

**Aprovechamiento Inteligente
de Datos Masivos**

Carrera:

Doctorado en Ciencias Informáticas (plan2001)

Docente Responsable: Dr. José Angel Olivas Varela

Docente: Dr. Andrés Montoro Montarroso

Duración total: 100hs

(75hs de interacción pedagógica y 25hs de trabajo autónomo del alumno)

Modalidad: Virtual sincrónico

Créditos: 4 (cuatro)

OBJETIVO

Presentar el aprovechamiento inteligente de datos masivos o entornos Big Data, desde de punto de vista del estudio detallado de los datos disponibles, las técnicas inteligentes a aplicar para su aprovechamiento y las diferentes salidas esperables como resultado del análisis, partiendo de la definición de 'Sistema Inteligente' de ACT (Comisión Europea 2025)

MODALIDAD DE EVALUACION

Para aprobar el curso se requiere un 80% de asistencia y la realización de un trabajo final que se definirá una vez completada la exposición de los contenidos teóricos. Dicho proyecto deberá ser desarrollado en forma individual y tendrá por objetivo profundizar uno o varios de los conceptos vistos en clase. La entrega consistirá en un reporte técnico de calidad científica producto de un estudio teórico o experimental específico.

PROGRAMA

Parte I: Conceptos

- Sistemas Inteligentes: Definición, objetivos, retos, límites, ejemplos.
- Aprovechamiento inteligente de datos masivos: Aprendizaje Automático y Conocimiento.
- Minería de datos, textos, opiniones y señales: Metodologías.
- Sistemas Inteligentes en distintos ámbitos: empresas, gobiernos, economía, leyes, finanzas, contadurías, recursos humanos, manejo de calidad, organización de proyectos, manejo de recursos, industria, deportes, redes sociales.

Parte II: Introducción al Aprendizaje Automático

- Técnicas inteligentes para el aprovechamiento de datos masivos: modelos supervisados y no supervisados. Problemas de clasificación, predicción, clustering y asociación. Modelos basados en reglas, árboles, redes neuronales, similitud, algoritmos evolutivos y genéticos. Conjuntos de entrenamiento, prueba y validación, generalización, sobreajuste.
- Introducción al Aprendizaje Supervisado: Regresión, Naïve Bayes, Decision Trees, KNN, Redes Neuronales Artificiales, Ensembles.
- Introducción al Aprendizaje No Supervisado: K-means Hierarchical Clustering, DBSCAN,
- Introducción al Procesamiento de Lenguaje Natural: Topic Modelling: BERTopic.
- Aplicación a problemas reales:
 - Segmentación, predicción/pronóstico/estimación en ciudades inteligentes.
 - Análisis de Series Temporales.
 - Sistemas recomendadores.
 - Observación en Redes Sociales.

BIBLIOGRAFIA

- ACT- Comisión Europea: <https://artificialintelligenceact.eu/the-act/>
- Gareth, J., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). An Introduction to Statistical Learning with Applications in R. Springer. <https://www.statlearning.com>
- Han J, Kamber M, Pei J. (2012). Data mining : concepts and techniques. 3rd ed.: Morgan Kaufmann.
- Tan PN, Steinbach M, Karpatne A, Kumar V. (2019). Introduction to Data Mining. 2nd ed.: Pearson.
- Witten, I. H. and Frank, E., Hall, M. A., & Pal, C. J. (2017). Data mining: practical machine learning tools and techniques. 4th ed.: Morgan Kaufmann.
- Goodfellow I, Bengio Y, Courville A. (2016). Deep Learning Cambridge: MIT Press.
- Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python. O'Reilly Media, Inc.
- <https://www.oreilly.com/library/view/natural-language-processing/9780596803346/>