



IFCT107. Responsable experto de data.

Sku: PSIFCT107

Horas: 240

Formato: HTML

OBJETIVOS

- Dominar el dato para entenderlo, transformarlo y capturar su máximo potencial con inteligencia artificial para optimizar procesos y simplificar el desarrollo de los proyectos y, comenzar con la creación de modelos de machine learning.

CONTENIDOS

- **1. Introducción: la importancia del dato**
 - Conocimientos/ Capacidades cognitivas y prácticas
 - Concienciación en relación a cómo la ciencia del dato afecta a todos los sectores y puede ser el aliado perfecto para el mundo de los negocios.
 - Conocimiento de los diferentes modelos de negocio basado en datos:
 - Nuevos modelos de negocio en torno al dato.
 - Tipología de datos y cómo se obtienen.
 - Maneras de rentabilizar el dato.
 - Data Business Model Canvas.
 - Adquisición de conocimientos sobre los avances en Big Data & Inteligencia Artificial (IA).
 - Estado actual.
 - Posible evolución futura.
 - Dominio de Conceptos básicos:
 - Qué es Big Data.
 - Qué es Machine Learning.
 - Qué es Deep Learning.
 - Concienciación sobre el poder de los datos en las organizaciones tanto para mejorar la toma de decisiones como para crear modelos de negocio basado en datos.

- Visión transversal sobre cómo Data Science se aplica en las diferentes áreas de la organización para impulsar las capacidades.
 - Habilidades de gestión, personales y sociales
 - Capacidades numéricas y analíticas.
 - Habilidad para trabajar en equipo y con las diferentes áreas de la compañía: Tecnologías de la Información, Seguridad, Legal, Ventas, etc.
 - Conocimiento e interés por cómo evolucionan las áreas de Tecnología y Seguridad.
 - Inquietud y alta motivación por seguir las novedades que se producen de forma continua en el entorno de los datos.
 - Visión de negocio para buscar y entender el encaje de los datos en los productos y servicios a definir.
- **2. Nociones iniciales de Python, data engineering y estadística. Módulo de formación**
 - Conocimientos/ Capacidades cognitivas y prácticas
 - Desarrollo de los fundamentos de Big Data.
 - Qué es el Big Data.
 - Perfiles técnicos: Data Analyst, Engineer, Architect, QA...
 - El proceso de construir un proyecto de Big Data.
 - Arquitecturas de Big Data.
 - Las claves del éxito del big data.
 - Identificación de los fundamentos de Machine Learning.
 - Introducción al Machine Learning: Datos, algoritmos, modelos, producción...
 - Aproximación a los algoritmos: clasificación, predicción y clustering.
 - Creación de modelos de Machine Learning.
 - Las claves del aprendizaje automático.
 - La productivización de modelos.
 - Análisis de los fundamentos de IA: Visión, NPL...
 - Introducción al Deep Learning y sus aplicaciones.
 - Tipos de redes neuronales y su entrenamiento.
 - Funcionamiento y retos del Data Science aplicado al Computer Vision y al NPL.
 - Conocimientos avanzados de Python Crash Course.
 - Uso de pycharm como entorno de trabajo.
 - Uso de notebooks.
 - Conocimiento de la sintaxis del lenguaje: bucles, variables.
 - Librerías.
 - Funciones.
 - Programación orientada a objetos en Python.
 - Aproximación a la arquitectura de la información y al SQL.
 - Qué es la arquitectura de datos.
 - Modelo relacional tradicional.
 - Modelo estrella.
 - Modelo copo de nieve.
 - Bases de la normalización de datos.

- Definiciones esenciales: base de datos, tablespace, tabla, vista, vista materializada, clave primaria, índice y partición.
 - SQL.
 - Sintaxis del lenguaje.
- Identificación del Crash course de python para data engineers: SQL y dataframes.
 - Consulta de datos.
 - Cruce de tablas.
 - Inserción y borrado.
 - Dataframes:
 - Qué es una serie.
 - Qué es un dataframe.
 - Cruces con dataframe.
 - Funciones lambda con columnas.
 - Leer y salvar datos desde y a dataframe desde distintos formatos.
- Diferenciación entre ETL/ELT y preproceso de datos.
 - ELT o ETL.
 - Pipelines de datos.
 - ¿Cómo estructurar los distintos pasos del workflow de datos?
 - Automatización de procesos.
- Aplicación del proceso del análisis exploratorio.
 - Arranque de un proceso exploratorio.
 - Los objetivos de un análisis exploratorio.
 - Uso del proceso exploratorio para validación de hipótesis. Naturaleza iterativa del proceso.
- Utilización de herramientas de visualización exploratoria y librerías.
 - Análisis descriptivo gráfico.
 - Análisis descriptivo basado en estadísticos.
- Realización de análisis univariante y multivariante.
 - Análisis multivariante.
 - Análisis univariante.
- Conocimiento de la Estadística descriptiva.
 - Media, mediana, momentos, etc.
 - Desviación, varianza.
 - Sesgos y medidas de homogeneidad de la información.
- Aproximación breve al álgebra lineal.
 - Operaciones matriz-escalar.
 - Operaciones matriz-matriz.
 - Propiedades de las matrices: no conmutativa, asociativa, distributiva, matriz identidad.
 - Trasposición e inversa de una matriz.
- Análisis de correlación de variables: fundamentos de estadística y probabilidad.
 - Ingeniería de características: descarte y selección de características.
 - Cómo analizar la correlación entre variables.
 - Efecto de las correlaciones.

- Eliminación de la correlación.
 - Deducción estadística y contraste de hipótesis.
 - Definición de una distribución una variable aleatoria.
 - Definición de una función de probabilidad.
 - Habilidades de gestión, personales y sociales
 - Competencias necesarias para el nuevo entorno digital:
 - La influencia digital.
 - La colaboración en el entorno.
 - La integración de la diversidad.
 - La gestión emocional.
 - La agilidad en toma de decisiones.
 - La anticipación en contextos digitales.
 - La flexibilidad para la transformación.
 - La asunción de incertidumbre y riesgos.
 - La elaboración, gestión y difusión de contenidos.
- **3. Conocimiento avanzado de machine learning & artificial intelligence**
 - Conocimientos/ Capacidades cognitivas y prácticas
 - Aproximación a la modelización.
 - Introducción a la modelización.
 - Tipos de modelos.
 - Identificación de los Modelos de regresión.
 - Modelos de regresión simple.
 - Modelos de regresión múltiple.
 - Modelos de regresión generalizado.
 - Clasificación de los distintos modelos de Árboles:
 - Modelos de inferencia de árboles de decisión y regresión:
 - Clasificadores binarios.
 - Regresión.
 - Modelos de inferencia de árboles de decisión: clasificadores multiclase.
 - Modelos avanzados de árboles:
 - Boosting.
 - Random forest.
 - Conocimientos de Algoritmos alternativos de clasificación.
 - Conocimientos de Eager y Lazy classifiers
 - Utilización de Clustering.
 - Introducción a los modelos no supervisados.
 - Análisis clúster.
 - Aproximación al método científico:
 - Evaluación de modelos.
 - Evaluación y optimización de modelos:
 - Control de outliers y análisis de residuos.
 - Modelos no supervisados.
 - Modelos supervisados.
 - Creación de Ingeniería de variables:
 - Principios de la ingeniería de variables en la creación de variables.

- La dimensionalidad.
 - Técnicas: PCA y SVD.
- Ensamblado de modelos:
 - Definición de model ensembles.
 - Modelos débiles y modelos fuertes.
 - Creación de pipelines para crear modelos ensamblados.
- Gestión del ciclo de vida de los modelos.
- Interpretabilidad.
- Aproximación a modelos heurísticos de optimización.
- Conocimiento de los algoritmos genéticos.
 - Principios de optimización basado en heurísticos.
 - Gradiente descendente.
 - Modelos basados en comportamientos animales.
 - El algoritmo genético.
 - Aplicaciones de los algoritmos genéticos.
 - Cromosoma y función de fitness.
 - Resolución de un problema usando algoritmos genéticos.
- Utilización de series temporales y forecasting.
 - Bases sobre componentes de las series temporales: tendencia, ciclo y estacionalidad.
 - Series estacionarias y no estacionarias.
 - Análisis de anomalías.
 - Suavizado exponencial.
 - Modelos autoregresivos.
 - Modelos univariantes y multivariantes.
 - Modelos arima.
 - Modelos con parametrización automática.
- Gestión de proyectos de Data Science.
- Habilidades de gestión, personales y sociales
 - Inteligencia emocional: habilidad humana de comprender y gestionar las emociones en el ámbito profesional.
 - Trabajo en equipo. Las claves del mejor trabajo en equipo:
 - Confianza: cada colaborador debe confiar en los demás tanto como en sí mismo. El apoyo mutuo es imprescindible para garantizar la fluidez en los procesos.
 - Equilibrio: dentro de cada equipo, cada persona tiene su propio rol y debe ser capaz de desarrollarlo de forma individual, en coherencia con los objetivos comunes.
 - Pertenencia: sentirse involucrado con el equipo incentiva una mayor implicación, las ganas de hacer aportaciones valiosas y ver crecer cada proyecto.
 - Adaptabilidad: el talento más flexible, que mejor se adapta a los nuevos escenarios es el más capacitado para materializar las soluciones más novedosas y originales.
 - Pensamiento crítico: Es una competencia clave para complementar la lógica pura, detectar otras variables que pueden afectar a la toma de decisiones y no dar siempre por válida la primera opción.

- Resolución de problemas: Si la inteligencia artificial es una tecnología orientada a la inmediatez en los resultados, las personas deben ser capaces de crear las estrategias más adecuadas a cada situación para lograr obtener las mejores soluciones.
- **4. Aplicaciones del deep learning**
 - Conocimientos/ Capacidades cognitivas y prácticas
 - Aproximación a redes neuronales artificiales (ANN).
 - ¿Qué es una red neuronal artificial?
 - Tipos de redes: profundas y superficiales (shallow).
 - Técnica del gradiente descendiente.
 - Conocimiento de los campos de aplicación de ANN.
 - Visión por computador.
 - Análisis y síntesis del lenguaje.
 - Análisis de secuencias.
 - GAN y deepfake.
 - Aproximación al Shallow & Deep neural networks. Introducción CNN.
 - Perceptrón simple.
 - Topologías de redes neuronales.
 - MNIST.
 - Regularización: L1, L2, dropout y otros.
 - Creación de una CNN con keras.
 - Comprensión del diseño de redes neuronales artificiales (ANN), redes convolucionales (CNN) y redes recurrentes (RNN).
 - Habilidades de gestión, personales y sociales
 - Inteligencia emocional: habilidad humana de comprender y gestionar las emociones en el ámbito profesional.
 - Trabajo en equipo. Las claves del mejor trabajo en equipo:
 - Confianza: cada colaborador debe confiar en los demás tanto como en sí mismo. El apoyo mutuo es imprescindible para garantizar la fluidez en los procesos.
 - Equilibrio: dentro de cada equipo, cada persona tiene su propio rol y debe ser capaz de desarrollarlo de forma individual, en coherencia con los objetivos comunes.
 - Pertenencia: sentirse involucrado con el equipo incentiva una mayor implicación, las ganas de hacer aportaciones valiosas y ver crecer cada proyecto.
 - Adaptabilidad: el talento más flexible, que mejor se adapta a los nuevos escenarios es el más capacitado para materializar las soluciones más novedosas y originales.
 - Pensamiento crítico: Es una competencia clave para complementar la lógica pura, detectar otras variables que pueden afectar a la toma de decisiones y no dar siempre por válida la primera opción.
 - Resolución de problemas: Si la inteligencia artificial es una tecnología orientada a la inmediatez en los resultados, las personas deben ser capaces de crear las estrategias más adecuadas a cada situación para lograr obtener las mejores soluciones.

- **5. Entornos Big Data & Cloud**

- Conocimientos/ Capacidades cognitivas y prácticas
 - Conocimientos de Big Data y de los principios de arquitecturas de computación distribuidas y altamente escalables.
 - Comprender el papel del Big data en la ciencia de datos.
 - Apache Spark.
 - Spark en modo batch y en semi-tiempo real (microbatches).
 - Lazy evaluation.
 - Utilización de Dataframes y Data pipelines en Spark.
 - Operaciones sobre dataframes (SQL o pyspark).
 - Creación de data pipelines con spark.
 - Transformación de dataframes.
 - Creación de modelos de Machine Learning en Spark.
 - Entrenamiento de modelos e integración en pipelines.
 - Aplicación de la inferencia de modelos en Spark.
 - Despliegue de pipelines de modelos con Spark.
 - Desarrollo de modelos con proveedores de Cloud Machine Learning.
- Habilidades de gestión, personales y sociales
 - Capacidades numéricas y analíticas.
 - Habilidad para trabajar en equipo y con las diferentes áreas de la compañía: IT, Seguridad, Legal, Ventas, etc.
 - Conocimiento e interés por cómo evolucionan las áreas de Tecnología y Seguridad.
 - Inquietud y alta motivación por seguir las novedades que se producen de forma continua en el entorno de los datos.
 - Visión de negocio para buscar y entender el encaje de los datos en los productos y servicios a definir.

- **6. Visualización de los datos**

- Conocimientos/ Capacidades cognitivas y prácticas
 - Comprensión de las claves del storytelling a través de los datos.
 - La ciencia de la visualización de datos y el storytelling: su valor en el mundo profesional.
 - Uso del storytelling con datos.
 - Principios de la comunicación visual de los datos.
 - Uso de los distintos controles visuales.
 - Comprensión de los conceptos Business Intelligence y Data Visualization.
 - Entender el proceso para trabajar en la herramienta de visualización en un contexto de Business Intelligence.
 - Fuentes de datos, tratamiento y preparación de datos, modelaje de datos (relaciones), análisis exploratorio y específicos, visualización y reporting.
 - Profundización en la herramienta PowerBI para poder conectar datos, prepararlos, modelarlos, explorarlos y visualizarlos.
 - Entorno de PowerBI: organización y componentes.
 - Importación de datos.
 - Visualizaciones básicas realizando agrupaciones y aplicando filtros.

- Trabajar y preparar los datos: columnas calculadas, medidas, fórmulas con DAX, tablas calculadas...
 - Trabajar con tablas matrices para aumentar la granularidad de los informes.
 - Filtrar los datos de manera dinámica con la segmentación de datos.
 - Aproximación a la herramienta Tableau.
 - Habilidades de gestión, personales y sociales
 - Habilidades analíticas y creativas para encontrar soluciones a problemas o necesidades de negocio, realizando el trabajo de manera exhaustiva, de un modo metódico, sistemático y creativo.
 - Capacidad de síntesis para observar los datos y extraer su información útil y relevante.
 - Capacidad crítica para ser capaz de cuestionar los datos y así desarrollar conclusiones en base a los mismos.
 - Conocimientos en informática, matemáticas y estadística: estos perfiles deben poder analizar bases de datos, construir modelos y realizar previsiones estadísticas, etc.
 - Facilidad para la comunicación para explicar los resultados del trabajo a gerentes y directores, generalmente sin formación técnica, de la empresa u organización.
- **7. Habilidades y competencias de gestión, personales y sociales para el entorno digital**
 - Conocimientos/ Capacidades cognitivas y prácticas
 - Impulso de habilidades digitales:
 - Liderazgo participativo.
 - Gestión del cambio.
 - Inteligencia emocional.
 - Storytelling.
 - Creación de marca personal.
 - Comunicación y negociación en entornos digitales.
 - Comprensión de las competencias necesarias para el nuevo entorno digital:
 - La influencia digital.
 - La colaboración en el entorno.
 - La integración de la diversidad.
 - La gestión emocional.
 - La agilidad en toma de decisiones.
 - La anticipación en contextos digitales.
 - La flexibilidad para la transformación.
 - La asunción de incertidumbre y riesgos.
 - La elaboración, gestión y difusión de contenidos.
 - Conocimiento de las herramientas imprescindibles para:
 - El trabajo colaborativo.
 - El trabajo en remoto.
 - La gestión de proyectos.
 - Automatización de flujos de trabajo.
 - Práctica en las claves del trabajo en equipo y de la productividad en remoto.

- Fomento del liderazgo participativo y la gestión del cambio.
- Dominio de la comunicación y la negociación en entornos digitales.
- Práctica del modelo de competencias para el entorno digital.
- Elaboración y gestión de la marca personal desde la importancia del storytelling.