

Bioinformática e ingeniería de datos aplicados a ciencias biomédicas Cod.: 01



DIRECTORES:

Federico Morán Abad (UCM) y José Jesús Cabrera Pantoja (UCB).

FECHAS Y HORARIO DEL CURSO:

Del 21 de octubre al 1 de noviembre de 2024.
Mañanas de 9:00 a 14:00 horas, de lunes a viernes.

PERFIL DEL ALUMNADO:

El curso está dirigido a Estudiantes de Grado y de Posgrado de Ingeniería biomédica, Ingeniería en Biotecnología, Ingeniería en sistemas, Farmacia, Bioquímica, Medicina, Biología Sanitaria, Nutrición, Enfermería, Fisioterapia, entre otros, así como a Profesionales del Área de la Salud e Ingeniería que estén interesados en adquirir y/o actualizar conocimientos relacionados con la bioinformática y los contenidos objeto del curso. Los estudiantes deberán poseer conocimientos previos de Bioquímica, Biología Molecular y Fisiología, y bien una noción básica o interés en la programación.

INTERÉS DEL CURSO:

La bioinformática y la biología computacional y de sistemas representan campos emergentes y de rápida evolución que están revolucionando nuestra comprensión de los sistemas biológicos complejos. El curso **Bioinformática e Ingeniería de Datos aplicados a ciencias biomédicas** ofrece una formación integral que abarca desde los fundamentos teóricos hasta las aplicaciones prácticas más avanzadas, integrando conocimientos de biología, informática, física y matemáticas.

Las tareas más importantes de las que se ocupa esta disciplina consisten en entender las correlaciones, las estructuras y los patrones en los datos biológicos. En los últimos años, la Bioinformática ha atraído la conjugación de varias disciplinas, entre las que están la informática, las matemáticas, la estadística, la química, biomédica y las ciencias biológicas no tradicionales. El curso que se propone dentro de la **Escuela Complutense Latinoamericana** pretende mostrar técnicas y metodologías computacionales inteligentes que sirvan para la creación de proyectos relacionados con las ciencias biomédicas. Esto implica el uso de tecnologías informáticas y métodos estadísticos para manejar y analizar un gran volumen de datos biológicos.

Este curso ofrece una formación integral y avanzada en bioinformática y biología de sistemas, preparando a los estudiantes para abordar los desafíos actuales en biomedicina, biotecnología e investigación biológica. La combinación de teoría, práctica y herramientas tecnológicas avanzadas hace de este curso una oportunidad invaluable para adquirir competencias críticas en un campo de creciente importancia y demanda.

En concreto, el curso es una apuesta para familiarizarse en el uso de las metodologías y técnicas computacionales y matemáticas para la integración y análisis de datos de sistemas biológicos. Se va a profundizar en las metodologías de ingeniería bioinformática mediante las tecnologías de análisis de la información biológica y biomédica. En el curso se integran diferentes aproximaciones para este propósito, entre otras: manejo de bases de datos ómicos, minería de datos, métodos de clustering no supervisado y predicción de metabolismos a partir de secuenciación de genomas completos y dinámica de poblaciones de virus tipo SARS CoVID19. El objetivo principal del curso es familiarizar al alumnado con estas metodologías desde un punto de vista multidisciplinar. El curso se plantea como teórico-práctico con el fin de que el alumno pueda practicar en el aula los diferentes métodos desarrollados en las introducciones teóricas. Para facilitar este objetivo se han diseñado en formato de taller práctico las clases de los viernes.

OBJETIVOS DEL CURSO:

- **Objetivo general:**
 1. Proporcionar una formación integral y avanzada en bioinformática, biología de sistemas e ingeniería de datos, abarcando desde los fundamentos teóricos hasta las aplicaciones prácticas más avanzadas. Los estudiantes adquirirán conocimientos interdisciplinarios en biología, informática, física y matemáticas, y desarrollarán habilidades en el análisis y gestión de grandes volúmenes de datos biológicos. A través de módulos específicos, aprenderán técnicas de alto rendimiento, métodos de minería de datos, análisis de secuencias, visualización de datos, y la integración de datos multiómicos, preparándolos para enfrentar desafíos en biomedicina, biotecnología e investigación biológica. Además, el curso incluye talleres prácticos que refuerzan la aplicación de conocimientos teóricos en situaciones reales, proporcionando una formación completa y práctica en bioinformática y biología de sistemas.

- **Objetivos específicos:**
 1. Proporcionar una formación integral y avanzada:
 - Capacitar a los estudiantes en bioinformática, biología de sistemas e ingeniería de datos.
 - Cubrir desde los fundamentos teóricos hasta las aplicaciones prácticas más avanzadas.
 2. Adquirir conocimientos interdisciplinarios:
 - Instruir en las áreas de biología, informática, física y matemáticas.
 - Desarrollar habilidades en el análisis y gestión de grandes volúmenes de datos biológicos.
 3. Familiarizarse con las técnicas de alto rendimiento:
 - Enseñar metodologías de minería de datos.
 - Instruir en análisis de secuencias.
 - Capacitar en visualización de datos y la integración de datos multiómicos.
 4. Conocimiento y manejo de las bases de datos biológicos:
 - Formar a los estudiantes en el conocimiento de las bases de datos disponibles.
 - Acceso y manejo de estas bases de datos con aplicación a ejemplos concretos.
 5. Familiarizar en el uso de metodologías y técnicas computacionales y matemáticas:
 - Integrar y analizar datos de sistemas biológicos.

- Profundizar en las metodologías de ingeniería bioinformática mediante tecnologías de análisis de la información biológica y biomédica.
- 6. Integrar diferentes aproximaciones metodológicas:
 - Manejo de bases de datos ómicos.
 - Aplicación de minería de datos.
 - Implementación de métodos de clustering no supervisado.
 - Predicción de metabolismos a partir de secuenciación de genomas completos.
 - Análisis de la dinámica de poblaciones de virus como SARS-CoV-2.
- 7. Preparar para enfrentar desafíos en biomedicina y biotecnología:
 - Equipar a los estudiantes con habilidades necesarias para la investigación biológica avanzada.
- 8. Enfoque multidisciplinar:
 - Familiarizar al alumnado con metodologías desde un punto de vista multidisciplinar.
 - Plantear el curso como teórico-práctico, permitiendo la práctica en aula de los métodos desarrollados en introducciones teóricas.
- 9. Diseño de talleres prácticos:
 - Facilitar la aplicación de conocimientos mediante clases en formato de taller práctico, especialmente en las clases de los viernes.
 - Reforzar la aplicación de conocimientos teóricos en ejemplos reales.
 - Familiarizarse con el manejo de bases de datos biológicos.

PROGRAMA:

- 1. Introducción a la modelización de sistemas biológicos complejos.**
Introducción a los sistemas biológicos complejos. Frontera entre Biología, Informática, Física y Matemáticas. Hacia una teoría general de la Biología. Desde la Biología Teórica a la Biología de Sistemas.
- 2. Ingeniería Bioinformática.**
Introducción a la Ingeniería Bioinformática. Aplicaciones de la Bioinformática. Ciencias de la computación + Biología + Matemática. Información y modelos de datos.
- 3. Introducción a la Biología de sistemas.**
Introducción a las técnicas de alto rendimiento en Biología Molecular: Genómica, Transcriptómica, Proteómica, Metabolómica y Fluxómica. Integración de datos y modelos de “whole-cell”.
- 4. Tecnologías para la gestión de información.**
Modelos tecnológicos para la modelación de información. Modelos de procesos de transformación de bioinformación. Tecnologías para gestión de la información en la bioinformática. Motores de base de datos.
- 5. Evolución molecular y sistemas autorreplicativos.**
Evolución prebiótica y mundo RNA en el contexto de origen de la vida. Modelos de evolución de replicones. Introducción a los conceptos de quasiespecies e hiperciclos. Aplicación a poblaciones actuales de virus RNA como SARS COVID19, HIV, HCV.

6. Minería de datos.

Minería de datos en la Bioinformática. Metodologías y guías para la transformación de la información.

7. Métodos de análisis de datos masivos y clasificación no supervisada basada en Redes Neuronales Artificiales.

Introducción a las Redes Neuronales Artificiales. Aprendizaje supervisado y no supervisado. Clasificación y clustering mediante algoritmo SOM. Ejemplos.

8. Taller práctico y de desarrollo y exposición de trabajos y bibliografía fundamental.

Práctica de clasificación y clustering mediante algoritmo SOM.

9. Fuentes de datos biológicos y herramientas básicas de análisis

Introducción a bases de datos biológicas y tipos de formatos comunes.

10. Análisis de Secuencias y Herramientas Estructurales

Análisis de secuencia BLAST. Algoritmos de búsqueda y alineamiento. Visualización y análisis de estructuras de proteínas.

11. Bases de datos en Bioinformática

Estudio de algunas de las bases de datos biológicas más utilizadas.

12. Herramientas de Visualización y Análisis de Datos.

Visualización de Datos Genómicos con IGV. Carga, interpretación y visualización de archivos BAM y VCF.

13. Evolución molecular y comparación de secuencias.

Diversidad y complejidad genómica. Alineamiento de secuencias.

14. Integración de Datos y Análisis Multiómico.

Análisis multiómico y su aplicación en Bioinformática. Uso de recursos para la bioinformática.

15. Métodos de análisis transcriptómico y su aplicación al diagnóstico en biomedicina.

Flujo de trabajo en estudios transcriptómicos. Tipos de datos. Utilidad de la tecnología ómica en el diagnóstico de las enfermedades raras.

16. Técnicas de visualización y análisis de redes biológicas.

Análisis y Visualización de Redes de Interacción Proteica. Análisis de Datos de Expresión Génica con Python y Shiny.

17. Taller de desarrollo y exposición de trabajos y bibliografía fundamental II

Práctica de simulación de diagnóstico a partir de datos de RNASeq.

ACTIVIDADES PRÁCTICAS:

- Práctica de clasificación y clustering mediante algoritmo SOM.

- Práctica de acceso y manejo de bases de datos biológicos.
- Práctica análisis de secuencias y BLAST.
- Práctica de simulación de diagnóstico a partir de datos de RNASeq.
- Análisis de Datos de Expresión Génica con Python y Shiny.

*Todo el software necesario para el desarrollo de estas prácticas se encuentra preinstalado en los ordenadores del aula o bien es de licencia libre y será instalado previamente a la práctica.

SÍNTESIS CURRICULAR DEL PROFESORADO:



Prof. Dr. Federico Morán, UCM.

- ✚ Catedrático de Bioquímica y Biología Molecular en la Universidad Complutense de Madrid desde 2007, previamente Profesor Titular desde 1986. Actualmente director de la Fundación para el Conocimiento Madrimasd, agencia de acreditación de la calidad de los estudios universitarios de Madrid, desde enero de 2019. Desde 2012 a 2015 ha sido Secretario General de Universidades en el Ministerio de Educación, Cultura y Deporte. Ha sido Investigador Senior y miembro fundador del Centro de Astrobiología (INTA-CSIC) desde su creación en 1999 hasta 2012. Ha sido Director Adjunto del Instituto Nacional de Bioinformática de 2004 a 2011.
- ✚ Formación postdoctoral en la Université Libre de Bruxelles, en el Departamento de Chimie Physique II bajo la dirección del Prof. Ilya Prigogine. Posteriormente en Estados Unidos realizó estancias posdoctorales en la University of Southern California en Los Ángeles (Profs. Michael Arbib y Christoph von der Malsburg) y en la University of California Irvine (Prof. Gary Linch). Desde 1996 a 2012 colaboración permanente de investigación en Stanford University con el Grupo del Prof. John Ross, como Visiting Scholar e investigador asociado en proyectos de National Science Foundation (USA) y NATO Collaborative Research Grants.
- ✚ Docencia Licenciatura y Grado en la UCM de Biofísica, Técnicas Bioquímicas, Biología de Sistemas. Coautor del libro “Biofísica: Procesos de Autoorganización en Biología”. Ha dirigido y organizado diferentes cursos de post-grado y doctorado de Biología teórica y Diseño Molecular, y Rutas Metabólicas. Codirector del Máster de la UCM “Bioinformática y Biología Computacional”, de 2003 a 2012.
- ✚ Ha dirigido e impartido varios cursos de la European School of Theoretical Biology, en Francia. Entre 1998 y 2003 fue Coordinador de los Cursos de Verano del Área de Ciencias de la Universidad Complutense en El Escorial.
- ✚ Actividad permanente de investigación en el Grupo de Biofísica, Departamento de Bioquímica, en el campo de la Biología Teórica y de Sistemas. Investigador Principal en sucesivos proyectos competitivos (Plan Nacional, Consolider-Ingenio, Programas Marco Europeos, Proyectos Comunidad de Madrid) en dinámica de sistemas complejos, redes metabólicas, redes neuronales, evolución molecular, origen de la vida y dinámica molecular. En el Centro de Astrobiología, Director del Laboratorio de Evolución Molecular en investigación sobre evolución prebiótica y origen celular.

- ✚ Producción científica con más de cien publicaciones en revistas JCR. Coautor con Francisco Montero del Libro “Biofísica Procesos de Autoorganización en Biología” (EUDEMA 1992). Editor de varios libros, entre ellos “Advances in Artificial Life” y “Orígenes de la Vida”. Autor de artículos de divulgación y comunicación de la ciencia.
- ✚ Miembro de diferentes sociedades científicas nacionales e internacionales. Miembro fundador de la Sociedad de Biofísica de España. Ha organizado y presidido múltiples congresos y reuniones científicas internacionales.



Prof^a. Dra. Sara Pérez-Luz, UCM.

- ✚ Profesora Ayudante Doctor en la Universidad Complutense de Madrid (UCM) desde 2023 aunque su relación con la UCM empezó en 2020 como Profesora Asociada.
- ✚ Licenciada en CC Biológicas por la Universidad de Alicante y doctora en Bioquímica y Biología Molecular por la Universidad Miguel Hernández de Elche, realizó su etapa predoctoral en la empresa Labaqua S.A. (beca Ministerio de Ciencia) enfocando su estudio en la obtención de un método rápido de diagnóstico molecular que permitiera acortar el tiempo de detección de microorganismos patógenos de transmisión hídrica (Pérez-Luz, et al., 2001; Perez-Luz et al., 1998; Perez-Luz et al., 2002).
- ✚ Su formación postdoctoral transcurre en Londres en el “Medical Research Council” y el “Imperial College London” (beca postdoctoral Ministerio de Educación) dando el salto a las enfermedades humanas, y más concretamente a las enfermedades raras, centrándose el interés no solo en la caracterización fisiopatológica de las mismas, sino también en aproximaciones terapéuticas basadas en la terapia génica. Colaboró en la generación de nuevos vectores para terapia génica que permiten un mantenimiento episomal del gen terapéutico de interés (Kotzamanis et al., 2005) y su proyecto postdoctoral se centró en la creación de vectores de terapia génica para el tratamiento de la hemofilia A (Perez-Luz et al., 2007). De vuelta en España se incorpora al grupo del Dr. Javier Díaz-Nido (CBM-UAM) para continuar en el campo de la terapia génica aplicada a la ataxia de Friedreich, una enfermedad rara neurodegenerativa. Durante este tiempo es responsable de varias líneas de investigación una de las cuales culmina en la generación de un nuevo modelo celular inducible y neuronal de deficiencia de frataxina (Moreno-Lorite et al., 2021) del que derivó una patente (ES 2387787). Asimismo, también es responsable de la línea de investigación dirigida a la generación y mejora de vectores genómicos para la terapia génica de la enfermedad que dio lugar a diferentes publicaciones y un trabajo de tesis doctoral.
- ✚ En 2020 entra a formar parte del Instituto de Investigación de Enfermedades Raras (ISCIH) dentro del grupo de Genética Humana enfocado en el estudio del Déficit de alfa 1 antitripsina usando organoides hepáticos como modelo celular, donde compagina su trabajo de investigación con la labor docente como profesora asociada en la UCM. Actualmente como profesora ayudante doctor en el departamento de Bioquímica y Biología Molecular de la Facultad de Veterinaria (UCM) compatibiliza

su labor docente con la investigación enfocada en la caracterización genética de enfermedades hematológicas raras.



Prof^a. Dra. Natalia Montellano Durán, UCB.

- ✚ Natalia Montellano Durán, PhD. Biotecnóloga y doctora en Ciencias Biológicas de la Universidad Nacional de Rosario en Argentina. Investigaciones en Ciencia de los alimentos, materia blanda, biofísica, propiedades funcionales de proteínas y bioactivos. Directora de Biotecnología en la Universidad Católica Boliviana San Pablo, Santa Cruz - Bolivia. Presidente del capítulo nacional Bolivia de la Organización para mujeres en ciencia en países en desarrollo (OWSD) y Early Career Fellow. Miembro del directorio del consejo de Investigación de la Academia Nacional de Ciencias de Bolivia y miembro fundador de Asociación de Emprendedores de Bolivia. Tutora de iGEM Bolivia.



Prof. José Jesús Cabrera, UCB.

- ✚ Con una licenciatura en robótica y mecatrónica, y una maestría especializada en robótica de servicio e industrial de la Universidad Técnica Estatal de Bauman de Moscú, aporta una sólida formación académica al campo. También ha completado cursos especializados en visión por computadora del Laboratorio Huawei, que tuvieron lugar en BMSTU, lo que ha mejorado su experiencia en tecnologías de vanguardia.
- ✚ En cuanto a la experiencia profesional, ha tenido la oportunidad de trabajar en GAZ Group, una empresa líder en la industria automotriz en Rusia, donde aplicó sus habilidades en robótica y mecatrónica para desarrollar soluciones innovadoras para la automatización industrial.
- ✚ Además, ha servido como profesor en la Universidad Católica Boliviana, donde compartió su conocimiento y pasión por la robótica y la mecatrónica con los estudiantes. Actualmente, ocupa el cargo de coordinador en el departamento de ingeniería de software, supervisando el desarrollo del currículo y asegurando que los estudiantes reciban una educación integral.
- ✚ Con una sólida base en robótica y mecatrónica, junto con una experiencia práctica en la industria y una dedicación a la educación, está bien preparado para contribuir a proyectos de vanguardia en el campo. Sus habilidades abarcan lenguajes de programación como Python y C++, y tiene experiencia práctica con marcos y herramientas de software de robótica, como ROS, RobotStudio y programación de KUKA.



Prof. Iván Mendoza Moya, UCB.

- ✚ **Más de 15000 horas en gestión de proyectos:** Gestión de proyectos, diseño, desarrollo y control en tecnología de computadoras: tics, management hydrocarbons, seguridad, tic, almacenes, rrhh, hsi, gerenciales, académicos y otros. Metodologías y

estándares: ISO 9000, TOGAF, EA, ITIL v4, COBIT, gestión de riesgos, gestión de cambio, PMBOK v6, gestión de pérdidas, ISO 31000, ISO 31022.

- ✚ **25 años en la industria:** Director de Proyectos. Unidad Multidisciplinaria Agropecuaria y Gestión en coordinación con USAID Jefe TI. Consorcio Hidroeléctrico Misicuni – Construcción civil. Coordinador Logística. Empresa Minera Inti Raymi S.A - Industria Minera aurífera (Gestión de 1.500.000 \$us/mes). Jefe TI, Analista de Adquisiciones. Complejo Metalúrgico Vinto S.A. Industria Metalúrgica. Jefe TI. Allied Deals Minera Huanuni - Industria Minera. Jefe TI. Sinchi Wayra S.A. - Industria Minera. Estadista procesos industriales. Fábrica de Cemento EMISA S.A. - Industria Metalúrgica. Jefe TI. Embotelladora PEPSI S.A. - Industria Alimenticia.

- ✚ **20 años en universidades pregrado y postgrado:** Docente, Universidad Católica Boliviana Santa Cruz. Docente, Universidad UNIVALLE Santa Cruz, Cochabamba, La Paz. Docente, Universidad Simón y Patiño, Cochabamba. Docente, Universidad Escuela Militar de Ingeniería EMI Santa Cruz. Docente, INFOCAL - Santa Cruz. Docente, Universidad La Salle - La Paz. Docente Postgrado, Universidad La Salle - La Paz. Docente Postgrado. Universidad del Valle UNIVALLE Cochabamba. Docente. Empresa de certificación Microsoft – COGNOS. Docente. Universidad Franz Tamayo UNIFRANZ. Jefe de Carrera, Jefe Proyectos, Docente. Universidad Ciencias. Tecnológicas y Administrativas UCATEC. Docente. Universidad Latinoamericana ULAT. Docente. Tecnológico Nacional Andrés Bello. Docente. Colegio Don Bosco. Docente. Anglo American School. Docente. Colegio de la Unión Bolivia Japón. Docente. Colegio Evangélico Reekie. Docente. Instituto CATEC.