

# **First-hitting-time Threshold Regression in the Data Science Era**

Mei-Ling Ting Lee

*University of Maryland, College Park, U.S.A.*

## **Abstract**

Degradation of an engineering system or disease progression in a patient can be described mathematically as a stochastic process. The system or the patient experiences a failure event when the wear and tear on the system or the patient's disease progression first reaches a critical threshold level. This happening defines a failure event and a first hitting time (FHT). First hitting time threshold regression (TR) models are based on an underlying stochastic process. They do not require the proportional hazards assumption, which is difficult to verify in a high-dimensional context, and represent a realistic alternative to the Cox model for capturing granular structure in a prediction model. Machine learning methods such as boosting and neural networks have been applied to FHT TR models for prediction and causal inference. They have also been extended to modeling cure rates and clinical trials. In addition to parametric models, FHT TR models have recently been extended to semiparametric applications.