

# Machine Learning para Bioinformática 2021 (Online vía Zoom)

## Descripción

El curso Machine Learning para Bioinformática refuerza conocimientos estadísticos, familiariza con el lenguaje de programación **Python** y capacita en la comprensión e implementación de técnicas de aprendizaje automático actualmente empleadas en el análisis de grandes volúmenes de datos, incluyendo desde datos ómicos como RNA-seq, Microbiomas y proteomas, hasta mediciones eléctricofisiológicas en Neurociencia.

La estructura general del curso incluye una Introducción a Python en la nube para personas sin experiencia en programación, con el fin de obtener los conocimientos mínimos para continuar con las secciones de Aprendizaje Automático No Supervisado, y finalmente Aprendizaje Automático Supervisado.

Cada sesión constará de una exposición teórica seguida de un laboratorio práctico, en donde se abordarán temas fundamentales requeridos para luego continuar con el curso de **Deep Learning** en el segundo semestre, donde se profundizará en técnicas avanzadas de Inteligencia Artificial basadas en redes neuronales artificiales aplicadas a problemas biológicos, incluyendo la traducción de datos multi-dominio (Imagen a Expresión Génica, y viceversa, por ejemplo).

## Objetivo General

Capacitar a sus asistentes en la comprensión teórica e implementación práctica de metodologías de Aprendizaje Automático aplicadas a la resolución de problemas en Ciencias Biológicas.

## Horario

Lunes 19:00 hasta las 21:30 horas. Vía Zoom.

## Público objetivo

Este curso intensivo está especialmente dirigido a personas trabajando en Ciencias Biológicas incluyendo Biología Molecular, Bioquímica, Biotecnología, Bioinformática, Ecología, Microbiología, Inmunología y Tecnología Médica, entre otras áreas afines.

## Programa

Semana 1 (lunes 5 de abril de 2021):

- Repaso conceptos generales de estadística clásica o frecuentista (valor p).
- Introducción a la programación en Python.
- Tipos de variables y estructuras de datos. Flujos de control y condicionales.
- Práctico en Python: Primeros pasos en la nube.

Semana 2 (lunes 12 de abril de 2021):

- Repaso conceptos generales de estadística Bayesiana (probabilidades a posteriori).
- Introducción a la programación en Python.
- Funciones y clases.
- Práctico en Python: Introducción a la computación científica con Python.

Semana 3 (lunes 19 de abril de 2021):

- Introducción a la programación en Python.
- Arreglos y operaciones matemáticas: Uso de paquetes NumPy y SciPy.
- Manipulación de Dataframes: Uso de paquete Pandas.
- Práctico en Python: Pre-procesamiento de datos tabulados. Casos Covid-19 Chile.

Semana 4 (lunes 26 de abril de 2021):

- Introducción a la programación en Python.
- Visualización de datos: Uso de paquetes Matplotlib, Plotly.
- Práctico en Python: Implementación de herramientas de visualización.

Semana 5 (lunes 3 de mayo de 2021):

- Aprendizaje Automático No Supervisado I.
- Clustering Jerárquico, K-means, DBSCAN.
- Práctico en Python: Exploración y visualización de Datos transcriptómicos I.

Semana 6 (lunes 10 de mayo de 2021):

- Aprendizaje Automático No Supervisado II.
- Reducción Dimensionalidad: PCA, tSNE, UMAP.
- Práctico en Python: Exploración y visualización de Datos transcriptómicos II.

Semana 7 (lunes 17 de mayo de 2021):

- Minería de Datos: Selección de Atributos.
- Métricas y estrategias para la selección de variables.
- Práctico en Python: Por qué y Cómo escoger las variables más “útiles” para nuestros modelos.

Semana 8 (lunes 24 de mayo de 2021):

- Automatización y búsqueda de hiperparámetros.
- Pipelines y GridSearch.
- Práctico en Python: Creación de flujos de trabajo y búsqueda automática de hiperparámetros.

Semana 9 (lunes 31 de mayo de 2021):

- Aprendizaje Automático Supervisado I.
- Métricas y Métodos para evaluar desempeño.
- Accuracy, Precision, Recall, F1, ROC.
- Holdout, Muestreo Aleatorio, Validación cruzada. Evaluando overfitting.
- Práctico en Python: Cómo comparar y escoger modelos.

Semana 10 (lunes 07 de junio de 2021):

- Aprendizaje Automático Supervisado II.
- Support Vector Machines (SVM), Naive Bayes, k-Nearest Neighbors (KNN).
- Práctico en Python: Clasificación de enfermedades basados en datos de Microbioma I.

Semana 11 (lunes 14 de junio de 2021):

- Aprendizaje Automático Supervisado III: .
- Árboles de Decisión, Random Forest, XGBoost.
- Práctico en Python: Clasificación de enfermedades basados en datos de Microbioma II.

**Instructor:**

Luis Valenzuela Villa, PhD Biotecnología Molecular, Universidad de Chile.  
Investigador Postdoctoral, INTA, Universidad de Chile, y Fundador Omics Lab SpA.  
<https://www.linkedin.com/in/luisvalenzuelavilla>

Costo Curso: \$60.000.

Inscripciones y más información: [luis.valenz.v@gmail.com](mailto:luis.valenz.v@gmail.com)