

B.4 Programas de los Cursos.

Formulario N°

B-4

PROGRAMA:	: Doctorado en Ciencias e Innovación en Medicina
CURSO	:Temas Avanzados de Bioestadística y Bioinformática
TRADUCCIÓN	: Advanced Topics in Biostatistics and Bioinformatics
CÓDIGO	: DCM3S232
CRÉDITOS	: 4
HORAS ACADÉMICAS	: 40 hrs cronológicas
REQUISITOS	: Bioestadística
CARÁCTER	: Electivo
RÉGIMEN	: Semestral
CALIFICACIÓN	: Nota calificación de 1.0 a 7.0, nota mínima de aprobación 4.0.
ASISTENCIA	: Igual o mayor al 70% de las actividades planificadas en el curso.

DOCENTE RESPONSABLE

- Boris Rebolledo, Investigador Docente Centro de Genética y Genómica

I. DESCRIPCIÓN

La manipulación eficiente de grandes cantidades de datos y su análisis integrativo se han convertido en una necesidad y una habilidad deseada en el científico y la científica actuales. Este curso busca entregar a sus estudiantes herramientas básicas e intermedias de programación para que implementen protocolos de análisis estadísticos y bioinformáticos de manera autónoma.

II. OBJETIVOS

- a. Objetivo General:
Desarrollar autonomía en la implementación de rutinas de análisis de datos.
- b. Objetivos Específicos:
 1. Entregar acceso a los estudiantes a la infraestructura computacional del ICIM, para que realicen el curso en un ambiente Linux.
 2. Familiarizar a los estudiantes con el ambiente Linux y los lenguajes de programación R y Python.
 3. Fomentar una cultura de reproducibilidad de resultados, a través de la documentación de protocolos de análisis.
 4. Implementar análisis de llamado de variantes para el diagnóstico de enfermedades poco frecuentes.
 5. Implementar análisis de transcriptómica.
 6. Implementar análisis de clasificación con *machine learning*.

III. CONTENIDOS

(Indicar programa con fecha y horario de cada clase, además incluir el listado de profesores nacionales/internacionales invitados que participarán del curso UDD/Externo internacional)

CLASE	Tema	Contenidos	Horas
1	Linux basics	Sistema de archivos, conda, instalación de programas, preparación de ambientes de trabajo	3
2	Formatos en bioinformática	FASTA, FASTQ, BED, GFF, VCF, SAM, BAM, TSV, CSV, otros	3
3	Llamado, anotación e interpretación clínica de variantes	Alineamiento de secuencias (BWA, Bowtie, HISAT, STAR)	3
4		Llamado de variantes con (GATK, Samtool, Freebayes)	3
5		Comparar y consolidar variantes	3
6		Franklin, Annovar, VEP	3
Evaluacion 01			
7	Transcriptómica	Data mining GEO and SRA archive	3
8		Recuento de transcritos. Visualización PCA	3
9		Expresión diferencial. Heatmaps	3
10		Enriquecimiento de vías	3
Evaluación 02			
11	Machine Learning	Introducción a machine learning	3
12		Árboles de decisión y clasificación	3
13		tidymodels (paquete de R)	3
14		Tutorial de clasificación	3
Evaluación 03			

Horas totales 42

En primera instancia las clases serán impartidas todos los martes de 13:30 a 16:30 a contar del martes 21 de marzo del 2023. Sin embargo, el horario puede variar según la disponibilidad de los estudiantes y profesor, pudiendo incluso separarse en dos bloques de 1.5 horas por semana.

IV. METODOLOGÍA

Las clases tendrán un componente expositivo y un componente de trabajo en el computador. En general, el contenido se entregará en formato de tutoriales que plantean un problema y los pasos a seguir para resolverlo.

V. EVALUACIÓN

- Se promediará la nota de tres presentaciones de implementación y documentación de un análisis de datos biomédicos y/o genómicos.
- Estudiantes que no obtengan la nota mínima de aprobación de 4.0 al finalizar todas las evaluaciones del curso, tendrán la opción de rendir un examen de recuperación, el cual se promediará con la nota obtenida originalmente.

VI. BIBLIOGRAFÍA

Dada la naturaleza de la disciplina, múltiples recursos existen en internet. Un listado de recursos recomendado será entregado al iniciar el curso.

Materiales presentados en la versión anterior del curso

<https://tinyurl.com/TAAB22>

Tutorial análisis de exoma.

<https://galaxyproject.github.io/training-material/topics/variant-analysis/tutorials/exome-seq/tutorial.html>

Tutorial RNA-Seq

<https://www.bioconductor.org/packages/release/workflows/vignettes/RnaSeqGeneEdgeRQL/inst/doc/edgeRQL.html>

Tutorial Machine Learning

https://www.rebeccabarter.com/blog/2020-03-25_machine_learning/

VII. DISPOSICIONES FINALES

Las situaciones excepcionales no previstas en el presente formulario serán resueltas por el Académico responsables de la signatura en conjunto con el Comité Académico del programa.