Universidad de Navarra

Machine Learning (MINT)
Guía docente 2025-26

PRESENTACIÓN

Breve descripción:

The recent increase in computing power, together with new parallel computing techniques, have opened up a huge range of potential applications in multiple industries.

This subject will study, from a theoretical-practical point of view, the newest technologies of both machine and deep learning (ML and DL) research for the search for applications in different fields.

- **Titulación**: Máster en Innovación Tecnológica
- **Módulo/Materia**: Tecnologías Emergentes (MINT) / Fundamentos avanzados de análisis de datos y aprendizaje automático
- **ECTS**: 5 ECTS
- **Curso, semestre**: 1º, Primero
- **Carácter**: Obligatorio
- **Profesorado**:
  - Doncel Llamas, Ander / Invitado
  - Granada Echeverria, Imanol - Email: igranada@external.unav.es / Invitado (Colab.Docente)
  - Rubio Díaz-Cordovés, Ángel / Catedrático
  - Salcedo Gamarra, José Luis / Invitado
- **Idioma**: Inglés

RESULTADOS DE APRENDIZAJE (Competencias)

- CG1 - Integrar visión estratégica y tecnología para generar nuevos modelos de negocio.
- CG2 - Coordinar grupos de trabajo multidisciplinares para desarrollar procesos de transformación basados en tecnologías emergentes.
- CB6 - 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
- CB7 - 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
- CB9 - 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.
- CE6 - Conocer y aplicar las técnicas de análisis de datos y los principales modelos de aprendizaje automático en diferentes casos de uso.
- CE7 - Conocer las principales herramientas, tecnologías y entornos de desarrollo disponibles para implementar soluciones de inteligencia artificial.
- CE8 - Evaluar ventajas e inconvenientes de sistemas de captación, depuración, visualización de datos.

PROGRAMA

Machine Learning Course Syllabus:

Machine Learning Overview:

- Introduction to supervised and unsupervised learning
- Train, Validation, Test Paradigm
- Understanding the importance of data splitting
- Evaluating model performance

Introduction to Python and other AI tools

- Overview of Python programming language
- Basics of Python for machine learning
- Hugging face
- Gemini, chatGPT

Supervised Learning: Regression

- Understanding regression models
- Building and evaluating regression models

Supervised Learning: Classification

- Logistic Regression, SVM, and Multi-class Classification
  - Logistic regression for binary classification
  - Support Vector Machines (SVM) for binary and multi-class classification
- Classification Trees, Random Forests, and XGBoost
  - Decision tree-based algorithms for classification
  - Ensemble methods and boosting techniques

Unsupervised learning

- Introduction to unsupervised learning algorithms
- Clustering techniques for identifying patterns in data
- Dimensionality reduction techniques
- Introduction to recommender systems

Case Studies and Overarching Project

- Applying machine learning techniques to real-world scenarios

Natural Language Processing (NLP) Workshop

- Analyzing and processing textual data
- Computer Vision (CV) Workshop
  - Applying machine learning to visual data analysis
- Workshop on Large Language Models (chatGPT like)

ACTIVIDADES FORMATIVAS

EVALUACIÓN

General presentation:

In today's world, companies need to constantly innovate and adapt to new challenges and opportunities that arise. The opportunities should be clearly analyzed in order to propose strategic decisions that would define the future of the companies.

Once a business opportunity has been identified, most innovative companies develop digital prototypes that allow their innovative concepts to be tested early and pivot according to the feedback received by the market, translating them into changes in their business model and concepts.

The so-called Pitch Decks are usually included in this process before investors or structures within the companies to obtain the support and resources to launch these new business opportunities on the market. This project is divided in two bimesters:

- In the first bimester, you will learn and apply how to define and analyze potential business strategies that would create value to the companies in comparison to their competitors. Based on this knowledge, in the first part of this project you will propose and justify a data-driven business idea that would help a company to be successful in the future.
- In the second bimester you will have to propose an architecture and deployment plan for your concept using one or more specific technologies (Cloud Computing, IoT and Blockchain), prototype an application or web interface for it and present the pitch deck to investors.

All deliverables of this project must be presented in English.

Particular tasks in this subject:

The project will include at least one machine learning problem. The students are free to use programming language although Python is strongly recommended. The students can develop their own model using different libraries, using HuggingFace or accessing the API of chatGPT or Gemini.

This project will count 60% of the grade of the subject. There will be two sprint reviews and a final presentation:

1st sprint review (around 6-10th of october 2025): Explanation of the main idea in an online meeting (5 minutes presentation per group + discussion)

2nd sprint review (20-24th of october 2025): Detailed description of the problem(s) to be solved and sketch of the techniques to be used. Online meeting (5 minutes presentation per group + discussion)

Final deliverable 11-13th of november 2025): A 10-15 minute video presentation that contains the following aspects:

- Machine learning problem (20% of the grade)
- Justification of the appropriateness of the approach (from scratch, API, HuggingFace) to the software development employed (10% of the grade)
- Description of the optimization process (20% of the grade)
- Visualization and analysis of the results (40% of the grade)
- Conclusions (10% of the grade)

IMPORTANT NOTE 1: the precision and the clarity in the presentation will be of utmost importance for the grading.

This grading strongly overlaps with the Optimization subject. The students are free to share some parts of the video between both subjects.

CONVOCATORIA ORDINARIA

60% Overarching project.

- Grading of the cross-curricular project that the students carry out during the bimester.

40% Other deliveries and participation.

- After each topic, a short test will be asked to be taken in order to consolidate the knowledge.

CONVOCATORIA EXTRAORDINARIA

- Machine learning project (different form the previous one) (60%)
- Exam on Machine Learning (40%).

HORARIOS DE ATENCIÓN

Dr Angel Rubio (arubio@unav.es)

- Horario de tutoria: Concertar teleconferencia con el profesor.

BIBLIOGRAFÍA

Annex: (possible sources of data)

General sources:

https://github.com/awesomedata/awesome-public-datasets

https://www.kaggle.com/datasets (use the opendatasets library for downloading datasets)

https://course.fast.ai/datasets

https://www.kaggle.com/competitions (check the "Completed" tab)

https://archive.ics.uci.edu/ml/index.php

https://datasetsearch.research.google.com/

World Covid-19 Dataset

https://www.kaggle.com/imdevskp/corona-virus-report

AI for Social Good - Google AI

https://ai.google/social-good/