

UNIVERSIDAD NACIONAL DEL COMAHUE

FACULTAD DE INGENIERÍA
Secretaría de Posgrado

Doctorado/Maestría en Enseñanza de las Ciencias Exactas y Naturales

Denominación del Curso:

Fundamentos de Modelización Matemática y Machine Learning en Biología

Docentes a cargo*:

Dr. Néstor Hugo Biedma

* Adjuntar Curriculum Vitae del profesor coordinador

Modalidad: Virtual por aislamiento
COVID-19

Carga Horaria Total: 60 hs.

• **Fundamentación:**

La modelización matemática está siendo aplicada con éxito en diversas áreas de la ciencia y la tecnología, en particular en la biología y en las ciencias médicas.

En este curso se pretende poner de manifiesto la potencialidad de la matemática para plantear modelos clásicos y de machine learning, con hincapié particularmente, en la interacción entre la matemática y la biología en los procesos de modelización, pero sin alejarnos de una mirada global que nos da la modelización.

La presentación y aprendizaje de las técnicas numéricas, de minería de datos y computacionales, tanto en CPU como en GPU, requeridas para el planteo y la resolución de los modelos, es uno de los contenidos que se proponen abordar. Simultáneamente, se desea alentar el uso de estas herramientas con el fin de generar modelos específicos y obtener desarrollos innovadores.

En la actualidad, la matemática es el soporte de herramientas como SVM, Redes Neuronales y Estructuras de Arboles de decisiones, entre varias otras, que son parte constante del proceso de diseño de modelos, con deficiencia explicativas en algunos casos, pero alto poder predictivo en la mayoría de las aplicaciones exitosas. Estos permiten afrontar estudios que antes eran casi inaccesibles.

Si bien en tema central de los modelos está basado en nociones biológicas, utilizaremos ejemplos y set de datos provenientes de la vida cotidiana que hacen que el abordaje pueda realizarse desde cualquier ciencia natural o exacta. Por lo que el curso puede ser tomado

tanto por egresados de carreras de grado en Matemática, Física, Biología, Ingeniería, o cualquier otra disciplina (que satisfagan los requisitos indicados más abajo) y que se interesen en el desafío de resolver problemas difíciles e interesantes, especialmente, provenientes del campo de la biología.

- **Requisitos**

Se necesita que los cursantes tengan:

Conceptos de métodos numéricos, Nociones de Ecuaciones Diferenciales, Nociones de programación y manejo de entornos computacionales – Nociones Algebra Lineal, Probabilidad y Estadística. En cuanto a lenguajes de programación, utilizaremos algo de Python y R, traduciremos algunas instrucciones de otros lenguajes para adaptarnos a los entornos. Si no se conocen estos lenguajes, no es problema, mientras tenga nociones básicas de programación, pues daremos una breve introducción a su manejo y entornos.

- **Objetivos**

Se propicia y espera que los cursantes sean capaces de:

- Tener nociones del diseño de modelos matemáticos en entornos computacionales orientados a la Biología.
- Apliquen y manejen los conceptos de Machine Learning básicos con una perspectiva Matemática.
- Hagan uso y comprendan el papel de los modelos computacionales y las estrategias de aprendizaje estadístico en la investigación actual.
- Profundicen las ideas de modelos matemáticos y los sustentos conceptuales asociados a este proceso para los entornos de ciencias naturales.
- Tengan un ámbito de discusión e investigación sobre estos tópicos en la enseñanza y el aprendizaje mediante modelización matemática.

- **Modalidad de Cursado**

Se realizarán encuentros virtuales durante 6 semanas, dos por cada una. En cada semana tendremos un encuentro principal durante el sábado de trabajo teórico práctico de 6 horas dividido en bloques desde las 9:30 a 15:30 (la duración de los bloques se pautará en clase) y un encuentro de practica los lunes por la tarde en horario a estipular de 2 horas máximo, para consultas y desarrollo de tópicos específicos. Utilizaremos la plataforma Google Meet. Se hará llegar el enlace por grupo virtual en WhatsApp que se organizará para el cursado. Se requiere uso de computadora y conexión a internet estable.

- **Programa Analítico**

Unidad I – Modelización Matemática y la población como un sistema dinámico

Modelización Matemática – Modelización y simulaciones – Computadoras como bases de simulaciones – Teoría de sistemas y su importancia - Sistemas de Álgebra Computacional - Las escalas en los modelos – Sistemas dinámicos- Sistemas biológicos como sistemas dinámicos – Modelos continuos y discretos – Modelos deterministas y no deterministas – El tiempo como factor de escala en los modelo biológicos - La dinámica poblacional - Ejemplos de modelos simples deterministas - Análisis de modelos lineales – Aplicaciones matriciales – Técnicas de autovalores y autovectores - Análisis de modelos no lineales: desarrollo de sistemas de ecuaciones diferenciales – El modelo de Lotka Volterra - Equilibrio y estabilidad – Análisis de sensibilidad y robustez - Análisis de límites – Variantes estocásticas – Recolección y tratamiento de Datos – Modelos basados en Machine Learning – Modelos Deterministas Vs. Modelos de Aprendizaje Estadístico.

Bibliografía: [1,4, 16, 17, 18]

Unidad II – Modelos computacionales y la genética

Dogma central de la biología molecular – DNA, RNA, proteínas, sus estructuras generales y algunas funciones - Bioinformática y su relación con la matemática – Biología y bases de datos – Modelado y acceso a datos – Lectura de datos y su tratamiento - El papel de la optimización y la minería de datos – Uso de estrategias estadísticas para inferencias sobre modelos basados en datos –El papel de los algoritmos – Big Data y Biomatemática – Uso de software y programación para modelados – Rudimentos de Programación en R – Programación en Python - Paquetes basados en R y Phyton – Nociones de álgebra de Tensores - Tensor Flow – SciKit Learn – Perspectiva de GPU vs CPU - Aplicaciones.

Bibliografía: [1,2,3,4,5,6,7, 16,17, 18]

Unidad III – Modelos de enfermedades contagiosas y el uso de elementos estadísticos

Modelos epidémicos elementales – número básico de reproducción – Difusión de la epidemia y su duración – Tasa de contacto – Variaciones en la epidemia y cambios en la dinámica de comportamiento – Estrategias de inmunización – Poblaciones múltiples e infecciones diferenciadas- Uso de ajustes de curvas con y sin polinomios – Método de mínimos cuadrados.

Bibliografía: [1,8]

Unidad IV- Modelando Casos particulares – Introducción a la matemática vinculada al estudio de situaciones Biológicas y Médicas

Conceptos básicos aspectos biológicos y médicos de temas seleccionados – Modelado matemático de tumor génesis – Introducción al modelado mediante ecuaciones diferenciales - Modelos estocásticos y basados en análisis de datos – Uso de casos clásicos abordados con R y Python - Nociones sobre el uso de autómatas celulares – SVM - Arboles de Decisión - Modelos basados en Redes Neuronales y otras estrategias de machine learning - Modelos híbridos.

Bibliografía: [9,10, 11,12,13,14,15, 16, 17,18]

- **Modalidad de Evaluación**

- Resolución y defensa de problemas seleccionados asignados
- Examen Final en formato Virtual

- **Bibliografía**

Bibliografía (Se estipulará bibliografía básica según los preconceptos de los participantes, así como nueva bibliografía de ser necesario)

- [1] ALLMAN E. S., RHODES J. A. – Mathematical Model in Biology: An Introduction – Cambridge University Press – New York 2004
- [2] CHEN, J., LONARDI, S., Biological Data Mining, CRC Press, Minnesota USA, 2010
- [3] HOFACKE, I., SCHUSTER, P., STADLER, P., Combinatorics of RNA secondary structures. Discrete Applied Mathematics, V. 88, Issues 1-3, pp. 207 – 237, 1998.
- [4] ISAEV, A., Introduction to Mathematical Methods in Bioinformatics, Springer, Berlin, 2006
- [5] MOULTON, V., ZUCKER, M., STEEL, M., POINTON, R., PENNY, D., Metrics on RNA secondary structures, Journal of Computational Biology, V. 7, Numbers ½, pp. 277-292, Mary Ann Liebert Inc.
- [6] MOUNT, D., Bioinformatic: Sequence and Genome Analysis, Cold Springer Harbor Laboratory
- [7] NUSSINOV, R., PIECZENIK, G., GRIGGS, J., KLEITMAN, D., Algorithms for Loop Matching, SIAM Journal on Applied Mathematics, V. 35, Issue 1, pp 68-82, 1978.
- [8] MURRAY, J.D. – Mathematical Biology: An Introduction – Vol 1. – Third Edition – Springer-Washington 2002
- [9] QUIROGA, A.A.I. ; FERNANDEZ, D.; TORRES, G.A.; TURNER, C. -Adjoint method for a tumor invasion PDE-constrained optimization problem in 2D using adaptive finite element method – Applied Mathematics and Computation, Volume 270, 1 November 2015, Pages 358–368
- [10] KNOPOFF, D.A.; FERNÁNDEZ, D.R.; TORRES, G.A.; TURNER, C.V. - Adjoint method for a tumor growth PDE-constrained optimization problem - Computers & Mathematics with Applications - Volume 66, Issue 6, October 2013, Pages 1104–1119
- [11] AGNELIA, J.P.; PADRA, C.; TURNER, C.V. - Shape optimization for tumor location - Computers & Mathematics with Applications - Volume 62, Issue 11, December 2011, Pages 4068–4081
- [12] MATERI, W.; WISHART, D. S. - Computational Systems Biology in Cancer: Modeling Methods and Applications - Gene Regul Syst Bio. 2007; 1: 91–110.

[13] JACKSON, T.; KOMAROVA, N.; SWANSON, K. - Mathematical Oncology: Using Mathematics to Enable Cancer Discoveries, The American Mathematical Monthly, Vol. 121, No. 9 (November 2014), pp. 840-856

[14] WODARZ, D.; KOMAROVA, N. L. –Computational Biology of Cancer: Lectures Notes and Mathematical Modeling –World Scientific Publishing Co. Pte. Ltd. 2005

[15] WODARZ, D.; KOMAROVA, N. L. – Dynamics of Cancer: Mathematical Foundations of Oncology – World Scientific Publishing Co. Pte. Ltd. 2014

[16] ALAN M. MOSES - Statistical Modeling and Machine Learning for Molecular Biology – Chapman and Hall.

[17] RACHEL SCHUTT – CATHY O’NEIL – Doing Data Science – O’Reilly, 2014

[18] AURÉLIEN GÉRON - Hands-On Machine Learning with, Scikit-Learn and TensorFlow – O’Reilly, Second Edition 2019

- **Correo electrónico del profesor Coordinador:**

nestbi7@gmail.com