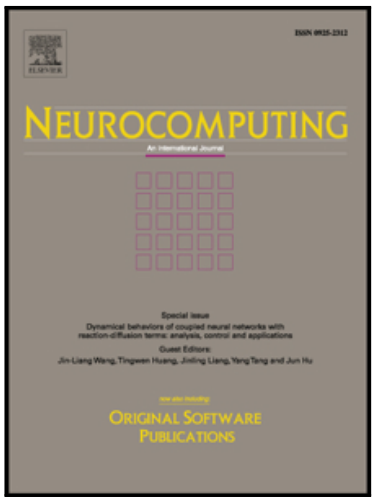
Accepted Manuscript




# A Hybrid Feature Selection Algorithm for Gene Expression Data Classification

Huijuan Lu , Junying Chen , Ke Yan , Qun Jin , Yu Xue , Zhigang Gao

PII: S0925-2312(17)30415-0
DOI: 10.1016/j.neucom.2016.07.080
Reference: NEUCOM 18165

To appear in: *Neurocomputing*

Received date: 14 May 2016
Revised date: 16 June 2016
Accepted date: 16 July 2016

Please cite this article as: Huijuan Lu , Junying Chen , Ke Yan , Qun Jin , Yu Xue , Zhigang Gao , A Hybrid Feature Selection Algorithm for Gene Expression Data Classification, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2016.07.080

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

# A Hybrid Feature Selection Algorithm for Gene Expression Data Classification

Huijuan Lu[a], Junying Chen[a], Ke Yan[a,*], Qun Jin[a,b], Yu Xue[c], Zhigang Gao[d]

[a] College of Information Engineering, China Jiliang University, 258 Xueyuan Street, Hangzhou, China, 310018.

[b] Department of Human Informatics and Cognitive Sciences, Waseda University, 2-579-15 Mikajima, Tokorozawa-shi, Saitama 359-1192, Japan.

[c] Nanjing University of Information Science & Technology, Nanjing, China, 210044.

[d] College of Computer Science，Hangzhou Dianzi University，Hangzhou, China, 310018.

## Abstract

In the DNA microarray research field, the increasing sample size and feature dimension of the gene expression data prompt the development of an efficient and robust feature selection algorithm for gene expression data classification. In this study, we propose a hybrid feature selection algorithm that combines the mutual information maximization (MIM) and the adaptive genetic algorithm (AGA). Experimental results show that the proposing MIMAGA-Selection method significantly reduces the dimension of gene expression data and removes the redundancies for classification. The reduced gene expression dataset provides highest classification accuracy compared to conventional feature selection algorithms. We also apply four different classifiers to the reduced dataset to demonstrate the robustness of the proposed MIMAGA-Selection algorithm.

* Corresponding author. Tel.: +86-153-9700-8303; fax: +86-571-8691-4580. E-mail address: yanke@cjlu.edu.cn

Download English Version:

https://daneshyari.com/en/article/4947266

Download Persian Version:

https://daneshyari.com/article/4947266

Daneshyari.com