

**CICLO DE COMPLEMENTACIÓN
CURRICULAR
LICENCIATURA EN CIENCIA DE DATOS**



Ciclo Disciplinar Específico.

1º Año – 1º Cuatrimestre

Programa de Asignatura

**CLCD 162 - APRENDIZAJE
AUTOMÁTICO**

- 2025 -

FUNDAMENTACIÓN

El Aprendizaje Automático es una rama de la Inteligencia Artificial que se dedica al estudio de algoritmos y programas que “aprenden” a realizar una tarea en base a evidencia conocida. El proceso de aprendizaje en base a evidencia implica la selección y entrenamiento de un determinado modelo de modo que pueda dar respuesta a nuevas situaciones no conocidas.

Y conlleva un proceso iterativo que requiere que cada paso sea revisado a medida que se aprende más sobre el problema bajo investigación. Este proceso iterativo puede requerir el uso de diferentes herramientas para cada paso.

La materia propone el estudio de los diferentes algoritmos de machine learning, comprendiendo el proceso, su utilización, la elección del algoritmo más adecuado según el caso de negocio, así como la evaluación y mejora del modelo.

Para la aplicación práctica basado en casos se utilizará el lenguaje Python mediante "notebooks" de Jupyter¹, entendida como una herramienta ampliamente utilizada por la comunidad a nivel mundial, que permite la combinación de código, textos, gráficas, figuras, etc.

OBJETIVOS

OBJETIVOS GENERALES

El aprendizaje automático es la ciencia de conseguir que las computadoras actúen sin haber sido explícitamente programadas. En esta materia se desarrollarán los fundamentos teóricos del aprendizaje automático y se

¹ https://es.wikipedia.org/wiki/Proyecto_Jupyter

resolverán problemas concretos. Al finalizar el curso el alumno conocerá los principales modelos y algoritmos de aprendizaje automático y podrá seleccionar el más adecuado para los casos prácticos que se presenten.

OBJETIVOS ESPECÍFICOS

Al finalizar la cursada el alumno alcanzará saberes conceptuales, procedimentales y actitudinales para:

- Comprender los fundamentos del aprendizaje automático, y conocer los tipos supervisado y no supervisado. Conocer conceptos básicos de aprendizaje por refuerzo y aprendizaje profundo.
- Comprender y aplicar el proceso de aprendizaje automático. Detectar que etapas requieren mayor esfuerzo para lograr resultados significativos. esto puede incluir técnicas estadísticas y herramientas de análisis de datos.
- Identificar que modelos utilizar de acuerdo al caso de negocio. Esto implica comprender los objetivos de negocio, seleccionar las variables relevantes, seleccionar el algoritmo, determinar el tamaño de los conjuntos de entrenamiento, testeo y validación.
- Entrenar el algoritmo, evaluar el resultado del modelo y comprender si se requiere optimizarlo.
- Comunicar los resultados de manera efectiva. Esto implica la redacción de informes y la presentación de los resultados para transmitir la información de manera comprensible.

CONTENIDOS

Unidad 1: Fundamentos del Aprendizaje Automático

- 1: Introducción al Aprendizaje Automático
- 2: Tipos de Aprendizaje Automático: Supervisado, No supervisado y Reforzado
- 3: Entrenamiento y evaluación de modelos de Aprendizaje Automático
- 4: Sesgo y varianza en el Aprendizaje Automático

Unidad 2: Aprendizaje Supervisado Básico

5. Componentes de un sistema de Aprendizaje Supervisado
- 6: Clasificación con modelos lineales
- 7: Clasificación y regresión basadas en instancias (KNN)
- 8: Árboles de decisión y Random Forests

Unidad 3: Aprendizaje Supervisado Avanzado

9. Ensembles de árboles: Bagging y Boosting
10. Support Vector Machines (SVM) para clasificación y regresión
11. Redes neuronales básicas: Perceptrón y MLPs
12. Redes neuronales pre-entrenadas y transferencia de aprendizaje

Unidad 4: Aprendizaje No Supervisado

13. Análisis de componentes principales (PCA)
- 14: Clustering: K-Means y DBSCAN
- 15: Clasificador Naive Bayes

Unidad 5: Evaluación de Modelos

- 16: Selección de atributos
- 17: Matriz de confusión
- 18: Sensibilidad y especificidad
- 19: Curva ROC
- 20: MSE y R^2

Unidad 6: Conjuntos de Datos desbalanceados

- 21: Casos de Uso con conjuntos de datos desbalanceados
- 22: Técnicas de Oversampling y Undersampling

Unidad 7: Introducción al Aprendizaje Profundo

23: Introducción a las redes neuronales profundas

24: Casos de Uso de redes neuronales profundas

BIBLIOGRAFÍA

Bishop C. (2006) Pattern Recognition and Machine Learning Springer 2006.

Bruce, P., Bruce, A., & Gedeck, P. (2020). Practical statistics for data scientists: 50+ essential concepts using R and Python. O'Reilly Media.

Grus J. (2015) Data Science from Scratch: first principles with Python. O'Reilly Media, Inc.: Gravenstein Highway North, Sebastopol.

McKinney, W. (2023). Python para Análisis de Datos: Manipulación de datos con Pandas, Numpy y Jupyter" Anaya Multimedia".

Mitchell Tom M. (1997) Machine Learning. WCB McGraw-Hill

Murphy, K. P. (2012). Machine learning: a probabilistic perspective. MIT press.

VanderPlas, J. (2016). Python data science handbook: Essential tools for working with data. " O'Reilly Media, Inc." - Cap. 3

Witten I., Frank E. Hall, M. (2011) Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann. Elsevier.

METODOLOGÍA DE TRABAJO

- Aprendizaje basado en proyecto y aula invertida

- Al inicio de cada semana, se subirá material en distintos formatos para que los /as alumnos puedan abordar su lectura, visualización o incluso experimentación (en caso de ser notebooks)
- Cada semana se realizarán dos clases sincrónicas de 1 a 1,5 horas de duración. La primera de carácter práctico, donde se revisarán las ejercitaciones propuestas, así como trabajos específicos sobre el material puesto a disposición. La segunda tendrá un contenido más teórico.
- Las clases quedarán grabadas y subidas a la plataforma

Modalidad UGR Virtual

En la Plataforma Virtual de la UGR se desarrollarán las actividades sincrónicas y se dispondrán los diferentes recursos de uso asincrónico indicados para la modalidad de aula invertida. Los formatos de dichos recursos serán tanto de lectura, escucha o visualización (material bibliográfico, podcasts y videos), como de experimentación (notebooks). La interacción con la cátedra se desarrollará también en la Plataforma Virtual para la entrega de los trabajos y ejercicios y para el acceso a las correcciones y devoluciones. En el mismo marco serán respondidas las consultas que se realicen acerca de contenidos teóricos, enunciados de ejercicios, cronogramas y aspectos organizativos.

EVALUACIÓN

La aprobación de la asignatura se puede alcanzar en condición de alumno regular o alumno libre.

Las instancias de evaluación propuestas por la cátedra son las siguientes:

- 1 examen parcial de contenido teórico práctico

- 1 ejercitación grupal a desarrollarse mediante aplicaciones web tipo Jupyter Notebooks. Las ejercitaciones se relacionarán con los contenidos presentados en las clases.
- 1 trabajo práctico integral grupal realizado en grupos de entre 4 y 5 alumnos, detallado en el aula virtual.

Al finalizar el cuatrimestre habrá una instancia recuperatoria para cada una de las evaluaciones.

- Estudiante **regular** es aquel que apruebe durante el periodo de cursada con nota igual o superior a 6 (seis) cada una de las instancias de evaluación definidas anteriormente, en primera instancia o en recuperatorio.
- Estudiante **promovido** es aquel que apruebe durante el periodo de cursada con nota igual o superior a 8 (ocho) en primera instancia de cada una de las evaluaciones definidas anteriormente (sin recuperatorio).
- Estudiante **libre** es aquel que no apruebe alguna de las evaluaciones dispuestas para el período de cursada, ni en primera instancia ni en recuperatorio.
Para aprobar la materia, los estudiantes libres deberán resolver en la instancia de **mesas de examen** un examen teórico práctico oral y escrito, nota mayor o igual a 6 (seis) en cada una de las evaluaciones (oral y escrito).

ACTIVIDADES DE EXTENSIÓN

La Asignatura no tiene previstas actividades de extensión curricularizadas.