

Cuerpo Académico de Bioinformática y Biología Computacional

En las últimas décadas, avances en la biología molecular y en la infraestructura disponible para el desarrollo de investigaciones en este campo han dado lugar a una revolución “ómica” en la biología molecular, con explosión de bases de datos: desde **genómica**, a **transcriptómica**, **proteómica**, **interactómica** y **metabolómica**. No obstante, esta enorme cantidad de información biológica ha portado a un cuello de botella en el procesamiento de datos. Este creciente flujo de información requiere el uso de técnicas innovadoras para su visualización, modelación, interpretación y análisis. Las ciencias de la computación, las ciencias de la información, la ingeniería, las matemáticas y la estadística han sido aplicadas a la biología, emergiendo así la disciplina hoy conocida como bioinformática.

Dos diferentes escenarios caracterizan la era post-genómica. De una parte, la enorme cantidad de conjuntos de datos disponibles de la investigación biológica contemporánea en todo el mundo demanda el uso de técnicas, herramientas y métodos apropiados para el modelado de procesos biológicos y análisis de secuencias biológicas. De otra parte, los sistemas biológicos trabajan como las fuentes de un amplio rango de nuevos modelos y paradigmas computacionales, los cuales pueden ser aplicados en el contexto de los sistemas computacionales.

La bioinformática puede ser vista como el trabajo interdisciplinario entre biólogos, informáticos, computólogos, matemáticos, estadísticos e ingenieros, el cual abarca tres aspectos. En primer lugar, la propuesta y discusión de problemas abiertos en biología molecular de la célula. En segundo la propuesta de métodos, modelos, técnicas y herramientas computacionales tanto para el análisis de datos biológicos como para el modelado coherente y robusto de sistemas biológicos, y, finalmente, el desarrollo de aplicaciones y sistemas computacionales inspirados en la biología, los cuales regresan nuevamente a ésta como soluciones.

La aplicación de técnicas y métodos pertenecientes a las áreas de estadística y en general de la matemática e inteligencia artificial a la biología molecular de la célula, ha hecho posible la propuesta de soluciones para problemas en la predicción de secuencias de proteínas, la estructura tridimensional de proteínas y de la función de proteínas; la simulación tanto de redes de interacción de proteínas como de redes de transducción de señales; el análisis de secuencias de DNA; el alineamiento de secuencias; el análisis de patrones de expresión de genes y de genomas; la predicción de genes y estructuras de genes, y la simulación de redes de expresión genética, entre otros.

Los problemas abiertos en esta área son variados y diversos. Por ejemplo, los métodos estadísticos tradicionales no resultan del todo adecuados para el análisis y modelado de datos y sistemas biológico. Ocurre con frecuencia que una significancia estadística no corresponde a una significancia biológica, o que una similitud estadística no conlleva a una similitud biológica: se requieren nuevos modelos, así como la integración de diferentes paradigmas para abordar los problemas actuales que ofrece la bioinformática. En particular se tiene en la estadística, el paradigma de problemas con “muestra pequeña dimensión grande”.

Por lo anterior, y debido al perfil e interés de varios de sus integrantes actuales, el DMAS ha formado el cuerpo académico de Bioinformática, estructurado en las siguientes tres líneas de generación y aplicación innovadora de conocimientos (LGAC):

- Modelos estocásticos en bioinformática
- Inteligencia artificial en bioinformática
- Biología de sistemas

Estas tres LGACs, las cuales por su naturaleza son fuertemente interactivas entre sí, estarán encaminadas hacia la búsqueda de nuevas soluciones para los problemas actuales que enfrenta la bioinformática y la biología computacional.

Por su naturaleza interdisciplinaria *per se*, este cuerpo académico promoverá la interacción con colegas de los otros departamentos de la División de Ciencias Naturales e Ingeniería, pues tiene el potencial de convertirse en un Cuerpo Académico Inter divisional.

LGAC Inteligencia artificial en Bioinformática

Los múltiples y novedosos resultados provenientes en los últimos años de áreas pertenecientes a ciencias de la computación, inteligencia artificial, ciencias de la información, matemáticas aplicadas y la ingeniería se han encontrado con la biología, guiándonos a una nueva área, emergiendo disciplinas tales como:

- Bioinformática
- Biología Computacional / Systems Biology

La bioinformática puede definirse como el desarrollo sistemático y la aplicación de métodos, técnicas y modelos computacionales (sistemas y soluciones computacionales) tanto al: (1) análisis de datos biológicos, provenientes de fuentes como la investigación básica, el modelado, las búsquedas en bases de datos biológicas, etc., como al (2) modelado de fenómenos y sistemas biológicos. Esta última área resulta comúnmente hoy conocida como biología computacional.

La inteligencia artificial ha jugado un importante rol como fuente inagotable de técnicas, métodos, modelos y algoritmos tanto para el análisis de datos biológicos como para el modelado y simulación de sistemas biológicos. Técnicas tales como redes neuronales artificiales, algoritmos evolutivos, autómatas celulares, redes Bayesianas y modelos ocultos de Markov, resultan ser enfoques ideales para dominios que se caracterizan por una explosión de datos y muy poca teoría, como es el caso de la bioinformática.

En el siguiente resumen panorámico se muestran los principales tópicos / dominios de aplicación en bioinformática y las técnicas de inteligencia artificial que han sido comúnmente usadas. Estos dominios de aplicación continúan siendo de especial interés dentro de la LGAC Inteligencia Artificial en Bioinformática del Departamento de Matemáticas Aplicadas y Sistemas, siendo aún variados y diversos los problemas abiertos que restan en esta área. De algunos de estos tópicos se cuenta ya con trabajos de investigación precedentes desarrollados por algunos de los miembros del cuerpo académico. Estos son los casos de los tópicos:

- modelado / simulación de redes de señalización biológica,

- modelado de redes genéticas, y
- análisis de microarreglos.

Dominio de aplicación	Técnica de Inteligencia Artificial
Predicción de la estructura y función de proteínas	Sistemas multi-agentes, redes neuronales artificiales, algoritmos evolutivos, modelos Bayesianos, modelos ocultos de Markov
Modelado / simulación de redes de señalización biológica	Redes Booleanas, redes neuronales artificiales, redes Petri, autómatas celulares, sistemas multi-agentes, sistemas expertos
Identificación de blancos terapéuticos	Sistemas multi-agentes, sistemas expertos, algoritmos evolutivos, redes neuronales evolutivas
Análisis de secuencias biológicas	Algoritmos evolutivos, redes neuronales artificiales
Modelado de redes genéticas	Redes neuronales artificiales, algoritmos evolutivos, redes Bayesianas, redes Booleanas, sistemas multi-agentes
Identificación de regiones de regulación transcripcional	Algoritmos genéticos, otros algoritmos evolutivos
Análisis de microarreglos	Mapas auto-organizativos, redes neuronales evolutivas, redes Bayesianas
Text mining y extracción de información	Redes neuronales adaptativas, algoritmos evolutivos, algoritmos genéticos