



Catálogo de Especialidades Formativas

PROGRAMA FORMATIVO

DATA ENGINEERING & INTELIGENCIA ARTIFICIAL

Marzo 2023

IDENTIFICACIÓN DE LA ESPECIALIDAD Y PARÁMETROS DEL CONTEXTO FORMATIVO

Denominación de la especialidad:	DATA ENGINEERING & INTELIGENCIA ARTIFICIAL
Familia Profesional:	INFORMÁTICA Y COMUNICACIONES
Área Profesional:	DESARROLLO
Código:	IFCD0077
Nivel de cualificación profesional:	3

Objetivo general

Formar a Técnicos en Ingeniería de datos y en proyectos de Inteligencia Artificial

Relación de módulos de formación

Módulo 1	Herramientas prácticas para la Inteligencia Artificial e Introducción al Machine Learning	60 horas
Módulo 2	Introducción al Deep Learning y Visión Artificial	80 horas
Módulo 3	Procesamiento del Lenguaje Natural y Modelos Temporales Profundos	50 horas
Módulo 4	Modelos Generativos y aprendizaje por refuerzo en Machine Learning	60 horas
Módulo 5	MLOps: Implantación y creación de producto basado en IA e Inteligencia artificial explicable (XIA)	60 horas
Módulo 6	Introducción a Python para Data Science y Big Data. Gestión de datos en Python	70 horas
Módulo 7	Necesidad y principios de las arquitecturas Big Data y Fundamentos de Apache Spark con Python	80 horas
Módulo 8	Spark Streaming y Analítica avanzada y Machine Learning con Spark MLLIB	50 horas
Módulo 9	Gestión de Proyectos de Business Intelligence	30 horas
Módulo 10	Power BI como Herramienta de Inteligencia Empresarial y ventaja competitiva	30 horas
Módulo 11	Tableau en Data Science	80 horas
Módulo 12	Power BI y Qlik Sense Business Intelligence	50 horas
Módulo 13	Arquitecturas Transaccionales Avanzadas	40 horas
Módulo 14	Microservicios y Arquitecturas Basadas en Eventos	40 horas
Módulo 15	Arquitecturas de Datos Empresariales	40 horas
Módulo 16	Adquisición, Extracción y Manipulación de Datos	40 horas
Módulo 17	Procesamiento Big Data y Streaming	40 horas
Módulo 18	Gobierno del Dato	40 horas
Módulo 19	Explotación y Análisis de Datos en IA	20 horas

Modalidad de impartición

Presencial

Duración de la formación

Duración total

960 horas

Requisitos de acceso del alumnado

Acreditaciones / titulaciones	Cumplir como mínimo alguno de los siguientes requisitos: Título de Técnico Superior (FP Grado superior) o equivalente Título de Grado o equivalente
Experiencia profesional	No se requiere
Otros	Inglés Nivel B1 mínimo. En relación con la titulación requerida de Título de Técnico Superior (FP Grado Superior) o necesario que dicha titulación corresponda a la familia profesional de Informática y Comunicaciones. Los títulos de grado o licenciaturas deben ser en carreras técnicas.

Prescripciones de formadores y tutores

Acreditación requerida	Cumplir como mínimo alguno de los siguientes requisitos: Licenciado, Ingeniero, Arquitecto o el Título de Grado correspondiente otros títulos equivalentes preferiblemente en la rama de conocimiento de Arquitectura e Ingeniería. Diplomado, Ingeniero Técnico, Arquitecto Técnico o el Título de Grado correspondiente u otros títulos equivalentes preferiblemente en la rama de conocimiento de Arquitectura e Ingeniería. FP de grado superior relacionados de la familia profesional de Informática y Comunicaciones.
Experiencia profesional mínima requerida	Experiencia mínima en dirección de proyectos de 5 años.
Competencia docente	Experiencia mínima de 2 años como docentes en áreas IT. Haber impartido formación bajo metodología Learning by Doing.
Otros	Conocimientos en metodologías Ágile. Manejo de herramientas para la dirección de proyectos IT (Kanban, dirección de proyectos, etc...)

Requisitos mínimos de espacios, instalaciones y equipamientos

Espacios formativos	Superficie m ² para 15 participantes	Incremento Superficie/ participante (Máximo 30 participantes)
Aula de gestión	45.0 m ²	2.4 m ² / participante

Espacio formativo	Equipamiento
Aula de gestión	- Mesa y silla para el formador - Mesas y sillas para el alumnado - Material de aula - Pizarra - PC instalado en red con posibilidad de impresión de documentos, cañón con proyección e Internet para el formador - PCs instalados en red e Internet con posibilidad de impresión para los participantes - Software específico para el aprendizaje de cada acción formativa. • Licencia de aprendizaje Cloud (Azure, Google y AWS) • BBDD

La superficie de los espacios e instalaciones estarán en función de su tipología y del número de participantes. Tendrán como mínimo los metros cuadrados que se indican para 15 participantes y el equipamiento suficiente para los mismos.

En el caso de que aumente el número de participantes, hasta un máximo de 30, la superficie de las aulas se incrementará proporcionalmente (según se indica en la tabla en lo relativo a m²/participante) y el equipamiento estará en consonancia con dicho aumento.

No debe interpretarse que los diversos espacios formativos identificados deban diferenciarse necesariamente mediante cerramientos.

Las instalaciones y equipamientos deberán cumplir con la normativa industrial e higiénico-sanitaria correspondiente y responderán a medidas de accesibilidad y seguridad de los participantes.

En el caso de que la formación se dirija a personas con discapacidad se realizarán las adaptaciones y los ajustes razonables para asegurar su participación en condiciones de igualdad.

Ocupaciones y puestos de trabajo relacionados

31391085 PROGRAMADORES DE MÁQUINAS CON CONTROL NUMÉRICO, EN GENERAL

31391094 PROGRAMADORES-CONTROLADORES DE ROBOTS INDUSTRIALES

38201017 PROGRAMADORES DE APLICACIONES INFORMÁTICAS

27131024 ANALISTAS-PROGRAMADORES WEB Y MULTIMEDIA

Requisitos oficiales de las entidades o centros de formación

Estar inscrito en el Registro de entidades de formación (Servicios Públicos de Empleo).

MÓDULO DE FORMACIÓN 1: Herramientas prácticas para la Inteligencia Artificial e Introducción al Machine Learning**OBJETIVO**

Aprender las bases científicas y las herramientas tecnológicas necesarias en la construcción de modelos de IA y explorar las técnicas de Machine Learning para desarrollar soluciones de analítica avanzada, así como conocer Python para la implementación de procesos de aprendizaje y la evaluación de modelos desde una perspectiva práctica.

DURACIÓN TOTAL:

60 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Conocimiento del lenguaje Python para el tratamiento de datos y el desarrollo de modelos.
- Manejo de los fundamentos matemáticos y estadísticos necesarios para la IA.
- El dominio de la manipulación, exploración y limpieza de datos.
- Conocimiento del análisis descriptivo y estadístico de los datos.
- Conocimiento del análisis exploratorio y geoespacial de la oferta de alojamientos turísticos en determinados programas sobre las ciudades españolas.
- Conocer el análisis de preferencias de usuarios sobre lista de reproducciones de programas musicales.
- Conocimiento de los principios de Machine learning o aprendizaje automático, Fuzzy logic o lógica difusa, el concepto de Vida artificial, Sistemas expertos, Data Mining o minería de datos y de Redes Bayesianas.
- Conocimiento inicial en áreas relacionadas con la Inteligencia Artificial: matemáticas básicas, fundamentos de informática, estadística, probabilidad, vectores, álgebra lineal y cálculo.
- Casos de uso para la visualización de desarrollos en coches automatizados, robótica de alta gama, gestión del tráfico basada en IA, mantenimiento de la red inteligente etc.
- Conocimiento de los principios del Machine Learning.
- Comprensión del aprendizaje supervisado, aplicaciones de clasificación y regresión.
- Comprensión del aprendizaje no supervisado, aplicaciones de segmentación y clustering.
- Desarrollo de técnicas de aprendizaje y evaluación, interpretación de principales métricas de calidad.
- Creación de modelos Avanzados de Machine Learning y técnicas específicas de ajuste y mejora.
- Comprensión del algoritmo de predicción de rotación de clientes de un operador de telecomunicaciones.
- El trabajo con la segmentación de clientes de un portal de e-commerce a partir del comportamiento en la web.
- Conocimiento de las bases para desarrollar, hacer pruebas y aplicar algoritmos de análisis predictivo sobre diferentes tipos de datos para predecir el futuro.
- Cómo acelerar el análisis de datos y hacer que sea más preciso, sabiendo asignar a las máquinas tareas centrales del análisis de datos como la clasificación, el clustering o la detección de anomalías.
- Conocimiento del aprendizaje supervisado, el aprendizaje no supervisado y el

aprendizaje por refuerzo.

Habilidades de gestión, personales y sociales

- Interés por la visión lógica y computación.
- Interés por el trabajo en equipo en proyectos tecnológicos.
- Implicación en el desarrollo de la capacidad matemática para fortalecer y entender nuevos conceptos.

MÓDULO DE FORMACIÓN 2: Introducción al Deep Learning y Visión Artificial

OBJETIVO

Aprender las bases de la Inteligencia Artificial, mediante Deep Learning, y construir modelos capaces de identificar objetos, personas y emociones, así como comprender la ciencia utilizada en el diseño de coches y robots autónomos, etiquetado de fotografías o cámaras de seguridad inteligentes.

DURACIÓN TOTAL:

80 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Tipos de modelos y aplicaciones
 - Conocimiento de casos de éxito reales y cómo funcionan: coches autónomos, asistentes inteligentes, Inteligencia en videojuegos, deepfakes.
 - El trabajo del ecosistema especializado Python para Deep Learning con Tensorflow y Keras.
 - Estudio del entorno de predicción del precio de la vivienda en España con datos del sector inmobiliario e indicadores socioeconómicos.
 - La detección del fraude bancario con datos de transacciones entre cuentas de particulares.
 - Estudio del entendimiento de los procesos internos de las funciones de activación en Deep Learning.
 - Conocimiento de Sigmoid, la Tanh y la ReLU.
 - Conocimiento de una red neuronal artificial y sus niveles jerárquicos.
 - Conocimiento de IBM Watson Developer Cloud, Amazon Machine Learning, Azure Machine Learning, TensorFlow o BigML.
 - Entendimiento y aplicabilidad del Deep Learning como elemento de la aproximación a la percepción humana.
 - Estudio de los modelos para clasificación y segmentación de imágenes.
 - Desarrollo de aplicaciones y modelos de visión artificial.
 - El trabajo con las BBDD y modelos disponibles para su integración.
 - El trabajo de la segmentación y detección de peatones y obstáculos desde la cámara de un coche autónomo.
 - Creación de un modelo de diagnóstico asistido a partir de imágenes médicas.
- La extracción de información significativa a partir de imágenes digitales, videos y otras entradas visuales, y la toma de medidas o realización de recomendaciones en función de esa información.

- Técnicas de inspección visual y la detección de defectos, la colocación y medición de piezas, además de identificar, clasificar y rastrear productos.
- Orientación operativa a los dispositivos en la ejecución de sus funciones basándose en la captura y el procesamiento de imágenes.
- Conocimiento de la visión artificial en la Industria 4.0.
- Dominio de los diferentes sistemas y herramientas existentes en el mercado para la visión artificial.

Habilidades de gestión, personales y sociales

- Trabajo en equipo en proyectos tecnológicos, entendiendo que cada proyecto exige perfiles diferentes.
- Mejorar las relaciones en equipos multidisciplinares.
- Toma de conciencia de la gran aplicabilidad que tiene Inteligencia Artificial para el desarrollo de soluciones empresariales.

MÓDULO DE FORMACIÓN 3: Procesamiento del Lenguaje Natural y Modelos Temporales Profundos

OBJETIVO

Aprender cómo la Inteligencia Artificial es capaz de entender y manipular el lenguaje humano, explorando la interacción entre el Aprendizaje Profundo y el lenguaje natural y dominar el primero con el fin de crear estructuras avanzadas para representar conceptos abstractos como las relaciones temporales, así como crear modelos que superaren las técnicas estadísticas clásicas.

DURACIÓN TOTAL:

50 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Las aplicaciones y fundamentos del Procesamiento de Lenguaje Natural.
- El trabajo de los modelos para la transcripción de audio a texto.
- El trabajo de los modelos para el análisis semántico del texto, detección de sentimiento o entidades.
- La práctica con modelos para generación de sentencias y diálogo.
- Entendimiento de ecosistema tecnológico de soluciones, Keras/TensorFlow, HuggingFace, BERT, GPT-3.
- El desarrollo y creación de una arquitectura de asistente virtual automatizado para un call center, generación de diálogos y sentencias.
- La realización del análisis automatizado del sentimiento sobre una marca a través de Twitter.
- La creación de un modelo que aprenda a escribir fake-news de manera autónoma.
- Conocimiento avanzado en Modelos Lógicos: gramáticas.
- Conocimiento avanzado en Modelos probabilísticos del lenguaje natural: basados en datos.
- Procesamiento del lenguaje: Análisis morfológico o léxico, Análisis sintáctico, Análisis semántico y Análisis pragmático.

- Entendimiento de la aplicabilidad del aprendizaje automático para procesar e interpretar textos y datos.
- Profundización en la psicología humana para proyectar en las máquinas.
- Conocimiento de las Redes Neuronales Recurrentes
- El Dominio de los modelos con memoria LSTM y GRU.
- Procesos de implementación en Keras.
- Desarrollo de un Taller de optimización de Logística para estimar la demanda y el stock de una planta de producción simulada.
- Realización de una predicción de aforo en partidos de fútbol de la Liga.
- Conocimiento avanzado en Análisis de series temporales univariantes.
- Predicción de series temporales con redes neuronales.
- Predicción de series temporales con Deep Learning.
- Manejo de datos perdidos en series temporales.
- Introducción a Series temporales con R.
- Conocimiento de las Series Temporales Avanzadas: Aplicación de Redes Neuronales para el Pronóstico de Series de Tiempo

Habilidades de gestión, personales y sociales

- Trasladar los conocimientos técnicos a la realidad diaria.
- Asimilación de la importancia de los desarrollos de ingeniería informática con la aplicación del aprendizaje profundo.
- Mejora de la comprensión analítica de los problemas que tienen las organizaciones.

MÓDULO DE FORMACIÓN 4:	Modelos Generativos y aprendizaje por refuerzo en Machine Learning
-------------------------------	---

OBJETIVO

Trabajar con los DeepFakes; cómo generar imágenes o vídeos falsos con los modelos generativos o GANs y conocer las innovaciones empresariales como Deep Mind o Boston Dynamics en el campo de la IA, aprendiendo técnicas de aprendizaje reforzado que permiten enseñar a las aplicaciones, comportamientos en base a la simulación de situaciones reales.

DURACIÓN TOTAL:

60 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Aprendizaje de los fundamentos y aplicaciones de los modelos generativos.
- Trabajo con la arquitectura y solución mediante aprendizaje profundo.
- Ejemplos reales de aplicación.
- Aplicación de una GAN básica para resolver problemas complejos.
- Generación de contenido multimedia sintético, caras inventadas o videos de famosos tipo DeepFake.
- Realización de entrenamiento de un modelo para detectar comportamientos anómalos en una máquina industrial.
- Entendimiento de Redes Neuronales Generativas Adversaria y su aplicación.

- Marketing digital y los modelos generativos.
- Visibilidad completa en la generación de caras con GANs.
- Arquitectura profunda con algoritmos de aprendizaje
- Principales técnicas de entrenamiento de un modelo generativo.
- El trabajar con Variational Autoencoders.
- Desarrollos con Generative Adversarial Networks.
- Desarrollo del entrenamiento simultáneo de dos arquitecturas y aplicaciones.
- El aprendizaje reforzado, principios y aplicaciones.
- Exposición de ejemplos de casos de éxito y sus fundamentos.
- Aprendizaje de técnicas estadísticas y basadas en modelos.
- Realización del entrenamiento de un modelo capaz de jugar de manera autónoma a videojuegos ATARI.
- Entrenamiento de un modelo capaz de optimizar campañas de publicidad en redes sociales de manera autónoma.
- Conocimiento y trabajo para desarrollar Q-Learning.
- El juego en Python.
- Obtención de soluciones en el mantenimiento predictivo o la personalización de experiencias de cliente.
- La optimización como concepto de mejora y el Machine Learning.
- Criterios en algoritmos para el control de aprendizaje.
- Principales métodos (Método de Montecarlo, Métodos de diferencias temporales).
- Introducción a los mecanismos de exploración inteligente.

Habilidades de gestión, personales y sociales

- Implicación en el aprendizaje de modelos generativos para dar soluciones a problemas complejos.
- Mejora de la capacidad espacial y analítica.
- Mejorar el potencial profesional aumentando la visión de negocio con la aplicación de tecnologías de nueva generación.

OBJETIVO

Conocer MLOps para para implantar con éxito la Inteligencia Artificial en proyectos. Crecer y generar retorno económico a partir de soluciones de IA, conociendo y entendiendo XIA como base a los desarrollos que surgirán en el futuro.

DURACIÓN TOTAL:

60 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- El uso de la visión transversal del MLOps.
- Profundización en aplicaciones y casos de éxito de la IA.
- El manejo de metodologías Ágiles aplicadas a Inteligencia Artificial.
- Conocimiento avanzado en tecnologías Cloud (AWS, Google Cloud y Azure) para la implantación de la IA.
- Integración de un modelo de IA para recomendaciones en una aplicación web.
- Desarrollo de un modelo de IA para detección del fraude en un sistema de transacciones.
- Implantación de un modelo de reconocimiento facial en una aplicación móvil.
- El proceso de desarrollo de productos en 6 etapas.
- Desarrollo de instrucciones paso a paso; preguntas planteadas en lenguaje matemático y código de programación.
- El trabajo con Bots: programa de software automatizado diseñado.
- Aplicación de sistemas autónomos.
- Tipos arquitectónicos de alto nivel: Inteligencia artificial pregenerada, Modelos de inteligencia artificial pregenerados de AI Builder e Inteligencia artificial personalizada.
- Introducción a lenguajes de código.
- Aportaciones prácticas. Actores relacionados.
- Metodología XIA: principales técnicas empleadas.
- Modelos basados en percepción visual, en interpretación matemática y en el procesamiento y comparación de datos.
- Sectores e industrias involucradas.
- Factores condicionantes.
- Predicciones y decisiones.
- Justificación de resultados.
- Explicabilidad de los algoritmos.
- Explicaciones locales y globales.
- Privacidad de los datos.
- Trabajo en el aprendizaje de las máquinas como un proceso comprensible.
- Reglamento General de Protección de Datos (GDPR).
- Técnicas complejas de IA, como el aprendizaje profundo y algoritmos genéticos.
- Aprendizaje de combinación de algoritmos planteados con el propósito de crear máquinas que presenten las mismas capacidades que el ser humano.

Habilidades de gestión, personales y sociales

- Autonomía para desarrollar modelos basados en IA
- Actitud positiva ante los nuevos elementos disruptivos del mercado.
- Sensibilización ante los aspectos éticos y de compliance en los desarrollos de

MÓDULO DE FORMACIÓN 6: Introducción a Python para Data Science y Big Data. Gestión de datos en Python

OBJETIVO

Realizar una introducción al ecosistema Python, orientada a desarrolladores para la transferencia de conocimientos al ecosistema de este lenguaje, así como manipular datos con lenguajes Python, Pandas y el modelo de Dataframes para la construcción de una interfaz de procesamiento de datos escalable.

DURACIÓN TOTAL:

70 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Evaluación de las necesidades de big data y analítica de datos.
- Jupyter Notebook.
- Sintaxis básica de Python:
 - o Valores, tipos, variables Operandos y expresiones.
 - o Condicionales y bucles para control de flujos
 - o Estructuras de datos: Listas y diccionarios
 - o Entrada y salida.
 - o Funciones.
 - o Orientación a objetos.
 - o Módulos.
- Gestión de proyectos en Python.
- Evaluación de la eficiencia del código.
- El trabajo con las funciones comunes en Python.
- Aprendizaje de desarrollo de programas de forma completa con Python.
- Aprendizaje de trabajo con módulos de código que intercambiables.
- Las tareas del interprete en Python.
- Aprendizaje del procesamiento e ingeniería de datos en Python:
 - o Introducción a Pandas y Numpy.
 - o Conocimiento de álgebra relacional basada en dataframes: proyección, filtrado, agrupación y unión de conjuntos de datos.
 - o Serialización y deserialización de datos.
 - o Integración de bases de datos relacionales y entornos cloud.
 - o Limpieza y validación de conjuntos de datos, tratamiento de duplicados y valores perdidos.
- Analítica de datos con Python.
 - o Cálculo estadístico descriptivo e inferencial.
 - o Visualización con matplotlib y seaborn.
- El trabajo con listas en Python, los condicionales “if” y los bucles “for”.
- Funciones en Python y la librería “Pandas”.
- Comprensión de las aplicaciones de analítica avanzada que utilizan Python.
- Aplicación de Python en el conocimiento de los fundamentos y técnicas de Web Scraping.

- Aplicación de Python en el conocimiento de los fundamentos y técnicas de escucha en RRSS.
- Conocimiento de los fundamentos de MongoDB como herramienta de gestión de base de datos no relacional.
- Conocimiento de los fundamentos del Big Data y su aplicación al procesamiento distribuido.

Habilidades de gestión, personales y sociales

- Predisposición en el uso del visionado de los proyectos desde parámetros y métricas.
- Autonomía como profesional experto en programación.
- Ganar en autoconfianza al asumir nuevos retos

MÓDULO DE FORMACIÓN 7: Necesidad y principios de las arquitecturas Big Data y Fundamentos de Apache Spark con Python

OBJETIVO

Definir los componentes del ecosistema Spark, los conceptos básicos de Apache Spark y su interfaz de programación principal, DataFrames conociendo las operaciones para el tratamiento de datos con Spark; extracción, transformación y almacenamiento., así como creando un catálogo de datos centralizado y gestionando metadatos.

DURACIÓN TOTAL:

80 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Conocimiento sobre la necesidad y adopción de las tecnologías Big Data.
- Estudio de arquitecturas de referencia: Data Lake y Lakehouse
- Examen a las limitaciones y soluciones de la arquitectura de análisis de datos existente en un caso de uso.
- Identificación del Ecosistema Hadoop y soluciones de computación escalable distribuida relacionadas con el ecosistema Spark.
 - o Componentes principales de hadoop: MapReduce, HDFS, HIVE.
 - o Principales servicios y proveedores.
 - o Limitaciones.
- Fundamentos y ecosistema Apache Spark.
 - o Necesidad y ejemplo de casos de uso en la industria.
 - o Arquitectura y componentes (rdd, sql , streaming y Machine Learning).
 - o Principales proveedores e infraestructuras: despliegue en Cloud privado, híbrido y público.
- Funciones del arquitecto de datos.
- Principales productos y herramientas que una organización utiliza para gestionar los datos.
- Modelo de interacciones entre sistemas de datos.
- Entendimiento y práctica de datos almacenados, datos en uso, los datos en movimiento, descripciones de los almacenes de datos, los grupos de datos y los elementos de datos; así como práctica con las asignaciones de artefactos de datos a las calidades de datos, las aplicaciones y las ubicaciones.

- El trabajo con Apache Spark, componentes y detalles técnicos de la arquitectura.
- Utilización de la Sintaxis principal de Pyspark,
- Caracterización de la carga de datos y funcionalidad básica de un DataFrame en Pyspark.
- Transformaciones básicas en PySpark.
- Manipulación y transformación con Pyspark.
 - o Agregaciones
 - o Uniones
 - o Funciones de ventana
- Serialización y almacenamiento de datos dentro de un Data Lake.
 - o Apache Parquet, particionado de ficheros, Delta Lake.
- Ejemplos y casos prácticos de soluciones empresariales.
- Depuración y optimización de procesos.

Habilidades de gestión, personales y sociales

- Desarrollo de autonomía para la creación de proyectos de Big Data.
- Reforzar el perfil del profesional de programador y Data Scientist.
- Mejorar las habilidades personales enfocadas en proyectos.

MÓDULO DE FORMACIÓN 8: Spark Streaming y Analítica avanzada y Machine Learning con Spark MLLIB

OBJETIVO

Trabajar con la capa rápida de una arquitectura Big Data, introduciendo alternativas de almacenamiento y procesamiento de datos en streaming, especialmente Apache Kafka, mediante la utilización de componentes de alto nivel de Apache Spark y estudiando modelos de Aprendizaje Automático a partir de grandes volúmenes con Spark MLlib, así como aplicando técnicas de aprendizaje supervisado y no supervisado sobre problemas reales.

DURACIÓN TOTAL:

50 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- El estudio avanzado de las arquitecturas basadas en eventos.
- La utilización de casos de uso de procesamiento continuado.
- Introducción a Apache Kafka.
- Procesamiento en streaming con Pyspark.
- Conocimiento avanzado en el procesamiento escalable, de alto rendimiento y con tolerancia a fallos de flujos de datos en streaming.
- Structured Spark Streaming: interfaz basado en Spark SQL
- Aplicación del motor unificado de analíticas para procesar datos a gran escala que integra módulos para SQL, streaming, aprendizaje automático y procesamiento de grafos.
- La ejecución de forma independiente o en Apache Hadoop, Apache Mesos, Kubernetes, la nube y distintas fuentes de datos.
- Spark Stack y Spark Core.

- Conocimiento del Aprendizaje Automático.
- Identificación de los problemas de escalabilidad en proyectos de aprendizaje.
- Aplicación de Spark MLlib a proyectos de analítica avanzada.
- Ejemplos de problemas predictivos de regresión y clasificación con Pyspark MLlib.
- Conocimiento del procesamiento en memoria.
- Introducción a la Analítica avanzada: cómo realizar consultas SQL y su uso para Machine Learning con librerías de data science como MLlib y GraphX.
- Abstracción RDD (Resilient Distributed Dataset).
- Evaluación perezosa y resultados.

Habilidades de gestión, personales y sociales

- Ganar confianza en el uso de las herramientas de programación.
- Demostración de autonomía para el trabajo en IA.
- Interés en los lenguajes de programación de éxito.

MÓDULO DE FORMACIÓN 9: Gestión de Proyectos de Business Intelligence

OBJETIVO

Conocer el valor añadido que aportan las herramientas de Business Analytics en la toma de decisiones, aprendiendo cómo transformar los datos que llegan de diferentes fuentes, en soluciones con inputs de valor en cada proyecto.

DURACIÓN TOTAL:

30 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Conocimiento de principales herramientas del Business Intelligence.
- El desarrollo de cuadros de mandos
- Dominio de métricas, indicadores y KPI's.
- Inicio en la metodología Agile: Scrum y Kanban en sus proyectos.
- Dominio de artifacts en Proyectos Agile.
- Aprendizaje de los principales procesos, patrones de actuación y cuadros de mando de mayor utilidad para cada proyecto.
- Los problemas al tratar con diversas fuentes de datos, al usar diferentes herramientas y la pérdida de tiempo cuando estos no tienen las mismas métricas.
- Análisis en profundidad de datos con el objetivo de localizar patrones no visibles.
- La traducción de complejos volúmenes de datos provenientes de diversas herramientas (ERP, Web, BBDD, CRM, ...). bajo paneles únicos para crear visualizaciones interactivas y comprensibles.
- Diferenciación de los datos necesarios para recoger y realizar un análisis enfocado en la toma de decisiones.

Habilidades de gestión, personales y sociales

- Asimilación de la importancia en adaptar el perfil profesional a metodologías Agile.
- Interés en proyectos desde parámetros y métricas.
- Capacidad de liderazgo en la gestión empresarial bajo un análisis inteligente de los datos.

MÓDULO DE FORMACIÓN 10:

Power BI como Herramienta de Inteligencia Empresarial y ventaja competitiva

OBJETIVO

Conocer los aspectos técnicos de Power Bi como herramienta de estrategia empresarial para la toma de decisiones y el desarrollo de presentaciones aportando valor a la organización.

DURACIÓN TOTAL:

30 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Aprender las principales componentes de la solución Power BI.
- Obtención y carga de datos en Power BI.
- Modelo lógico Power BI y Manejo de relaciones.
- La transformación de los datos
- Asumir la seguridad y los roles
- Aprendizaje del trabajo en entornos colaborativos.
- Dominio de la exportación de datos y creación de alarmas.
- Conocimiento de las principales herramientas de minería de datos con el objeto de localizar y recopilar los datos, procesarlos y clasificarlos.
- Conocimiento de las principales herramientas OLAP (Online Analytical Processing) para comenzar a trabajar con tablas de bases de datos relacionales.
- Conocimiento avanzado en herramientas de cuadros de mando para aprender a visualizar la evolución del rendimiento de los principales KPIs con gráficos e iconos para su propia necesidad o al servicio de la dirección de la empresa.
- Conocimiento de las principales herramientas de modelización para poder crear los planes de negocio basados en datos y simular el comportamiento en el mercado o en una situación hipotética.

Habilidades de gestión, personales y sociales

- Interés por el desarrollo de análisis crítico.
- Capacidad de autovaloración en el visionado de proyectos desde parámetros y métricas.
- Implicación e interés por desarrollar y manipular datos para realizar presentaciones impactantes.

OBJETIVO

Conocer y practicar con BI y Tableau para la gestión de datos con el fin de desarrollar habilidades multitask y disponer de una visión práctica total la herramienta.

DURACIÓN TOTAL:

80 horas

RESULTADOS DE APRENDIZAJE**Conocimientos / Capacidades cognitivas y prácticas**

- Aprendizaje del conjunto dinámico y estático.
- Combinación de conjuntos.
- Conjunto con parámetros.
- Dashboard: uniformidad en las fuentes, líneas de referencia, Droplines y Tooltips.
- Conocimiento del análisis interactivo enriquecido con acciones de conjunto de Tableau.
- El trabajo con las principales técnicas de mapeo avanzado con Tableau.
- Aprendizaje de la base del análisis avanzado con los primeros pasos en R, Python y MATLAB.
- La mejora de procesos de visualización de datos en forma de varios gráficos y mapas.
- Aprendizaje de los múltiples Joins en Tableau.
- Comparación entre Campos Calculados vs Cálculos de Tabla.
- El análisis de Cálculos de Tabla Avanzados.
- Explicación específica de la dirección de un cálculo.
- Cálculo de Tablas propios
- Media móvil de una segunda capa.
- Aseguramiento de la calidad para Cálculos de Tablas.
- Conocimiento de líneas de tendencia para Power-Insights.
- La importancia del Storyline.
- Aprendizaje del análisis interactivo enriquecido con acciones de conjunto de Tableau.
- El trabajo con las principales técnicas de mapeo avanzado con Tableau.
- Practicas con casos de uso de tablas avanzadas para alcanzar objetivos eficientes.
- Aprendizaje de Box Plots en Tableau.
- Conocimiento de Grandes Conjuntos de Datos en Tableau.
- La Descripción de Pivot Split.
- Las líneas de Tendencia.
- Blending avanzado de series de tiempo.
- Aprendizaje del análisis interactivo enriquecido con acciones de conjunto de Tableau a un nivel profesional
- Trabajo con aspectos iniciales en la predicción de situaciones con el fin de proyectarlos en los aprendido con el Machine Learning.
- Practica con casos de uso con datos avanzados.
- Aprendizaje de las relaciones de Blending.
- Conocimiento de los procesos de visualización.
- Animaciones
- Datos de un Blend (manual y avanzado).
- Conocimiento avanzado en el análisis interactivo enriquecido con acciones de conjunto de Tableau a un nivel profesional.
- El trabajo con las principales técnicas de desarrollo de dashboard.

- Preferencias en animaciones.
- Practicas con casos de uso con datos avanzados
- Aprendizaje de Workbook.
- Intuición de Cálculos LOD.
- ATTR() en Tableau.
- INCLUDE/EXCLUDE.
- Múltiples Campos en un Cálculo LOD.
- FIXED.
- Conocimiento Avanzado en la configuración de la visualización.
- Creación de la expresión LOD.
- Utilización de la expresión LOD en la visualización.
- Expresión LOD rápida.
- Tipos de expresiones LOD.
- Sintaxis de las expresiones LOD.
- Aprendizaje de los mapas de fondo.
- Herramientas de dibujo.
- Cálculos de distancias.
- Mapas con Mapbox.
- Cómo conectar y unir datos a fuentes de datos espaciales comunes en Tableau, como archivos de forma y SQL Server.
- Aprendizaje de los diferentes modos de análisis en la distribución espacial de datos.
 - o Uso de operaciones avanzadas para comprender las relaciones espaciales y agregar datos espacialmente.
 - o Ejemplos de análisis y exploración basados en la interacción que permitan una máxima aplicabilidad.
 - o Opciones que se integran en ArcGIS Maps y proporcionan el contexto adecuado para el análisis geoespacial

Habilidades de gestión, personales y sociales

- Asimilación de cómo impactar con BI.
- Ampliar la visión de soluciones empresariales a partir del análisis de Business Intelligence.
- Interés por ser experto en el Business Intelligence para proyectos de tamaño medio.
- Ganar autoconfianza para la toma de decisiones basadas en la explotación de datos.
- Mejora de la capacidad Multitask
- Ampliar el uso de herramientas disruptivas para completar el perfil profesional con el dominio de la visualización y realizar el análisis de la distribución espacial de datos.

OBJETIVO

Trabajar con BI, Power BI, Qlik Sense ampliando el dominio de herramientas de gestión de proyectos con el fin de desarrollar habilidades multitask y colaborar con equipos multidisciplinares con metodologías Agile.

DURACIÓN TOTAL:

50 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Aprendizaje de las funciones y medidas DAX en el modelo de datos.
- Visualizaciones avanzadas de datos.
- Segmentación de datos y opciones de configuración adicional.
- El servicio en la nube
- Configuración y opciones de trabajo colaborativo.
- Aprendizaje de las principales herramientas de minería de datos con el objeto de localizar y recopilar los datos, procesarlos y clasificarlos a través de Power BI.
- Estudio de las herramientas de cuadros de mando para aprender a visualizar la evolución del rendimiento de los principales KPIs con gráficos e iconos.
- Aprendizaje de las principales herramientas de modelización para poder crear los planes de negocio basados en datos y simular el comportamiento en el mercado o en una situación hipotética y así poder cruzarlos con soluciones de Machine Learning.
- Aprendizaje base de la herramienta.
- Script y carga de datos: Mapping, Crosstable, Range, cargas personalizadas, Variables, y Bucle For.
- Desarrollo de aplicaciones: variables y su aplicación, objetivos personalizados.
- La visualización de datos de autoservicio para la exploración libre de datos.
- Aprendizaje de una analítica guiada con el fin de alinear a los usuarios en un proceso o flujo de trabajo estándar.
- Mejora en los procesos de analítica integrada, enfocado a mejorar los sitios web y las aplicaciones y aplicaciones de analítica personalizada, para respaldar procesos y usos específicos de negocio.
- Conocimiento necesario para mejorar los sitios web y las aplicaciones y aplicaciones de analítica personalizada, para respaldar procesos y usos específicos de negocio.

Habilidades de gestión, personales y sociales

- Mejorar la autoestima con la adquisición de un perfil profesional en ingeniería de datos.
- Interés en el aprendizaje de nuevas herramientas tecnológicas que permiten una mayor empleabilidad.
- Interés por conocer herramientas que permitan la colaboración con equipos multitask.

OBJETIVO

Conocer la ingeniería de datos y comprender la evolución de las arquitecturas de datos hasta su estado más moderno, identificando cada componente clave, así como descubrir nuevas tecnología y patrones para almacenar y consultar datos en BBDD y componentes de última generación.

DURACIÓN TOTAL:

40 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Programación avanzada en Python para la creación de aplicaciones modernas.
- Dominio de las arquitecturas de datos para procesos intensivos en datos.
- Creación de bases de datos NoSQL y principales patrones de aplicación, ejemplos con MongoDB, Cassandra, Neo4j y DynamoDB.
- Arquitecturas Aplicadas y diseño de APIs REST.
- Introducción al Cloud Computing y primeros pasos con Amazon Web Services.
- Creación del backend de una red social a gran escala desplegada en Cloud.
- Implementación de un backend Cloud para el almacenamiento y búsqueda geoespacial de una aplicación de alquiler turístico.
- Aspectos iniciales de la arquitectura de microservicios (MSA).
- Las operaciones transaccionales donde participan múltiples microservicios.
- Prácticas con ejemplos en los que los modelos transaccionales avanzados suponen una alternativa para la implementación de transacciones de larga duración en microservicios.
- Entendimiento de ACID (atomicidad, coherencia, aislamiento y durabilidad) y su aplicación.
- Aprendizaje de diseños de flujos de trabajo transaccionales mediante Cloud Run, Pub/Sub, Workflows y Datastore.

Habilidades de gestión, personales y sociales

- Ampliar el registro personal a nivel profesional como programador de arquitecturas de microservicios (MSA),
- Interés por el diseño de flujos de trabajo transaccionales.

OBJETIVO

Aprender a desplegar arquitecturas de comunicación distribuidas y escalables, creando soluciones de sistemas centrados en computación Cloud distribuida para que fragmente sus componentes en diversas piezas y sistemas de comunicación avanzados entre las mismas.

DURACIÓN TOTAL:

40 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Conocimiento de arquitecturas basadas en microservicios.
- El desarrollo de arquitecturas con contenedores Docker y Kubernetes.
- Conocimiento de las arquitecturas desacopladas:
 - o CQRS
 - o y Event Sourcing.
- El trabajo con sistemas de comunicación basados en eventos, colas, pub/sub y brokers:
 - o . Ejemplos con Apache Kafka, AWS SNS y RabbitMQ.
- Implementación de una arquitectura en tiempo real de gestión de stock y pedidos a gran escala.
- Creación de un sistema de adquisición de datos de vehículos a gran escala para simular una aplicación de movilidad urbana.
- Aspectos más importantes de EDA (Event-Driven Architecture).
- Conocimiento de todos los procesos y etapas de un evento.
- El desarrollo de un patrón de arquitectura a través de un mecanismo que garantice la comunicación entre sus componentes.
- El inicio de creación de una aplicación a través del desarrollo de un modelo y una arquitectura de software.
- Conocimiento inicial de las primeras nociones de Apache Kafka para la transmisión de datos distribuida en el procesamiento de eventos.
- La gestión de la publicación, el almacenamiento y el procesamiento de flujos de eventos de inmediato.

Habilidades de gestión, personales y sociales

- Ampliar el registro personal a nivel profesional como ingeniero programador para la creación de aplicaciones.
- Mejora de la autoestima con el desarrollo de la capacidad de creación de soluciones a medida

OBJETIVO

Entender la arquitectura paralela para el almacenamiento y explotación de datos, descubriendo los patrones principales que implementan las organizaciones para disponer de su información de forma eficiente.

DURACIÓN TOTAL:

40 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Aprendizaje de la necesidad y motivación para la creación de una arquitectura de datos empresarial en la nube.
- El desarrollo de patrones modernos de implantación:
 - o Data Warehouse,
 - o Data Lake
 - o y Data Lakehouse.
- Conocimiento del modelado y acceso a la información con SQL, patrones avanzados y modelado multidimensional.
- El diseño y la implementación de un Data Lake moderno en una empresa retail de Internet a gran escala, identificación de entidades de negocio y creación del modelo de datos.
- La descripción de la estructura de los activos de datos lógicos y físicos y los recursos de gestión de datos de una organización.
- El trabajo con modelos, políticas, reglas y estándares que rigen la recopilación, el almacenamiento, la disposición, la integración y el uso de datos en las organizaciones.
- El trabajo inicial con canalizaciones de datos, con el almacenamiento en la nube, computación en la nube.
- El Uso de API y desarrollo de Modelos de IA y ML básicos.
- Conocimiento en Transmisión de datos y en Orquestación de contenedores.
- Los principales marcos de arquitectura de datos (DAMA-DMBOK 2, Marco Zachman para arquitectura empresarial y El Marco de Arquitectura de Grupo Abierto (TOGAF).
- Prácticas recomendadas de la arquitectura de datos moderna.
- Profundización de la Inteligencia Artificial (IA), la automatización, el Internet de las Cosas (IoT) y la cadena de bloques a través de las arquitecturas de datos modernas.

Habilidades de gestión, personales y sociales

- Ampliar el registro personal a nivel profesional como programador orientado a la Inteligencia Artificial.
- Mejora de la autoestima por el desarrollo de la capacidad de iniciar procesos avanzados de ingeniería de datos.

OBJETIVO

Trabajar la naturaleza y los tipos de datos, profundizando en patrones existentes para la extracción y manipulación de datos para que estos sean adecuados a cada tipo de proyecto.

DURACIÓN TOTAL:

40 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Entender el procesamiento ETL, ELT o EL para el movimiento y transformación de datos entre sistemas transaccionales e informacionales
- Adquisición y procesamiento de datos semi estructurados o no estructurados mediante APIs o Web Scraping
- Extracción de patrones mediante Expresiones Regulares y Motores de Búsqueda como ElasticSearch.
- Extracción automatizada de datos de activos de bolsa y criptomonedas para crear un sistema de análisis para inversores.
- Web scraping sobre datos en tiempo real de reservas de vuelos y precios.
- Aprendizaje de técnicas de control de la extracción de datos y su automatización, disminuyendo el tiempo empleado en el descubrimiento de procesos no documentados, minimizando el margen de error y permitiendo mayor flexibilidad.
- La gestión integrada del Data Warehouse y los Data Marts; integrando la extracción, transformación y carga para la construcción del Data Warehouse corporativo y de los Data Marts.
- Aprendizaje del uso de la arquitectura de metadatos, facilitando la definición de los objetos de negocio y las reglas de consolidación.

Habilidades de gestión, personales y sociales

- Ampliar el registro personal a nivel profesional en la gestión de proyectos y en tareas de dirección de equipos/proyectos.
- Disponer de autonomía para la toma de decisiones con el desarrollo de un proyecto de IA.

OBJETIVO

Conocer tecnologías que combinan el uso de múltiples servidores en paralelo y tecnologías específicas para procesar miles de millones de datos, entendiendo la naturaleza y el uso de las tecnologías para la resolución de problemas de alta complejidad por el volumen o la velocidad de los datos.

DURACIÓN TOTAL:

40 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Entender el ecosistema y fundamentos de tecnologías Big Data: Apache Hadoop.
- Prácticas de Transformación de datos masivos con Apache Spark.
- Dominio del procesamiento en tiempo real con Apache Spark Streaming.
- El trabajo con las arquitecturas y soluciones cloud para trabajar con grandes volúmenes de datos, AWS Athena y Google Big Query.
- Desarrollo de patrones avanzados para almacenar y procesar grandes volúmenes de datos.
- Cómo analizar datos de un dataset masivo de reviews de productos con más de 1000 millones de reviews para una empresa de retail.
- Desarrollo de análisis en tiempo real de twitter en trending topics a nivel internacional.
- Dominio del procesamiento de datos de forma continua.
- Realización de la monitorización de sistemas, de redes y de aplicaciones.
- Aprendizaje y trabajo con dispositivos Internet of Things (IoT).
- Desarrollo de sistemas de recomendación y optimización de resultados.
- Ejemplos sobre transacciones financieras, detección de fraude y trading.
- Cómo desarrollar el seguimiento de usuarios en páginas web y comercio electrónico.
- Creación de notificaciones en dispositivos y aplicaciones móviles en tiempo real.

Habilidades de gestión, personales y sociales

- Habilidades de participación de equipos multidisciplinares.
- Visión holística del análisis de datos. .
- Capacidad de análisis crítico para discutir aspectos de un proyecto.

OBJETIVO

Crear mecanismos de gobierno y control para superar las propias arquitecturas de datos a medida que crece la complejidad de los procesos y de la comunicación de grandes volúmenes de datos.

DURACIÓN TOTAL:

40 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- El trabajo con habilidades duras sobre el Gobierno del Dato orientada a perfiles técnicos.
- Dominio de roles y procesos en el gobierno de datos.
- Las dimensiones a gobernar; calidad, trazabilidad, estandarización y regulación de los datos.
- Cómo sobrevivir e implantar correctamente políticas de gobierno en contextos de Big Data y Cloud Computing.
- La identificación y creación de los distintos componentes de gobierno del dato en un escenario real de un proyecto de Big Data.
- La adopción de medidas necesarias para que los datos sean seguros, privados, precisos, disponibles y utilizables.
- El establecimiento de estándares internos de políticas de datos que regulen la recopilación, el almacenamiento, el proceso y la eliminación de los datos.
- El diseño de procesos para la recogida, proceso, protección y almacenamiento de datos de una organización, para ser utilizados en la toma de decisiones estratégicas que permitan mejorar los resultados de negocio.
- Conocimiento de los principales procesos para implementar una infraestructura de gobierno de datos con el fin de aumentar el valor de los datos de una organización.
- Cómo impulsar los conocimientos sobre datos y escala, así como garantizar la integridad de los datos, su precisión, exhaustividad y coherencia.

Habilidades de gestión, personales y sociales

- Ampliar el registro personal a nivel profesional como ingeniero de datos orientado a la recogida y el uso de datos.
- Concienciación de la necesidad de adquisición de soft skills centrada en la gestión de proyectos.

OBJETIVO

Dominar el uso de técnicas de visualización o analítica avanzada e Inteligencia Artificial y aprender los principales procesos de analítica y reporting implantados en las organizaciones, conectando el ecosistema de Ingeniería de datos a los distintos componentes y procesos.

DURACIÓN TOTAL:

20 horas

RESULTADOS DE APRENDIZAJE

Conocimientos / Capacidades cognitivas y prácticas

- Conocimiento del Machine Learning y a la Inteligencia Artificial.
- Principales procesos de analítica y puesta en producción de modelos.
- Principales tecnologías y entornos Cloud.
- Conocimiento de las principales técnicas de explotación de los datos masivos.
- Análisis de aspectos críticos.

Habilidades de gestión, personales y sociales

- Aumento de la autoestima al convertirse en un experto del análisis de datos en proyectos end-to-end.
- Desarrollo de la capacitación de participar en proyectos de diferentes índoles y sectores, aportando un background completo para trabajar en una consultora IT.

EVALUACIÓN DEL APRENDIZAJE EN LA ACCIÓN FORMATIVA

- La evaluación tendrá un carácter teórico-práctico y se realizará de forma sistemática y continua, durante el desarrollo de cada módulo y al final del curso.
- Puede incluir una evaluación inicial de carácter diagnóstico para detectar el nivel de partida del alumnado.
- La evaluación se llevará a cabo mediante los métodos e instrumentos más adecuados para comprobar los distintos resultados de aprendizaje, y que garanticen la fiabilidad y validez de la misma.
- Cada instrumento de evaluación se acompañará de su correspondiente sistema de corrección y puntuación en el que se explicita, de forma clara e inequívoca, los criterios de medida para evaluar los resultados alcanzados por los participantes.
- La puntuación final alcanzada se expresará en términos de Apto/ No Apto.

Se desarrollarán plantillas de puntuación para realizar seguimientos sistemáticos y análisis de la evolución de cada participante con el fin de ir ajustando las desviaciones

en caso de existir.