Title: Classification with multiscale hidden variables using Expectation Reflection

PI(s) Grant: N/A

Institution(s): Laboratory of Biological Modeling, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD.

Grant Number: N/A

Title of Grant: N/A

Abstract Authors:

Danh-Tai Hoang

Sujay Singh

Vipul Periwal

Abstract Text:

Classifying instances into one of two or more classes is a fundamental problem, not only in quantitative biology but more generally in data science. Mechanistic algorithms identify the influence (coupling weights) from observed variables to targets based on a training set, then make a prediction for a new test set. We recently developed a new data-driven method, Expectation Reflection (ER), that outperforms current state-of-the-art methods in inferring the network interactions between binary variables, especially in small sample sizes [1]. By introducing effective hidden variables, often necessary in multi-scale systems, ER still works in partially observed systems [2]. In this work, we first extend ER to infer interactions from observed features to targets in which the features are not restricted to binary but can be continuous or categorical variables. We then apply this extension to classify biomedical data. We demonstrated the performance of ER in 20 different biomedical data sets and showed that ER works better than logistic regression (LR) in 10 data sets, as well as LR in 6 data sets, and worse than LR in 4 data sets. Our results suggest that ER can be an effective method for multiscale problems where hidden variables may correspond to effective interactions induced from variables observed at different scales.

References:

[1] D. T. Hoang, J. Song, V. Periwal, and J. Jo (2019), Network inference in stochastic systems from neurons to currencies: Improved performance at small sample size, Physical Review E, 99, 023311.
[2] D. T. Hoang, J. Jo, and V. Periwal (2019), Data-driven inference of hidden nodes in networks, Physical Review E, 99, 042114.