



ITINERARIO ONLINE



300 horas  
5 Bloques



# MACHINE LEARNING Y CIENCIA DE DATOS CON PYTHON Y R

## INTRODUCCIÓN

---

La diferencia entre la automatización y la inteligencia artificial es que, la primera, sigue unas reglas y pasos preconfigurados, mientras que la segunda va mucho más allá y es capaz de realizar tareas y tomar decisiones sin definición previa del proceso, mejorando además con el tiempo.

La inteligencia artificial y, en particular, el aprendizaje automático (*Machine Learning*) están contribuyendo de forma importante a la revolución industrial 4.0, haciendo posible que los datos de la empresa se puedan explotar para innovar, crecer, ahorrar costes, mejorar la calidad, generar nuevos productos, etc.

Pero existe una gran escasez de personal cualificado que está limitando su adopción. Además, esta falta de técnicos cualificados está empeorando debido a la adopción imparable de estas tecnologías en todos los sectores, que aumenta la competencia por el talento.

La mejor estrategia para no perder el tren del futuro es sacar partido a la gente que ya tienes, que conocen la empresa, sus procesos y sus productos/servicios, y formarlos para explotar este conocimiento y los datos de la empresa con las técnicas punteras de inteligencia artificial.

El OBJETIVO de esta formación es que, partiendo de cero, los profesionales que la cursen puedan comprender, aplicar y desarrollar todas las técnicas de *Machine Learning* importantes, incluyendo las más avanzadas (*Deep Learning, ensembles...*), de modo que, en un tiempo razonable y con un esfuerzo controlado, incorporen los conceptos fundamentales subyacentes y la práctica necesaria para sacarles partido según las necesidades de su empresa.

## DESTINATARIOS

---



Si bien el Programa es de aplicación en todos los sectores del ámbito industrial, la presente convocatoria se dirige exclusivamente a profesionales del Sector Gallego de Automoción.

Para realizarlo con aprovechamiento, solo es necesario contar con un nivel de matemáticas de bachillerato (aritmética y geometría básicas). El perfil que más provecho le puede sacar es el de una persona con cierto interés por los números y por el pensamiento abstracto y que desarrolle, actual o potencialmente, funciones relacionadas con la Ciencia de Datos.

## PLAZO Y REQUISITOS PARA SUPERAR LA FORMACIÓN

---

En el plazo que se indique en la plataforma (alrededor de 10 semanas) se deberán cumplir los siguientes requisitos para superar el curso:

- Visualización del 75% del contenido.
- Conectarse al menos al 30% de las horas del curso.
- Superar el 100% de los controles periódicos.

## METODOLOGÍA

- Formación impartida **íntegramente online** que incluye: teoría, vídeos prácticos, ejemplos de código, referencias cruzadas, actividades prácticas y evaluaciones.
  - Los alumnos tienen **acceso 24x7** al material desde el primer día, lo que permite realizar la formación desde cualquier lugar y a cualquier hora.
  - **Tutorías: contacto directo con los propios autores** del curso por **mensajería interna privada** en la propia plataforma. Estos responderán sus dudas acerca en menos de 24 horas laborables.
  - **Foros privados** en cada parte, para poner en común dudas e inquietudes.
  - Este itinerario formativo, equivalente al Máster del mismo título, está dividido en **5 acciones formativas (bloques)**, que se deben cursar en secuencia y que forman un todo único. **Para superar la formación** es necesario realizarlos todos, incluida la **actividad practica final**.
- La duración total de la formación es de **300 horas**. La **dedicación estimada es de entre 12 y 15 horas semanales**.



## AUTORES / TUTORES

### DAVID CHARTE

Ingeniero informático, licenciado en matemáticas y doctor en Inteligencia Artificial por la Universidad de Granada, especializado en Deep Learning. Ha participado en multitud de proyectos e investigaciones para varios sectores en el campo de la ciencia de datos y el ML para empresas privadas y públicas (optimización de procesos industriales, análisis de estructuras microscópicas o asistencia al diagnóstico con métodos de aprendizaje automático ..)

### FRANCISCO CHARTE

Ingeniero informático y Doctor en tecnologías de la información y la Comunicación. Reconocido autor de numerosos libros y centenares de artículos. Docente del departamento de Inteligencia Artificial de la Universidad de Jaén, entre otras instituciones. Amplia experiencia en proyectos de ciencia de datos y ML como detección de anomalías en piezas a partir de fotografías, desarrollo de sistemas inteligentes de procesamiento de imágenes o detección de amenazas en escenarios de conflicto bélico entre muchos otros.

## MEDIOS NECESARIOS



Para seguir la formación es **necesario contar con:**

- Ordenador con Windows, Linux o macOS, con un mínimo de 8GB de memoria RAM.
- Sistema de audio para escuchar los videos.
- Conexión a internet de banda ancha.

Para acelerar los cálculos, es recomendable que el PC tenga menos de 6 años de antigüedad, que sea potente (procesador reciente, deseables 16GB RAM) y que cuente con una tarjeta gráfica dedicada (pero no es indispensable).

Todo el software utilizado es Open Source y gratuito.

# ITINERARIO FORMATIVO: CONTENIDO DE LOS BLOQUES

## Bloque 1: Fundamentos de Machine Learning y Ciencia de Datos

60h

### Fundamentos del ML y la Ciencia de Datos

- MÓDULO 1: Introducción al Machine Learning: conceptos, técnicas y herramientas

### Fundamentos de análisis exploratorio y preparación de datos

- MÓDULO 2: Carga y exportación de datos
- MÓDULO 3: Análisis Exploratorio

## Bloque 2: Análisis Predictivo Fundamental

60h

### Adecuación de los datos

- MÓDULO 1: Limpieza, valores anómalos, redundancia, valores perdidos, datos sintéticos

### Análisis predictivo fundamental

- MÓDULO 2: Predicción de valores numéricos
- MÓDULO 3: Clasificación binaria

## Bloque 3: Análisis Predictivo Avanzado y Análisis Descriptivo

60h

- MÓDULO 1: Clasificación multiclase
- MÓDULO 2: Agrupamiento automático de datos

## Bloque 4: Técnicas avanzadas de ML

60h

- MÓDULO 1: Sistemas de Recomendación
- MÓDULO 2: Ensembles. Uso de múltiples modelos para mejorar los resultados
- MÓDULO 3: Predicción de series temporales con ARIMA
- MÓDULO 4: Técnicas evolutivas y algoritmos genéticos

## Bloque 5: Deep Learning

60h

- MÓDULO 1: Eliminación de ruido en imágenes con autoencoders
- MÓDULO 2: Detección de objetos en imágenes
- MÓDULO 3: Modelado generativo de datos
- MÓDULO 4: Aplicación práctica a un proyecto acotado