessions report and homework TI: Técnica del tipo Trabajo Individual Evaluación Global No presencial Duración: 00:00 Final exam EX: Técnica del tipo Examen Escrito Evaluación Global Presencial Duración: 02:00 |

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 |
|------|--------------------------------------|--|------------|----------|-----------------|-------------|--|
| 2 | Lab session report and homework 1 | EP: Técnica del tipo Examen de Prácticas | Presencial | 03:00 | 6.25% | 4 / 10 | CB06 CB07 CB09 CB10 |
| 5 | Lab session report and homework 2 | EP: Técnica del tipo Examen de Prácticas | Presencial | 03:00 | 6.25% | 4 / 10 | CB06 CB07 CB09 CB10 |
| 7 | Lab session report and homework 3 | EP: Técnica del tipo Examen de Prácticas | Presencial | 03:00 | 6.25% | 4 / 10 | CB06 CB07 CB09 CB10 |
| 9 | Lab session report and homework 4 | EP: Técnica del tipo Examen de Prácticas | Presencial | 03:00 | 6.25% | 4 / 10 | CB07 CB09 CB10 |
| 15 | Project development and presentation | TG: Técnica del tipo Trabajo en Grupo | Presencial | 00:30 | 35% | 5 / 10 | CB06 CB07 CB09 CB10 CT05 CE01 CE02 |
| 17 | Exam: Test / Short Questions | EX: Técnica del tipo Examen Escrito | Presencial | 02:00 | 40% | 4 / 10 | CB07 CB09 CB10 CT05 CE01 |

6.1.2. Prueba evaluación global

| Sem | Descripción | Modalidad | Tipo | Duración | Peso en la nota | Nota mínima | Competencias evaluadas |
|-----|-------------|-----------|------|----------|-----------------|-------------|------------------------|
|-----|-------------|-----------|------|----------|-----------------|-------------|------------------------|

| | | | | | | | |
|----|----------------------------------|---|---------------|-------|-----|--------|--|
| 17 | Lab sessions report and homework | TI: Técnica del tipo Trabajo Individual | No Presencial | 00:00 | 25% | 5 / 10 | CB06 CB07 CB09 CB10 |
| 17 | Final exam | EX: Técnica del tipo Examen Escrito | Presencial | 02:00 | 75% | 5 / 10 | CB06 CB07 CB09 CB10 CT05 CE01 CE02 |

6.1.3. Evaluación convocatoria extraordinaria

| Descripción | Modalidad | Tipo | Duración | Peso en la nota | Nota mínima | Competencias evaluadas |
|----------------------------------|---|------------|----------|-----------------|-------------|--|
| Extraordinary assessment | EX: Técnica del tipo Examen Escrito | Presencial | 02:00 | 75% | 5 / 10 | CB06 CB07 CB09 CB10 CT05 CE01 CE02 |
| Lab sessions report and homework | TI: Técnica del tipo Trabajo Individual | Presencial | 00:00 | 25% | 5 / 10 | CB06 CB07 CB09 CB10 |

6.2. Criterios de evaluación

Students will be assessed through **progressive evaluation**. This evaluation will determine if students have acquired all the competencies of the subject. Therefore, the final assessment will encompass all the evaluation techniques used in continuous evaluation (EX, ET, TG, etc.) and will take place during the exam period approved by the School Board for the current academic semester and year. Evaluation activities that measure learning outcomes not suitable for a single exam may be conducted throughout the semester. The **progressive evaluation** will be based on the following elements:

- Attendance and Participation: Attend and follow both theory and practical sessions. Four laboratory sessions will be held. Each lab will include questions that need to be solved during the session and submitted at the end (mandatory assistance). After the laboratory sessions, a homework assignment on different datasets must be completed and submitted by the provided date (mandatory practices). If a student doesn't submit some homework on time or doesn't pass the tasks with the minimum mark, they will be evaluated by global examination (described below). If a student cannot attend some mandatory laboratory sessions, it must be properly justified, and an extraordinary period can be arranged to submit the laboratory questions and answers.
- Final Project: Develop a final project in pairs or individually and present the results to the class. This activity includes reading and extracting the main ideas from relevant papers in the field. The project will include a report (code and results) and a presentation to the group (10-15 minutes plus 5 minutes of discussion) by each pair/individual. During the project, a presentation of the current results will be conducted weekly (project follow-up).
- Individual Exam: Questions on the theoretical content of the course and practical sessions.

All parts are required to pass the subject.

The weight of such activities, **all mandatory** are:

- Lab sessions and homework (individual) 25% - minimal threshold 4/10 in each one
- Project (pairs/individual) 35% - minimal threshold 5/10
- Exam [Test / Short questions] (individual) 40% - minimal threshold of 4/10

In the event that a student does not pass the lab sessions and practices or the project, the student will be evaluated by **global examination**, with the requirement of submitting the laboratory practice tasks (homework) at least two days before the exam. If a student has already passed some labs/homework, it is not necessary to submit them again, only the ones that were not passed. In this case, the percentages for the evaluations are as follows:

- Lab homework (individual) 25% - minimal threshold 5/10 in each one
- Exam [Test / Short questions] (individual) 75% - minimal threshold of 5/10

Extraordinary examination will be carried out by a final exam and the requirement of completing the laboratory practices tasks (homework) at least two days before the exam. If a student has already passed some homework, it is not necessary to submit them again, only the ones that were not passed. The minimum percentages to pass the subject in this case are:

- Lab homeworks (individual) 25% - minimal threshold 5/10
- Exam [Test / Short questions] (individual) 75% - minimal threshold of 5/10

7. Recursos didácticos

7.1. Recursos didácticos de la asignatura

| Nombre | Tipo | Observaciones |
|---|--------------|--|
| Pentreath, N. (2015). Machine Learning with Spark. Packt Publishing Ltd. | Bibliografía | Machine Learning with Spark |
| Segaran, T. (2007). Programming collective intelligence: building smart web 2.0 applications. " O'Reilly Media, Inc." | Bibliografía | Programming collective intelligence: building smart web 2.0 applications |
| Aggarwal, C. C., & Zhai, C. (2012). Mining text data. Springer Science & Business Media. | Bibliografía | Mining text data. Springer Science & Business Media. |
| Python Data Science Handbook - Jake VanderPlas | Bibliografía | Data analysis |
| Python for Data Analysis, Data Wrangling with Pandas, NumPy, and IPython - Wes Mckinney | Bibliografía | Data analysis |

| | | |
|--|--------------|------------------------------------|
| Hands On Machine Learning with Scikit Learn and TensorFlow - Aurelien Geron | Bibliografía | Machine Learning and Deep Learning |
| Data Analysis and Visualization Using Python Analyze Data to Create Visualizations for BI Systems - Dr. Ossama Embarak | Bibliografía | Data analysis and visualization |
| Introduction to Machine Learning with Python: A Guide for Data Scientists - Sarah Guido | Bibliografía | Artificial intelligence |

8. Otra información

8.1. Otra información sobre la asignatura

In this subject we align with the the Sustainable Development Goals (SDG) 4, 5 and 9.

Especifically this subject will support in the activities to be carried out:

4.4 By 2030, substantially increase the number of youth and adults who have relevant skills, including technical and vocational skills, for employment, decent jobs and entrepreneurship.

5.B Enhance the use of enabling technology, in particular information and communications technology, to promote the empowerment of women.

9.5 Enhance scientific research, upgrade the technological capabilities of industrial sectors in all countries, in particular developing countries, including, by 2030, encouraging innovation and substantially increasing the number of research and development workers per 1 million people and public and private research and development spending.