

3rd International Conference on

Systems and Synthetic Biology

July 20-21, 2017 Munich, Germany

A robust PCA algorithm for metagenomic biomarker detection

Erchin Serpedin, Mustafa Alshawaqfeh, Ahmad Bashaire and Jan Suchodolski
Texas A&M University, USA

We propose a novel consistency-classification framework that enables the assessment of consistency and classification performance of a biomarker discovery algorithm. The proposed evaluation protocol is based on random resampling those models for the variation in the experiment size. The metagenomic data matrix is modeled as a superposition of two matrices. The first matrix is a low-rank matrix that depicts the abundance levels of the irrelevant bacteria. The second matrix is a sparse matrix that describes the abundance levels of the bacteria that are differentially abundant between different phenotypes. We propose a novel Robust Principal Component Analysis (RPCA) based biomarker discovery algorithm to recover the sparse matrix. RPCA is a multivariate feature selection approach that processes the features collectively rather than individually. Comprehensive comparisons of RPCA with the state-of-the-art algorithms on two realistic datasets show that RPCA consistently outperforms the existing state-of-the-art algorithms in terms of classification accuracy and reproducibility performance. Thus, the proposed RPCA-based biomarker detection algorithm provides a high reproducibility performance irrespective of the complexity of the dataset and the number of selected biomarkers. RPCA selects also biomarkers with quite high discriminative accuracy. Therefore, RPCA appears to represent a very consistent and accurate methodology for selecting taxonomical biomarkers in microbial populations.

Biography

Erchin Serpedin is currently a Professor at Texas A&M University in College Station, TX. He is the author of more than 140 journal papers, 250 conference papers, and 4 books. His research interests lie in the areas of computational biology, systems biology, signal processing and machine learning.

eserpedin@tamu.edu

Notes: