



## **Ficha técnica de materia optativa**

<b>Nombre del curso:</b> Aplicaciones de la IA en las ciencias genómicas
<b>Docente:</b> Dr. Omar Fabián Hernández-Zepeda
<b>Día y horario:</b> Martes 8:00 a 11:00 horas y jueves de 08:00 a 10:00 horas
<b>Cupo máximo:</b> 20 personas
<b>Criterios de inscripción (si aplica):</b> Mostrar inquietud, iniciativa y ser propositivo para desarrollar y desarrollar trabajo e investigación en laboratorio, industria y bioinformática con datos de las ciencias Omica's (genómica, transcriptómica, proteómica, epigenética y metabolómica).
<b>Conceptos básicos:</b>  <b>General:</b> Analizar y comprender la relevancia del desarrollo, uso y aplicación de la inteligencia artificial en las ciencias genómicas y en la industria biotecnológica.  <b>Específicos:</b> <ol style="list-style-type: none"><li>1) Análisis y comprensión del uso de datos genómicos.</li><li>2) Análisis y comprensión de modelos de machine learning.</li><li>3) Análisis del desarrollo de APP con uso de IA en la industria biotecnológica.</li></ol>
<b>Justificación:</b>  El uso y aplicación de la inteligencia artificial (IA) y machine learning han revolucionado el uso y análisis de los datos Genómicos. En la actualidad el manejo adecuado de los datos masivos generado por estrategias de las ciencias Omica's resulta en una encrucijada mundial que incorporar aspectos tanto tecnológicos como éticos de suma relevancia. Además de que dentro de las industrias se ha optado por incorporar dicha tecnología para resolver problemas complejos. Países como Estados Unidos de América, Reino Unido y Japón por mencionar algunos han destinado gran capital de gubernamental para asegurar el desarrollo e implementación de la IA para facilitar los procesos actuales. Una de las industrias más importantes que han decidido incorporar las aplicaciones de IA son la industria Farmacéutica y Agro-Alimentaria. La industria Farmacéutica ha revolucionado el campo del uso de la IA incorporándola en: <ul style="list-style-type: none"><li>• <b>Identificación de dianas farmacológicas:</b> Los datos ómicos ayudan a identificar nuevas dianas terapéuticas al revelar genes, proteínas y vías moleculares clave involucradas en la enfermedad.</li><li>• <b>Desarrollo de fármacos más específicos:</b> Los datos ómicos permiten diseñar fármacos que se dirijan específicamente a las moléculas alteradas en la enfermedad, minimizando los efectos secundarios.</li></ul>



- **Predicción de la respuesta al fármaco:** Los datos ómicos pueden predecir cómo un individuo responderá a un fármaco en función de su perfil genético y molecular, lo que permite una medicina más personalizada.

La industria Agro- Alimentaria los aspectos más relevantes en los cuales se incorpora la IA son:

- **Resistencia a plagas y enfermedades:** La genómica permite identificar genes de resistencia en las plantas y animales, lo que facilita el desarrollo de variedades más resistentes y reduce la necesidad de pesticidas y antibióticos.
- **Incremento de la productividad:** El análisis ómico ayuda a comprender los mecanismos moleculares que controlan el crecimiento y desarrollo de plantas y animales, permitiendo optimizar las prácticas agrícolas y ganaderas para aumentar la producción.
- **Mejora de la calidad nutricional:** La metabolómica permite identificar y cuantificar los compuestos bioactivos presentes en los alimentos, lo que facilita el desarrollo de variedades con mayor valor nutricional y propiedades beneficiosas para la salud.

En México es esencial que los desarrollo de la IA se incorpore en la producción agrícola sustentable y así como en la industria farmacéutica.

**Objetivo general:**

Analizar y comprender la relevancia del desarrollo, uso y aplicación de la inteligencia artificial en las ciencias genómicas y en la industria biotecnológica.

**Objetivos específicos:**

Analizar y comprender los conceptos principales para el análisis de datos Omica's

Conocer los fundamentos de IA y su relación con la genómica.

Análisis de variantes genéticas con herramientas de IA.

Analizar la integración de la IA a la industria biotecnológica y sus derivados.

**Método de trabajo: Teórico – Práctico.**

Herramientas y Software Para Usar

- Python: Para introducir librerías como Scikit-learn, TensorFlow o Keras.
- Google Colab: Entorno gratuito para ejecutar modelos en la nube.
- Bases de datos genómicas: 1000 Genomas, GTEx, o bases públicas como BioProjects.
- AlphaFold para estructuras proteicas y DeepVariant para variantes genómicas.



### **Prácticas propuestas para desarrollar:**

Predicción de enfermedades genéticas:

1. Usar datos simulados o públicos (como 1000 genomas) para entrenar un modelo ML básico que prediga la probabilidad de una enfermedad en plantas y en humanos.
2. Visualización de datos genómicos:
  - Crear un modelo de clustering que agrupe individuos según similitudes genéticas usando datos reales o simulados.
3. Clasificación de variantes genómicas:
  - Implementar un modelo para diferenciar variantes patogénicas y benignas utilizando herramientas básicas de IA.
4. Análisis de expresión génica:
  - Aplicar IA para interpretar datasets de RNA-seq y encontrar genes diferencialmente expresados.
5. Propuesta para el desarrollo de APPs con la incorporación del IA para el análisis de datos Omico's.
  - Analizar a profundidad los modelos de desarrollo de APPs para la incorporación de IA en el análisis de datos Omico's.

### **Criterios de evaluación:**

Diagnóstica: Análisis y evaluación de los conocimientos previos.

Formativa: Para recomendar los materiales correspondientes a cada estudiante.

Sumativa: Ejercicios de asimilación conceptual, así como un ejercicio integrador final.

Diagnóstica: 10 %

Formativa: 30 %

Sumativa: 50 %

Autoevaluación y coevaluación: 10 %

## **Temario**

### **Unidad 1**

Introducción a la IA en Genómica

- ¿Qué es la inteligencia artificial y cómo se aplica en genómica?
- Conceptos básicos de IA, machine learning (ML) y deep learning (DL).
- Casos de uso en la investigación genómica: desde la predicción de variantes funcionales hasta el diseño de fármacos.
- Bases de datos y big data en genómica:
- Cómo la IA ayuda a analizar grandes volúmenes de datos genómicos (p. ej., datos de secuenciación masiva).
- Introducción a bases de datos genómicas que usan IA: ClinVar, Ensembl y más.



## **Unidad 2**

### Predicción de Variantes Genéticas con IA

- Análisis funcional de variantes genéticas:
- Cómo modelos de ML identifican variantes asociadas a enfermedades.
- Ejemplos prácticos: herramientas como DeepVariant de Google y AlphaFold de DeepMind.
- Redes neuronales para clasificar mutaciones:
- Entrenamiento de un modelo básico para diferenciar variantes benignas y patogénicas.
- Breve taller usando Python y bibliotecas como TensorFlow o PyTorch.

## **Unidad 3**

### IA en la Interpretación de Datos Genómicos

- Medicina personalizada con IA:
- Predicción de respuestas a medicamentos según el perfil genético (farmacogenómica).
- Algoritmos que priorizan genes candidatos en enfermedades complejas.
- Metagenómica y microbiomas:
- Uso de IA para analizar comunidades microbianas y su relación con la salud humana.
- Proyectos como QIIME2 y la integración de modelos predictivos.

## **Unidad 4**

### IA y Secuenciación de Nueva Generación (NGS)

- Procesamiento de datos de secuenciación:
- Cómo la IA mejora la precisión en el ensamblaje de genomas.
- Algoritmos para detectar errores en datos NGS.
- Descubrimiento de biomarcadores:
- Uso de IA para identificar biomarcadores genéticos en enfermedades como el cáncer o desórdenes raros.

## **Unidad 5**

### Ética y Limitaciones de la IA en Genómica

- Privacidad y seguridad de los datos genéticos:
- Riesgos asociados con el uso de datos genómicos y herramientas basadas en IA.
- Regulaciones como el RGPD y su impacto en la investigación genómica.
- ¿Qué pasa con los sesgos en la IA?



- Cómo los sesgos en datasets genéticos pueden afectar a poblaciones subrepresentadas.
- Estrategias para desarrollar IA inclusiva en genómica.

## Unidad 6

- Aplicaciones más utilizadas en la industria biotecnológica y sus derivados para el análisis de datos Omico's.
- Principales desarrolladores de IA y aplicaciones para el análisis de datos Omico's
- Industria Agro- alimentaria.
- Industria farmacéutica.
- Industria de desarrollo de software.

## Bibliografía básica:

- Alharbi, W. S., & Rashid, M. (2022). A review of deep learning applications in human genomics using next-generation sequencing data. *Human Genomics*, 16(1), 1–20. <https://doi.org/10.1186/s40246-022-00396-x>
- Chevallereau, A., Blasdel, B. G., De Smet, J., Monot, M., Zimmermann, M., Kogadeeva, M., Sauer, U., Jorth, P., Whiteley, M., Debarbieux, L., & Lavigne, R. (2016). Next-Generation “-omics” Approaches Reveal a Massive Alteration of Host RNA Metabolism during Bacteriophage Infection of *Pseudomonas aeruginosa*. *PLoS Genetics*, 12(7), 1–20. <https://doi.org/10.1371/journal.pgen.1006134>
- Deka, P. C. (2024). Genomics, proteomics, and metabolomics. In *Research Anthology on Bioinformatics, Genomics, and Computational Biology*. <https://doi.org/10.4018/979-8-3693-3026-5.ch037>
- Dias, R., & Torkamani, A. (2019). Artificial intelligence in clinical and genomic diagnostics. *Genome Medicine*, 11(1), 1–12. <https://doi.org/10.1186/s13073-019-0689-8>
- Ferguson, J. F., Allayee, H., Gerszten, R. E., Ideraabdullah, F., Kris-Etherton, P. M., Ordovás, J. M., Rimm, E. B., Wang, T. J., & Bennett, B. J. (2016). Nutrigenomics, the microbiome, and gene-environment interactions: New directions in cardiovascular disease research, prevention, and treatment. *Circulation: Cardiovascular Genetics*, 9(3), 291–313. <https://doi.org/10.1161/HCG.0000000000000030>
- Ferguson, L. R., De Caterina, R., Görman, U., Allayee, H., Kohlmeier, M., Prasad, C., Choi, M. S., Curi, R., De Luis, D. A., Gil, Á., Kang, J. X., Martin, R. L., Milagro, F. I., Nicoletti, C. F., Nonino, C. B., Ordovas, J. M., Parslow, V. R., Portillo, M. P., Santos, J. L., ... Martinez, J. A. (2016). Guide and Position of the International Society of Nutrigenetics/Nutrigenomics on Personalised Nutrition: Part 1 - Fields of Precision Nutrition. *Journal of Nutrigenetics and*



- Nutrigenomics*, 9(1), 12–27. <https://doi.org/10.1159/000445350>
- Gajare, S. R., Deshmukh, A. S., & Shinde, C. K. (2021). Personalized Medicine: A Review. *International Journal of Pharmaceutical Sciences Review and Research*, 69(1). <https://doi.org/10.47583/ijpsrr.2021.v69i01.023>
- Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics*, 16(6), 321–332. <https://doi.org/10.1038/nrg3920>
- Maqsood, K., Hagra, H., & Zabet, N. R. (2024). An overview of artificial intelligence in the field of genomics. *Discover Artificial Intelligence*, 4(1). <https://doi.org/10.1007/s44163-024-00103-w>
- Raza, S. (2020). *Artificial intelligence for genomic medicine*. 1–50. [www.phgfoundation.org](http://www.phgfoundation.org)
- Robert, J. (2006). Pharmacogenetics and pharmacogenomics. *Bulletin Du Cancer*, 93(SPEC. ISS. APR.), 113–123. <https://doi.org/10.1385/1-59259-850-1:235>
- Sales, N. M. R., Pelegrini, P. B., & Goersch, M. C. (2014). Nutrigenomics: Definitions and advances of this new science. *Journal of Nutrition and Metabolism*, 2014. <https://doi.org/10.1155/2014/202759>
- Sergey Nurk, Sergey Koren<sup>1</sup>, Arang Rhie<sup>1</sup>, Mikko Rautiainen<sup>1</sup>, Andrey V. Bzikadze, A., Mitchell R. Vollger, Nicolas Altemose, Lev Uralsky, Ariel Gershman, S. A., Savannah J. Hoyt, Mark Diekhans<sup>1</sup>, Glennis A. Logsdon<sup>4</sup>, Michael Alonge, S. E. A., mosaic of haplotypes. As a result, several G., Matthew Borchers, Gerard G. Bouffard, Shelise Y. Brooks, Gina V. Caldas, N.-C. C., In-, assembly gaps are unsolvable because of, Haoyu Cheng, Chen-Shan Chin<sup>18</sup>, William Chow, Leonardo G. de Lima, P. D., Richard Durbin, Tatiana Dvorkina<sup>3</sup>, Ian T. Fiddes, Giulio Formenti, R. S., Arkarachai Functammasan, Erik Garrison, Patrick G. S. Grady, T. A. G.-L., Ira M. Hall, Nancy F. Hansen, Gabrielle A. Hartley, Marina Haukness<sup>11</sup>, K. H., Michael W. Hunkapiller, Chirag Jain, Miten Jain<sup>11</sup>, Erich D. Jarvis<sup>22, 23</sup>, P. K., Contains, T. Grc. reference assembly, Melanie Kirsche, Mikhail Kolmogorov, Jonas Korlach, Milinn Kremitzki, Heng Li, Se, Valerie V. Maduro, Tobias Marschall, Ann M. McCartney<sup>1</sup>, Jennifer McDaniel, Danny E. Miller, Genome, quence distributed throughout the, James C. Mullikin, Eugene W. Myers, Nathan D. Olson, Benedict Paten<sup>11</sup>, P. P., Subtelomeric, including pericentromeric and, Pavel A. Pevzner, David Porubsky, Tamara Potapova, Evgeny I. Rogae, J. A. R., ... Michael C. Schatz, Evan E. Eichler, Karen H. Miga A. M. P. (2022). The complete sequence of a human genome. *Science*, 376 (1 April 2022), 44–53. <https://doi.org/DOI: 10.1126/science.abj6987>
- Zeeshan, S., Xiong, R., Liang, B. T., & Ahmed, Z. (2023). Editorial: Artificial intelligence for personalized and predictive genomics. *Frontiers in Genetics*, 21(3), 885–905. <https://doi.org/https://doi.org/10.3389/fgene.2023.1162869>

