

Statistics Seminar  
June 14, 2024, 11 a.m, Ensai

## Machine learning with functional data: near-perfect classification

Alberto Suárez  
*Universidad Autónoma de Madrid (UAM)*

**Abstract:** Functional data consists of observations that depend on a continuous parameter, such as curves, surfaces, and volumes. These types of data appear in numerous areas of application, such as medicine (for example, diagnosis of heart conditions from electrocardiograms), the environment (e.g., characterization of meteorological patterns), or economics (e.g., prediction of economic indicators). Due to their continuous nature, the statistical analysis of functional data presents specific challenges. In the general case, they have infinite dimensions. Because of this characteristic, some quantities, such as the probability density or the likelihood function are ill-defined. Therefore, standard methods of statistical inference and machine learning, most of which are based on multivariate statistics, must be adapted, extended, or reformulated to deal with these types of data. In this work, we present different methods to derive optimal classification rules for binary classification problems in which the trajectories are sampled from two different Gaussian processes (GPs), depending on the class. These methods rely on discretization, spectral analysis of the covariance functions of the GPs, or the theory of reproducing kernel Hilbert spaces. Especial attention is paid to classification problems in which the GPs are orthogonal. In such cases, the classification rules involve singular terms, and zero classification error is obtained asymptotically.

**Vitae:** Alberto Suárez received the degree of Licenciado (BSc) in Chemistry, specialization in Quantum Chemistry, from the Universidad Autónoma de Madrid, Spain, in 1988, and the PhD in Physical Chemistry from the Massachusetts Institute of Technology (MIT), Cambridge, MA, in 1993. After holding postdoctoral positions at Stanford University (USA), at Université Libre de Bruxelles (Belgium), as a research fellow financed by the European Commission within the Marie Curie “Training and Mobility of Researchers” program, and at the Katholieke Universiteit Leuven (Belgium), he is currently Professor of Computer Science and Artificial Intelligence in the Computer Science Department at the Universidad Autónoma de Madrid (UAM), where he co-directs the Machine Learning Group- Grupo de Aprendizaje Automático (MLG-GAA) [[www.eps.uam.es/gaa](http://www.eps.uam.es/gaa)]. He has also held appointments as “Senior Visiting Scientist” at the International Computer Science Institute (Berkeley, CA) and at MIT (Cambridge, MA). He has worked on relaxation theory in condensed media, stochastic and thermodynamic theories of nonequilibrium systems, lattice-gas automata, and automatic induction from data. His current research interests include artificial intelligence, in particular, machine learning, computational statistics, functional data analysis, and causality. He is a member of IEEE, of the European Laboratory for Learning and Intelligent Systems (ELLIS) and a founding member of ELLIS Unit Madrid.

*FunMathStat network*