



Especialización en Bioinformática

Taller de aplicaciones avanzadas de bioinformática

Año 2022

Duración total: 70hs.

Responsable: Facultad de Ciencias Exactas
y Facultad de Informáticas

Docente Responsable: Dr. Mauricio Lozano

Docentes: Dr. Franco Ronchetti
Dr. Enzo Rucci

OBJETIVO GENERAL

El objetivo general del curso es integrar y aplicar los conocimientos biológicos e informáticos hacia la resolución de problemas bioinformáticos complejos.

Objetivos específicos:

- Revisar casos de aplicación de técnicas avanzadas de procesamiento de datos (en especial, las vinculadas a sistemas inteligentes y paralelismo) en la resolución de problemas biológicos complejos.
- Brindar a los alumnos las herramientas para comprender los algoritmos aplicados en: estudios filogenéticos y filogenómicos; ensamblado de genomas; clustering de diferente tipo de datos; aprendizaje automático para la estimación de estructura de proteínas.

COMPETENCIAS A DESARROLLAR EN RELACION CON EL OBJETIVO DE LA CARRERA

C. 2 - Utilizar distintas técnicas de procesamiento de datos biológicos para su representación/visualización y análisis eficiente, mediante algoritmos de software ejecutados sobre plataformas adecuadas para el tipo y volumen de datos en cuestión.

C. 3 - Tener capacidad para evaluar la eficiencia (en tiempo, recursos utilizados y consumo energético) así como la calidad de las soluciones obtenidas, considerando el ámbito crítico de la aplicación de las mismas.

PROGRAMA

Módulo 1

Evolución molecular. Estudios filogenéticos. Modelos evolutivos paramétricos y no paramétricos. Maximum Likelihood, LRT. Detección de selección positiva. Reconstrucción ancestral. Filogenómica. CONCAT. ASTRAL. ANI, dDDH y métodos no basados en alineamiento secuencial.

Módulo 2



Ensamblado de genomas: Grafos solapados (Overlap Graphs) y Grafos de De Bruijn. Algoritmos. Mapeo de secuencias en genomas de referencia. Árboles de sufijos y transformación de Burrows-Wheeler. Anotación de genomas. Búsqueda de genes (En procariontas y eucariotas). Anotación funcional.

Módulo 3

Análisis de datos de expresión de experimentos ÓMICOS. Normalización. Expresión diferencial. Clustering (Jerárquico, PCA, singular value decomposition (SVD), Self Organizing Maps (SOM), supportvector machines (SVMs)).

Módulo 4

Aplicación de técnicas inteligentes para la predicción del nivel de expresión de proteínas a partir del sesgo de uso de codones, la predicción de péptidos señal, la predicción de estructura secundaria de proteínas y a la predicción de secuencias plasmídicas (o de otros orígenes).

METODOLOGIA Y MODALIDAD DE EVALUACION

La metodología se basa en clases presenciales combinadas con actividades experimentales en el Aula y/o Laboratorio para aplicar los conceptos teóricos y que así el alumno adquiera las competencias y habilidades sobre cada uno de los temas que forman parte del contenido de la asignatura.

Para aprobar el curso se requiere un 80% de asistencia a las clases presenciales (teóricas y prácticas).

La evaluación se realizará mediante un examen escrito al final de las clases del curso para evaluar el grado de conocimientos del alumno (20%), el proyecto/desarrollo experimental que deberá entregar el alumno al final de las horas programadas (70%) y la participación y aportaciones de calidad/excelencia a las soluciones propuestas (10%). El proyecto/desarrollo experimental consistirá en un reporte técnico (símil artículo científico) producto de un estudio teórico o experimental específico.

ACTIVIDADES EXPERIMENTALES y DE INVESTIGACION

Los trabajos experimentales se desarrollarán en una o varias clases. Parten de una consigna explicada en clase por el docente, seguida del trabajo individual, o en grupos (máximo 3 alumnos, y priorizando la mezcla de alumnos provenientes de las áreas informáticas y biológicas), en el que los alumnos resuelven un problema experimental concreto relacionado con la temática. Estos trabajos pretenden desarrollar y/o fortalecer las aptitudes de opinión crítica en los temas relativos del curso. Los alumnos deberán sintetizar su comprensión de los temas, al realizar correctamente la tarea experimental propuesta. También se pretende desarrollar la capacidad de comunicar y transmitir los resultados, en presentaciones pautadas a lo largo del curso.

En general, finalizada cada actividad, hay una sesión de discusión conjunta donde los participantes comunicarán sus opiniones e intercambiarán los distintos puntos de vista.

La evaluación se realizará mediante la entrega de un trabajo final.



BIBLIOGRAFÍA BASICA

- Biological sequence analysis. Probabilistic models of protein and nucleic acids (1998). R. Durbin, S. Eddy, A. Krogh and G. Mitchison. Cambridge University Press.
- Bioinformatics: The Machine Learning Approach (2001). Pierre Baldi and Søren Brunak, MIT Press.
- Bioinformatics: Sequence and Genome Analysis (2004). 2da edición. CBS Publishers.
- Bioinformatics and molecular evolution (2005). Paul G. Higgs and Teresa K. Attwood. Blackwell Science.
- Introduction to bioinformatics (2008). Lesk, A.M. 3rd Ed., Oxford University Press.
- Bioinformatics: high performance parallel computer architectures (2011).- Bertil Schmidt (ed). CRC Press. 2011.
- Bioinformatics Algorithms. Design and Implementation in Python (2018). Miguel Rocha, Pedro G. Ferreira. Elsevier.