

Titulació	Tipus	Curs
Bioinformàtica / Bioinformatics	OB	1

Professor/a de contacte

Nom : Elisa Ruth Heymann Pignolo

Correu electrònic : elisa.heyman@uab.cat

Equip docent

Lidia Garrido Sanz

Idiomes dels grups

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

Prerequisits

Per al desenvolupament general del curs, es recomana tenir un nivell B2 d'anglès o similar.

Per al mòdul de programació, és molt recomanable tenir nocions bàsiques de Linux com a usuari (conèixer les eines bàsiques d'edició d'arxius de text i de gestió de carpetes i permisos)

Objectius

Els objectius generals d'aquest mòdul són l'aplicació de les eines i tècniques bàsiques per al desenvolupament en aquesta àrea. Es treballen les capacitats de resoldre reptes i adaptar-se a les tecnologies i paradigmes de la bioinformàtica.

Resultats d'aprenentatge

- CA01 (Programar amb eficàcia i precisió per a la resolució de problemes bioinformàtics de diferent complexitat.) Programar amb eficàcia i precisió per a la resolució de problemes bioinformàtics de diferent complexitat.
- CA02 (Proposar solucions bioinformàtiques a problemes complexos i participar en equips multidisciplinaris.) Proposar solucions bioinformàtiques a problemes complexos i participar en equips multidisciplinaris.
- KA01 (Descriure els conceptes fonamentals de programació, algorítmica i organització de la informació que s'apliquen en bioinformàtica.) Descriure els conceptes fonamentals de programació, algorítmica i organització de la informació que s'apliquen en bioinformàtica.
- KA02 (Identificar els principals esquemes algorítmics i els seus casos d'aplicació en bioinformàtica.) Identificar els principals esquemes algorítmics i els seus casos d'aplicació en bioinformàtica.

- KA03 (Explicar els fonaments i procediments dels mètodes d'accés, anàlisi i mineria de dades en bioinformàtica i destacar-ne la utilitat per a la gestió eficient i la interpretació de grans volums de seqüències biològiques.) Explicar els fonaments i procediments dels mètodes d'accés, anàlisi i mineria de dades en bioinformàtica i destacar-ne la utilitat per a la gestió eficient i la interpretació de grans volums de seqüències biològiques.
- SA01 (Resoldre problemes d'índole biològica aplicant-hi conceptes de programació i l'ús de llibreries bioinformàtiques especialitzades (com BioPython o Bioconductor), i garantint la fiabilitat del codi i la interpretació biològica dels resultats.) Resoldre problemes d'índole biològica aplicant-hi conceptes de programació i l'ús de llibreries bioinformàtiques especialitzades (com BioPython o Bioconductor), i garantint la fiabilitat del codi i la interpretació biològica dels resultats.
- SA02 (Aplicar els esquemes algorítmics principals i algunes de les seves variants a la resolució de problemes bioinformàtics.) Aplicar els esquemes algorítmics principals i algunes de les seves variants a la resolució de problemes bioinformàtics.
- SA03 (Integrar l'accés a grans bases de dades biològiques amb infraestructures locals i sistemes Cloud i combinar-hi els algorismes necessaris.) Integrar l'accés a grans bases de dades biològiques amb infraestructures locals i sistemes Cloud i combinar-hi els algorismes necessaris.

Continguts

1. Linux (Comandes i shell scripting)

Comandes bàsiques, gestió de usuari, gestió de software i sistema d'arxius

Eines de processament de text i de processament de text

Redireccions, pipes, i filters

Programació de shell scripts amb Bash

2. Llenguatges de programació

Introducció a R

Introducció a la programació amb Python dins Bioinformatics

Variables, expressions, tipus de dades, operadors, construccions programàtiques i contextos

Functions, modulis, i subroutines

Programació recursiva

Entrada/Sortida

Depuració de codi

Altres llenguatges de programació: R

3. Estructures de dades i processament de dades

Estructures de dades bàsiques (strings, lists, tuple, sets i dictionaries)

Estructures de dades niades i objectes

Arbres i grafs

Modelatge i representació de dades en bioinformàtica

Formats bàsics en bioinformàtica (FASTQ, SAM, VCF)

Expressions regulars

5. Eines i llibreries en bioinformàtica

Eines per a la visualització de dades

Introducció a Biopython

Introducció a NumPy i Pandas

Activitats formatives i Metodologia

Títol	Hores	ECTS	Resultats d'aprenentatge
Treball realitzat a l'aula	20	0,8	
Solució de problemes a l'aula	14	0,56	
Treball realitzat de forma setmanal sobre els entregables i materials proporcionats	83	3,32	
Treball realitzat en el laboratori	12	0,48	
Treball realitzat al laboratori a partir de les lectures recomenades	15	0,6	

La metodologia combinarà treball a l'aula, solució de problemes supervisada a classe, treball no supervisat en el laboratori i treball a realitzar de forma individual a partir de les lectures recomanades i entregues a realitzar. S'utilitzarà una plataforma virtual per a l'entrega dels informes dels treballs.

Ús restringit de la IA: Per a aquesta assignatura, es permet l'ús de tecnologies d'Intel·ligència Artificial (IA) de manera limitada. Per a les pràctiques de laboratori, els estudiants poden resoldre dubtes genèrics amb eines d'IA, però NO poden lliurar fragments de codi generats per aquestes eines. En la memòria de les pràctiques de laboratori, s'haurà d'identificar clarament l'ajuda rebuda per part d'aquestes eines, especificar les eines utilitzades i incloure els prompts utilitzats. La manca de transparència en l'ús de la IA en les activitats avaluable es considerarà una falta d'honestedat acadèmica i podrà comportar una penalització parcial o total en la nota de l'activitat, o sancions més greus en casos de gravetat.

Nota: es reservaran 15 minuts d'una classe, dins del calendari establert pel centre/titulació, perquè l'alumnat completi 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
Treball al laboratori	30%	3	0,12	CA01, CA02, KA01, KA02, SA01, SA02, SA03
Avaluació del treball realitzat per l'alumnat durant el mòdul	10%	1	0,04	CA01, CA02, KA01, KA02
Exàmen final	60%	2	0,08	CA01, CA02, KA01, KA02, KA03, SA02

La metodologia combinarà treball a classe, solució de problemes a classe, treball autònom en el laboratori i fora del laboratori. S'utilitzaran les plataformes virtuals pel seguiment del curs. Cap de les activitats d'avaluació individuals representarà més del 50% de la nota final.

Examen de recuperació

Per poder participar en el procés de recuperació, l'alumnat haurà d'haver participat prèviament en com a mínim l'equivalent a dos terços de la nota final del mòdul en activitats d'avaluació. El professorat informará dels procediments i terminis per al procés de recuperació. Cal notar que les activitats realitzades dins de classe de forma contínua no poden recuperar-se.

No avaluable

L'alumnat serà qualificat com a "No avaluable" quan el pes de l'avaluació en què ha participat sigui inferior a l'equivalent al 67% de la nota final del mòdul.

Aquesta assignatura/mòdul no preveu el sistema d'avaluació única.

La realització de qualsevol irregularitat en un acte d'avaluació (fraud acadèmic, plagi o ús indegut de la IA, tret que aquest ús estigui autoritzat expressament a la guia docent), que pugui conduir a una variació significativa de la qualificació, suposa que aquest acte es qualificarà amb un 0. En cas que la guia docent prevegi que per superar l'assignatura sigui requisit imprescindible haver obtingut una nota mínima en aquest acte d'avaluació o que es produeixin diverses irregularitats en els actes d'avaluació d'una mateixa assignatura, la qualificació final d'aquesta assignatura és 0. Al marge d'això, es podrà instruir un procés disciplinari a l'estudiant que incorri en alguna d'aquestes irregularitats.

Bibliografia

- Bell, Charles; Kindahl, Mats; Thalmann, Lars. "MySQL High Availability". O'Reilly, 2010.
- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Rapp, B. A. & Wheeler, D. L. (2002).
- GenBank. *Nucl. Acids Res.*, 30(1):17-20. URL <http://nar.oupjournals.org/cgi/content/abstract/30/1/17>.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Res.*, 28(1):235-242.
- Bessant, C., Shadford, I., Oakley, D. "Building Bioinformatics Solutions with Perl, R and MySQL", Oxford University Press, 2009
- Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.-C., Estreicher, A., Gasteiger, E., Martin, M. J., Michoud, K., O'Donovan, C., Phan, I., Pilbout, S. & Schneider, M. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.*, 31(1):365-370.
- Christiansen, P., Wall, L., Orwant, J., "Programming Perl". 4th Edition, O'Reilly, 2012
- Mäkinen et al., Genome-Scale Algorithm Design: Biological Sequence Analysis in the Era of High-Throughput Sequencing. Cambridge Univ. Press, 2015.
- Matloff, N., "The Art of R Programming". No Starch Press Inc., 2011
- Lutz, M., "Learning Python", O'Reilly, 5th edition, 2013
- Siever, E., Figgins, S., "Linux in a nutshell" O'Reilly 2009.
- Sobell, M., "A Practical Guide to Linux. Commands, editors and shell programming". Prentice Hall,

2009.

- Tindall, James., Beginning Perl for Bioinformatics. O'Reilly 2012.

Webs de referència:

- <http://mscbioinformatics.uab.cat>
- <https://cv.uab.cat>

Recursos digitals a la biblioteca UAB:

- <http://www.uab.cat/biblioteques/trobador>
- <http://pagues.uab.cat/bctdigital/>

Programari

Linux (Ubuntu, Bash, linux-tools, etc)

Python 3.x

Jupyter Notebook / PyCharm

Matplotlib/Seaborn

Numpy/Pandas

Grups i idiomes de l'assignatura

La informació proporcionada és provisional fins al 30 de novembre.

A partir d'aquesta data, podreu consultar l'idioma de cada grup a través d'aquest [enllaç](#). Per accedir a la informació, caldrà introduir el CODI de l'assignatura

Tipus de docència	Grup	Idioma	Semestre	Torn
(TEm) Teoria (màster)	1	Anglès	primer quadrimestre	matí-mixt
(PLABm) Pràctiques de laboratori (màster)	1	Anglès	primer quadrimestre	matí-mixt