FICHA IDENTIFICATIVA

Datos de la Asignatura

Código	36434
Nombre	Estadística para datos ómicos
Ciclo	Grado
Créditos ECTS	6.0
Curso académico	2022 - 2023

Titulación(es)

Titulación	Centro	Curso	Periodo
1406 - Grado en Ciencia de Datos	Escuela Técnica Superior de Ingeniería	3	Segundo cuatrimestre

Materias

Titulación	Materia	Carácter
1406 - Grado en Ciencia de Datos	11 - Salud	Obligatoria

Coordinación

Nombre	Departamento
AYALA GALLEGO, GUILLERMO	130 - Estadística e Investigación Operativa
SILVA MORENO, FRANCISCO J.	194 - Genética

RESUMEN

Se propone una introducción al análisis estadístico de datos ómicos. Entre otros utilizaremos datos obtenidos con microarrays y RNA-Seq. Estudiamos los conceptos biológicos básicos previos. Se comentan las técnicas ómicas, el procesado previo de la información y los problemas estadísticos. En particular, estudiamos el problema de las comparaciones múltiples. Luego mostramos su aplicación al problema de la expresión diferencial. También consideramos análisis de grupos de genes. El software utilizado es R y Bioconductor.

Las clases de teoría se impartirán en castellano y las clases prácticas y de laboratorio según consta en la ficha de la asignatura disponible en la web del grado.

CONOCIMIENTOS PREVIOS

Relación con otras asignaturas de la misma titulación

No se han especificado restricciones de matrícula con otras asignaturas del plan de estudios.

Otros tipos de requisitos

Un primer curso en Estadística así como un dominio medio del lenguaje R.

COMPETENCIAS (RD 1393/2007) // RESULTADOS DEL APRENDIZAJE (RD 822/2021)

1406 - Grado en Ciencia de Datos

- (CG01) Conocimiento de materias básicas y tecnologías, que le capacite para el aprendizaje de nuevos métodos y tecnologías, así como que le dote de una gran versatilidad para adaptarse a nuevas situaciones.
- (CG03) Capacidad para la realización de modelos, cálculos, informes, planificación de tareas y otros trabajos análogos en el ámbito específico de la Ciencia de Datos.
- (CG07) Capacidad para tomar decisiones de forma autónoma, elaborando de forma adecuada y original, argumentos razonados, pudiendo obtener así hipótesis razonables y contrastables.
- (CT01) Ser capaces de acceder a herramientas de información (bibliográficas) y de utilizarlas apropiadamente en el desarrollo de sus tareas cotidianas.
- (CT03) Habilidad para defender su trabajo con rigor y argumentos, exponiéndolo de forma adecuada y precisa, apoyándose en los medios necesarios.
- (CE12) Capacidad para diseñar y poner en marcha soluciones basadas en análisis de datos en el ámbito de la medicina y de los negocios, teniendo en cuenta los requisitos específicos de este tipo de casos de uso.
- (CE15) Capacidad para modelizar y analizar la incertidumbre en estudios basados en datos así como saber interpretar y contextualizar los resultados obtenidos.
- (CB2) Que los estudiantes sepan aplicar sus conocimientos a su trabajo o vocación de una forma profesional y posean las competencias que suelen demostrarse por medio de la elaboración y defensa de argumentos y la resolución de problemas dentro de su área de estudio.
- (CB5) Que los estudiantes hayan desarrollado aquellas habilidades de aprendizaje necesarias para emprender estudios posteriores con un alto grado de autonomía.

RESULTADOS DE APRENDIZAJE (RD 1393/2007) // SIN CONTENIDO (RD 822/2021)

Aprender las características especiales de los datos en biomedicina. Conocer los datos ómicos (proteómicos, genómicos y metabólicos). Conocer las aplicaciones de la ciencia de datos en biomedicina. Aplicar los métodos conocidos en problemas de biomedicina.

En particular, realizar un análisis completo con microarrays y RNA-Seq con R/Bioconductor con especial interés en la utilización de herramientas de investigación reproducible como Rmarkdown.

(Evaluación de las competencias CB5, CG7, CT3, CE15)

DESCRIPCIÓN DE CONTENIDOS

1. Los seres vivos y los datos ómicos

Introducción a las ciencias ómicas. Las macromóleculas: ácidos nucleicos y proteínas. Los seres vivos y los virus. Las células. Las funciones celulares y las pequeñas moléculas. Expresión génica. La regulación de la expresión génica. La epigenética. La mutación y la variación poblacional. Los genomas de los seres vivos. Las mutaciones y la salud. La herencia de las enfermedades. El cáncer: las mutaciones génicas y cromosómicas. Las bases de datos generales de secuencias y los archivos de secuencia. El programa BLAST. Las técnicas ómicas.

2. Introducción a los datos ómicos

ntroducción a datos ómicos. R y Bioconductor. Anotación de genomas.

3. Datos

Microarrays. Obtención y preprocesado de datos de microarrays. Método RMA.
GEO y Array Express.
RNA-Seq. Obtención y preprocesado de datos de RNA-Seq. Alineamiento y normalización.

4. Expresión diferencial

Expresión diferencial marginal o gen a gen.
Microarrays. Selección no específica. Comparaciones múltiples. Método de Benjamini-Hochberg.
El q-valor. Método limma para datos de microarrays.
RNA-Seq. Métodos edgeR y DESeq2.

5. Análisis de grupos de genes

Expresión diferencial considerando grupos de genes.
Grupos de genes utilizando GO y Kegg. Análisis de sobre representación. Análisis basado en enriquecimiento.

VOLUMEN DE TRABAJO

ACTIVIDAD	Horas	% Presencial
Clases de teoría	32,00	100
Prácticas en laboratorio	20,00	100
Prácticas en aula	8,00	100
Estudio y trabajo autónomo	60,00	0
Preparación de actividades de evaluación	24,00	0
Resolución de cuestionarios on-line	6,00	0
TOTAL	**150,00**	

METODOLOGÍA DOCENTE

Se plantean clases magistrales así como clases prácticas en aula de Informática. Se proporciona el material desarrollado en un manual. Se propondrán cuestionarios online semanales.

(Evaluación de las competencias CB5, CT1)

EVALUACIÓN

En la primera convocatoria, la asignatura se evaluará atendiendo a dos aspectos:

SE1: Prueba objetiva, que se realizará al terminar la docencia y que constará de cuestiones teórico-prácticas y problemas. Este apartado de la evaluación contará un 70% de la nota final. De este 70%, el 50% (sobre el total de la asignatura) corresponde a la parte de Estadística y el 20% restante (sobre el total de la asignatura) corresponde a la parte de Biología (evaluación de las competencias CB4, CB5, CG1, CE9, CE15).

SE2: Realización de cuestionarios durante el curso evaluando, mediante preguntas de teoría y práctica, el trabajo del estudiante. Esta parte corresponde completamente a la parte de Estadística y supone el 30% sobre la nota final. Esta parte de la evaluación no es recuperable (evaluación de las competencias CB2, CB5, CG1, CT2, CE9, CE15).

El reparto de las valoraciones se hará de modo que el primer bloque de contenidos biológicos tendrá una valoración global del 20% y los restantes bloques del 80%.

El apartado SE2 no es recuperable en la segunda convocatoria.

En la segunda convocatoria se mantendrá el peso relativo de los contenidos de forma de modo que el primer bloque cuente un 20% y los restantes un 80%.

En cualquier caso, el sistema de evaluación se regirá por lo establecido en el Reglamento de Evaluación y Calificación de la Universidad de Valencia para Grados y Másteres (https://webges.uv.es/uvTaeWeb/MuestraInformacionEdictoPublicoFrontAction.do?accion = inicio & idEdictoSeleccionado = 5639).

REFERENCIAS

Básicas

- Bioinformática Estadística. Análisis estadístico de datos ómicos. Guillermo Ayala. https://www.uv.es/ayala/docencia/tami/tami13.pdf
- Pevsner J. (2015) Bioinformatics and Functional Genomics, 3rd Edition Wiley-Blackwell
- Genetics Home Reference (2020). Help Me Understand Genetics (https://ghr.nlm.nih.gov/)