

# Machine learning, el estado del arte

Seminario de 2 días - 14h

Ref.: MLE - Precio 2025: 1 490€ sin IVA

Este seminario detalla las cuestiones relacionadas con el tratamiento de datos mediante Inteligencia Artificial, y algoritmos de Aprendizaje Automático en particular. Muestra a los responsables de la toma de decisiones los principales algoritmos en este campo, soluciones prácticas y el enfoque de proyecto que debe aplicarse en función del caso de uso en la empresa.

## OBJETIVOS PEDAGÓGICOS

Al término de la formación, el alumno podrá:

Posicionamiento del aprendizaje automático en la cadena de procesamiento de datos

Distinguir entre las competencias requeridas y los perfiles a contratar

Identificar las claves del éxito de un proyecto de aprendizaje automático

Comprender los conceptos de aprendizaje automático y la evolución de Big Data hacia el aprendizaje automático

Comprender los retos que plantea el uso del aprendizaje automático, incluidos los beneficios esperados y ejemplos de uso.

## MÉTODOS PEDAGÓGICOS

Ilustrado con estudios de casos.

Presentación de los principales casos de uso por sector empresarial (automoción, industria, bienes de consumo, finanzas, sanidad, energía, agricultura, transporte, telecomunicaciones, etc.).

## PROGRAMA

última actualización: 02/2024

### 1) Historia del aprendizaje automático y el contexto de Big Data

- Ponga en perspectiva los conceptos de Inteligencia Artificial, aprendizaje automático, etc.
- El vínculo con las matemáticas, la estadística (inferencial), la minería de datos y la ciencia de datos.
- Pasar del análisis descriptivo al predictivo y luego al prescriptivo.
- Aplicaciones de aprendizaje automático (motores de búsqueda, detección de spam, lectura de cheques).
- Tipología de algoritmos de Dominique CARDON.
- La comunidad de la ciencia de datos y los retos de Kaggle (por ejemplo, Netflix).

*Estudios de caso : Estudios de aplicaciones concretas del aprendizaje automático (motores de búsqueda, detección de spam, lectura de cheques).*

### 2) Datos disponibles: recogida y preparación

- Datos estructurados, semiestructurados y no estructurados.
- Naturaleza estadística de los datos (cualitativos o cuantitativos).
- Objetos conectados (IoT) y streaming.
- Oportunidades y límites de los datos abiertos.
- Identificación de correlaciones, el problema de la multicolinealidad.
- Reducción de dimensiones mediante análisis de componentes principales.
- Detección y corrección de valores atípicos.
- ETL (extracción, transformación y carga).

## PARTICIPANTES

Directores de empresa (CEO, COO, CFO, SG, HRD, etc.), CIO, CDO, responsables de TI, consultores, responsables de proyectos de Big Data.

## REQUISITOS PREVIOS

Se recomiendan conocimientos generales de informática y de probabilidad y estadística.

## COMPETENCIAS DEL FORMADOR

Los expertos que imparten la formación son especialistas en las materias tratadas. Han sido validados por nuestros equipos pedagógicos, tanto en el plano de los conocimientos profesionales como en el de la pedagogía, para cada curso que imparten. Cuentan al menos con entre cinco y diez años de experiencia en su área y ocupan o han ocupado puestos de responsabilidad en empresas.

## MODALIDADES DE EVALUACIÓN

El formador evalúa los progresos pedagógicos del participante a lo largo de toda la formación mediante preguntas de opción múltiple, escenificaciones de situaciones, trabajos prácticos, etc. El participante también completará una prueba de posicionamiento previo y posterior para validar las competencias adquiridas.

## MEDIOS PEDAGÓGICOS Y TÉCNICOS

- Los medios pedagógicos y los métodos de enseñanza utilizados son principalmente: ayudas audiovisuales, documentación y soporte de cursos, ejercicios prácticos de aplicación y ejercicios corregidos para los cursillos prácticos, estudios de casos o presentación de casos reales para los seminarios de formación.
- Al final de cada cursillo o seminario, ORSYS facilita a los participantes un cuestionario de evaluación del curso que analizarán luego nuestros equipos pedagógicos.
- Al final de la formación se entrega una hoja de presencia por cada media jornada de presencia, así como un certificado de fin de formación si el alumno ha asistido a la totalidad de la sesión.

## MODALIDADES Y PLAZOS DE ACCESO

La inscripción debe estar finalizada 24 horas antes del inicio de la formación.

## ACCESIBILIDAD DE LAS PERSONAS CON DISCAPACIDAD

¿Tiene alguna necesidad específica de accesibilidad? Póngase en contacto con la Sra. FOSSE, interlocutora sobre discapacidad, en la siguiente dirección [psh-accueil@orsys.fr](mailto:psh-accueil@orsys.fr) para estudiar de la mejor forma posible su solicitud y su viabilidad.

- Web scraping.

*Demostración* : Demostración ETL (Extract Transform Load). Recogida de datos web.

### 3) Herramientas de mercado para el tratamiento de datos y el aprendizaje automático

- El software tradicional (SAS, SPSS, Stata, etc.) y su apertura al código abierto.
- Elija entre los dos líderes del código abierto: Python y R.
- Plataformas en la nube (Azure, AWS, Google Cloud Platform) y soluciones SaaS (IBM Watson, Dataiku).
- Nuevos empleos en empresas: ingeniero de datos, científico de datos, analista de datos, etc.
- Adecuar las competencias a las distintas herramientas.
- API en línea (IBM Watson, Microsoft Cortana Intelligence, etc.).
- Chatbots (agentes conversacionales).

*Demostración* : Demostración de un chatbot (agente conversacional) y Azure Machine Learning.

### 4) Los distintos tipos de aprendizaje automático

- Aprendizaje supervisado: repetir un ejemplo.
- Aprendizaje no supervisado: descubrir los datos.
- Aprendizaje (automático) en línea frente a técnicas por lotes.
- Aprendizaje por refuerzo: optimización de una recompensa.
- Otros tipos de aprendizaje (transferencia, secuencial, activo, etc.).
- Ilustraciones (motores de recomendación, etc.).

*Demostración* : Demostraciones de los distintos tipos de aprendizaje automático posibles.

### 5) Algoritmos de aprendizaje automático

- Regresión lineal simple y múltiple. Limitaciones de los enfoques lineales.
- Regresión polinómica (LASSO). Series temporales.
- Regresión logística y aplicaciones de puntuación.
- Clasificación jerárquica y no jerárquica (KMeans).
- Clasificación mediante árboles de decisión o el método Naïve Bayes.
- Random Forest (desarrollo de árboles de decisión).
- Refuerzo gradual. Redes neuronales. Máquina de soporte vectorial.
- Deep Learning: ejemplos y razones del éxito actual.
- Minería de textos: análisis de corpus de datos textuales.

*Demostración* : Demostración de los distintos algoritmos básicos en R o Python.

### 6) Procedimiento de entrenamiento y evaluación del algoritmo

- Separación de conjuntos de datos: formación, prueba y validación.
- Técnicas Bootstrap (bagging).
- Ejemplo de validación cruzada.
- Definición de una métrica de rendimiento.
- Descenso de gradiente estocástico (minimización métrica).
- Curvas ROC y de elevación para evaluar y comparar algoritmos.
- Matriz de confusión: falsos positivos y falsos negativos.

*Demostración* : Demostrar la elección del mejor algoritmo.

### 7) Elaboración de un algoritmo de aprendizaje automático

- Descripción de una plataforma de Big Data.
- Cómo funcionan los PLC.
- Del desarrollo a la producción.
- Estrategia de mantenimiento correctivo y evolutivo.
- Evaluación de los costes de explotación de la producción.

*Demostración* : Demostración de las API de geolocalización y análisis de sentimientos.

### 8) Aspectos éticos y jurídicos de la Inteligencia Artificial

- Competencias de la CNIL y evolución futura.

- Derecho de acceso a los datos personales.
- La cuestión de la propiedad intelectual de los algoritmos.
- Nuevas funciones en la empresa: Director de Datos y Responsable de Protección de Datos.
- La cuestión de la imparcialidad de los algoritmos.
- Cuidado con el sesgo de confirmación.
- Sectores y profesiones afectados por la automatización.

*Reflexión colectiva : Lluvia de ideas para identificar las claves del éxito.*

## FECHAS

---

Contacto