Theory and applications of Bayesian machine learning in paleoclimatology, nuclear physics, soil science, hospital management, and agricultural economics

Frederi Viens[1]

[1] Michigan State University

The framework of Bayesian statistics traces back more than 250 years, but it is only in the most recent decades that computational advances, particularly in hardware, have made it possible to unleash its full power on data analysis. In this presentation, we will attempt to show, through several examples of diverse applications, how this framework gives rise, in a rather intuitive fashion, to computational supervised machine learning schemes. These allow specific participatory modes of collaboration via prior elicitation, which are particularly useful to situations which are challenging to frequentist frameworks, such as observational studies, environmental space-time data analysis, and studies with multiple competing models. Some of the examples we will cover include: the use of validation metrics for model selection in paleoclimatology, Bayesian model averaging for competing models with non-identical application domains in nuclear physics, attribution of factors in hydrology and agro-ecology, predictive analytics and cropping recommendations based on farmer surveys, and in the context of workplace satisfaction for nurses' work environment, and hierarchical modeling with time lag distribution in global agricultural total factor productivity. Time permitting, we will also discuss theoretical questions using analysis on Wiener space which may give rise to highly simplified approximations of these machine-learning algorithms. This is joint work with collaborators and Ph.D. students at Michigan State University, Purdue University, and the US Department of Agriculture, and is partially supported by these three entities and by the National Science Foundation.