

Fundamentos de la ciencia de datos

Curso de síntesis de 2 días - 14h

Ref.: DTX - Precio 2026: 1 420€ sin IVA

La ciencia de datos, una cuestión estratégica de primer orden para las organizaciones, utiliza herramientas matemáticas para revelar el comportamiento de los datos y analizar los acontecimientos que describen. Este curso abarca los fundamentos de la ciencia de datos y enseña a analizarlos.

OBJETIVOS PEDAGÓGICOS

Al término de la formación, el alumno podrá:

Comprender los principios básicos de la ciencia de datos y cómo organizar el proceso

Comprender la aplicación de la ciencia de datos para resolver cuestiones y sus limitaciones

Desarrollar la capacidad de analizar e interpretar cifras mediante la representación gráfica.

Comprender cómo utilizar las herramientas de la ciencia de datos y desarrollar modelos con fines profesionales.

Abrirse al reto de explotar los datos en un contexto competitivo y de mejora continua

Comprender la organización y la infraestructura de los servicios y proyectos de ciencia de datos

MÉTODOS PEDAGÓGICOS

Durante este curso resumido, el formador realiza demostraciones que cada participante reproduce para poner en práctica los conceptos principales.

Práctica guiada de los fundamentos mediante ejercicios. Los MCQ y los cuadros sinópticos le ayudarán a situarse.

PROGRAMA

última actualización: 02/2024

1) ¿Qué es la ciencia de datos?

- Conceptos básicos: big data, lago de datos, minería de datos, inteligencia artificial, aprendizaje automático y profundo, minería de textos.
- Los nuevos retos: la aparición y multiplicación de nuevas fuentes de datos.
- Hay que tener en cuenta la heterogeneidad de los datos, los flujos en tiempo real y la explosión de los volúmenes de datos.
- El ecosistema tecnológico de Big Data.
- Desmitificando el mundo de la ciencia de datos: análisis descriptivo, predictivo y prescriptivo.
- El trabajo, las herramientas y los métodos del científico de datos.
- Introducción al aprendizaje automático, análisis supervisado y análisis no supervisado.
- Nociones de sobreaprendizaje y subaprendizaje.

Demostración : Casos de uso de la ciencia de datos en una cadena de valor empresarial (comportamiento del cliente, oferta de productos, etc.).

2) Métodos y modelos de la ciencia de datos

- Recogida, preparación y exploración de datos.
- La importancia del enfoque de la calidad de los datos (limpieza, transformación, enriquecimiento).
- Definición de métrica.
- Métodos estadísticos básicos.

PARTICIPANTES

Gestores de SI, gestores de proyectos de análisis de datos, gestores de investigación estadística.

REQUISITOS PREVIOS

No se requieren conocimientos especiales.

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 para estudiar de la mejor forma posible su solicitud y su viabilidad.

- Las principales clases de algoritmos supervisados: árboles de decisión, vecinos más próximos a K, regresión, Naive Bayes.
- Las principales clases de algoritmos no supervisados: clustering, PCA, CAH, redes neuronales.
- Minería de textos y otras familias de algoritmos.

Intercambios : Análisis sencillos con R o Python para ilustrar las técnicas de análisis supervisado (regresión y clasificación) y no supervisado (agrupación, segmentación y detección de anomalías).

3) Representación gráfica y recuperación de datos

- Lenguajes de análisis estadístico R y Python.
- Sus entornos de desarrollo (R-Studio, Anaconda, PyCharm) y bibliotecas (Pandas, aprendizaje automático).
- Herramientas DataViz (Power BI, Qlik, Tableau, etc.).
- Modelización de datos: representación de procesos, flujos, controles y condiciones.
- Comunicación de resultados mediante la narración de datos: organización de elementos visuales (diagramas, clasificaciones, mapas).
- Modelización de datos: herramientas (Orange, Power BI).
- Comunicación de resultados mediante la narración de datos: transmitir la importancia de los resultados.

Intercambios : Ejercicios de exploración gráfica de datos, análisis de la posición y extensión de los datos (nubes, histogramas, etc.).

4) Modelización de un problema de ciencia de datos

- Resumen del proceso.
- Estudio de caso 1: relaciones con los clientes en el sector de los seguros.
- Dirigir las campañas de marketing. Comprender las causas de la pérdida de clientes. ¿Qué productos para qué clientes?
- Caso práctico 2: detección del fraude.
- Comparar la investigación mediante estadísticas tradicionales y minería de datos.
- Detección por método supervisado. Detección no supervisada.
- Análisis de dos casos empresariales, las relaciones con los clientes y la detección de fraudes, por ejemplo, pero también son posibles otros.

Estudios de caso : Aplicación práctica del método storytelling a casos empresariales.

FECHAS

Contacto