

{tag}

{/tag}

IJCA Proceedings on National Conference on
Communication Technologies & its impact on Next Generation Computing 2012

© 2012 by IJCA Journal

CTNGC - Number 3

Year of Publication: 2012

Authors:

G. Baskar

P. Ponmuthuramalingam

{bibtex}ctngc1032.bib{/bibtex}

Abstract

Microarray is a powerful technology for biological exploration which enables to simultaneously measure the level of activity of thousands genes in various cancer study . clustering is important data mining technique to extract useful information from various high dimensional datasets. A wide range of clustering algorithm is available and still in an open area of research k-Means algorithm is one of the basic and most simple partitioning clustering technique is given by Mac Queen in 1967. In this paper a sample weighting and efficient margin based sample weighting algorithm to improve the stability of feature selection. We proposed a weighted k-means to improve the cluster stability and presented an experimental evaluation of the proposed method,

the experiment of microarray dataset show the feature selection algorithm such as SVM-RFE are more stable in gene selection.

Refer

ences

- T. R. Golub, D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R. Downing, M. A. Caligiuri, C. D. Bloomfield, and E. S. Lander, "Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring," Science, vol. 286, pp. 531-537, 1999.
- T. Li, C. Zhang, and M. Ogihara, "A Comparative Study of Feature Selection and Multiclass Classification Methods for Tissue Classification Based on Gene Expression," Bioinformatics, vol. 20, pp. 2429-2437, 2004.
- Y. Saeys, I. Inza, and P. Larranaga, "A Review of Feature Selection Techniques in Bioinformatics," Bioinformatics, vol. 23, no. 19, pp. 2507-2517, 2007.
- H. Liu, J. Li, and L. Wong, "A Comparative Study on Feature Selection and Classification Methods Using Gene Expression Profiles and Proteomic Patterns," Genome Informatics, vol. 13, pp. 51-60, 2002.
- P. A. Mundra and J. C. Rajapakse, "SVM-RFE with MRMR Filter for Gene Selection," IEEE Trans. NanoBioscience, vol. 9, no. 1, pp. 31- 37, Mar. 2010
- I. H. Witten and E. Frank, Data Mining - Practical Machine Learning Tools and Techniques. Morgan Kaufmann Publishers, 2005.
- B. Y. Rubinstein, Simulation and the Monte Carlo Method. John Wiley & Sons, 1981.
- Y. Tang, Y. Q. Zhang, and Z. Huang, "Development Two-Stage SVM-RFE Gene Selection Strategy for Microarray Expression Data Analysis," IEEE/ACM Trans. Computational Biology and Bioinformatics, vol. 4, no. 3, pp. 365-381, July 2007.
- Pawan Lingras, Chad West. Interval set Clustering of Web users with Rough k-Means, submitted to the Journal of Intelligent Information System in 2002.
- Yeung K. Y, Haynor D. R, Ruzzo W. L. Validating clustering for gene expression data. Bioinformatics. 2001.

Index Terms

Computer Science

Bio-sciences

Keywords

Feature Selection Classification Clustering Gene Expression Microarray

