

Técnicas de aprendizaje automático y minería de datos: nivel iniciación

TÉCNICAS DE APRENDIZAJE AUTOMÁTICO Y MINERÍA DE DATOS: MODELOS PREDICTIVOS - SOFTWARE STATA (microcredencial)

Eje estratégico: Formación en competencias para la investigación

Modalidad de formación: Presencial

Colectivo al que se dirige: Personal docente e investigador

Plazas ofertadas: 40

La realización de esta acción formativa está condicionada a que haya un mínimo de 10 participantes.

Requisitos previos:

- Conocimientos básicos de programación (cualquier lenguaje).
- Fundamentos de estadística, incluyendo:
 - Indicadores estadísticos básicos (media, desviación típica, mediana, cuantiles).
 - Concepto de muestreo.
 - Nociones básicas de estimación de modelos (regresión lineal y regresión logística).

Es necesario que las personas participantes lleven sus equipos personales.

Duración estimada de la formación: 10 horas

Calendario y horario: 26 y 27 de junio de 2025, de 9:00 a 14:00 horas.

Lugar: Edificio 14 Aula de informática 1

Objetivos de la acción formativa:

- Visión general

El aprendizaje automático y la minería de datos son herramientas que nos ayudan a extraer información valiosa que está oculta en grandes cantidades de datos. Con estas técnicas podemos identificar patrones y relaciones que no son evidentes a simple vista, permitiéndonos tomar mejores decisiones basadas en evidencia.

- Relevancia y aplicaciones

El Machine Learning representa una herramienta innovadora con amplio potencial de aplicación:

- Marketing y finanzas: Segmentación de clientes, detección de fraudes, análisis de riesgo crediticio.
- Medicina y ciencias de la salud: Análisis de datos clínicos, identificación de factores de riesgo, estudios epidemiológicos.
- Ciencias del deporte: Análisis de rendimiento deportivo, predicción de lesiones, optimización de entrenamientos.

- Estudios medioambientales: Predicción de calidad del aire, análisis de patrones climáticos, evaluación de impacto ambiental.
- Economía y políticas públicas: Predicción de indicadores económicos, evaluación de intervenciones sociales.

- **Objetivos específicos**

1. Conceptos fundamentales: Introducción a los principios del Machine Learning.
2. Técnicas de análisis: Dominio de métodos de minería de datos y procesamiento.
3. Modelos predictivos: Aplicación de técnicas típicas de predicción.
4. Aplicación práctica: Implementación mediante STATA en casos reales.

- **Público objetivo**

Profesionales/Investigadores interesados en técnicas actuales de análisis y procesamiento de datos que buscan desarrollar habilidades prácticas para aplicar en el mundo real. Este curso está diseñado para ser accesible a principiantes y mejorar las competencias en el trabajo con datos.

Programa de contenidos:

1. Conceptos fundamentales en el aprendizaje automático y minería de datos

Objetivo: Establecer las bases conceptuales del Machine Learning

Contenidos clase:

- Funciones de pérdida: Métricas de evaluación de modelos
- Sobreajuste (Overfitting): Identificación y prevención
- Matriz de confusión: Interpretación de resultados de clasificación
- El problema de la dimensión: Curse of dimensionality
- Validación cruzada: Técnicas de evaluación robusta

2. Aprendizaje automático supervisado

Objetivo: Dominar técnicas de predicción y clasificación.

2.1 Regresión paramétrica

- Fundamentos de regresión supervisada
- Técnicas de penalización: LASSO, Elastic Net, Ridge Regression
- Aplicaciones prácticas con STATA

2.2 Regresión no paramétrica

- Árboles de decisión: Construcción e interpretación
- Bosques aleatorios (Random Forests): Bagging y técnicas de combinación

2.3 Técnicas de clasificación

- Diferenciación entre clasificación y predicción de valores
- Algoritmos de regularización aplicados a clasificación
- Evaluación y comparación de técnicas

3. Aprendizaje automático no supervisado

Objetivo: Familiarización con técnicas de descubrimiento de patrones

3.1 Técnicas de Agrupamiento (Clustering)

- Clustering jerárquico: Dendrogramas e interpretación
- Clustering secuencial: Métodos iterativos
- K-means: Algoritmo y aplicaciones
- Evaluación de similitud: Métricas de distancia

Enfoque práctico: Distinción entre tipos de clustering, análisis de problemas de similitud y evaluación de resultados.

Metodología de evaluación:

Evaluación diseñada para determinar competencias adquiridas tras el proceso de aprendizaje, correlacionando criterios con objetivos establecidos.

- Instrumentos de evaluación

- Ejercicios prácticos: implementación de técnicas con datos reales

- Criterios de evaluación

- Habilidades procedimentales: aplicación correcta de técnicas
- Destrezas de ejecución: implementación práctica eficiente

Resultados de aprendizaje:

- Competencias fundamentales

Al finalizar este curso, las personas participantes dominarán los conceptos básicos del Machine Learning y las técnicas necesarias para evaluar el rendimiento de algoritmos y modelos.

- Habilidades específicas adquiridas

1. Distinción de paradigmas:
 - o Diferenciación entre aprendizaje supervisado y no supervisado
 - o Identificación del tipo de problema según contexto
2. Aplicación práctica supervisada:
 - o Selección de técnicas óptimas para predicción
 - o Implementación en entornos de programación reales
 - o Interpretación de resultados
3. Dominio de técnicas no supervisadas:
 - o Aplicación de métodos de clustering
 - o Análisis de patrones ocultos en datos
 - o Evaluación de agrupamientos
4. Evaluación y comparación:
 - o Análisis crítico de resultados obtenidos
 - o Comparación sistemática entre técnicas
 - o Selección justificada de metodologías

5. Competencia tecnológica:

- o Utilización de STATA para análisis de datos
- o Generación de modelos de aprendizaje automático
- o Interpretación de outputs estadísticos

Criterios para acreditar la formación realizada: Superación de la evaluación.

Plazo de presentación de solicitudes: Hasta el día **20 de junio de 2025**.

Persona formadora: D.^a Yolanda Rebollo Sanz

ORCID: 0000-0003-1876-7093

- Formación Académica:
Doctora en Economía - Universidad Pablo de Olavide
Máster en Análisis Económico - Universidad Pompeu Fabra e Instituto de Estudios Fiscales
- Experiencia Profesional:
Profesora Titular - Departamento de Economía, Universidad Pablo de Olavide de Sevilla
Investigadora Principal - Grupos especializados en métodos cuantitativos y evaluación de políticas públicas
Experiencia docente: Más de 25 años (desde 1998)
- Especialización Investigadora:
Métodos cuantitativos avanzados: Microeconometría Avanzada e Inferencia Causal
Publicaciones: Más de 20 artículos de investigación
Aplicación: Comprensión de la realidad económica mediante análisis estadístico avanzado