SEMINARIO

Jean-Michel Loubes

Research Director, INRIA & Institut de Mathématiques de Toulouse

Minority & Majority in Machine Learning

Abstract: Despite growing empirical evidence of bias amplification in machine learning, its theoretical foundations remain poorly understood. We develop a formal framework for majority-minority learning tasks, showing how standard training can favor majority groups and produce stereotypical predictors that neglect minority-specific features. Assuming population and variance imbalance, our analysis reveals three key findings: (i) the close proximity between "full-data" and stereotypical predictors, (ii) the dominance of a region where training the entire model tends to merely learn the majority traits, and (iii) a lower bound on the additional training required.

Seminario del Departamento de Estadística e Inv. Operativa
Miércoles 15 de Abril de 2026 (10:00)
Organiza: G.I.R. Probabilidad y Estadística Matemática