

IFCT159 Introducción al Big Data e IA

Sku: PC980

Horas: 40

Formato: HTML

OBJETIVOS

Conocer las principales características del procesamiento masivo de datos en su relación con aplicaciones y algoritmos de Inteligencia Artificial, así como facilitar una aproximación a la planificación y desarrollo de programas de inteligencia artificial y Big Data en el entorno empresarial.

CONTENIDOS

Adquisición y dominio de conceptos básicos y de conocimientos sobre los avances en Big Data.

Para comenzar.

Para activar el aprendizaje, se abordará qué es el Big Data y su importancia en el manejo masivo de datos. Asimismo, se invitará y motivará a explorar los contenidos de la unidad, destacando su evolución desde el Business Intelligence tradicional hasta las tecnologías modernas. Se resolverán una serie de actividades diseñadas para activar ideas y conocimientos previos sobre el impacto, las aplicaciones y los paradigmas fundamentales del Big Data.

Cómo evoluciona el BI tradicional al Big Data.

Para contextualizar y activar el aprendizaje, los estudiantes revisarán cómo ha evolucionado el Business Intelligence hacia el Big Data, identificando las limitaciones del BI tradicional y las oportunidades que brinda el Big Data en el análisis de datos. Además, aprenderán sobre su aplicación en distintos sectores, lo que les permitirá interpretar y utilizar herramientas analíticas para la toma de decisiones en entornos empresariales y tecnológicos modernos.

El Big Data como solución al tratamiento masivo de datos: definición e historia desde su invención.

Aprende sobre el Big Data como solución para el tratamiento masivo de datos, comprendiendo su definición, evolución histórica y relevancia actual.

Hadoop como revolución para el tratamiento paralelo de datos masivo.

Comprende cómo Hadoop transformó el procesamiento de datos masivos, permitiendo el almacenamiento y análisis distribuido a gran escala. Explorará su origen, componentes clave (HDFS, MapReduce, YARN) y su impacto en sectores como finanzas, tecnología y retail. También analizará su evolución y el surgimiento de herramientas más avanzadas como Apache Spark.

Características del Big Data (4 V's y más).

Identifica las 7 V's del Big Data y su impacto en la gestión y análisis de datos masivos. Explorará cómo la volumetría, velocidad, variedad, veracidad, valor, variabilidad y visualización influyen en la toma de decisiones estratégicas y el uso de herramientas avanzadas para optimizar procesos y extraer valor de la información.

Nuevos paradigmas del Big Data: procesos en Real Time y Cloud Computing.

Analiza cómo estas tecnologías optimizan la gestión del Big Data, comparará herramientas como Apache Kafka, Flink y Spark Streaming, y evaluará el impacto de plataformas en la nube como AWS, Azure y Google Cloud en diferentes industrias. También aplicará estos conocimientos a través de casos prácticos y reflexionará sobre su importancia en la transformación digital y la toma de decisiones estratégicas.

Ejercicio práctico de libre expresión escrita.

En esta lección tendrás la oportunidad de poner en práctica lo aprendido sobre los fundamentos del Big Data, su evolución, tecnologías clave y aplicaciones prácticas. A través de ejercicios de expresión escrita, podrás reflexionar, analizar y utilizar tu creatividad para proponer ideas y explorar retos actuales de este campo en expansión. Estos ejercicios están diseñados para animarte a pensar de manera original, conectar conceptos y desarrollar opiniones propias, uniendo el conocimiento técnico con casos reales y escenarios que reflejen el impacto del Big Data en la sociedad.

Role playing y estado de avance.

En esta lección de repaso podrás consolidar los conceptos clave estudiados en la unidad, usando tarjetas de memorización y un ejercicio interactivo basado en situaciones reales. Repasarás las bases del Big Data, sus tecnologías y paradigmas modernos, así como la diferencia con el BI tradicional. Con la actividad de role play practicarás aplicar estos conocimientos en una interacción profesional.

Evaluación de la unidad.

En esta lección pondrás a prueba tus conocimientos adquiridos sobre los fundamentos y la evolución del Big Data, sus tecnologías principales, sus características clave y los nuevos paradigmas. Demuestra lo que has aprendido respondiendo a preguntas seleccionadas

sobre los temas principales de la unidad.

Conocimiento de nociones básicas sobre arquitectura Big Data y principales tecnologías.

Para comenzar.

Para activar el aprendizaje, se introduce el Big Data, sus tecnologías clave y su impacto en la toma de decisiones, complementado con ejercicios para explorar conocimientos previos.

El ecosistema Hadoop: HDFS y MapReduce.

Explora cómo Hadoop, a través de HDFS y MapReduce, permite almacenar y procesar grandes volúmenes de datos de manera distribuida y eficiente. Se analiza su funcionamiento, optimizaciones y aplicaciones en sectores como banca, streaming e inteligencia artificial, comprendiendo su impacto en la escalabilidad y análisis de datos masivos.

Principales lenguajes de programación utilizados: Java, Scala, SQL y Python.

Se aborda los lenguajes clave en Big Data: Java, Scala, SQL y Python, destacando su integración con herramientas como Hadoop, Spark y Kafka. Se exploran sus aplicaciones en procesamiento distribuido, consultas y machine learning, mostrando cómo optimizan el análisis de datos masivos para la toma de decisiones.

Procesos ETL (extracción, transformación y carga): Flume, Sqoop y HIVE.

Para estimular el aprendizaje, el estudiante aplica conocimientos en la identificación y diseño de flujos de datos eficientes. A través del análisis de casos reales, evalúa estrategias de extracción, transformación y carga en entornos distribuidos, diferenciando entre ETL y ELT. Además, desarrolla habilidades prácticas en la implementación de consultas y procesos de integración de datos, optimizando su almacenamiento y procesamiento en sistemas de Big Data.

Procesos Real Time y bases de datos de alta disponibilidad: Kafka, HBASE y Redis.

Aprende a gestionar flujos de datos en tiempo real con Kafka, optimizar el almacenamiento distribuido con HBase y acelerar accesos con Redis, aplicando estos conocimientos en detección de fraudes, análisis de logs y monitoreo en tiempo real.

Procesamiento y analítica avanzada con Spark.

Usa Apache Spark para el procesamiento y análisis eficiente de datos masivos, aplicando Spark SQL, Streaming y MLlib en machine learning, análisis en tiempo real y optimización de consultas en Big Data.

Seguridad y gobierno del dato.

Gestiona la seguridad y gobernanza de datos, asegurando su disponibilidad, integridad y cumplimiento normativo mediante estrategias y herramientas especializadas.

Ejercicio práctico de libre expresión escrita.

En esta lección te proponemos una serie de ejercicios de expresión escrita, pensados para que apliques de manera creativa y personal los conocimientos adquiridos sobre las tecnologías y conceptos clave de la arquitectura Big Data. Tendrás la oportunidad de reflexionar, analizar situaciones reales y proponer tus propias ideas, conectando la teoría con ejemplos y posibles soluciones en la vida profesional y social. Estos ejercicios fomentan tu pensamiento original, tu capacidad de síntesis y argumentación, mientras demuestras comprensión de los temas vistos durante la unidad. Aprovecha esta oportunidad para mostrar tu perspectiva única dentro del apasionante mundo del Big Data y las tecnologías asociadas.

Role playing y estado de avance.

En esta lección de repaso, podrás consolidar los conocimientos clave sobre la arquitectura y tecnologías más importantes en el mundo del Big Data. Comenzarás repasando los conceptos, ventajas, desventajas y ejemplos prácticos a través de flashcards. Luego, participarás en una simulación de rol interactiva para poner en práctica lo aprendido en un contexto realista de análisis y gestión de datos.

Evaluación de la unidad.

En esta lección tendrás la oportunidad de evaluar tus conocimientos sobre los conceptos clave de la arquitectura Big Data y las principales tecnologías que estudiaste en la unidad. Las preguntas cubren puntos esenciales sobre Hadoop, lenguajes de programación, procesos ETL, procesamiento en tiempo real, alta disponibilidad, Spark, seguridad y gobernanza del dato.

Comprensión de los principales conceptos sobre la «ciencia de datos» e IA.

Para comenzar.

Para contextualizar y activar el aprendizaje, el estudiante comprenderá la importancia del crecimiento exponencial de los datos y su impacto en diversas industrias. Explorará cómo la ciencia de datos y la inteligencia artificial (IA) permiten analizar grandes volúmenes de información para generar valor, resolver problemas complejos y mejorar la toma de decisiones en múltiples sectores.

Introducción a la «ciencia de datos» y la Inteligencia Artificial.

Adquiere conocimientos sobre ciencia de datos e inteligencia artificial (IA), explorando su impacto en el análisis y gestión de grandes volúmenes de información. Desarrolla habilidades en recolección, limpieza y análisis de datos, así como en el uso de modelos de aprendizaje automático para la toma de decisiones inteligentes. Aplica técnicas como

minería de datos, modelos predictivos y algoritmos de IA en sectores como salud, finanzas y marketing, abordando también los desafíos éticos y tecnológicos de estas innovaciones.

Principales lenguajes de programación utilizados: R y Python.

Adquiere habilidades en R y Python para análisis de datos, estadística y aprendizaje automático. Aprende a utilizar sus principales librerías, optimizar modelos y aplicar ambas herramientas en ciencia de datos e inteligencia artificial.

Algoritmos supervisados.

Desarrolla habilidades en algoritmos supervisados, diferenciando clasificación y regresión. Explora modelos como regresión lineal, regresión logística y Random Forest, aplicándolos a problemas reales en ciencia de datos e inteligencia artificial.

Algoritmos no-supervisados.

Explora los algoritmos no supervisados, identificando patrones ocultos en datos sin etiquetas. Aprende técnicas como clustering, reducción de dimensionalidad y detección de anomalías, aplicando modelos como K-Means, PCA y DBSCAN en casos reales.

Introducción al Deep Learning y el Aprendizaje por Refuerzo.

Descubre el Deep Learning y el Aprendizaje por Refuerzo, aplicando redes neuronales como CNN, RNN y GAN en visión, lenguaje y generación de contenido. Desarrolla habilidades en entrenamiento y optimización de modelos para resolver problemas complejos.

Procesamiento de información no estructurada: imágenes y textos.

Identifica el procesamiento de datos no estructurados, como imágenes y textos, comprendiendo su complejidad y las técnicas avanzadas necesarias para analizarlos. Aprende sobre Procesamiento del Lenguaje Natural (NLP), Visión por Computador (Computer Vision) y modelos multimodales, aplicando herramientas como CNN, BERT y OpenCV. Desarrolla habilidades para extraer información de grandes volúmenes de datos no estructurados y aplicarlos en casos reales.

Visualización de datos: visualizaciones interactivas y dashboards.

Domina la visualización de datos, creando gráficos y dashboards interactivos con herramientas como Plotly, Power BI y Tableau para analizar y presentar información de forma efectiva.

Ejercicio práctico de libre expresión escrita.

En esta lección tendrás la oportunidad de aplicar y expresar tus conocimientos sobre ciencia de datos, inteligencia artificial y los conceptos asociados que hemos tratado durante la unidad. Se trata de una actividad de expresión libre, en la que cultivarás tu creatividad escribiendo sobre escenarios reales, desafíos y aplicaciones innovadoras de Big Data e IA. Aprovecha este espacio para reflexionar y comunicar tus ideas con originalidad.

Role playing y estado de avance.

En esta lección de repaso, podrás consolidar los principales conocimientos adquiridos sobre ciencia de datos, inteligencia artificial, algoritmos y lenguajes de programación. Primero, repasarás con tarjetas de memoria los conceptos fundamentales, técnicas y herramientas estudiadas en la unidad. Después, participarás en una actividad de rol interactiva donde aplicarás tus conocimientos en una conversación práctica sobre un caso realista de ciencia de datos e IA.

Evaluación de la unidad.

Esta evaluación te permitirá comprobar tu comprensión de los principales conceptos de ciencia de datos e inteligencia artificial vistos en esta unidad. Enfrentarás preguntas sobre algoritmos supervisados y no supervisados, procesamiento de datos no estructurados, lenguajes y herramientas, así como visualización de datos, entre otros temas clave.

Adquisición de una visión trasversal sobre el futuro del Big Data y cómo se aplica actualmente en diferentes áreas.

Para comenzar.

Para activar el aprendizaje, el estudiante descubrirá cómo el Open Data y el Big Data modernizan la gestión pública, impulsando la transparencia y eficiencia gubernamental. Aprenderá cómo el análisis de grandes volúmenes de datos optimiza decisiones, mejora políticas y genera impacto social, fomentando innovación y desarrollo.

Ejemplos en las instituciones públicas: OpenData Aprende cómo distintos países usan Open Data

Ejemplos en el mundo empresarial: ejemplos de aplicabilidad del Big Data a la eficiencia de las operaciones de una compañía.

Descubre cómo el Big Data optimiza procesos, reduce costos y mejora la toma de decisiones en empresas, con aplicaciones en logística, mantenimiento, personalización, seguridad y gestión del talento.para mejorar la transparencia, optimizar servicios y fomentar la innovación, aplicándolo en movilidad, salud, clima y lucha contra la corrupción.

«Data for Good»: Big Data para el bien social.

Analiza cómo el Big Data se usa para resolver problemas sociales, optimizar recursos y mejorar la salud, el clima, la educación y la seguridad, generando un impacto positivo en la sociedad.

Reflexiones finales sobre el impacto del Big Data en los años venideros.

El estudiante explora el futuro del Big Data, su relación con la inteligencia artificial, la automatización, la ciberseguridad y la computación cuántica, así como su impacto en la

privacidad y el empleo. También analiza cómo el Big Data influye en la sociedad a través del metaverso, Blockchain y nuevas regulaciones, comprendiendo los desafíos y oportunidades que trae su evolución.

Ejercicio práctico de libre expresión escrita.

En esta lección tendrás la oportunidad de reflexionar y expresar tus propias ideas sobre el impacto y la aplicabilidad del Big Data en diferentes ámbitos, tanto públicos como privados. A través de ejercicios de escritura creativa, pondrás en práctica los conocimientos adquiridos sobre el uso social y empresarial de los datos, y desarrollarás tu capacidad para analizar, proponer e imaginar futuros posibles relacionados con el Big Data y la inteligencia artificial.

Role playing y estado de avance.

En esta lección de repaso, tendrás la oportunidad de consolidar los conocimientos adquiridos sobre el impacto y el futuro del Big Data, su integración con Open Data, sus aplicaciones empresariales, sociales y gubernamentales, y los principales retos asociados. Empezarás con una actividad de tarjetas de memoria para reforzar conceptos clave y luego participarás en un role play interactivo en el que simularás un escenario realista vinculado a la aplicación y desafíos del Big Data.

Evaluación de la unidad.

En esta lección podrás revisar tus conocimientos sobre los temas tratados en la unidad acerca del futuro del Big Data y su aplicación en diferentes áreas. Se incluyen preguntas sobre gobierno abierto, eficiencia empresarial, proyectos sociales y tendencias tecnológicas.

Evaluación final.

Evaluación final.

Esta es la evaluación final del curso. Pondrás a prueba los conocimientos y habilidades adquiridos durante todas las unidades, incluyendo los fundamentos de Big Data, inteligencia artificial, arquitecturas, algoritmos, herramientas, aplicaciones empresariales, sociales y desafíos éticos. Lee atentamente cada indicación y responde con criterio y reflexión.