

Titulació	Tipus	Curs
Modelització per a la Ciència i l'Enginyeria / Modelling for Science and Engineering	OP	1

Professor/a de contacte

Nom: Alvaro Wong Gonzalez

Correu electrònic: alvaro.wong@uab.cat

Equip docent

Daniel Franco Puentes

Idiomes dels grups

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

Prerequisits

Es recomana tenir coneixements de programació en Python i conèixer l'ús de sistemes Linux per projectes de desenvolupament.

Objectius

Els objectius del mòdul són els següents:

- Donar solucions a problemes d'anàlisi amb eines de codi obert com ara Linux i eines de middleware pròpies d'aquest sistema operatiu.
- Entendre les limitacions de les eines de gestió de dades per seleccionar les eines necessàries per a un determinat problema.
- Aprendre metodologies de consulta a gestors de dades de cada tecnologia
- Utilitzar eines de Cloud Computing per solucionar problemes d'anàlisi de dades
- Aplicar una metodologia d'anàlisi de dades per resoldre problemes pràctiques

Al final de les sessions de teoria i de laboratori els estudiants haurien de tenir suficients coneixements per entendre els requeriments de un problema d'anàlisi de dades en un contexte acadèmic o industrial. Haurien de poder triar una combinació d'eines i dissenyar una solució per a un problema de dades concret.

L'assignatura està orientada a adquirir coneixements de resolució de problemes.

Resultats d'aprenentatge

1. CA24 (Competència) Aplicar les eines computacionals de l'anàlisi de les grans bases de dades a la resolució de problemes en l'àmbit industrial o de la recerca.
2. CA25 (Competència) Integrar les eines computacionals de l'anàlisi de les grans bases de dades en entorns de treball multidisciplinaris.
3. CA26 (Competència) Treballar en equips multidisciplinaris amb l'objectiu de desenvolupar projectes on s'apliquin les tècniques d'anàlisi de grans bases de dades.
4. KA19 (Coneixement) Caracteritzar l'estructura i el rendiment de les diferents arquitectures i estructures d'organització o gestió de bases de dades.
5. KA20 (Coneixement) Descriure les eines de computació en el núvol utilitzades per a l'anàlisi de dades i avaluar-ne els avantatges i les limitacions.
6. SA24 (Habilitat) Fer servir programari específic en la resolució de problemes de tractament de dades en sistemes distribuïts.
7. SA25 (Habilitat) Desenvolupar aplicacions informàtiques encaminades a modelitzar un procés concret mitjançant sistemes distribuïts i avaluar-ne el rendiment computacional.
8. SA26 (Habilitat) Interpretar els informes i resultats que ofereix l'anàlisi d'una base de dades concreta.

Continguts

T1: Introduction to Distributed Systems and large data processing systems (2 hours)

T2: Cloud computing (2 hours)

- Introduction to cloud computing
- Data analysis with a cloud computing provider: AWS / Azure

T3: Cluster and supercomputer infrastructures (14 hours)

- Principles of job execution under batch queue systems (SLURM).
- Advanced control of jobs: array jobs, dependencies, process binding, heterogeneous resources (GPUs)
- Creation of large jobs and workflows
- Virtualization and environments

T4: Cloud Networking and Virtual Private Clouds (8 hours)

- Intro to VPC
- Build our VPC and launch a web server tutorial
- VPC Lab

T5: Fault tolerance systems (4 hours)

- Availability zones
- Load balancing and autoscaling tutorial
- ELB Lab

T6: Database Cloud project: relational and DynamoDB implementations (8 hours)

- Intro to RDS and DynamoDB
- Build a database server tutorial
- Distributed database Lab

T7: Serverless services and Lambda (2 hours)

- Intro to Lambda services and serverless computing
- Lambda tutorial
- Lambda Lab

Activitats formatives i Metodologia

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Laboratori	24	0,96	CA24, CA25, SA24
Teoria	38	1,52	KA19, KA20, SA25
Tipus: Autònomes			
Desenvolupament d'exercicis pràctics	62	2,48	CA26, SA26

El desenvolupament de l'assignatura es basarà fonamentalment en el concepte d'"aprendre fent". Hi haurà sessions teòriques inicials per a cada tema, on el professor presentarà els conceptes claus i es proporcionarà als estudiants materials d'estudi complementaris (llibres, recursos didàctics on-line, articles i altra documentació tècnica). Amb aquesta informació, els estudiants treballaran a les sessions pràctiques en la resolució d'exercicis i petits projectes de forma individual o en grups de dues persones. Els alumnes elaboraran informes escrits sobre el treball pràctic realitzat en cada tema.

Nota: es reservaran 15 minuts d'una classe, dins del calendari establert pel centre/titulació, perquè els alumnes completin les enquestes d'avaluació de l'actuació del professorat i d'avaluació de l'assignatura.

Avaluació

Activitats d'avaluació continuada

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
ELB Lab	20%	6	0,24	CA26, SA24, SA25, SA26
Laboratori Infraestructura	30%	6	0,24	CA24, CA26, SA25, SA26
Lambda Lab	10%	4	0,16	CA26, KA20, SA25, SA26
RDS Lab	20%	6	0,24	CA25, CA26, SA25, SA26
VPC Lab	20%	4	0,16	CA26, KA19, SA25, SA26

L'avaluació de l'assignatura es realitzarà amb la combinació del treball desenvolupat en les sessions de laboratori i l'entrega dels corresponents informes.

Bibliografia

A. Wittig, M. Wittig. "Amazon Web Services in Action", Manning, 2nd Edition, 2018.

G. Coulouris, J. Dollimore and T. Kinderg, "Distributed Systems. Concepts and design ", Addison-Wesley, 5th edition, 2012.

Bell, Charles; Kindahl, Mats; Thalmann, Lars. "MySQL High Availability". O'Reilly, 2010.

Chang, Fay, et al. "Bigtable: A Distributed Storage System for Structured Data." OSDI, 2006

Dewitt, David, and Jim Gray. "Parallel Database Systems: The Future of High Performance Database Processing." Communications of the ACM 35, no. 6 (1992): 85-98

Schwartz, Baron; Zaitsev, Peter; Tkachenko, Vadim; Zawodny, Jeremy D.; Lentz, Arjen; Balling, Derek J. "High Performance MySQL", O'Reilly, 2008.

Seyed M. M. "Saied" Tahaghoghi and Hugh E. Williams. Learning MySQL. O'Reilly, 2006

Nathan Haines. "Beginning Ubuntu for Windows and Mac Users". Apress 2015. *recurs electrònic a la biblioteca de la UAB*

William E. Shotts. "The Linux Command Line". Second Internet Edition. 2013. <http://linuxcommand.org/tlcl.php>

Dan C. Marinescu. "Cloud Computing. Theory and Practice". Morgan-Kaufmann. 2018.

R. Buyya, R. N. Calheiros, A. V. Dastjerdi. "Big data. Principles and paradigms". Morgan-Kaufmann. 2016.

Programari

En l'assignatura es treballarà amb les versions més actualitzades dels sistemes i eines:

-Rocky Linux

-SLURM

-Linux development environment

Grups i idiomes de l'assignatura

La informació proporcionada és provisional fins al 30 de novembre de 2025. A partir d'aquesta data, podreu consultar l'idioma de cada grup a través daquest [enllaç](#). Per accedir a la informació, caldrà introduir el CODI de l'assignatura

Nom	Grup	Idioma	Semestre	Torn
(PLABm) Pràctiques de laboratori (màster)	1	Anglès	primer quadrimestre	tarda
(TEm) Teoria (màster)	1	Anglès	primer quadrimestre	tarda