

PROGRAMA FORMATIU

Data Scientist

Febrer 2022

G146NCEF-001-00



IDENTIFICACIÓ DE L'ESPECIALITAT I PARÀMETRES DEL CONTEXT FORMATIU

Denominació de l'especialitat:	DATA SCIENTIST
Família Professional:	INFORMÀTICA I COMUNICACIONS
Àrea Professional:	DESENVOLUPAMENT
Codi:	IFCD66
Nivell de qualificació professional:	5

Objectiu general

Extreure coneixement d'utilitat per a un propòsit concret a partir d'enormes volums de dades de fonts diverses disponibles en format digital.

Relació de mòduls de formació

Mòdul 1	Sistemes de suport a la presa de decisions i gestió de dades	100 hores
Mòdul 2	Gestió i processament de dades	80 hores
Mòdul 3	Aprenentatge automàtic i visualització	130 hores

Modalitats d'impartició

Presencial
Teleformació

Durada de la formació

Durada total en qualsevol modalitat d'impartició	310 hores
Teleformació	Durada total de las tutories presencials: 0 hores

Requisits d'accés de l'alumnat

Acreditacions/ titulacions	<p>Complir com a mínim alguns dels següents requisits:</p> <ul style="list-style-type: none"> - Títol de Grau o equivalent de la família professional Informàtica i Comunicacions. - Títol de Postgrau (Màster) o equivalent de la família professional Informàtica i Comunicacions.
Experiència professional	No es requereix.
Altres	<p>Han de tenir coneixements de:</p> <ul style="list-style-type: none"> - Programació bàsica - Disseny d'algoritmes - Metodologia SCRUM - Coneixements de llenguatge SQL per BBDD relacionals - Tenir coneixements bàsics d'àlgebra lineal, matemàtica discreta, probabilitat i estadística. Es valorarà acreditar coneixements de matèries com: Fonaments matemàtics, Probabilitat i estadística, Grafs i geometria discreta.

	Quan l'alumnat no disposi de l'acreditació o titulació requerida demostrarà els coneixements i les competències suficients mitjançant una prova competencial pràctica de nivell consistent en exercicis de programació d'aplicacions a nivell pràctic (bàsic), disseny d'algorismes, coneixements de llenguatge SQL per a BBDD relacionals i experiència en metodologies àgils de desenvolupament d'aplicacions.
Modalitat de teleformació	A més de l'indicat amb anterioritat, els alumnes han de tenir les destreses suficients per ser usuari de la plataforma virtual en la que es recolza l'acció formativa.

Prescripcions de formadors i tutors

Acreditació requerida	Complir com a mínim algun dels següents requisits: - Llicenciat, enginyer, arquitecte o el títol de grau corresponent o altres títols equivalents.
Experiència professional mínima requerida	No es requereix
Competència docent	Experiència docent o investigadora en l'àmbit de la disciplina acreditable d'almenys 60 hores en modalitat presencial.
Modalitat de teleformació	A més de complir amb les prescripcions establertes anteriorment, els tutors-formadors han d'acreditar una formació, de com a mínim 30 hores, o experiència, de com a mínim 60 hores, en aquesta modalitat i en la utilització de les tecnologies de la informació i comunicació.

Requisits mínims d'espai, instal·lacions i equipaments

Espais formatius	Superfície m ² per 15 alumnes	Increment Superfície/alumne (Màxim 30 alumnes)
Aula de gestió	45 m ²	2,4 m ² / alumne

Espai formatiu	Equipament
Aula de gestió	<ul style="list-style-type: none"> - Taula i cadires pel formador - Taula i cadires per l'alumnat - Material d'aula - Pissarra - Ordinador personal instal·lat en xarxa amb possibilitat d'impressió de documents, canó amb projecció i Internet per al formador - Ordinadors personals instal·lats en xarxa i Internet amb possibilitat d'impressió per als alumnes. - Programari específic per a l'aprenentatge de cada acció formativa: <ul style="list-style-type: none"> o Models de dades: HDFS, Apache Avro, Apache Parquet, Apache HBase, MongoDB, Neo4j, Cypher, GraphDB o Eines modelització multidimensional: PowerBI, Indyco o Eines ETL: Talend, Apache NiFi o Eines analítica de dades: Spark, Kafka, Airflow o Eines de visualització: Grafana, PowerBI, Tableau

La superfície dels espais i instal·lacions estaran en funció de la seva tipologia i del número d'alumnes. Tindran com a mínim els metres quadrats que s'indiquen per 15 alumnes i l'equipament suficient pels mateixos.

En el cas, de que augment el nombre d'alumnes, fins un màxim de 30, la superfície de les aules s'incrementarà proporcionalment (Segons s'indica a la taula en el relatiu a m²/alumne) i l'equipament estarà en consonància amb dit augment.

No s'ha d'interpretar que els diversos espais formatius identificats hagin de diferenciar-se necessàriament mitjançant tancaments.

Les instal·lacions i equipaments hauran de complir amb la normativa industrial i higiènic-sanitària corresponent i respondran a mesures d'accessibilitat i seguretat dels alumnes.

En el cas que la formació es dirigeixi a persones amb discapacitat, es realitzaran les adaptacions i ajustaments raonables per assegurar la seva participació en condicions d'igualtat.

A més, en el cas de **teleformació** s'ha de disposar del següent equipament.

Plataforma de teleformació:

La plataforma de teleformació que s'utilitzi per impartir accions formatives, haurà d'allotjar el material virtual d'aprenentatge corresponent, posseir capacitat suficient per desenvolupar el procés d'aprenentatge i gestionar i garantir la formació de l'alumnat, permetent la interactivitat i el treball cooperatiu, i reunir els següents requisits tècnics d'infraestructura, software i serveis:

- **Infraestructura**

- Tenir un rendiment, entès com número d'alumnes que suporti la plataforma, velocitat de resposta del servidor als usuaris, i temps de carrega de les pàgines web o de descarrega d'arxius, que permeti:

- a) Suportar un número d'alumnes equivalent al número total d'alumnes en les accions formatives de formació professional per l'ocupació que estigui impartint el centre o entitat de formació, garantint un allotjament mínim igual al total de l'alumnat de dites accions, considerant que el número màxim d'alumnes per tutor és de 80 i un número d'usuaris concurrents del 40% d'aquests alumnat.
- b) Disposar de la capacitat de transferència necessària perquè no es produeixin efecte

retardar en la comunicació audiovisual en temps real, havent de tenir el servidor en el que s'allotja la plataforma un ample de banda mínim de 300 Mbs, suficient en baixada i pujada.

- Estar en funcionament 24 hores al dia, els 7 dies de la setmana.
- **Software:**
 - Compatibilitat amb el estàndard SCORM i paquets de continguts IMS.
 - Nivells d'accessibilitat i interactivitat amb els continguts disponibles mitjançant tecnologies web, que com a mínim compleixin les prioritats 1 i 2 de la Norma UNE 139803:2012 o posteriors actualitzacions, segons l'estipulat al capítol III del Reial Decret 1494/2007, de 12 de novembre.
 - El servidor de la plataforma de teleformació ha de complir amb els requisits establerts a la Llei Orgànica 3/2018, de 5 de desembre, de Protecció de Dades Personals i Garantia dels Drets Digitals, pel que el responsable de dita plataforma ha d'identificar la localització física del servidor i el compliment de l'establert sobre transferències internacionals de dades als articles 40 a 43 de la citada Llei Orgànica 3/2018, de 5 de desembre, així com, en el que resulti d'aplicació, en el Reglament (UE) 2016/679 del Parlament Europeu i del Consell, de 27 d'abril de 2016, relatiu a la protecció de les persones físiques respecte del tractament de dades personals i la lliure circulació d'aquestes dades i pel que es deroga la Directiva 95/46/CE.
 - Compatibilitat tecnològica i possibilitats d'integració amb qualsevol sistema operatiu, base de dades, navegador d'Internet dels més usuals o servidor web, havent de ser possible utilitzar les funcions de la plataforma com complements (plug-in) i visualitzadors compatibles. Si es requereix la instal·lacions d'algun suport per funcionalitats avançades, la plataforma ha de facilitar l'accés al mateix sense cost.
 - Disponibilitat del servei web de seguiment (operatiu i en funcionament) de les accions formatives impartides, conforme al model de dades i protocol de transmissió establerts en l'annex V de l'Ordre TMS/369/2019, de 28 de març.
 - **Serveis i suport**
 - Sustentar el material virtual d'aprenentatge de l'especialitat formativa que a través d'ella s'imparteixi.
 - Disponibilitat d'un servei d'atenció a usuaris que doni suport tècnic i mantingui la infraestructura tecnològica i que, de forma estructurada i centralitzada, atengui i resolgui les consultes i incidències tècniques de l'alumnat. Les formes d'establir contacte amb aquest servei, que seran mitjançant telèfon i missatgeria electrònica, han d'estar disponibles per l'alumnat des de l'inici fins al finalitzar l'acció formativa, mantenint un horari de funcionament de matí tarda i un temps de demora en la resposta no superior a 48 hores laborables.
 - Personalització amb la imatge institucional de l'administració laboral corresponent, amb les pautes d'imatge corporativa que s'estableixin.

Amb l'objecte de gestionar, administrar, organitzar, dissenyar, impartir i avaluar accions formatives a través d'Internet, la plataforma de teleformació integrarà les eines i recursos necessaris a tal fi, disposant, específicament d'eines de:

- Comunicació, que permetin que cada alumne pugui interaccionar a través del navegador amb el tutor-formador, el sistema i la resta d'alumnes. Aquesta comunicació electrònica ha de portar-se a terme mitjançant eines de comunicació síncrones (aula virtual, xat, pissarra

electrònica) i asíncrones (correu electrònic, fòrum, calendari, tauló d'anuncis, avisos). Serà obligatori que cada acció formativa en modalitat de teleformació disposi, com a mínim d'un servei de missatgeria, un fòrum i xat.

- Col·laboració, que permeti tant el treball cooperatiu entre els membres d'un grup, com la gestió de grups. Mitjançant les eines ha de ser possible realitzar operacions d'alta, modificació o esborrament de grups d'alumnes, així com la creació d' "escenaris virtuals" pel treball cooperatiu dels membres d'un grup (directoris o "carpetes" per l'intercanvi d'arxius, eines per la publicació dels continguts, i fòrums o xats privats pels membres de cada grup).
- Administració, que permeti la gestió d'usuaris (altes, modificacions, esborrat, gestió de la llista de classe, definició, assignació i gestió de permisos, perfils i rols, autenticació i assignació de nivells de seguretat) i la gestió d'accions formatives.
- Gestió de continguts, que possibiliten l'emmagatzematge i la gestió d'arxius (visualitzar arxius, organitzar-los en carpetes – directoris- i subcarpetes, copiar, enganxar, eliminar, comprimir, descarregar o carregar arxius), la publicació organitzada i selectiva dels continguts de dits arxius i la creació de continguts.
- Avaluació i control de progrés de l'alumnat, que permeti la creació, edició i realització de proves d'avaluació i autoavaluació i de les activitats i treballs avaluables, la seva autocorrecció o correcció (amb retroalimentació), la seva qualificació, assignació de puntuacions i la ponderació de les mateixes, el registre personalitzar i la publicació de qualificacions, la visualització d'informació estadística sobre els resultats i el progrés de cada alumne i l'obtenció d'informes de seguiment.

Material virtual d'aprenentatge:

El material virtual d'aprenentatge per l'alumnat mitjançant el que s'imparteixi la formació es concretarà en el curs complet en format multimèdia (que mantingui una estructura i funcionalitat homogènia), havent d'ajustar-se a tots els elements de la programació (objectius i resultats d'aprenentatge) d'aquest programa formatiu que figura en el Catàleg d'Especialitats Formatives i contingut el qual, compleixi aquests requisits:

- Com a mínim, ser l'establert en el citat programa formatiu del Catàleg d'Especialitats Formatives.
- Referit tant als objectius com als coneixements/capacitats cognitives i pràctiques, i habilitats de gestió, personals i socials, de manera que en el seu conjunt permetin aconseguir les resultats d'aprenentatge previstos.
- Organitzar-se a través d'índex, mapes, taules de continguts, esquemes, epígrafs o titulars de fàcil discriminació i seqüenciar-se pedagògicament de tal manera que permetin la seva comprensió i retenció.
- No ser merament informatiu, promovent la seva aplicació pràctica a través d'activitats d'aprenentatge (autoavaluables o valorades pel tutor-formador) rellevants per l'adquisició de competències, que serveixi per verificar el progrés de l'aprenentatge de l'alumnat, fer un seguiment de les seves dificultats d'aprenentatge i donar el suport adequat.
- No ser exclusivament textuals, incloent recursos variats (necessaris i rellevants), tant estàtics com interactius (imatges, gràfics, àudio, vídeo, animacions, enllaços, simulacions, articles, fòrums, xat, etc) de forma periòdica.
- Poder ser ampliat o complementat mitjançant diferents recursos addicionals als que l'alumnat pugui accedir i consultar a voluntat.
- Donar lloc a resums o síntesi i a glossaris que identifiquin i defineixin els termes o vocabulari bàsic, rellevant o claus per la comprensió dels aprenentatges.
- Avaluar l'adquisició durant i a la finalització de l'acció formativa a través d'activitats d'avaluació (exercicis, preguntes, treballs, problemes, cassos, proves, etc.), que permetin mesurar el rendiment o desenvolupament de l'alumnat.

Aula virtual

Tecnologia i equips	Plataforma d'aprenentatge que permeti la connexió síncrona de docents i alumnes, amb sistema incorporat d'àudio, vídeo i possibilitat de compartir arxius, la mateixa pantalla o altres aplicacions tant pel docent com per els / les alumnes, amb registre dels temps de connectivitat.
----------------------------	--

Ocupacions i llocs de treball relacionats

- 2712 Analistes i dissenyadors de programari
- 2713 Analistes, programadors i dissenyadors web i multimèdia
- 2719 Analistes i dissenyadors de programari i multimèdia no classificats sota altres epígrafs

Requisits oficials de les entitats o centres de formació

Estar inscrit en el Registre d'entitats de formació (Serveis Públics d'Ocupació)

DESENVOLUPAMENT MODULAR

MÒDUL DE FORMACIÓ 1: SISTEMES DE SUPORT A LA PRESA DE DECISIONS I
GESTIÓ DE DADES

OBJECTIU

Desenvolupar aplicacions informàtiques per a realitzar un tractament de les dades bàsic amb el llenguatge Python, identificant els mètodes per a l'exploració de grans fons de dades i els sistemes de gestió de dades relacionals i no relacionals (NoSQL).

DURADA EN QUALSEVOL MODALITAT D'IMPARTICIÓ: 100 hores

Teleformació: Durada de les tutories presencials: 0 hores

RESULTATS D'APRENTATGE

Coneixements/ Capacitats cognitives i pràctiques

- Caracterització de l'aplicació del llenguatge Python
 - Llenguatge Python
 - Execució de programes Python
 - Objectes en Python
 - Tipus numèrics i dinàmics
 - Gestió de cadenes de text: llistes, diccionaris, tuples i fitxers
 - Sentències Python: assignacions, expressions i imprimir resultats
 - Tests de variables, regles de sintaxi
 - Bucles for i while
- Interpretació l'aplicació de protocols API
 - Ús d'APIs remotes
 - Integració de les aplicacions amb APIs remotes
 - Exemples d'aplicació d'APIs remotes en llenguatge Python
- Programació d'un algoritme modular en llenguatge Python
 - Programació de mòduls
 - Fonaments de programació de classes
 - Utilització d'APIs i integració amb aplicacions Python
- Distinció dels conceptes Cloud bàsics
 - Principis de computació en el núvol (Cloud Computing)
 - Enginyeria de serveis: software as a service, Platform as a service, Infrastructure as a Service
 - Exemples d'aplicacions rellevants en la indústria
- Ús de BBDD NoSQL i nous models de dades (estructurats i no estructurats)
 - Fonaments de l'paradigma NoSQL
 - Distribució de les dades i processament en paral·lel
 - Principals models de dades en el món NoSQL: clau-valor, orientació a documents, grafs de propietat, grafs de coneixement
- Coneixement de l'emmagatzematge Big Data i les eines de processament massiu
 - Aplicacions basades en la gestió i l'anàlisi de grans volums de dades

- Fonaments arquitectònics dels sistemes distribuïts
 - Principals arquitectures de referència
 - Nous models de dades
 - Sistemes de fitxers distribuïts
 - Document stores
 - Bases de dades de grafs
- Avaluació de les metodologies i tècniques aplicades en la resolució de problemes i justificació dels plantejaments, decisions i propostes realitzades
 - Sistemes de suport a la presa de decisions
 - Anàlisi de les dades: anàlisi descriptiva, predictiu i prescriptiu
 - Casos d'ús: gestió i anàlisi de grans volums de dades
- Identificació dels factors clau d'un problema complex en el context d'un projecte d'anàlisi.
 - Context de la societat / economia de les dades i el paradigma de les aplicacions orientades a les dades
 - Fonaments de bases de dades relacionals: llenguatge SQL.
 - Necessitat d'un canvi de paradigma: NoSQL. El principi 'one size does not fit all'.
 - Principals models de dades en el món NoSQL: Key-Value, Document-oriented, Property Graphs i Knowledge Graphs
 - Fonaments arquitectònics: sistemes distribuïts, escalabilitat, paral·lelisme. Principals arquitectures de referència (shared nothing, shared disk, shared memory)
- Distinció i aplicació dels nous models de dades
 - Sistemes d'arxius distribuïts: conceptes i principis (distribució, replicació, particionament horitzontal vs. Vertical, formats d'arxius especialitzats)
 - Coneixement i utilització de Hadoop File System (HDFS), Apache Avro, Apache Parquet, Key-value stores: Apache HBase
 - Document stores: conceptes i principis (mecanismes de rèplica, sharding, consultes espacials)
 - Immersió a MongoDB i el Aggregation Framework
 - Graph databases: property i knowledge graphs. Conceptes i principis Modelització en graf, consultes regulars. Introducció a Neo4j i Cypher
 - Knowledge graphs. Conceptes i principis: el paradigma open / linked data, RDF i SPARQL. Introducció a GraphDB
- Identificació i anàlisi de problemes complexos en l'àrea d'anàlisi de dades i plantejament de solucions
 - Principals conceptes dels fluxos de processament de dades en sistemes de gran volum
 - Fases principals de la gestió de grans volums de dades i reptes associats
 - Rols de l'enginyer de dades en les fases principals de la gestió de dades
 - Limitacions principals dels models tradicionals de gestió de les dades
 - Nous models de dades
- Planificació i execució d'un treball d'anàlisi de dades amb una proposta metodològica
 - Definició d'un conjunt de dades de partida i un seguit de necessitats de negoci que requereixin una agregació de les dades, una captura de dades externa, un procés ETL, anàlisi de dades i una visualització final dels resultats obtinguts

- Implementació d'un sistema d'arxius distribuït
 - Ús de Hadoop per emmagatzemar un conjunt de dades d'activitat de xarxa social.
 - Emmagatzematge d'un conjunt de dades en un entorn HDFS
 - Modelització de grafs: emmagatzemar un conjunt de dades en una base de dades documental o orientada a grafs.
- Elecció d'un repositori adequat per a les dades de el problema i definició d'una estratègia d'emmagatzematge.
 - Cicle de vida de les dades: disseny de bases de dades, gestor dels fluxos de dades, arquitectura dels sistemes d'extracció, càrrega i transformació de les dades i sistemes d'emmagatzematge i processament distribuït
 - Gestió de les dades: límits de el model relacional i distribució de les dades

Habilitats de gestió, personals i socials

- Efectivitat en la resolució de problemes complexos en el desenvolupament d'un coneixement per dissenyar prototips de solucions programari en Python en fases sense perdre de vista la complexitat de el problema global.
- Capacitat per analitzar els elements importants del projecte de desenvolupament d'una solució de gestió de dades.
- Desenvolupament de pensament i raonament crític de les diverses tècniques a aplicar en el marc del problema a resoldre, fent balanç entre complexitat de la solució i el seu funcionament real.
- Identificació de les eines a aplicar i el seu cost i les necessitats del cicle de dades requerit.
- Desenvolupament d'una actitud positiva davant l'aprenentatge i millora contínua, amb l'objectiu de conèixer i revisar els proveïdors de les eines i els mètodes d'instal·lació i actualització.
- Demostració d'iniciativa i autonomia en la presentació de prototips i discussió de problemes i solucions per a ser discutits en grup, revisant requeriments i els seus costos.

MÒDUL DE FORMACIÓ 2: GESTIÓ I PROCESSAMENT DE DADES**OBJECTIU**

Identificar els principis de gestió de dades per a un projecte amb múltiples fonts d'entrada i aplicar tècniques d'organització de models de dades des d'un punt de vista lògic i físic.

DURADA EN QUALESEVOL MODALITAT D'IMPARTICIÓ: 80 hores

Teleformació: Durada de les tutories presencials: 0 hores

RESULTATS D'APRENTATGE**Coneixements/ Capacitats cognitives i pràctiques**

- **Avaluació crítica de les metodologies i tècniques a aplicar en la resolució de problemes i justificació dels plantejaments, decisions i propostes realitzades**
 - Fonaments de gestió de les dades per a un projecte amb múltiples fonts d'entrada de dades
 - Tècniques d'organització de models de dades des d'un punt de vista lògic i físic
- **Identificació dels fluxos de dades i ETL (Extract Transform Load)**
 - Fonaments de Data Warehousing i Business Intelligence
 - Conceptes d'OLAP i Extracció d'informació
 - Procés ETL: Extracció, Transformació i Càrrega de les dades
 - Tipus de fluxos i Operacions
 - Data cleaning
 - Data quality
 - Exemples d'aplicacions
- **Disseny d'un procés ETL i un model d'anàlisi multidimensional.**
 - Modelització multidimensional
 - DFM: Dimensional Fact Model
 - Esquema en estrella i derivats
 - Operadors OLAP
 - Implementació de cubs i operadors OLAP en entorns relacionals
 - Eines de modelització multidimensional
- **Disseny d'una càrrega de dades a un repositori NoSQL i anàlisi de les dades bàsic utilitzant Spark**
 - Disseny, implementació i manteniment de solucions Data Lake. Conceptes i principis (esquema-on-write vs. schema-on-read). Modelització i governança de dades
 - Conceptes i principis de processament distribuït de dades (solucions declaratives vs. no declaratives)
 - Models de processament distribuït de dades: Basats en disc i basats en memòria principal
 - MapReduce i a Apache Spark
 - Processament de dades en temps real (streaming). Conceptes i principis (models, finestres temporals, consultes temporals). Llenguatges de consultes sobre streams. Introducció a eines streaming: Apache Kafka, Apache Spark Streaming
 - Arquitectures BigData: Lambda, Kappa i orquestradors. Eines de gestió de workflows: Apache Airflow

- Identificació dels factors clau d'un problema complex en el context d'un projecte d'anàlisi.
 - Projecte de disseny i implementació ETL amb eines NoSQL
 - Procés d'Incorporació de dades batch amb eines Apache.
 - Anàlisi de dades i extracció de dades per a model de Negoci a partir del conjunt de dades amb Spark
 - Anàlisi de dades amb Apache Spark
 - Lectura i Exportació de dades
 - Revisió de la qualitat de les dades
 - Filtres i transformacions de les dades
 - Processament de les dades per obtenir resums i agrupacions
 - Combinacions, particions i reformulació de les dades.
 - Configuració, monitorització i gestió dels errors de les aplicacions Spark

Habilitats de gestió, personals i socials

- Demostració d'una actitud crítica d'un pensament estratègic, presentant esquemes de tractament de les dades i permetent la discussió amb grups d'interès interns i externs a l'empresa per formular actuacions orientades a el futur.
- Desenvolupament de les activitats de disseny i anàlisi de dades amb responsabilitat social, honestat intel·lectual i integritat científica.
- Conscienciació de la necessitat d'una actitud responsable i compromesa amb els resultats i la limitació dels recursos disponibles a la presa de decisions en entorns professionals complexos.
- Valoració de la importància de l'adaptació a les restriccions de cost, disponibilitat, temps de desenvolupament o implantació en la revisió d'un disseny inicial de gestió de dades.

MÒDUL DE FORMACIÓ 3: APRENTATGE AUTOMÀTIC I VISUALITZACIÓ

OBJECTIU

Aplicar els fonaments d'aprenentatge automàtic i de la visualització per a l'anàlisi dels resultats del processament de dades.

DURADA EN QUALSEVOL MODALITAT D'IMPARTICIÓ: 130 hores

Teleformació: Durada de les tutories presencials: 0 hores

RESULTATS D'APRENTATGE

Coneixements/ Capacitats cognitives i pràctiques

- Identificació dels fonaments d'anàlisi de dades i aprenentatge automàtic (Machine Learning)
 - Tipologia de tasques i algorismes d'aprenentatge (supervisat, no supervisat, semi-supervisat)
 - Mètodes principals d'aprenentatge
 - Validació i avaluació de resultats

- Distinció dels mètodes classificadors.
 - Models predictius
 - Mètodes no supervisats. Agrupament jeràrquic. Agrupament de partició (k-means i derivats). Reducció de la dimensionalitat (PCA i altres)
 - Mètodes supervisats. K-NN. Arbres de decisió. SVM. Xarxes neuronals
 - Validació i avaluació de resultats
- Aplicació de les tècniques d'aprenentatge automàtic i la integració de diverses fonts de dades
 - Anàlisi de sentiments i polaritat sobre el conjunt de tweets recollits.
 - Construcció d'una anàlisi de perfils mitjançant l'ús d'algoritmes d'agrupament no supervisats (clustering).
 - Implementació d'una anàlisi de polaritat (sentiment analysis) sobre el conjunt de missatges recollits.
 - Implementació de dos enfocaments alternatius per poder comparar el rendiment obtingut: Aproximació basada en diccionaris. Aproximació a vectorització (Word2Vec) i ús d'un model supervisat d'aprenentatge automàtic.
- Disseny, desenvolupament i avaluació dels mètodes d'aprenentatge automàtic.
 - Processament de dades
 - Fonaments d'aprenentatge automàtic
 - Tipologia de tasques i algoritmes d'aprenentatge
 - Validació i avaluació de resultats
- Disseny i desenvolupament de dashboards.
 - Principis de visualització de dades.
 - Disseny de panells de control i dashboards per definir alarmes i transmetre resultats
 - Integració de la visualització amb eines d'anàlisi i consultes de dades
 - Documentació visual i escrita dels resultats dels projectes d'anàlisi de dades per a audiències no especialitzades
- Utilització d'una eina de visualització de dades per al disseny i càrrega de dades a un panell de control
 - Eines de visualització de dades: Grafana, MS PowerBar, Tableau
 - Visualització de consultes de negoci i panell de control de resultats en eines de visualització de dades
- Elecció, aplicació i avaluació de la qualitat d'un algoritme d'aprenentatge automàtic per a un problema donat a partir d'un conjunt de dades.
 - Processament de textos (NLP)
 - Anàlisi de polaritat basat en diccionaris
 - Anàlisi basat en models predictius supervisats
 - Extracció de característiques (Word2Vec)

Habilitats de gestió, personals i socials

- Ús d'habilitats de comunicació amb els grups d'interès per mostrar els aspectes més rellevants dels resultats obtinguts en els resultats del procés i la seva adaptació a les necessitats de el projecte.
- Aplicació de solucions innovadores i adaptació als entorns canviants.
- Capacitat de desenvolupament continu de projectes i comunicació dels resultats i decisions amb tècniques i eines de visualització.
- Coordinació i comunicació amb especialistes, no especialistes, supervisors i clients amb

l'ús de les eines de comunicació per al disseny d'informació rellevant sobre els aspectes claus de l'aplicació.

AVALUACIÓ DE L'APRENTATGE EN L'ACCIÓ FORMATIVA

- L'avaluació tindrà un caràcter teòric-pràctic i es realitzarà de forma sistemàtica i continua, durant el desenvolupament de cada mòdul i al final del curs.
- Pot incloure's una avaluació inicial i de caràcter diagnòstic per detectar el nivell de partida de l'alumnat.
- L'avaluació es portarà a terme mitjançant els mètodes i instruments més adequats per comprovar els diversos resultats d'aprenentatge, i que garanteixin la fiabilitat i validesa de la mateixa.
- Cada instrument d'avaluació s'acompanyarà del seu corresponent sistema de correcció i puntuació en el que s'expliciti, de forma clara i inequívoca, els criteris de mesura per avaluar els resultats assolits pels alumnes.
- La puntuació final assolida s'expressarà en termes d'Apte/No apte.