

Machine Learning



Modalidad Online en Streaming

20 horas



#formaciónCEL

BONIFICABLE FUNDAE



Centro Español de
Logística

Descripción del curso

El programa realiza un viaje por el campo del Machine Learning. Partiendo de una perspectiva histórica y funcional de cómo las actuales fuerzas del cambio analítico (Big Data, la nueva Algoritmia, y el actual Procesamiento de Información) producen la actual transformación digital, se describirán múltiples casos de uso cognitivos endiferentes industrias que inspiren al alumno en cuanto qué puede esperar de esta tecnología.

Se describirá de una manera práctica todas las fases de la metodología open source CRISP-DM (Cross-Industry Process for Data Mining) de desarrollo de modelos de Ciencia de Datos. La metodología cubre el ciclo completo del desarrollo de un modelo: desde el enunciado del problema de negocio a resolver y posterior especificación del modelo, pasando por la descripción del dataset y su análisis de calidad, la modelización y su validación, y finalmente, la descripción de resultados y el despliegue del modelo. Continuaremos con el detalle del desarrollo de un modelo.

Comenzaremos conociendo algunos conceptos básicos estadísticos para describir datos, y emplearemos otro instrumental para comprender el dataset. Analizaremos la calidad de los datos. A continuación, depuraremos el dataset y, transformaremos los datos, normalizándolos o creando nuevas variables que mejoren el rendimiento del modelo.

Con un dataset adecuado para la modelización, comenzaremos a describir las técnicas de aprendizaje supervisado para los problemas de estimación y clasificación, y también arrancaremos los métodos no supervisados con los algoritmos de clustering.

El programa finalizará introduciendo las redes neuronales y su aplicación en el Deep Learning. Concluiremos con métodos avanzados analíticos como el bootstrapping, boosting, bagging, métodos de conjunto, modelos apilados, máquinas de comité/mezclas de expertos.

Contenidos

Sesión 1. Introducción: Datos y Ciencia

- El alumno obtiene una comprensión del valor del dato hoy día, qué es el Big Data y qué tipo de utilización se le puede dar, cuáles son las distintas capas de la analítica de negocio, tipos de de datos y utilidades, así como de una arquitectura completa de Big Data & Business Analytics entendiendo el tipo de procesamiento a realizar en los datos
- Recorreremos la metodología de desarrollo de modelos predictivos open source más empleada en el mundo, CRISP-DM
 - Comprensión de Negocio
 - Comprensión de Datos
 - Preparación de Datos
 - Modelización
 - Validación
 - Despliegue
- Caso Práctico: Ejercicios de Descripción y Preparación de Datos con datasets de Fabricantes de Productos/Retailers/Operadores Logísticos

Sesión 2. Algoritmia de Machine Learning I

- Recorreremos la metodología de nuevo con la herramienta de Ciencia de Datos, pasando por cada una de las actividades especificación del modelo, carga de datos, descripción de datos, análisis de calidad, modelización, validación y descripción de resultados
- Explicaremos la base matemática de la algoritmia más empleada en los principales procesos de negocio Algoritmos de Aprendizaje Supervisado que incluyen métodos de estimación como el algoritmo de regresión o redes neuronales o métodos de clasificación como regresión logística o árboles de decisión
- Se comprenderán y realizarán múltiples ejercicios sencillos con los estudiantes durante sesión
- Caso práctico: aplicación de la algoritmia a problemas comerciales reales, como el abandono de clientes, modelos de riesgos, o el mantenimiento predictivo de máquinas

Contenidos

Sesión 3. Algoritmia de Machine Learning II

- Ampliamos la base de la algoritmia con Algoritmos de Aprendizaje No Supervisado que incluyen los métodos de clustering o las reglas de asociación
- Añadimos a la lista de algoritmos otro muy interesante para encontrar gemelos en la industria del automóvil: Vecino K más cercano (kNN)
- Se realizarán múltiples ejercicios sencillos que recorran la metodología
- Caso Práctico: ejercicios de segmentación estratégica de clientes con técnicas de clustering, identificación y predicción de gemelos, y competición de las distintas técnicas en diferentes problemas de Ciencia de Datos aplicados a diferentes procesos de negocio

Sesión 4. Redes Neuronales y Modelos de Conjunto

- Redes neuronales, topologías, tipos y aplicaciones. Ejemplos prácticos
- Modelos de Conjunto
 - Validación cruzada
 - Bagging - Bosques aleatorios
 - Boosting - XGBoost
 - Modelos Apilados y Mixtos
- Máquinas de Comité y Mezclas de Expertos
- Casos práctico: ejercicios de estimación de demanda con redes neuronales, y desarrollo de un modelo apilado de abandono de clientes

Contenidos

Wolfram Rozas es economista cuantitativo y doctorando en Tecnologías Industriales. Es experto en el empleo de todas las tecnologías exponenciales, Big Data & Business Analytics, Inteligencia Artificial, Cloud, Internet of Things, Blockchain y Quantum Computing, a la consecución de objetivos estratégicos de negocio

Experiencia como consultor: Ha dedicado su carrera profesional de más de 25 años de experiencia en empresas como PWC o IBM a dirigir proyectos de Análisis o Ciencia de Datos en los sectores Financiero, Telecomunicaciones, Distribución, Turismo y Transporte, Gran Consumo, Energía y Utilities, Químicas y Petróleo. En todos estos sectores industriales ha construido soluciones de negocio para mejora de procesos relacionados con las funciones de Marketing & Comercial, Finanzas y Riesgos, Cadena de Suministro, Operaciones y Recursos Humanos.

Experiencia académica: Ha desarrollado su carrera docente desde 2003, colaborando como experto en Análisis de Datos en Escuelas de Negocio y Universidades como Escuela de Organización Industrial, Instituto de Empresa, U. Pablo de Olavide, U. Carlos III, UNIR, U. Salamanca, CUNEF, SPEGC, AFI, UNED, U. La Laguna, U. Politécnica Valencia, U. Alicante, U. Navarra, Deusto, U. País Vasco, FUNDESEM, Next IBS, entre otras. Recibió el Premio a una Trayectoria Académica de manos de la Ministra de Industria, Comercio y Turismo, Reyes Maroto, el 16 de Julio de 2019. Nombrado por el INE en Julio 2020, miembro del Grupo de Trabajo sobre el papel de la Estadística Oficial en la Administración y Gestión de Datos (Data Stewardship) del Consejo Superior de Estadística.

Inscripción

Formulario de inscripción:
<https://cel-logistica.org/curso/machine-learning/>

Más información

Centro Español de Logística
formacion@cel-logistica.org
Teléfono 91 781 14 70