

Anàlisi de Big Data en Bioinformàtica

Codi: 104886

Crèdits: 6

Titulació	Tipus	Curs	Semestre
2503852 Estadística Aplicada	OT	4	0

La metodologia docent i l'avaluació proposades a la guia poden experimentar alguna modificació en funció de les restriccions a la presencialitat que imposin les autoritats sanitàries.

Professor/a de contacte

Nom: Natalia Isabel Vilor Tejedor

Correu electrònic: Nataliasabel.Vilor@uab.cat

Utilització d'idiomes a l'assignatura

Llengua vehicular majoritària: català (cat)

Grup íntegre en anglès: No

Grup íntegre en català: Sí

Grup íntegre en espanyol: No

Equip docent

Natalia Isabel Vilor Tejedor

Angel González Wong

Prerequisits

Cap. Recomanable haver cursat l'assignatura Bioinformàtica.

Objectius

L'assignatura pretén donar una visió sobre les possibilitats de l'anàlisi de big data focalitzat en bioinformàtica. L'assignatura constarà de dos blocs: 1) metodologies computacionals aplicades al descobriment de fàrmacs i 2) anàlisi de dades òmiques. L'assignatura forma part de la Menció en Estadística per a les Ciències de la Salut.

Competències

- Analitzar dades mitjançant l'aplicació de mètodes i tècniques estadístiques, treballant amb dades de diverses tipologies.
- Aplicar l'esperit crític i el rigor per validar o refutar arguments tant propis com d'altres persones.
- Avaluar de manera crítica i amb criteris de qualitat el treball realitzat.
- Formular hipòtesis estadístiques i desenvolupar estratègies per confirmar-les o refutar-les.
- Identificar la utilitat i la potencialitat de l'estadística en les diferents àrees de coneixement i saber aplicar-la adequadament per extreure conclusions rellevants.
- Interpretar resultats, extreure conclusions i elaborar informes tècnics en el camp de l'estadística.
- Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.
- Que els estudiants sàpiguin aplicar els coneixements propis a la seva feina o vocació d'una manera professional i tinguin les competències que se solen demostrar per mitjà de l'elaboració i la defensa d'arguments i la resolució de problemes dins de la seva àrea d'estudi.

- 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'indole social, científica o ètica.
- Seleccionar els models o tècniques estadístiques per aplicar-los en estudis i problemes reals, així com conèixer-ne les eines de validació.
- Treballar cooperativament en un context multidisciplinari assumint i respectant el rol dels diferents membres de l'equip.
- Utilitzar correctament un ampli espectre del programari i llenguatges de programació estadístiques, escollint el més apropiat per a cada anàlisi i ser capaç d'adaptar-lo a noves necessitats.
- Utilitzar eficaçment la bibliografia i els recursos electrònics per obtenir informació.

Resultats d'aprenentatge

1. Aplicar l'esperit crític i el rigor per validar o refutar arguments, tant propis com d'altres.
2. Aplicar mètodes estadístics a l'anàlisi de dades d'expressió gènica.
3. Avaluar de manera crítica i amb criteris de qualitat la feina feta.
4. Dissenyar i dur a terme tests d'hipòtesi en els diferents camps d'aplicació estudiats.
5. Elaborar informes tècnics que expressin clarament els resultats i les conclusions de l'estudi utilitzant vocabulari propi de l'àmbit d'aplicació.
6. Extreure conclusions coherents amb el context experimental propi de la disciplina a partir dels resultats obtinguts.
7. Interpretar els resultats estadístics en contextos aplicats.
8. Justificar l'elecció de cada mètode particular dins del context en què s'aplica.
9. Que els estudiants puguin transmetre informació, idees, problemes i solucions a un públic tant especialitzat com no especialitzat.
10. Que els estudiants sàpiguin aplicar els coneixements propis a la seva feina o vocació d'una manera professional i tinguin les competències que se solen demostrar per mitjà de l'elaboració i la defensa d'arguments i la resolució de problemes dins de la seva àrea d'estudi.
11. 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'indole social, científica o ètica.
12. Reconèixer els mètodes d'inferència estadística més utilitzats en bioinformàtica.
13. Reconèixer la importància dels mètodes estadístics estudiats dins de cada aplicació particular.
14. Treballar cooperativament en un context multidisciplinari assumint i respectant el rol dels diferents membres de l'equip.
15. Utilitzar diferents programes (tant lliures com comercials) associats a les diferents branques aplicades.
16. Utilitzar eficaçment bibliografia i recursos electrònics per obtenir informació.

Continguts

BLOC 1. Big Data en Disseny de Fàrmacs

1. Introducció al big data en disseny de fàrmacs
2. Estructura de proteïnes i espai químic en molècules petites (tipus fàrmacs).
3. Interaccions proteïna-fàrmac
4. Cribatge virtual
5. Dinàmica molecular

BLOC 2. Big Data en Anàlisi de Dades Òmiques

1. Introducció a Bioconductor i eines bioinformàtiques per l'anàlisi de dades òmiques.
2. Estudis d'associació genètica i GWAS (Estudis de genoma complet).
3. Mètodes Multivariants per la Integració de dades òmiques i big data.

*Llevat que les restriccions imposades per les autoritats sanitàries obliguin a una priorització o reducció d'aquests continguts.

Metodologia

L'assignatura està organitzada en sessions de 3 hores on cada sessió consta d'una part teòrica (aules de teoria) on s'introduirà el temari nou seguida d'una part pràctica (aules d'informàtica) on es treballarà l'aplicació dels conceptes explicats en la part teòrica. A cada sessió el professor/a indicarà als estudiants algunes tasques a fer de manera autònoma, com ara lectura d'articles o l'elaboració d'informes de pràctiques. El material utilitzat pels professors estarà disponible al Campus Virtual de l'assignatura.

*La metodologia docent proposada pot experimentar alguna modificació en funció de les restriccions a la presencialitat que imposin les autoritats sanitàries.

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.

Activitats formatives

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Clases teòriques	21	0,84	9, 11, 12, 13
Clases pràctiques	21	0,84	2, 3, 4, 6, 7, 15, 16
Presentació Projecte de Recerca	3	0,12	9, 10
Tipus: Supervisades			
Tutoritzacions	10	0,4	6
Tipus: Autònomes			
Estudi	70	2,8	10
Preparació Projecte de Recerca	20	0,8	1, 4, 14

Avaluació

BLOC 1. Big Data en Disseny de Fàrmacs (50%):

- exercicis de pràctiques (10%)
- prova teòrico-pràctica (20%)
- presentació treball bioinformàtic davant d'una comissió (20%)

BLOC 2. Big Data en Anàlisi de Dades Òmiques (50%):

- exercicis de pràctiques (30%)
- prova teòrico-pràctica (20%)

La qualificació mínima global necessària per superar l'assignatura serà de 5 punts. Per a fer mitjana cal que la nota mínima de cadascuna de les activitats avaluables sigui igual o superior a 4 punts. Els estudiants que tinguin alguna de les proves suspeses o no presentades podran fer l'examen de recuperació on es podran examinar del bloc suspès.

*L'avaluació proposada pot experimentar alguna modificació en funció de les restriccions a la presencialitat que imposin les autoritats sanitàries.

Activitats d'avaluació

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Examens teòrico-pràctiques	40	4	0,16	2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 15
Presentació Projecte de Recerca	20	0,5	0,02	1, 5, 9, 14, 16
Presentació informes de pràctiques	40	0,5	0,02	5

Bibliografia

- Lesk A.M. *Introduction to Bioinformatics*. Oxford University Press 2005.
- Attwood, T.K., Parry-Smith, D.J., *Introducción a la Bioinformática*. Pearson Education, 2002.
- Foulkes A.S. *Applied Statistical Genetics with R. For Population-based Association Studies*. Springer Dordrecht Heidelberg London New York. ISBN 978-0-387-89553-6
- Gonzalez JR, Cáceres A. *Omic association studies with R and Bioconductor*. Chapman and Hall/CRC, ISBN 9781138340565, 2019.
- <https://www.bioconductor.org/>

La docència virtual ha posat de manifest la importància de poder disposar de recursos *online*. Tenint en compte que durant aquests mesos les editorials han posat en obert molt contingut, i que a més es disposa de la plataforma a prova de llibres digitals (50.000 llibres accessibles - <https://mirades.uab.cat/ebs/>), us demanem que, en la mesura de les vostres possibilitats, feu un esforç per tal que les Guies Docents del curs 2020/2021 incrementin les referències de documents digitals. En aquest enllaç, trobareu una infografia que ha preparat el Servei de Biblioteques per facilitar la localització de llibres electrònics: <https://ddd.uab.cat/record/22492>

Programari

R