


Research Article

Improving the performance of the iterative signature algorithm for the identification of relevant patterns

A. Freitas , V. Afreixo, M. Pinheiro, J. L. Oliveira, G. Moura, M. Santos

First published: 13 January 2011

<https://doi.org/10.1002/sam.10104>

Cited by: 4



About



Access

JLS



ire

Abstract

The iterative signature algorithm (ISA) has become very attractive to detect co - regulated genes from microarray data matrices and can be a useful tool for the identification of similar patterns in many other kinds of numerical data matrices. Nevertheless, its algorithmic strategy exhibits some limitations since it is based on statistical behavior of the average and considers averages weighted by scores not necessarily positive. Hence, we propose to take the median instead of the average and to use absolute scores in ISA's structure. Furthermore, a generalized function is also introduced in the algorithm in order to improve its algorithmic strategy for detecting high value or low value biclusters. The effects of these simple modifications on the performance of the biclustering algorithm are evaluated through an experimental comparative study involving synthetic data sets and real data from the organism *Saccharomyces cerevisiae*. The experimental results show that the proposed variations of ISA outperform the original version in many situations. Absolute scores in ISA are shown to be essential for the correct interpretation of the biclusters found by the algorithm. The median instead of the average turns the biclustering algorithm more resilient to outliers in the data sets. © 2011 Wiley Periodicals, Inc. Statistical Analysis and Data Mining 4: 71–83 2011

Citing Literature



About Wiley Online Library



Help & Support



Opportunities



Connect with Wiley



Copyright © 1999-2018 John Wiley & Sons, Inc. All rights reserved

WILEY