

Titulació	Tipus	Curs
2503758 Enginyeria de Dades	OB	3

## Professor/a de contacte

Nom: Ramon Baldrich Caselles

Correu electrònic: ramon.baldrich@uab.cat

## Idiomes dels grups

Podeu consultar aquesta informació al [final](#) del document.

## Prerequisits

Es fonamental haver adquirit una bona base matemàtica així com tenir un bon nivell de programació, principalment en Python.

## Objectius

L'assignatura pretén introduir els conceptes de la intel·ligència artificial que es basa en l'obtenció del coneixent, conceptes i tendències a partir de les dades. Es tracta de formar a l'alumne per ser un "enginyer de dades", i és una de les professions amb més futur i més demandades en l'actualitat per les grans empreses i start-ups tecnològiques. De fet, es preveu que el creixement de la demanda d'aquests professionals en enginyeria de dades sigui exponencial a nivell europeu, sobretot degut al creixement en la generació de dades massives. Així, el principal objectiu de l'assignatura és que l'alumne sàpiga trobar una bona solució (a vegades la millor és impossible) a problemes en contextes diferents dels tractats, a partir d'identificar les necessitats de representació del coneixement i, segons sigui aquest, aplicar la/les tècnica/ques més adequada/es per generar automàticament bons models matemàtics que expliquin les dades amb un error acceptable.

Els continguts escollits per aquesta assignatura són les tècniques i conceptes que s'usen extensivament en la indústria, entenent-la en el seu concepte més ampli. La base algorítmica serà fonamental durant el desenvolupament de l'assignatura que vol tenir un enfocament eminentment enginyeril, focalitzant-se en el ús dels algorismes i tècniques que es mostren sense deixar de banda la comprensió dels fonaments matemàtics que les sustenten. Els algorismes i tècniques que es mostren són la base fonamental per a l'aprenentatge computacional 'tradicional' sense el que no es pot entendre les tècniques que es desenvoluparan en propers cursos. No per bàsiques són obsoletes, ans al contrari, cobreixen un gran ventall d'aplicacions i problemes on són fonamentals. L'alumne ha de ser conscient que aquest coneixement que és punta de llança de l'estat de l'art té una dificultat inherent, implicant un estudi i una dedicació considerables, quantificada en hores en la secció d'Activitats formatives d'aquesta guia. Això és perquè en aquesta assignatura no només s'ensenyen una part dels continguts més importants en matèria d'aprenentatge automàtic per esdevenir enginyer de dades, sinó a més es treballa una línia de currículum que permeti ampliar el ventall de llocs de treball als que podreu accedir després de la carrera, així com posar les bases metodològiques necessàries per fer un Màster en enginyeria de dades o en intel·ligència artificial.

Els objectius de l'assignatura es poden resumir en:

Coneixements:

- Descriure les tècniques bàsiques d'aprenentatge computacional.
- Enumerar els passos essencials dels diferents algorismes d'aprenentatge
- Identificar els avantatges i inconvenients dels algorismes d'aprenentatge que s'expliquen.
- Resoldre problemes computacionals aplicant diferents tècniques d'aprenentatge per trobar la solució òptima.
- Entendre el resultat i les limitacions de les tècniques d'aprenentatge en diferents casos d'estudi.
- Saber escollir l'algorisme d'aprenentatge més adequat per solucionar problemes contextualitzats.

Habilitats:

- Reconèixer les situacions en les quals l'aplicació d'algorismes d'aprenentatge computacional pot ser adient per solucionar un problema
- Analitzar el problema a resoldre i dissenyar la solució òptima aplicant les tècniques apreses
- Redactar documents tècnics relacionats amb l'anàlisi i la solució d'un problema
- Programar els algorismes bàsics per solucionar els problemes proposats
- Avaluar els resultats de la solució implementada i valorar les possibles millores
- Defensar i argumentar les decisions preses en la solució dels problemes proposats

## Competències

- Analitzar les dades de forma eficient per al desenvolupament de sistemes intel·ligents amb capacitat d'aprenentatge autònom i/o per mineria de dades.
- Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins de la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes destacats d'índole social, científica o ètica.
- Treballar cooperativament, en entorns complexos o incerts i amb recursos limitats, en un context multidisciplinari, assumint i respectant el rol dels diferents membres de l'equip.

## Resultats d'aprenentatge

1. Decidir el mètode d'aprenentatge de dades més adequat segons les característiques de les dades que cal analitzar.
2. Escollir l'algoritme de cerca i el paradigma de programació per a un problema d'optimització de paràmetres o estats.
3. Que els estudiants tinguin la capacitat de reunir i interpretar dades rellevants (normalment dins de la seva àrea d'estudi) per emetre judicis que incloguin una reflexió sobre temes destacats d'índole social, científica o ètica.
4. Treballar cooperativament, en entorns complexos o incerts i amb recursos limitats, en un context multidisciplinari, assumint i respectant el rol dels diferents membres de l'equip.

## Continguts

## Continguts

## TEMA 1: INTRODUCCIÓ

- 1.1 Conceptes bàsics i els paradigmes bioinspirats
- 1.2 Història de l'aprenentatge computacional

## TEMA 2: REGRESSIÓ I CLASSIFICACIÓ

- 2.1 Regressió de dades numèriques: descens del gradient
- 2.2 Regularització i regressió logística
- 2.3 Classificació de dades numèriques: màquines de vectors de suport
- 2.4 Arbres de Decissió i Random Forest
- 2.5 Classificació Bayesiana

## TEMA 3: AGRUPACIÓ I CERCA

- 3.1 Memorització: aprenentatge mandrós
- 3.2 Sistemes recomanadors: Content-based vs. Collaborative filtering
- 3.3 Clustering: k-means i Expectation-Maximization

## Activitats formatives i Metodologia

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Contingut teoric	22	0,88	2, 1, 3
Tipus: Supervisades			
Classes pràctiques	16	0,64	2, 1, 3, 4
serminaris	10	0,4	2, 1
Tipus: Autònomes			
Preparació i realització dels projectes pràctics	52	2,08	2, 1, 3, 4
estudi	28	1,12	2, 1, 3

Tota la informació de l'assignatura i els documents relacionats que els estudiants necessitin es trobaran a la pàgina de Campus Virtual (<http://cv.uab.cat/>).

Les diferents activitats que es duran a terme en l'assignatura s'organitzen de la següent manera:

### Classes de teoria

S'exposaran els principals conceptes i algorismes de cada tema de teoria. Aquests temes suposen el punt de partida en el treball de l'assignatura.

## Seminaris de problemes

Seràn classes amb grups reduïts d'estudiants, que facilitin la interacció, o de caràcter individual, segons els casos. En aquestes classes es plantejaran casos pràctics que requereixin el disseny d'una solució en la qual es facin servir els mètodes vistos a les classes de teoria. És impossible seguir les classes de problemes si no se segueixen els continguts de les classes de teoria. El resultat d'aquestes sessions és la resolució dels problemes que s'hauran d'entregar de forma setmanal. El mecanisme específic per a l'entrega, així com el mecanisme d'avaluació, s'indicarà en la pàgina web de l'assignatura.

Aquest seminari de problemes seran tots ells pràctics i inclouran la programació d'una solució al problema plantejat. Aquesta solució s'haurà d'entregar setmanalment creant un portfoli de treball. Les entregues no són recuperables.

### Projecte en Grup:

es disposarà de tres setmanes mínim per realitzar un projecte de més abast on posar en pràctica el que s'haurà practicat en les sessions de problemes.

Tant l'avaluació de teoria com el treball grupal seran recuperables.

Nota: es reservaran 15 minuts d'una classe, dins del calendari establert pel centre/titulació, per a la complementació per part de l'alumnat de les enquestes d'avaluació de l'actuació del professorat i d'avaluació de l'assignatura/mòdul.

## Avaluació

### Activitats d'avaluació continuada

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Defensa de projecte	20%	5	0,2	2, 3, 4
Portfoli problemes	10	5	0,2	2, 1, 3
Projecte en grup	40	5	0,2	2, 1, 3, 4
Proves individuals	30	7	0,28	2, 1, 3

### Activitats i instruments d'avaluació:

Aquesta assignatura no preveu el sistema d'avaluació única.

Per avaluar l'assoliment dels coneixements i competències associats a l'assignatura s'estableix un mecanisme d'avaluació que combina l'assimilació dels coneixements, la capacitat de resolució de problemes, i de forma significativa, la capacitat de generar solucions computacionals a problemes complexos, tant grupal com individualment.

Amb aquest objectiu es divideix l'avaluació en tres parts:

- Avaluació de continguts

La nota final de continguts es calcularà a partir de varis exàmens parcials:

$$\text{Nota Continguts} = 1/N * \text{Prova}_i$$

El número de proves pot variar i s'establiran a l'inici de curs. Per a poder tenir una nota de continguts caldrà que les notes dels cada un dels proves sigui superior a 4.

Les proves parcials es faran durant el curs i eminentment seran de contingut conceptual on respondre a diferents preguntes sobre el contingut desenvolupat a les sessions 'teoriques'.

Aquestes proves pretenen ser una avaluació individualitzada de l'estudiant amb les seves capacitats d'entendre les tècniques explicades a classe així com avaluar el nivell de conceptualització que l'estudiant n'ha fet de les tècniques vistes.

Tests de recuperació. En cas que la nota de continguts no arribi al nivell adequat en alguna de les proves, per obtenir una nota final suficient per considerar l'assoliment dels coneixements, els estudiants es poden presentar a l'examen de la convocatòria de l'assignatura i tornar a fer un examen que avaluï els continguts vists a l'assignatura de la/les part/s no superada/es.

No hi ha convalidacions en cas que s'hagués superat la part teòrica en anys anteriors.

– Avaluació del treball en els seminaris de problemes

Els problemes tenen com objectiu provocar que l'estudiant entri en els continguts de l'assignatura de manera continuada i a partir de petits problemes que facin que es familiaritzi directament en l'aplicació de la teoria. Com a evidència d'aquest treball es demana la presentació d'un portfoli en el que haurà anat guardant els problemes que haurà anat realitzant. Aquest portfoli tindrà entrega setmanal digital. L'alumne podrà autoavaluar-se continuament ja que disposarà de les solucions de cada un dels conjunts de problemes un cop finalitzat el període d'entrega. Juntament amb les hores de tutoria per si apareixen dubtes, és suficient per a que cada alumne identifiqui els seus punts febles.

– Avaluació del projecte en grup.

En les darreres setmanes del semestre es realitzarà un projecte de més abast que els exercicis fets durant el curs en les sessions de problemes. Aquest projecte s'avaluarà tant grupal com individualment. Els mecanismes d'avaluació seran codi, informe, presentació i seguiment del projecte en les sessions de classe assignades.

La nota final de l'assignatura s'obté combinant l'avaluació d'aquestes 3 activitats de la manera següent:

$$\text{Nota Final} = (0.3 * \text{Continguts}) + (0.1 * \text{portfoli de problemes}) + (0.6 * \text{Projecte})$$

el projecte tindrà una nota de la seva defensa i una altra del seu desenvolupament i profunditat de la solució

Condicions per a superar l'assignatura:

La nota final de continguts ha de ser més gran o igual que 4 per poder aprovar l'assignatura.

La nota del projecte i la seva defensa ha de ser més gran o igual que 6 per poder aprovar l'assignatura.

En cas que la nota, aplicant la fórmula de l'apartat anterior ("nota final de l'assignatura"), fos superior a 5 però no s'hagués superat el mínim exigint en alguna de les parts, la nota final en l'expedient serà un 4,5.

S'assignaran tantes matrícules d'honor com la normativa vigent permeti sempre i quan la nota sigui superior a 9.0. L'assignació de les matrícules es farà seguint l'ordre de notes. En cas d'haver-hi múltiples candidats amb la mateixa avaluació susceptibles de rebre Md'H es proposaran activitats suplementàries per a determinar el/s millor/s candidat/s.

L'alumne es qualificarà com "No Avaluable" si no té cap part avaluada ni dels continguts teòrics ni dels continguts pràctics.

En cada publicació de notes s'especificaran els mecanismes a seguir per a la recuperació del contingut avaluat si es dona el cas que aquest contingut sigui re-avaluable.

Avisos importants:

- Les dates d'avaluació continuada i lliurament de treballs, així com tot el material docent es publicaran al campus virtual (<http://cv.uab.cat/>), a l'espai d'aquesta assignatura i poden estar subjectes a canvis de programació per motius d'adaptació a possibles incidències. Sempre s'informarà a [cerbero.uab.cat](mailto:cerbero.uab.cat) sobre aquests canvis ja que s'entén que Caronte esdevindrà el mecanisme habitual d'intercanvi d'informació entre professor i estudiants.
- Per a cada activitat d'avaluació, s'indicarà un lloc, data i hora de revisió en la que l'estudiant podrà revisar l'activitat amb el professor. En aquest context, es podran fer reclamacions sobre la nota de l'activitat, que seran avaluades pel professorat responsable de l'assignatura. Si l'estudiant no es presenta a aquesta revisió, no es revisarà posteriorment aquesta activitat.
- Sense perjudici d'altres mesures disciplinàries que s'estimin oportunes, i d'acord amb la normativa acadèmica vigent, les irregularitats comeses per un estudiant que puguin conduir a una variació de la qualificació en una activitat avaluable es qualificaran amb un zero (0). Les activitats d'avaluació qualificades d'aquesta forma i per aquest procediment no seran recuperables. Si és necessari superar qualsevol d'aquestes activitats d'avaluació per aprovar l'assignatura, aquesta assignatura quedarà suspesa directament, sense oportunitat de recuperar-la en el mateix curs. Aquestes irregularitats inclouen, entre d'altres:

- la còpia total o parcial d'una pràctica, informe, o qualsevol altra activitat d'avaluació;
- deixar copiar;
- presentar un treball de grup no fet íntegrament pels membres del grup (aplicat a tots els membres, no solament els que no han treballat);
- presentar com a propis materials elaborats per un tercer, encara que siguin traduccions o adaptacions, i en general treballs amb elements no originals i exclusius de l'estudiant;
- tenir dispositius digitals i/o de comunicació (com telèfons mòbils, smart watches, bolígrafs amb càmera, etc.) accessibles durant les proves d'avaluació teòrico-pràctiques individuals (exàmens).
- parlar amb companys durant les proves d'avaluació teòrico-pràctiques individuals (exàmens).
- observar/mirar les proves d'avaluació teòrico-pràctiques (exàmens) d'altres companys durant la realització de la mateixa, encara que no s'hagi procedit a la còpia.
- observar/mirar en la taula, fulls, paret etc escrits relacionats amb la matèria durant la realització de les proves d'avaluació teòrico-pràctiques (exàmens) encara que no s'hagi procedit a la còpia.

La nota numèrica de l'expedient serà el valor menor entre 3.0 i la mitjana ponderada de les notes en cas que l'estudiant hagi comès irregularitats en un acte d'avaluació (i per tant no serà possible l'aprovat per compensació). En resum: copiar, deixar copiar o plagiar (o l'intent de) en qualsevol de les activitats d'avaluació equival a un SUSPENS amb nota inferior a 3,5.

## Bibliografia

Enllaços web

- Pàgina web de l'assignatura : <http://cv.uab.cat>
- Artificial Intelligence: A Modern Approach. <http://aima.cs.berkeley.edu/>

## Bibliografia bàsica

- S. Russell, P. Norvig. Artificial Intelligence: A Modern Approach. Ed. Prentice Hall, Second Edition, 2003. (Existeix traducció al castellà: Inteligencia artificial: Un Enfoque Moderno)
- T. Mitchell. Machine Learning. McGraw Hill. 1997.

## Bibliografia complementària

- C. Bishop. Pattern Recognition and Machine Learning. Springer-Verlag New York, Inc. 2006

## Programari

No s'usarà cap programari especial apart dels habituals en aquests estudis (python, etc). Es podra treballar en la plataforma colab.

## Llista d'idiomes

Nom	Grup	Idioma	Semestre	Torn
(PAUL) Pràctiques d'aula	811	Català	primer quadrimestre	tarda
(PAUL) Pràctiques d'aula	812	Català	primer quadrimestre	tarda