

Formulario de aprobación de curso de posgrado/educación permanente

Asignatura: Técnicas de Aprendizaje Automático y Minería de Datos: Modelos Predictivos Nivel iniciación (Software: Stata)

| | | |
|---|-----------------------------|-------------------------------------|
| Modalidad: (posgrado, educación permanente o ambas) | Posgrado | <input checked="" type="checkbox"/> |
| | Educación permanente | <input checked="" type="checkbox"/> |

Profesor de la asignatura ¹: Dra. Yolanda Rebollo, Profesora Titular, Departamento de Economía de la Universidad Pablo de Olavide
(título, nombre, grado o cargo, instituto o institución)

Profesor Responsable Local ¹: Dra. Ing. Aiala Rosá, grado 4, InCo
(título, nombre, grado, instituto)

Otros docentes de la Facultad:
(título, nombre, grado, instituto)

Docentes fuera de Facultad:
(título, nombre, cargo, institución, país)

¹ Agregar CV si el curso se dicta por primera vez.
(Si el profesor de la asignatura no es docente de la Facultad se deberá designar un responsable local)

[Si es curso de posgrado]

Programa(s) de posgrado: Especialización en Ciencia de Datos, Maestría en Ciencia de Datos Aplicada

Instituto o unidad: Instituto de Computación

Departamento o área: Grupo Procesamiento de Lenguaje Natural

Horas Presenciales: 15
(se deberán discriminar las horas en el ítem Metodología de enseñanza)

Nº de Créditos: 2
[Exclusivamente para curso de posgrado]
(de acuerdo a la definición de la UdelaR, un crédito equivale a 15 horas de dedicación del estudiante según se detalla en el ítem Metodología de enseñanza)

Público objetivo: Interesados en la aplicación de técnicas avanzadas de análisis y procesamiento de datos, como el Aprendizaje Automático y la Inteligencia Artificial que quieran desarrollar habilidades prácticas para aplicarlas en el mundo real.

Cupos: 10
(si corresponde, se indicará el número de plazas, mínimo y máximo y los criterios de selección. Asimismo, se adjuntará en nota aparte los fundamentos de los cupos propuestos. Si no existe indicación particular para el cupo máximo, el criterio general será el orden de inscripción, hasta completar el cupo asignado)

Objetivos:

En la actualidad, el Big Data se ha convertido en un aspecto fundamental en el mundo empresarial y tecnológico. La capacidad de recopilar, procesar y analizar grandes cantidades de datos es esencial para la toma de decisiones estratégicas y la innovación en diversas áreas, desde la investigación científica hasta la inteligencia de negocios. El presente curso está diseñado para presentar los métodos y técnicas básicos para el procesamiento y análisis de grandes conjuntos de datos y aplicar técnicas de Aprendizaje Automático para resolver problemas del mundo real. Entre sus objetivos se encuentran los siguientes: i) Comprender los fundamentos del Big Data y cómo se diferencia del análisis de datos tradicional; Familiarizarse con las herramientas y tecnologías utilizadas en el procesamiento y análisis de Big Data; Aprender técnicas de análisis y procesamiento de datos, como la minería de datos ; Aprender los conceptos básicos del Aprendizaje Automático; Aplicar técnicas de Aprendizaje Automático a conjuntos de datos de Big Data, y comprender los desafíos y consideraciones específicas al trabajar con conjuntos de datos grandes y complejos; Desarrollar habilidades prácticas mediante la implementación de proyectos y ejercicios prácticos en entornos reales de Big Data.

Las técnicas de minería de datos (data mining) y aprendizaje automático (Machine Learning) tienen como objetivo esencial el descubrimiento automático del conocimiento contenido en la información almacenada de modo ordenado en conjuntos de datos que pueden tener una dimensión elevada. Estas técnicas tienen por objeto descubrir patrones, perfiles, tendencias y otras relaciones presentes en la información, pero ocultas si no se tratan adecuadamente. Cuando la dimensión de la información es muy alta, proviene de múltiples orígenes y no está estructurada, entran en juego las herramientas de tratamiento masivo de datos (Big Data).

El curso está destinado a aquellos que están interesados en la aplicación de técnicas avanzadas de análisis y procesamiento de datos, como el Aprendizaje Automático y la Inteligencia Artificial, y quieren desarrollar habilidades prácticas para aplicarlas en el mundo real. Está diseñado para ser accesible a principiantes en este área temática y les puede ayudar a mejorar sus habilidades para trabajar con datos. En resumen, cualquier persona interesada en el procesamiento y análisis de grandes cantidades de datos y la aplicación de técnicas avanzadas de análisis y procesamiento de datos puede beneficiarse de este curso.

El presente curso está diseñado para explicar a un nivel introductorio las técnicas de Aprendizaje Automático para resolver problemas del mundo real. Entre sus objetivos se encuentran los siguientes: i) Aprender los conceptos básicos del Aprendizaje Automático; ii) Aprender técnicas de análisis y procesamiento de datos, como la minería de datos; iii) Aprender Técnicas de Aprendizaje Automático típicamente usadas en los modelos de predicción. La temática de este curso se centra en las técnicas de minería de datos y Machine Learning y su implementación mediante software específico (STATA) que se puede usar en diferentes áreas de conocimiento. El objetivo general es capacitar a los alumnos para llevar a cabo análisis de información con métodos y herramientas de trabajo muy actuales de modo sencillo y automatizado. Se hará hincapié en la parte de aplicación de casos e interpretación de los resultados de los análisis.

Conocimientos previos exigidos: No tiene

Conocimientos previos recomendados: Familiaridad básica con el análisis de datos -estadística y econometría-, y el uso de software estadístico

Metodología de enseñanza:

(comprende una descripción de la metodología de enseñanza y de las horas dedicadas por el estudiante a la asignatura, distribuidas en horas presenciales -de clase práctica, teórico, laboratorio, consulta, etc.- y no presenciales de trabajo personal del estudiante)

Descripción de la metodología:

La metodología del curso está diseñada para dotar a los participantes de las habilidades necesarias para implementar y evaluar modelos de aprendizaje automático en diversas situaciones. El enfoque del curso se centra en la aplicación práctica de las técnicas presentadas y en el razonamiento detrás de su selección, asegurando una comprensión profunda y funcional de los conceptos.

1. Combinación de Teoría y Práctica:

- Se presentarán los fundamentos básicos de las herramientas analíticas y técnicas de aprendizaje automático.
- Los conceptos teóricos se complementarán con la práctica de herramientas de análisis de datos, asegurando una integración efectiva del conocimiento.

2. Aplicación en Problemas Económicos Reales:

- Se plantearán diferentes problemas económicos reales para que los participantes puedan aplicar las técnicas aprendidas en contextos prácticos.
- Se utilizarán bases de datos relevantes para la economía, permitiendo a los participantes trabajar con datos reales y adquirir experiencia práctica.

3. Uso de Software Específico (STATA):

- Todas las técnicas y herramientas de análisis se implementarán utilizando el software STATA.
- Se proporcionará formación específica sobre el uso de STATA para el análisis de datos y la aplicación de técnicas de aprendizaje automático.

4. Participación Activa y Aprendizaje Basado en Ejemplos:

- Se fomentará la participación activa de los alumnos a través de ejemplos reales y ejercicios prácticos.
- Los participantes tendrán la oportunidad de colaborar, discutir y resolver problemas en equipo, facilitando un aprendizaje colaborativo y enriquecedor.

Este curso proporcionará una base sólida en los conceptos y técnicas de aprendizaje automático supervisado, equipando a los participantes con las herramientas necesarias para abordar problemas del mundo real mediante la aplicación práctica de estos conocimientos. La metodología está diseñada para garantizar que los participantes no solo comprendan los conceptos teóricos, sino que también sean capaces de aplicar estas técnicas de manera efectiva en su trabajo diario.

Detalle de horas:

- Horas de clase (teórico): 14
- Horas de clase (práctico):
- Horas de clase (laboratorio):
- Horas de consulta:
- Horas de evaluación: 1
 - Subtotal de horas presenciales: 15
- Horas de estudio: 15
- Horas de resolución de ejercicios/prácticos:
- Horas proyecto final/monografía:
 - Total de horas de dedicación del estudiante: 30

Forma de evaluación: Prueba escrita individual

[Indique la forma de evaluación para estudiantes de posgrado, si corresponde]

[Indique la forma de evaluación para estudiantes de educación permanente, si corresponde]

Temario:

1. Introducción: Conceptos fundamentales en el aprendizaje automático y minería de Datos.
 - Funciones de pérdida, Sobreajuste, Matriz de confusión, el problema de la dimensión, Validación Cruzada, etc.
 2. Iniciación al Aprendizaje automático supervisado (predicción y clasificación).
 - Técnicas y Aplicaciones con Software STATA
 - Clasificación y la predicción de valores, técnicas básicas o algoritmos de Penalización (Regularized Regression), árboles de decisión y bosques aleatorios. Lasso, Elastic Net y Ridge Regression.
 3. Iniciación al Aprendizaje automático no supervisado (predicción y clasificación)
 - Clustering jerárquico, Clustering secuencial, K-means
-

Bibliografía:

- Athey, S & Imbens, GW (2019) : ANNUAL REVIEW OF ECONOMICS, VOL 11, Page 685-725, DOI 10.1146/annurev-economics-080217-053433
 - Colin, A.C. & Trivedi, P.K. (2022) : Microeconometrics using stata Volumen I. ISBN-13. 978-1-59718-359-8
 - Colin, A.C. & Trivedi, P.K. (2022) : Microeconometrics using stata Volumen II. ISBN-13. 978-1-59718-361-1
 - (textbook) James, G., Witten, D., Hastie, T., and R. Tibshirani, R. (2013): An Introduction to Statistical Learning with Applications in R. Springer.
-

Datos del curso

Fecha de inicio y finalización: 2 al 4 de setiembre de 2024

Horario y Salón: De 08:30 a 11:30 y de 13:00 a 15:00hs., salón en FCEA a determinar.

Arancel: \$ 9.600

[Si la modalidad no corresponde indique "no corresponde". Si el curso contempla otorgar becas, indíquelo]

Arancel para estudiantes inscriptos en la modalidad posgrado: \$ 9.600

Arancel para estudiantes inscriptos en la modalidad educación permanente: \$ 9.600
