# A Novel Exponential Loss Function for Multi-class Classification for Complex Imbalanced Data

Yiming Li[1], Wei-Wen Hsu[2]

[1]*Amazon, Seattle, U.S.A.*
[2]*Division of Biostatistics and Bioinformatics, Department of Environmental and Public Health Sciences, University of Cincinnati, U.S.A.*

**Abstract**

Multi-class classification has been gaining popularity in many areas such as disease screening, image classification, and object identification. Traditional classifiers typically assume a balanced class distribution and tend to misclassify instances from minority classes as ones from majority classes, thus resulting in low predictive accuracy for minority classes. It could be even worse in a longitudinal and high dimensional setting. It is expected that the overall classification performance for multi-class imbalanced data will be compromised due to the aforementioned complications. In this paper, we propose a novel two-stage classification framework for multi-class imbalanced data in longitudinal and high-dimensional settings. Specifically, the techniques of natural cubic spline are first employed to efficiently extract features in a longitudinal data structure. Coupled with the group LASSO penalty, a weights-adjusted margin-based exponential loss is then proposed for the multi-class classification in high-dimensional settings. The empirical performance of the proposed method in finite sample size is evaluated by extensive simulations. The method is further illustrated with real applications of Alzheimer's disease early detection, cancer type discrimination, and gait-based person identification.

Keywords: Functional data; Loss function; Cancer type screening.