



PROGRAMA FORMATIVO

Data Scientist

Febrero 2022

IDENTIFICACIÓN DE LA ESPECIALIDAD Y PARÁMETROS DEL CONTEXTO FORMATIVO

Denominación de la especialidad:	DATA SCIENTIST
Familia Profesional:	INFORMÁTICA Y COMUNICACIONES
Área Profesional:	DESARROLLO
Código:	IFCD66
Nivel de cualificación profesional:	5

Objetivo general

Extraer conocimiento de utilidad para un propósito concreto a partir de enormes volúmenes de datos de fuentes diversas disponibles en formato digital.

Relación de módulos de formación

Módulo 1	Sistemas de apoyo a la toma de decisiones y gestión de datos	100 horas
Módulo 2	Gestión y procesamiento de datos	80 horas
Módulo 3	Aprendizaje automático y visualización	130 horas

Modalidades de impartición

Presencial

Teleformación

Duración de la formación

Duración total en cualquier modalidad de impartición 310 horas

Teleformación Duración total de las tutorías presenciales: 0 horas

Requisitos de acceso del alumnado

Acreditaciones/ titulaciones	Cumplir como mínimo alguno de los siguientes requisitos: <ul style="list-style-type: none">- Título de Grado o equivalente de la familia profesional Informática y Comunicaciones.- Título de Postgrado (Máster) o equivalente de la familia profesional Informática y Comunicaciones.
Experiencia profesional	No se requiere.
Otros	Deben poseer conocimientos de: <ul style="list-style-type: none">- Programación básica- Diseño de algoritmos- Metodología SCRUM- Conocimientos de lenguaje SQL para BBDD relacionales- Tener conocimientos básicos de álgebra lineal, matemática discreta, probabilidad y estadística. Se valorará acreditar conocimientos de materias como: Fundamentos matemáticos, Probabilidad y estadística, Grafs y geometría discreta.

	Cuando el alumnado no disponga de la acreditación o titulación requerida demostrará los conocimientos y competencias suficientes mediante una prueba competencial práctica de nivel consistente en ejercicios de programación de aplicaciones a nivel práctico (básico), diseño de algoritmos, conocimientos de lenguaje SQL para BBDD relacionales y experiencia en metodologías ágiles de desarrollo de aplicaciones.
Modalidad de teleformación	Además de lo indicado anteriormente, los/as participantes han de tener las destrezas suficientes para ser usuarios de la plataforma virtual en la que se apoya la acción formativa.

Prescripciones de formadores y tutores

Acreditación requerida	Cumplir como mínimo alguno de los siguientes requisitos: <ul style="list-style-type: none"> - Licenciado, Ingeniero, Arquitecto o el Título de Grado correspondiente u otros títulos equivalentes.
Experiencia profesional mínima requerida	No se requiere.
Competencia docente	Experiencia docente o investigadora en el ámbito de la disciplina acreditable de al menos 60 horas en modalidad presencial.
Modalidad de teleformación o mixta	Además de cumplir con las prescripciones establecidas anteriormente, los tutores-formadores deben acreditar una formación, de al menos 30 horas, o experiencia, de al menos 60 horas, en esta modalidad y en la utilización de las tecnologías de la información y comunicación.

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 m ²	2,4 m ² / participante

Espacio Formativo	Equipamiento
Aula de gestión	<ul style="list-style-type: none"> - 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 alumnos. - Software específico para el aprendizaje de cada acción formativa (se proporcionan indicaciones sobre herramientas equivalentes) <ul style="list-style-type: none"> • Modelos de datos: HDFS, Apache Avro, Apache Parquet, Apache HBase, MongoDB, Neo4j, Cypher, GraphDB • Herramientas modelización multidimensional: PowerBI, Indycó • Herramientas ETL: Talend, Apache NiFi • Herramienta analítica de datos: Spark, Kafka, Airflow • Herramientas de visualización: Grafana, PowerBI, Tableau

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

En el caso de que aumente el número de alumnos/as, hasta un máximo de 30, la superficie de las aulas se incrementará proporcionalmente (según se indica en la tabla en lo relativo a m²/ alumno) 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/as alumnos/as.

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.

Para impartir la formación en **modalidad de teleformación**, se ha de disponer del siguiente equipamiento.

Plataforma de teleformación:

La plataforma de teleformación que se utilice para impartir acciones formativas deberá alojar el material virtual de aprendizaje correspondiente, poseer capacidad suficiente para desarrollar el proceso de aprendizaje y gestionar y garantizar la formación del alumnado, permitiendo la interactividad y el trabajo cooperativo, y reunir los siguientes requisitos técnicos de infraestructura, software y servicios:

• **Infraestructura**

- Tener un rendimiento, entendido como número de alumnos que soporte la plataforma, velocidad de respuesta del servidor a los usuarios, y tiempo de carga de las páginas Web o de descarga de archivos, que permita:
 - a) Soportar un número de alumnos equivalente al número total de alumnos en las acciones formativas de formación profesional para el empleo que esté impartiendo el centro o entidad de formación, garantizando un hospedaje mínimo igual al total del alumnado de dichas acciones, considerando que el número máximo de alumnos por tutor es de 80 y un número de usuarios concurrentes del 40% de ese alumnado.
 - b) Disponer de la capacidad de transferencia necesaria para que no se produzca efecto retardo en la comunicación audiovisual en tiempo real, debiendo tener el servidor en el que se aloja la plataforma un ancho de banda mínimo de 300 Mbs, suficiente en bajada y subida.
- Estar en funcionamiento 24 horas al día, los 7 días de la semana.

• **Software:**

- Compatibilidad con el estándar SCORM y paquetes de contenidos IMS.
- Niveles de accesibilidad e interactividad de los contenidos disponibles mediante tecnologías web que como mínimo cumplan las prioridades 1 y 2 de la Norma UNE 139803:2012 o posteriores actualizaciones, según lo estipulado en el capítulo III del Real Decreto 1494/2007, de 12 de noviembre.
- El servidor de la plataforma de teleformación ha de cumplir con los requisitos establecidos en la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales, por lo que el responsable de dicha plataforma ha de identificar la localización física del servidor y el cumplimiento de lo establecido sobre transferencias internacionales de datos en los artículos 40 a 43 de la citada Ley Orgánica 3/2018, de 5 de diciembre, así como, en lo que resulte de aplicación, en el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas respecto del tratamiento de datos personales y la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE.
- Compatibilidad tecnológica y posibilidades de integración con cualquier sistema operativo, base de datos, navegador de Internet de los más usuales o servidor web, debiendo ser posible utilizar

las funciones de la plataforma con complementos (plug-in) y visualizadores compatibles. Si se requiriese la instalación adicional de algún soporte para funcionalidades avanzadas, la plataforma debe facilitar el acceso al mismo sin coste.

- Disponibilidad del servicio web de seguimiento (operativo y en funcionamiento) de las acciones formativas impartidas, conforme al modelo de datos y protocolo de transmisión establecidos en el anexo V de la Orden/TMS/369/2019, de 28 de marzo.

- **Servicios y soporte**

- Sustentar el material virtual de aprendizaje de la especialidad formativa que a través de ella se imparta.
- Disponibilidad de un servicio de atención a usuarios que de soporte técnico y mantenga la infraestructura tecnológica y que, de forma estructurada y centralizada, atienda y resuelva las consultas e incidencias técnicas del alumnado. Las formas de establecer contacto con este servicio, que serán mediante teléfono y mensajería electrónica, tienen que estar disponibles para el alumnado desde el inicio hasta la finalización de la acción formativa, manteniendo un horario de funcionamiento de mañana y de tarde y un tiempo de demora en la respuesta no superior a 48 horas laborables.
- Personalización con la imagen institucional de la administración laboral correspondiente, con las pautas de imagen corporativa que se establezcan.

Con el objeto de gestionar, administrar, organizar, diseñar, impartir y evaluar acciones formativas a través de Internet, la plataforma de teleformación integrará las herramientas y recursos necesarios a tal fin, disponiendo, específicamente, de herramientas de:

- Comunicación, que permitan que cada alumno pueda interaccionar a través del navegador con el tutor-formador, el sistema y con los demás alumnos. Esta comunicación electrónica ha de llevarse a cabo mediante herramientas de comunicación síncronas (aula virtual, chat, pizarra electrónica) y asíncronas (correo electrónico, foro, calendario, tablón de anuncios, avisos). Será obligatorio que cada acción formativa en modalidad de teleformación disponga, como mínimo, de un servicio de mensajería, un foro y un chat.
- Colaboración, que permitan tanto el trabajo cooperativo entre los miembros de un grupo, como la gestión de grupos. Mediante tales herramientas ha de ser posible realizar operaciones de alta, modificación o borrado de grupos de alumnos, así como creación de «escenarios virtuales» para el trabajo cooperativo de los miembros de un grupo (directorios o «carpetas» para el intercambio de archivos, herramientas para la publicación de los contenidos, y foros o chats privados para los miembros de cada grupo).
- Administración, que permitan la gestión de usuarios (altas, modificaciones, borrado, gestión de la lista de clase, definición, asignación y gestión de permisos, perfiles y roles, autenticación y asignación de niveles de seguridad) y la gestión de acciones formativas.
- Gestión de contenidos, que posibiliten el almacenamiento y la gestión de archivos (visualizar archivos, organizarlos en carpetas –directorios- y subcarpetas, copiar, pegar, eliminar, comprimir, descargar o cargar archivos), la publicación organizada y selectiva de los contenidos de dichos archivos, y la creación de contenidos.
- Evaluación y control del progreso del alumnado, que permitan la creación, edición y realización de pruebas de evaluación y autoevaluación y de actividades y trabajos evaluables, su autocorrección o su corrección (con retroalimentación), su calificación, la asignación de puntuaciones y la ponderación de las mismas, el registro personalizado y la publicación de calificaciones, la visualización de información estadística sobre los resultados y el progreso de cada alumno y la obtención de informes de seguimiento.

Material virtual de aprendizaje:

El material virtual de aprendizaje para el alumnado mediante el que se imparta la formación se concretará en el curso completo en formato multimedia (que mantenga una estructura y funcionalidad homogénea), debiendo ajustarse a todos los elementos de la programación (objetivos y resultados de aprendizaje) de este programa formativo que figura en el Catálogo de Especialidades Formativas y cuyo contenido cumpla estos requisitos:

- Como mínimo, ser el establecido en el citado programa formativo del Catálogo de Especialidades Formativas.
- Estar referido tanto a los objetivos como a los conocimientos/ capacidades cognitivas y prácticas, y habilidades de gestión, personales y sociales, de manera que en su conjunto permitan conseguir los resultados de aprendizaje previstos.
- Organizarse a través de índices, mapas, tablas de contenido, esquemas, epígrafes o titulares de fácil discriminación y secuenciarse pedagógicamente de tal manera que permitan su comprensión y retención.
- No ser meramente informativos, promoviendo su aplicación práctica a través de actividades de aprendizaje (autoevaluables o valoradas por el tutor-formador) relevantes para la adquisición de competencias, que sirvan para verificar el progreso del aprendizaje del alumnado, hacer un seguimiento de sus dificultades de aprendizaje y prestarle el apoyo adecuado.
- No ser exclusivamente textuales, incluyendo variados recursos (necesarios y relevantes), tanto estáticos como interactivos (imágenes, gráficos, audio, video, animaciones, enlaces, simulaciones, artículos, foro, chat, etc.). de forma periódica.
- Poder ser ampliados o complementados mediante diferentes recursos adicionales a los que el alumnado pueda acceder y consultar a voluntad.
- Dar lugar a resúmenes o síntesis y a glosarios que identifiquen y definan los términos o vocablos básicos, relevantes o claves para la comprensión de los aprendizajes.
- Evaluar su adquisición durante y a la finalización de la acción formativa a través de actividades de evaluación (ejercicios, preguntas, trabajos, problemas, casos, pruebas, etc.), que permitan medir el rendimiento o desempeño del alumnado.

Aula virtual

Tecnología y equipos	Plataforma de aprendizaje que permita la conexión síncrona de docentes y alumnos, con sistema incorporado de audio, video y posibilidad de compartir archivos, la propia pantalla u otras aplicaciones tanto por el docente como por los/as alumnos/as, con registro de los tiempos de conectividad.
-----------------------------	--

Ocupaciones y puestos de trabajo relacionados

- 2712 Analistas y diseñadores de software
- 2713 Analistas, programadores y diseñadores Web y multimedia
- 2719 Analistas y diseñadores de software y multimedia no clasificados bajo otros epígrafes

Requisitos oficiales de las entidades o centros de formación

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

DESARROLLO MODULAR

MÓDULO DE FORMACIÓN 1: SISTEMAS DE SOPORTE A LA TOMA DE DECISIONES Y GESTIÓN DE DATOS

OBJETIVO

Desarrollar aplicaciones informáticas para realizar un tratamiento de los datos básico con el lenguaje Python, identificando los métodos para la exploración de grandes fondos de datos y los sistemas de gestión de datos relacionales y no relacionales (NoSQL).

DURACIÓN EN CUALQUIER MODALIDAD DE IMPARTICIÓN: 100 horas

Teleformación: Duración de las tutorías presenciales: 0 horas

RESULTADOS DE APRENDIZAJE

Conocimientos/ Capacidades cognitivas y prácticas

- Caracterización de la aplicación del lenguaje Python
 - Lenguaje Python
 - Ejecución de programas Python
 - Objetos en Python
 - Tipos numéricos y dinámicos
 - Gestión de cadenas de texto: listas, diccionarios, tuplas y ficheros
 - Sentencias Python: asignaciones, expresiones e imprimir resultados
 - Tests de variables, reglas de sintaxis
 - Bucles for y while
- Interpretación la aplicación de protocolos API
 - Uso de APIs remotas
 - Integración de las aplicaciones con APIs remotas
 - Ejemplos de aplicación de APIs remotas en lenguaje Python
- Programación de un algoritmo modular en lenguaje Python
 - Programación de módulos
 - Fundamentos de programación de clases
 - Utilización de APIs e integración con aplicaciones Python
- Distinción de los conceptos Cloud básicos
 - Principios de computación en la nube (Cloud Computing)
 - Ingeniería de servicios: software as a service, Platform as a service, Infrastructure as a Service
 - Ejemplos de aplicaciones relevantes en la industria
- Uso de BBDD NoSQL y nuevos modelos de datos (estructurados y no estructurados)
 - Fundamentos del paradigma NoSQL
 - Distribución de los datos y procesamiento en paralelo
 - Principales modelos de datos en el mundo NoSQL: clave-valor, orientación a documentos, grafos de propiedad, grafos de conocimiento
- Conocimiento del almacenamiento Big Data y las herramientas de procesamiento masivo
 - Aplicaciones basadas en la gestión y el análisis de grandes volúmenes de datos

- Fundamentos arquitectónicos de los sistemas distribuidos
 - Principales arquitecturas de referencia
 - Nuevos modelos de datos
 - Sistemas de ficheros distribuidos
 - Document stores
 - Bases de datos de grafos
- Evaluación de las metodologías y técnicas aplicadas en la resolución de problemas y justificación de los planteamientos, decisiones y propuestas realizadas
 - Sistemas de soporte a la toma de decisiones
 - Análisis de los datos: análisis descriptivo, predictivo y prescriptivo
 - Casos de uso: gestión y análisis de grandes volúmenes de datos
- Identificación de los factores clave de un problema complejo en el contexto de un proyecto de analítica.
 - Contexto de la sociedad /economía de los datos y el paradigma de las aplicaciones orientadas a los datos
 - Fundamentos de bases de datos relacionales: lenguaje SQL.
 - Necesidad de un cambio de paradigma: NoSQL. El principio 'one size does not fit all'.
 - Principales modelos de datos en el mundo NoSQL: Key-Value, Documento-oriented, Property Graphs y Knowledge Graphs
 - Fundamentos arquitectónicos: sistemas distribuidos, escalabilidad, paralelismo. Principales arquitecturas de referencia (shared nothing, shared disk, shared memory)
- Distinción y aplicación de los nuevos modelos de datos
 - Sistemas de archivos distribuidos: conceptos y principios (distribución, replicación, particionamiento horizontal vs. Vertical, formatos de archivos especializados)
 - Conocimiento y utilización de Hadoop File System (HDFS), Apache Avro, Apache Parquet, Key-value stores: Apache HBase
 - Document stores: conceptos y principios (mecanismos de réplica, sharding, consultas espaciales)
 - Inmersión a MongoDB y el Aggregation Framework
 - Graph databases: property y knowledge graphs. Conceptos y principios Modelización en grafo, consultas regulares. Introducción a Neo4j y Cypher
 - Knowledge graphs. Conceptos y principios: el paradigma open / linked data, RDF y SPARQL. Introducción a GraphDB
- Identificación y análisis de problemas complejos en el área de análisis de datos y planteamiento de soluciones
 - Principales conceptos de los flujos de procesamiento de datos en sistemas de gran volumen
 - Fases principales de la gestión de grandes volúmenes de datos y retos asociados
 - Roles del ingeniero de datos en las fases principales de la gestión de datos
 - Limitaciones principales de los modelos tradicionales de gestión de los datos
 - Nuevos modelos de datos
- Planificación y ejecución de un trabajo de análisis de datos con una propuesta metodológica
 - Definición de un conjunto de datos de partida y una serie de necesidades de negocio que requieran una agregación de los datos, una captura de datos externa, un proceso ETL, análisis de datos y una visualización final de los resultados obtenidos
 - Implementación de un sistema de archivos distribuido
 - Uso de Hadoop para almacenar un conjunto de datos de actividad de red social. Almacenamiento de un conjunto de datos en un entorno HDFS

- Modelización de grafos: almacenar un conjunto de datos en una base de datos documental u orientada a grafos.
- Elección de un repositorio adecuado para los datos del problema y definición de una estrategia de almacenamiento.
 - Ciclo de vida de los datos: diseño de bases de datos, gestor de los flujos de datos, arquitectura de los sistemas de extracción, carga y transformación de los datos y sistemas de almacenamiento y procesamiento distribuido
 - Gestión de los datos: límites del modelo relacional y distribución de los datos

Habilidades de gestión, personales y sociales

- Efectividad en la resolución de problemas complejos en el desarrollo de un conocimiento para diseñar prototipos de soluciones software en Python en fases sin perder de vista la complejidad del problema global.
- Capacidad para analizar los elementos importantes del proyecto de desarrollo de una solución de gestión de datos.
- Desarrollo de pensamiento y razonamiento crítico de las diversas técnicas a aplicar en el marco del problema a resolver, haciendo balance entre complejidad de la solución y su funcionamiento real.
- Identificación de las herramientas a aplicar y su coste y las necesidades del ciclo de datos requerido.
- Desarrollo de una actitud positiva frente al aprendizaje y mejora continua, con el objetivo de conocer y revisar los proveedores de las herramientas y los métodos de instalación y actualización.
- Demostración de iniciativa y autonomía en la presentación de prototipos y discusión de problemas y soluciones para ser discutidos en grupo, revisando requerimientos y sus costes.

MÓDULO DE FORMACIÓN 2: GESTIÓN Y PROCESAMIENTO DE DATOS

OBJETIVO

Identificar los principios de gestión de datos para un proyecto con múltiples fuentes de entrada y aplicar técnicas de organización de modelos de datos desde un punto de vista lógico y físico.

DURACIÓN EN CUALQUIER MODALIDAD DE IMPARTICIÓN: 80 horas

Teleformación: Duración de las tutorías presenciales: 0 horas

RESULTADOS DE APRENDIZAJE

Conocimientos/ Capacidades cognitivas y prácticas

- Evaluación crítica de las metodologías y técnicas a aplicar en la resolución de problemas y justificación de los planteamientos, decisiones y propuestas realizadas
 - Fundamentos de gestión de los datos para un proyecto con múltiples fuentes de entrada de datos
 - Técnicas de organización de modelos de datos desde un punto de vista lógico y físico
- Identificación de los flujos de datos y ETL (Extract Transform Load)
 - Fundamentos de Data Warehousing y Business Intelligence
 - Conceptos de OLAP y extracción de información
 - Proceso ETL: extracción, transformación y carga de los datos

- Tipos de flujos y operaciones
 - Data cleaning
 - Data quality
 - Ejemplos de aplicaciones
- Diseño de un proceso ETL y un modelo de análisis multidimensional.
 - Modelización multidimensional
 - DFM: Dimensional Fact Model
 - Esquema en estrella y derivados
 - Operadores OLAP
 - Implementación de cubos y operadores OLAP en entornos relacionales
 - Herramientas de modelización multidimensional
- Diseño de una carga de datos a un repositorio NoSQL y análisis de los datos básico utilizando Spark
 - Diseño, implementación y mantenimiento de soluciones Fecha Lake. Conceptos y principios (schema-on-write vs. schema-on-read). Modelización y gobernanza de datos
 - Conceptos y principios de procesamiento distribuido de datos (soluciones declarativas vs. no declarativas)
 - Modelos de procesamiento distribuido de datos: Basados en disco y basados en memoria principal
 - MapReduce y a Apache Spark
 - Procesamiento de datos en tiempo real (streaming). Conceptos y principios (modelos, ventanas temporales, consultas temporales). Lenguajes de consultas sobre streams. Introducción a herramientas streaming: Apache Kafka, Apache Spark Streaming
 - Arquitecturas BigData: Lambda, Kappa y orquestadores. Herramientas de gestión de workflows: Apache Airflow
- Identificación de los factores clave de un problema complejo en el contexto de un proyecto de analítica.
 - Proyecto de diseño e implementación ETL con herramientas NoSQL
 - Proceso de incorporación de datos batch con herramientas Apache.
 - Análisis de datos y extracción de datos para modelo de negocio a partir del conjunto de datos con Spark
 - Análisis de datos con Apache Spark
 - Lectura y exportación de datos
 - Revisión de la calidad de los datos
 - Filtros y transformaciones de los datos
 - Procesamiento de los datos para obtener resúmenes y agrupaciones
 - Combinaciones, particiones y reformulación de los datos.
 - Configuración, monitorización y gestión de los errores de las aplicaciones Spark

Habilidades de gestión, personales y sociales

- Demostración de una actitud crítica de un pensamiento estratégico, presentando esquemas de tratamiento de los datos y permitiendo la discusión con grupos de interés internos y externos a la empresa para formular actuaciones orientadas al futuro.
- Desarrollo de las actividades de diseño y análisis de datos con responsabilidad social, honestidad intelectual e integridad científica.
- Concienciación de la necesidad de una actitud responsable y comprometida con los resultados y la limitación de los recursos disponibles en la toma de decisiones en entornos profesionales complejos.
- Valoración de la importancia de la adaptación a las restricciones de coste, disponibilidad, tiempo de desarrollo o implantación en la revisión de un diseño inicial de gestión de datos.

MÓDULO DE FORMACIÓN 3: APRENDIZAJE AUTOMÁTICO Y VISUALIZACIÓN

OBJETIVO

Aplicar los fundamentos de aprendizaje automático y de la visualización para el análisis de los resultados del procesamiento de datos.

DURACIÓN EN CUALQUIER MODALIDAD DE IMPARTICIÓN: 130 horas

Teleformación: Duración de las tutorías presenciales: 0 horas

RESULTADOS DE APRENDIZAJE

Conocimientos/ Capacidades cognitivas y prácticas

- Identificación de los fundamentos de análisis de datos y aprendizaje automático (Machine Learning)
 - Tipología de tareas y algoritmos de aprendizaje (supervisado, no supervisado, semi-supervisado)
 - Métodos principales de aprendizaje
 - Validación y evaluación de resultados
- Distinción de los métodos clasificadores.
 - Modelos predictivos
 - Métodos no supervisados. Agrupamiento jerárquico. Agrupamiento particional (k-means y derivados). Reducción de la dimensionalidad (PCA y otros)
 - Métodos supervisados. K-NN. Árboles de decisión. SVM. Redes neuronales
 - Validación y evaluación de resultados
- Aplicación de las técnicas de aprendizaje automático y la integración de diversas fuentes de datos
 - Análisis de sentimientos y polaridad sobre el conjunto de tweets recogidos.
 - Construcción de un análisis de perfiles mediante el uso de algoritmos de agrupamiento no supervisados (clustering).
 - Implementación de un análisis de polaridad (sentimiento analysis) sobre el conjunto de mensajes recogidos.
 - Implementación de dos enfoques alternativos para poder comparar el rendimiento obtenido: Aproximación basada en diccionarios. Aproximación en vectorización (Word2Vec) y uso de un modelo supervisado de aprendizaje automático.
- Diseño, desarrollo y evaluación de los métodos de aprendizaje automático.
 - Procesamiento de datos
 - Fundamentos de aprendizaje automático
 - Tipología de tareas y algoritmos de aprendizaje
 - Validación y evaluación de resultados
- Diseño y desarrollo de dashboards.
 - Principios de visualización de datos.
 - Diseño de paneles de control y dashboards para definir alarmas y transmitir resultados
 - Integración de la visualización con herramientas de análisis y consultas de datos
 - Documentación visual y escrita de los resultados de los proyectos de analítica de datos para audiencias no especializadas

- Utilización de una herramienta de visualización de datos para el diseño y carga de datos a un panel de control
 - Herramientas de visualización de datos: Grafana, MS PowerBar, Tableau
 - Visualización de consultas de negocio y panel de control de resultados en herramientas de visualización de datos
- Elección, aplicación y evaluación de la calidad de un algoritmo de aprendizaje automático para un problema dado a partir de un conjunto de datos.
 - Procesamiento de textos (NLP)
 - Análisis de polaridad basado en diccionarios
 - Análisis basado en modelos predictivos supervisados
 - Extracción de características (Word2Vec)

Habilidades de gestión, personales y sociales

- Uso de habilidades de comunicación con los grupos de interés para mostrar los aspectos más relevantes de los resultados obtenidos en los resultados del proceso y su adaptación a las necesidades del proyecto.
- Aplicación de soluciones innovadoras y adaptación a los entornos cambiantes.
- Capacidad de desarrollo continuo de proyectos y comunicación de los resultados y decisiones con técnicas y herramientas de visualización
- Coordinación y comunicación con especialistas, no especialistas, supervisores y clientes con el uso de las herramientas de comunicación para el diseño de información relevante sobre los aspectos claves de la aplicación.

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/as alumnos/as.
- La puntuación final alcanzada se expresará en términos de Apto/ No Apto.