

Aprendizaje Máquina Teórico

(Theoretical Machine Learning)

Descripción

En este curso se abordan temas clásicos y contemporáneos del aprendizaje máquina teórico, cuyo objetivo principal es establecer garantías estadísticas, de optimización y de aproximación para diversas metodologías. Además de los temas propios del aprendizaje máquina (ML, por sus siglas en inglés), en este curso se exponen de manera autocontenida las herramientas matemáticas necesarias, e.g., grandes desviaciones, análisis convexo.

Prerrequisitos

Se recomienda haber cursado con anterioridad al menos un curso de álgebra lineal, cálculo, probabilidad y estadística. El curso se enfoca en los aspectos teóricos del aprendizaje máquina, por lo que no es necesario contar con conocimientos de programación.

Objetivos

Al finalizar el curso los participantes deberán:

- aprender las herramientas matemáticas básicas utilizadas en aprendizaje máquina teórico;
- reconocer el arquetipo de garantías estadísticas, de aproximación y de optimización en ML;
- conocer los retos técnicos presentes en el análisis de diversas metodologías de ML.

Contenido

Módulo 1: Garantías Estadísticas (8 Semanas)

1. Aprendizaje Supervisado y Clasificación
2. Elementos de Grandes Desviaciones
3. Aprendizaje Probablemente Correcto (PAC)
4. Convergencia Uniforme
5. Teorema de No Free Lunch
6. Complejidad de Rademacher
7. Cotas Uniformes para Promedios Centrados
8. Complejidad de Rademacher y Convergencia Uniforme
9. Dimensión Vapnik–Chervonenkis (VC) y el Lema de Sauer
10. Teorema Fundamental del Aprendizaje Estadístico (TFAE)
11. Álgebra de la Función de Crecimiento y Aplicaciones
12. TFAE y Deep Learning
13. Generalización y Enfoques Contemporáneos

Módulo 2: Garantías de Optimización (4 Semanas)

1. Elementos de Análisis Convexo
2. Descenso por Gradiente y Garantías de Optimización
3. Descenso por Gradiente Estocástico y Redes Neuronales
4. Regularización, Estabilidad y Generalización
5. Aprendizaje en Línea (Online Learning)
6. Aprendizaje en Línea y Aprendizaje PAC

Módulo 3: Garantías de Aproximación de Funciones con Redes Neuronales (4 Semanas)

1. Redes Neuronales como Aproximadores Universales
2. Elementos de Análisis de Fourier
3. Teorema de Barron y Aplicaciones

Bibliografía Básica

1. Sanjeev Arora. Theoretical Foundations of Deep Learning. Lecture Notes for CS597G (Princeton), 2018.
2. Andrew Barron. Universal approximation bounds for superpositions of a sigmoidal function. *IEEE Transactions on Information Theory*, 1993.
3. Thomas Cover & Joy Thomas. *Elements of Information Theory*. John Wiley & Sons, 2012.
4. Mario Diaz. *Lecture Notes on Theoretical Machine Learning, Information Theory, and Privacy*. En preparación, 2022.
5. Elad Hazan. *Introduction to Online Convex Optimization*. Foundations and Trends in Optimization, 2016.
6. Yurii Nesterov. *Lectures on Convex Optimization*. Springer, 2018.
7. Shai Shalev-Shwartz & Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014.

Acerca del Curso

- Evaluación: tareas (70%) y proyecto final (30%)
- Horario: viernes de 12:00pm a 15:00pm
- Instructor: Mario Diaz, IIMAS - Oficina 115, mario.diaz@sigma.iimas.unam.mx