

# Inteligencia de Datos

Luis Norberto Zúñiga Morales

8 de enero de 2024

# Contenido

- 1 Introducción al curso
- 2 Actividades del Curso
- 3 Temario
- 4 Bibliografía
- 5 Instrumentos de Evaluación
- 6 Acerca de

# Introducción al curso

- La materia de Inteligencia de Datos contempla la segunda parte de un curso de Machine Learning.
- También cubre consideraciones que se presentan en ambientes prácticos que muchas veces suelen ignorarse.
- La materia considera un aspecto de aplicación muy importante. Por lo tanto, programaremos mucho.



# Actividades del Curso

- 1 Presentación de los temas contemplados en el curso por parte del profesor.

# Actividades del Curso

- 1 Presentación de los temas contemplados en el curso por parte del profesor.
- 2 Prácticas de laboratorio de cómputo mediante [Google Colab](#) para implementar los modelos y expandir los conceptos revisados en clase.

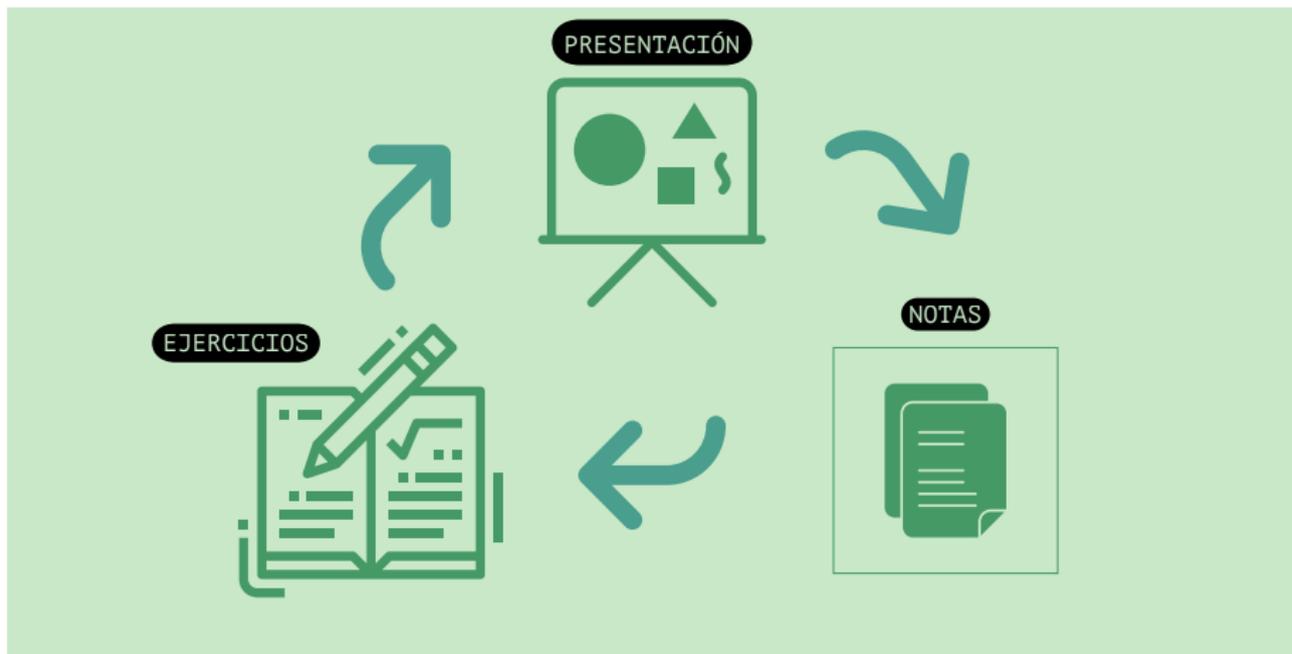
# Actividades del Curso

- 1 Presentación de los temas contemplados en el curso por parte del profesor.
- 2 Prácticas de laboratorio de cómputo mediante [Google Colab](#) para implementar los modelos y expandir los conceptos revisados en clase.
- 3 Lecturas de artículos científicos relativos a aplicaciones, paradigmas y filosofía de la Ciencia de Datos.

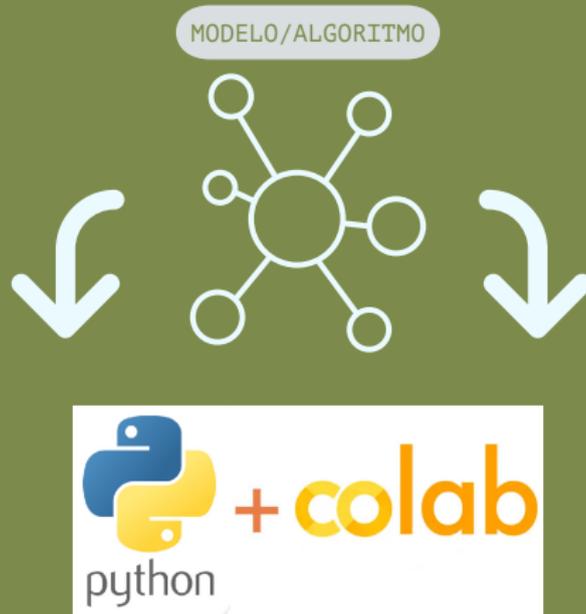
# Actividades del Curso

- 1 Presentación de los temas contemplados en el curso por parte del profesor.
- 2 Prácticas de laboratorio de cómputo mediante [Google Colab](#) para implementar los modelos y expandir los conceptos revisados en clase.
- 3 Lecturas de artículos científicos relativos a aplicaciones, paradigmas y filosofía de la Ciencia de Datos.
- 4 Creación de una página de Wikipedia en español sobre un tema relativo al Machine Learning.

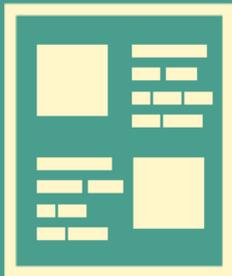
# Presentación de temas



# Prácticas de laboratorio



ARTÍCULOS



CASOS DE ESTUDIO



FILOSOFÍA



# Proyecto Wikipedia

- Crear una página de Wikipedia en español.
- El tema se deja a su elección, pero debe relacionarse con el Machine Learning.
- Revisen [guías y formatos](#) para escribir el artículo. Consideren las referencias, formatos de escritura y posibles traducciones.
- Se revisa en tres entregas. Consideren que el proceso incluye un periodo de aceptación: a planear con tiempo.



# Sobre las Tareas

- Las tareas y prácticas de laboratorio se deben entregar en un archivo PDF en la actividad correspondiente en Brightspace. Respeten las fechas de entrega.
- Dichos archivos se pueden crear en Word,  $\text{\LaTeX}$ , o similares.
- No olviden agregar sus referencias en las tareas.
- **Cualquier intento de copiar-traducir-pegar o copiar-pegar como entrega de una tarea o práctica tendrá como resultado la anulación de la misma.**
- Si la tarea lo amerita, incluyan fórmulas matemáticas. Verifiquen que se muestren correctamente en el documento final.
- Es posible volver a enviar una tarea si quieren atender los detalles para modificar su calificación.

# Temario del Curso

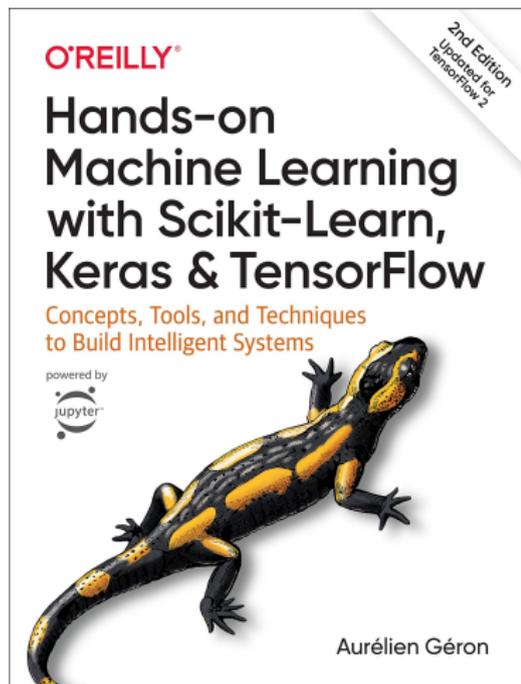
- 1 Repaso de Machine Learning
- 2 Máquinas de Vectores de Soporte
- 3 Árboles de Decisión
- 4 Ensemble Learning
- 5 Ciclo de vida de un proyecto de ciencia de datos
- 6 Proyecto de Aplicación PLN, Consideraciones Prácticas y Patrones de Diseño en ML.

## Bibliografía Sugerida

- [1] Yaser S. Abu-Mostafa, Malik Magdon-Ismail, and Hsuan-Tien Lin. *Learning from Data: A Short Course*. AMLBook, 2012.
- [2] Ovidiu Calin. *Deep Learning Architectures A Mathematical Approach*. Springer, 2020.
- [3] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.
- [4] Aurélien Géron. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc., 2nd edition, 2019.
- [5] Valliappa Lakshmanan, Sara Robinson, and Michael Munn. *Machine Learning Design Patterns: Solutions to Common Challenges in Data Preparation, Model Building, and MLOps*. O'Reilly Media, Incorporated, 2020.
- [6] Andreas C. Müller and Sarah Guido. *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media, Incorporated, 2016.

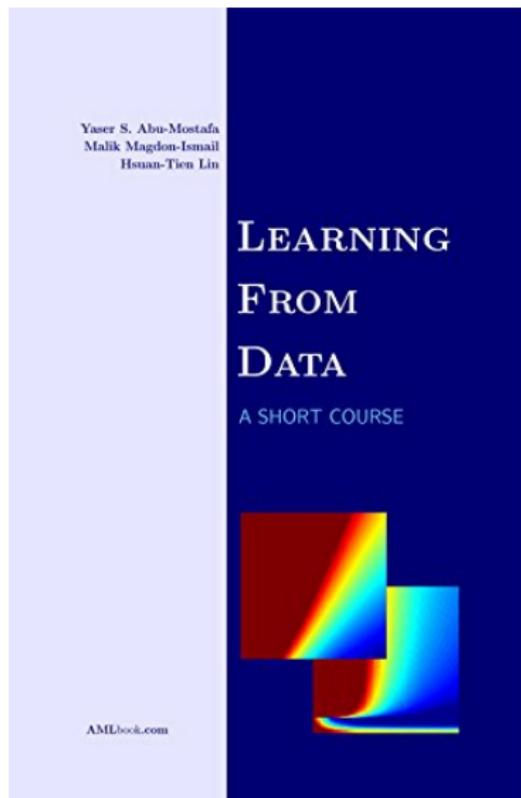
# Bibliografía Sugerida

- Este libro cubre el aspecto de implementación de modelos de Machine Learning y Deep Learning.
- En contraste, cubre lo mínimo de teoría. Lo deja a un lado para enfocarse en lo práctico.
- Es un excelente libro para aprender/practicar programación en Python.



# Bibliografía Sugerida

- Este libro contempla la historia que más me gusta para enseñar Machine Learning.
- Existen diversas formas de enseñar esta materia, pero a nivel licenciatura cubre (en mi criterio) lo necesario.
- Sin embargo, ya que es una introducción, no cubre métodos/teoría modernos.



# Bibliografía Sugerida

- La primera parte del curso se recopila en unas notas para los alumnos.
- Cubren los temas de una manera directa. La bibliografía en ellas es vital para que vean de primera mano el origen de algunos métodos.

## Máquinas de Vectores de Soporte

Luis Norberto Zúñiga Morales

15 de agosto de 2022

### Índice

1. Motivación	2
2. Máquina de Vectores de Soporte Lineal	3
2.1. Teoría	3
2.2. Implementación Práctica	7
3. Máquina de Vectores de Soporte y el Margen Suave	8
3.1. Teoría	8
3.2. Implementación Práctica	9
4. Máquinas de Vectores de Soporte para Casos No Lineales	10
4.1. Implementación Práctica	12
5. Máquinas de Vectores de Soporte para el Caso Multiclase	13
6. Ejercicios	14

Esta obra está bajo una licencia [Creative Commons "Atribución-NoComercial-CompartirIgual 4.0 Internacional"](https://creativecommons.org/licenses/by-nc-sa/4.0/).

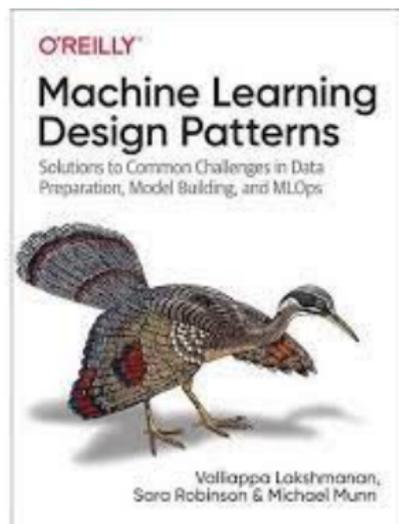


### Resumen

La Máquina de Vectores de Soporte (MVS) es un poderoso modelo de aprendizaje automático que es utilizado en aplicaciones que cubren un amplio espectro de áreas de estudio, tales como medicina, lingüística computacional, cómputo financiero, psicología, entre muchos otros. La MVS [1] es un algoritmo de clasificación binaria (i.e., separa objetos de dos clases distintas) que puede adaptarse a un problema de clasificación multiclase (más de dos clases) y que, gracias al truco del kernel, es posible salir de la idea inicial de separar clases por medio de funciones lineales, y utilizarse funciones no lineales para realizar separaciones con formas más complejas. Esto vuelve a las MVS en un modelo de clasificación muy flexible.

# Bibliografía Sugerida

- Otro libro interesante desde la perspectiva de entrenamiento de modelos.
- Cubre una gran parte de las buenas prácticas antes, durante y después del entrenamiento de modelos de ML.



# Instrumentos de Evaluación

<b>Instrumento</b>	<b>Porcentaje</b>
Tareas y Prácticas de Cómputo	20 %
Proyecto Wikipedia	20 %
Primera Evaluación	20 %
Segunda Evaluación	20 %
Proyecto Final	20 %
<b>Total</b>	<b>100 %</b>

Para el proyecto Wikipedia, se divide en tres partes:

- 1 Entrega del primer borrador del texto para la página (5 %).
- 2 Entrega del segundo borrador del texto para la página (5 %).
- 3 Publicación final en Wikipedia (10 %).

Redondeo de calificaciones:

- 6.0 - 6.49 → 6
- 6.5 - 7.49 → 7
- 7.5 - 8.49 → 8
- 8.5 - 9.49 → 9
- 9.5 - 10.0 → 10

# Fechas Especiales

- Proyecto Wikipedia (primer borrador): hasta el 20 de marzo de 2024.
- Proyecto Wikipedia (segundo borrador): hasta el 24 de abril de 2024.
- Proyecto Wikipedia (entrega final): hasta el 6 de mayo de 2024.
- Proyecto final: hasta el 6 de mayo de 2024.
- Tareas y prácticas: fecha límite 6 de mayo de 2024.
- Primer examen: 4 de marzo de 2024.
- Segundo examen: 10 de abril de 2024.

# Consideraciones especiales

- Toda la comunicación se realizará por medio de **Brightspace** o **correo institucional**.
  - Mi correo: p40887@correo.uia.mx
- Todos los archivos necesarios (notas, programas, lecturas, etc.) para la clase se subirán en la plataforma de Brightspace.
- Cualquier eventualidad se notificará por medio de Brightspace.
- No acostumbro faltar ni llegar tarde, pero si sucede, algo malo me sucedió. En dado caso, necesitaremos organizarnos para la reposición de clases.
- Firmar la GEP.

# Consideraciones adicionales

- **No tomo asistencia.** Si faltan, es su responsabilidad estar al tanto de cualquier tarea, proyecto, lectura y/o actualizaciones en las fechas de entrega o evaluaciones.
- Las prácticas de laboratorio se hacen en equipos de dos personas. Armar los equipos y notificarme para implementarlos en BS para la entrega de prácticas.
- Tengan cuidado al formar su equipo. **Una vez hecho, no hay cambios.**
- Las fechas de las evaluaciones dependen del avance del curso. Se notificará cualquier cambio con suficiente tiempo.
- **No fumar, no vaping.** No pueden comer ni beber en los laboratorios.
- Si necesitan salir del salón, háganlo discretamente.

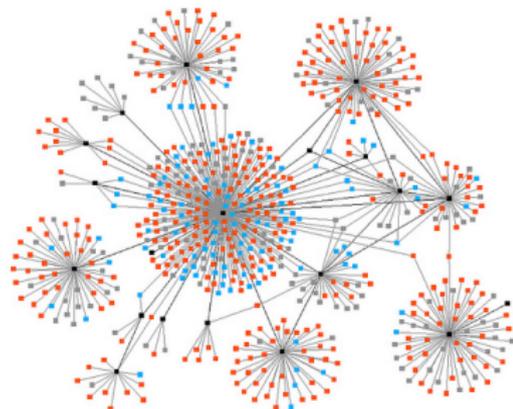
# About me...

- Ing. Matemático, Escuela Superior de Física y Matemáticas, Instituto Politécnico Nacional.
- Mtro. en Ciencias de la Computación, Centro de Investigación en Computación, Instituto Politécnico Nacional.
- Candidato a Doctor en Ciencias de la Ingeniería, Universidad Iberoamericana. Tema: *Multimodal Sentiment Analysis on Social Networks.*



# About me...

- Áreas de interés: Machine Learning, Deep Learning, Procesamiento de Lenguaje Natural, Network Science, Stock Trading, Análisis de Redes Sociales.
- Me interesa el área de las matemáticas educativas y la educación abierta masiva.



¿Y ustedes?