**Dr. Jorge Méndez Astudillo**
**Dr. Gustavo Carreón**
**Instituto de Investigaciones Económicas**

# Introducción a la ciencia de datos para economistas

**Objetivo:** Que el alumno aprenda conceptos básicos de las técnicas de aprendizaje automatizado (Machine Learning) y pueda aplicarlos en investigaciones de su área de conocimiento. El componente práctico del curso requiere acceso a una computadora con buenas capacidades de procesamiento, así como acceso a Python 3. Se recomienda que el estudiante tenga conocimientos básicos de estadística y de programación, así como un nivel básico de inglés (para poder leer documentación de Python).

**Horario:** 3 horas a la semana durante un semestre todos los jueves de 4 a 7pm

Se requiere acceso a un aula de cómputo y de preferencia a un ayudante para resolver dudas de los estudiantes.

Temas:

1. Conceptos básicos de ciencia de datos
   - 1.1 Definiciones de ciencia de datos (Data Science)
     - 1.1.1 Bases de datos
     - 1.1.2 Big Data System
     - 1.1.3 Estadística
     - 1.14 Minería de datos (Data Mining)
     - 1.1.5 Visualización de datos
   - 1.2 Inteligencia Artificial, (AI)
   - 1.3 Aprendizaje Automatizado, (ML)
   - 1.4 Aprendizaje Profundo, (DL)

     En todos los casos ejemplos de uso en la economía
2. Introducción a Python
   - 2.1 Introducción a estructuras de datos en Python
   - 2.2 Uso de Python Notebooks
   - 2.3 Técnicas para limpiar textos
   - 2.4 Redes de palabras y bigramas
   - 2.5 Nubes de etiquetas (wordclouds)
   - 2.6 Numpy y Pandas
   - 2.7 Geopandas
3. Regresión
   - 3.1 Conceptos teóricos de la regresión
   - 3.2 Regresión Lineal
   - 3.3 Regresión logística
   - 3.4 Support Vector Machine
   - 3.5 Ejercicio práctico de regresión.
4. Clasificación
   - 4.1 Conceptos teóricos de clasificación
   - 4.2 K-nearest neighbour
   - 4.3 Decision Tree
   - 4.4 Ejercicio práctico de clasificación
5. Clustering
   - 5.1 Conceptos teóricos de unsupervised learning
   - 5.2 K-means

5.3 DBSCAN
5.4 Ejercicio práctico de clustering

6. Redes Neuronales
   6.1 Teoría de RN
   6.2 Perceptor simple
   6.3 Perceptor multicapa
   6.4 Aplicaciones de RN
   6.5 Mapas auto organizables (Teórico y práctico)
7. El futuro
   7.1 Deep Learning
   7.2 Large Language Models

**Bibliografía sugerida:**


- Documentación de Python: https://docs.python.org/3/download.html

- Documentación Aprendizaje Automático: https://machinelearningmastery.com

- Hal Varian (2014), "Big Data: New Tricks for Econometrics", Journal of Economic Perspectives, Spring, 3-28.

- Jon Kleinberg, H. Lakkaraju, Jure Leskovec, Jens Ludwig, Sendhil Mullainathan (2018), "Human Decisions and Machine Predictions", Quarterly Journal of Economics, 237-293.

- Hugo Storm, Kathy Baylis, Thomas Heckelei, Machine learning in agricultural and applied economics, *European Review of Agricultural Economics*, Volume 47, Issue 3, July 2020, Pages 849–892, https://doi.org/10.1093/erae/jbz033


Semblanza del Dr. Jorge Méndez Astudillo

El Dr. Jorge Méndez Astudillo es investigador asociado C en el Instituto de Investigaciones Económicas de la UNAM desde marzo 2022. Es miembro del Sistema Nacional de Investigadores nivel candidato. Obtuvo su doctorado en ciencias digitales en la universidad de Nottingham (Reino Unido), donde aprendió a manejar grandes volúmenes de datos, así como a analizarlos usando técnicas de aprendizaje automatizado y estadística. Tiene experiencia postdoctoral en el Instituto de Geografía de la UNAM donde usó imágenes satelitales y técnicas de análisis geoespacial para estudiar la contaminación atmosférica en las principales ciudades de México. Ha publicado al menos 7 artículos en revistas indizadas como la revista “Aerosol and Air Quality Research” y la “Revista Internacional de Contaminación Ambiental”. Además, tiene tres capítulos de libro publicados. El Dr. Méndez Astudillo ha participado en congresos y coloquios nacionales e internacionales. Es tutor en el posgrado y docente en el área de estadística y ciencia de datos. Sus principales líneas de investigación son la aplicación de la ciencia de datos en la investigación económica y la exposición a climas extremos y sus consecuencias en la sociedad y economía observado con imágenes satelitales y otras bases de datos. Sus intereses de investigación son: ciencia de datos, aprendizaje automatizado (Machine Learning), urbanismo digital, cambio climático, inteligencia artificial y economía digital.

# Semblanza

**Gustavo Carreón Vázquez**
**Correo: gcarreon@unam.mx**

Es Doctor en Ciencias por parte del Posgrado en Ciencia e Ingeniería de la Computación de la UNAM (2018), Maestro en Dinámica no Lineal y Sistemas Complejos (UACM, 2007) y Licenciado en Ciencias de la Computación (Facultad de Ciencias-UNAM,2004). Pertenece al SNII (Nivel 1). Miembro del personal académico del Instituto de Investigaciones Económicas y Profesor de la Facultad de Ciencias. Sus líneas de investigación son: Sistemas Complejos, Sistemas de Transporte Público, Modelación Basada en Agentes y "Economía y complejidad". Coordinó el proyecto: #Metrevolución en el Metro de la Ciudad de México para regular el ascenso y descenso de pasajeros en la plataforma. Ha publicado en revistas nacionales e internacionales. Es coordinador de los cursos "Sistemas complejos aplicados a la economía y fenómenos sociales" del IIEc-UNAM, y del Grupo Interdisciplinario de Complejidad y Economía del IIEc-UNAM (ver http://www.complejidad.iiec.unam.mx/).