Universidad de Navarra

Machine learning (MCD)
Guía docente 2026-27

PRESENTACIÓN

Breve descripción:

Este curso sentaremos las bases sobre los conceptos de *machine learning* usados en la ciencia de datos. A través de una combinación de teoría y ejercicios prácticos, los estudiantes explorarán algoritmos clave, técnicas de modelado y métodos de aprendizaje automático aplicados en diversos campos, desde la biomedicina hasta el análisis de datos complejos. Los participantes no solo desarrollarán competencias teórico-técnicas en herramientas como Python, sino que también aprenderán a abordar problemas reales trabajando en casos de uso con datos reales. Para ello introduciremos conceptos de modelos estadísticos tradicionales, modelos lineales, regresión y modelos generalizados. Posteriormente introduciremos conceptos de aprendizaje automático, clasificaciones supervisadas y no supervisadas, así como modelos de tomas de decisión. Finalmente, usaremos estos conceptos de forma práctica para desarrollar proyectos integradores en los que los estudiantes podrán aplicar técnicas de preprocesamiento, modelado predictivo y validación de modelos en escenarios reales.

- **Titulación**: **MÁSTER UNIVERSITARIO EN CIENCIA DE DATOS PARA CIENCIAS EXPERIMENTALES**
- **Módulo/Materia**: **Módulo 2 (General) /Materia 2.3. (Introducción a la Bioinformática)**
- **ECTS**: 5
- **Curso, semestre**: 1º Curso, 1º Semestre
- **Carácter**: Obligatorio
- **Profesorado**:
- Iñaki Echeverría Huarte, email: iecheverriah@unav.es
- Ibon Tamayo Uria, email: itamayou@unav.es
- **Idioma**: Español. Se requieren conocimientos de inglés.
- **Aula, Horario**: Por determinar.

RESULTADOS DE APRENDIZAJE (Competencias)

- RA1: Aplicar soluciones computacionales para problemas científicos, usando una amplia gama de plataformas de análisis de datos.
- RA6: Evaluar las implicaciones éticas de los sistemas automáticos e inteligencia artificial en las ciencias experimentales, considerando sesgos algorítmicos, discriminación y transparencia en la toma de decisiones.
- RA7: Implementar algoritmos eficientes para la resolución de problemas científicos utilizando programación estructurada y pseudocódigo.
- RA8: Evaluar la aplicabilidad de diferentes lenguajes de programación (Python y Matlab) en el contexto de las ciencias experimentales.
- RA9: Emplear con soltura la sintaxis de Python y Matlab para la manipulación de datos y la implementación de algoritmos científicos.
- RA14: Manipular y procesar datos científicos de diversas fuentes (archivos de texto, bases de datos, etc.) utilizando técnicas de data wrangling en diferentes entornos de programación.
- RA15: Generar e interpretar representaciones gráficas adecuadas para comunicar resultados científicos de manera clara y efectiva en diferentes entornos de programación.
- RA17: Implementar técnicas de inferencia estadística para analizar datos científicos, extraer conclusiones y tomar decisiones informadas.
- RA 18: Desarrollar modelos de aprendizaje automático de extremo a extremo para resolver problemas de predicción, clasificación y agrupamiento en ciencias experimentales.

PROGRAMA

Introducción:

- ¿Qué es el Machine Learning?
- Tipos de aprendizaje
  - Aprendizaje Supervisado
  - Aprendizaje No Supervisado
  - Aprendizaje Auto Supervisado
  - Aprendizaje Por Refuerzo
- Notaciones y definiciones importantes.

Introducción a los modelos de aprendizaje automático supervisado: Regresión

- Modelo lineal simple.
- Modelos lineales generalizados.

Introducción a los modelos de aprendizaje automático supervisado: Clasificación

- Regresión logística.
- Support Vector Machine.
- Árboles de decisión y métodos de conjunto.

Introducción a los modelos de aprendizaje automático no supervisado

- Reducción de dimensionalidad: Principal Component Analysis (PCA).
- Clusterización: K-means. Clustering jerarquizado.

End-to-end machine learning:

- Seleccion de modelos.
- Bias/Variance trade-off - Overfitting.
- Cross-validation.
- Seleccion de variables.
- Caso practico.

ACTIVIDADES FORMATIVAS

Asignatura de 5 ECTS que equivalen a **50 horas presenciales.**

1. ACTIVIDADES PRESENCIALES (**50 horas**)

- Clases presenciales teóricas (30 horas)
- Clases presenciales teórico-prácticas (20 horas)

2. ACTIVIDADES NO PRESENCIALES (**75 horas**)

- Trabajos dirigidos (15 horas)
- Estudio personal (60 horas)

EVALUACIÓN

La evaluación de este módulo será continua. Los alumnos se expondrán a las siguientes evaluaciones:

- Intervención en clases, seminarios y clases prácticas (10%)

- Prueba escrita teoría y práctica con Python - Aprendizaje automatico supervisado (20%)
  - Regresión Lineal (10%)
  - Regresión Logística (10%)
- Prueba escrita teoría y práctica con Python - Aprendizaje automatico no supervisado (10%)
- Trabajo en práctica en Python – End-to-end machine learning (15%)
- Trabajo final de asignatura y exposición pública del trabajo (35%)

HORARIOS DE ATENCIÓN

No hay horario específico para la atención de estudiantes. Concertar cita previa por e-mail y se acuerda la fecha y hora de la tutoría:

Profesores:

- Iñaki Echeverría, email: iecheverriah@unav.es
- Ibon Tamayo, email: itamayou@unav.es

BIBLIOGRAFÍA Y RECURSOS

- **An Introduction to Statistical Learning**. Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, 2013. ISBN: 1-4614-7138-9 Localízalo en la Biblioteca [Recurso electrónico]
- **Machine Learning with R: Expert Techniques for Predictive Modeling to Solve All Your Data Analysis Problems**. Lantz, Brett. Birmingham: Packt Publishing, 2019. ISBN: 9781523125241 Localízalo en la Biblioteca
- **Introduction to machine learning with Python: a guide for data scientists.** Andreas C. Müller and Sarah Guido. Müller, Andreas C.. Sebastopol, CA: O'Reilly, 2016. ISBN: 978-1-449-36941-5 Localízalo en la Biblioteca
- **Python machine learning.** Lee, Wei-Meng. Indianapolis, IN: Wiley, 2019. Localízalo en la Biblioteca [Recurso electrónico]