

Inteligencia de Datos

Luis Norberto Zúñiga Morales

4 de enero de 2023

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 conceptos revisados en clase.

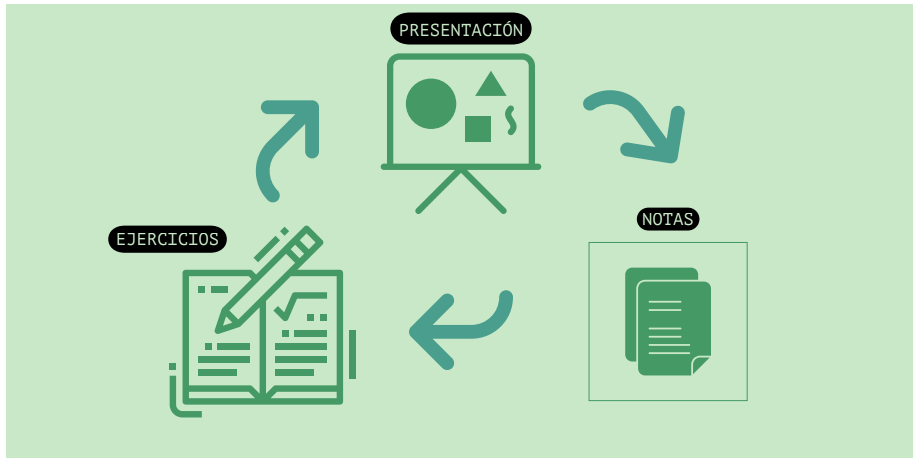
Actividades del Curso

- 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 conceptos revisados en clase.
- 3 Lecturas de artículos científicos relativos a aplicaciones, paradigmas y filosofía de la Ciencia de Datos.

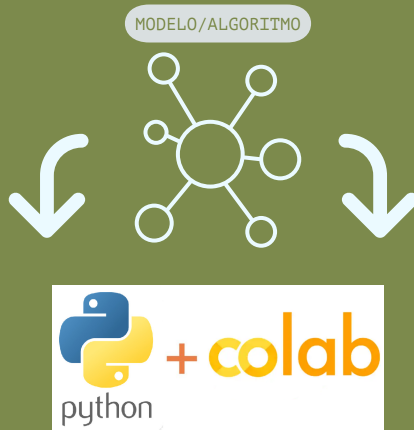
Actividades del Curso

- 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 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 a la Ciencia de Datos.

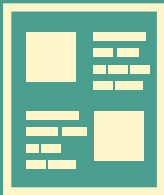
Presentación de temas



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 la Ciencia de Datos.
- Estudiar [guías y formatos](#) para escribir el artículo. Consideren las referencias, formatos de escritura y posibles traducciones.
- Se revisa a finales del semestre. Consideren que el proceso incluye un periodo de aceptación: a planear con tiempo.



Sobre las Tareas

- Las tareas y prácticas de laboratorio se deben recopilar en un archivo que se entrega al final del semestre como evidencias del curso.
- Dicho archivo debe ser en formato \LaTeX en el sitio de [Overleaf](#).
- Cada tarea debe ser una sección individual.
- En sí, este archivo es un pequeño libro. Es su oportunidad para mostrar su creatividad.
- No olvidar bibliografía y referencias.
- Fuera de estos requerimientos, pueden personalizar el archivo como ustedes lo deseen.

Temario del Curso

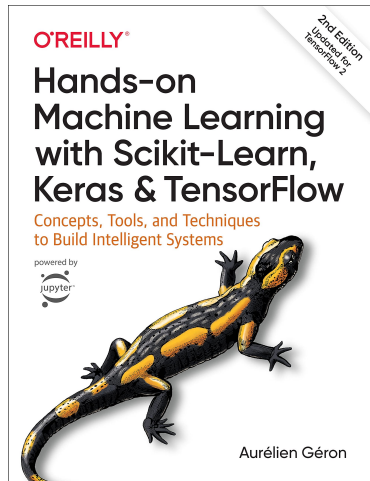
- 1 Repaso de Machine Learning
- 2 Máquinas de Vectores de Soporte
- 3 Árboles de Decisión
- 4 Ensemble Learning
- 5 Ciclo de vida de un proyecto de ciencia de datos
- 6 Proyecto de Aplicación PLN, Consideraciones Prácticas
- 7 Detección de Anomalías*
- 8 Deep Learning Práctico*

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] Marcos López de Prado. *Advances in Financial Machine Learning*. Wiley, 2018.
- [4] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.
- [5] Aurélien Géron. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc., 2nd edition, 2019.
- [6] Andreas C. Müller and Sarah Guido. *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media, Incorporated, 2016.

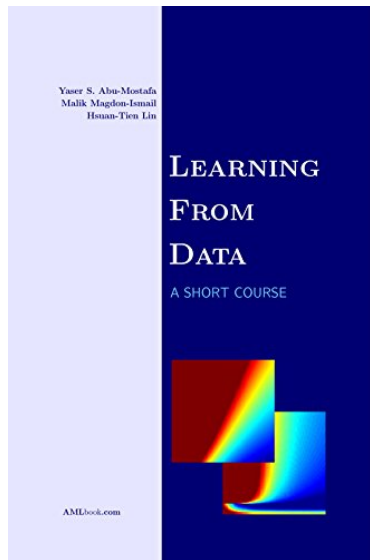
Bibliografía Sugerida

- 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.



Bibliografía Sugerida

- La mayor parte de los temas del curso está recopilado en unas notas que realicé para los alumnos.
- Cubren los temas de una manera amplia y 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"](#).



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.

Instrumentos de Evaluación

No se realizará una evaluación tradicional en el curso. En su lugar, se requiere un esfuerzo continuo de su parte. Sin embargo, sí aplicaré algunas evaluaciones sorpresa adicionales. ¡No se confíen!

Instrumento	Porcentaje
Tareas, Prácticas y Notas	25 %
Proyecto Wikipedia	25 %
Proyecto Final	25 %
Examen	25 %
Evaluación Adicional	10 %
Total	110 %

Redondeo de calificaciones:

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

Fechas Especiales

- Proyecto Wikipedia: Semana del 1 al 5 de mayo 2023.
- Recopilación de Notas: Hasta la semana del 1 al 5 de mayo 2023.
- Lecturas opcionales: A su discreción.
- Lecturas del curso: Siguiendo clase.
- Tareas/notas: Hasta una semana después, fecha límite dada en el sistema de Brighspace.

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.
- Posible plan de contingencia: nos vamos a Zoom o Teams.
- No acostumbro faltar, pero si sucede, piensen lo peor.
Necesitamos ponernos de acuerdo para la reposición de clases.

Consideraciones especiales

- No tomo asistencia. Es su responsabilidad estar al tanto de tareas, proyectos, lecturas y fechas de las evaluaciones.
- Las prácticas de laboratorio de hacen en equipos de dos personas. Hacer los equipos y notificarme para implementarlos en BS.
- Las fechas de las evaluaciones dependen del avance del curso. Se notificarán con suficiente tiempo de antelación.
- No fumar, no vaping.
- Si necesitan salir momentáneamente del salón, háganlo discretamente.

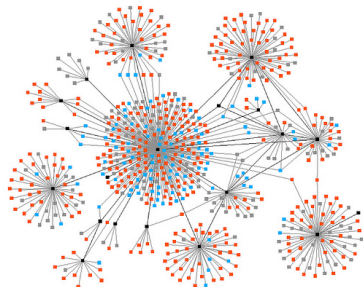
About me...

- 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?