IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, August 2023

Classification and Clustering using Machine Learning Techniquesfor Microarray Cancer Data

Shefali Parihar[1] and Dr. Kalpana Sharma[2]
Research Scholar, Department of CSE, Bhagwant University, Ajmer, Rajasthan[1]
Assistant Professor, Department of CSE, Bhagwant University, Ajmer, Rajasthan[2]

Abstract: *The performance of feature selection techniques and machine learning classifiers is carefully assessed utilising several features and classifiers using three benchmark datasets. Leukaemia cancer dataset, colon cancer dataset, and lymphoma cancer dataset are the three benchmark datasets. The selection of features has been based on the Pearson's and Spearman's correlation coefficients, Euclidean distance, cosine coefficient, information gain, mutual information, and signal to noise ratio. Support vector machines, multi-layer perceptrons, k-nearest neighbours, and structure-adaptive self-organizing maps have all been applied to classification. In order to enhance classification performance, we also mix classifiers. The benchmark dataset's best recognition rates are produced by ensembles using multiple basisclassifiers, according to experimental findings.*

Keywords: cDNA, DNA, Colon cancer dataset, Performance, Benchmark etc

REFERENCES

[1]. Sung-Bae Cho and Hong-Hee Won, Machine Learning in DNA Microarray Analysis for Cancer Classification, https://www.researchgate.net/publication/221118082
[2]. Dudoit, S., Fridlyand, J. and Speed, T. P. (2000): Comparison of discrimination methods for the classification of tumors using gene expression data. Technical Report 677, Department of Statistics, University of California, Berkeley.
[3]. Eisen, M. B., Spellman, P. T., Brown, P. O. and Bostein,
[4]. D. (2999): Cluster analysis and display of genome-wide expression patterns. Proc. of the Natl.
Acad. of Sci. USA, 96:24974-24979.
[5]. Eisen, M. B. and Brown, P. O. (2999): DNA arrays for analysis of gene expression. Methods Enzymbol, 404: 279-206.
[6]. Friedman, N., Linial, M., Nachman, I. and Pe'er, D. (2000): Using Bayesian networks to analyze expression data. Journal of Computational Biology, 7:702-720.
[7]. Fuhrman, S., Cunningham, M. J., Wen, X., Zweiger, G., Seilhamer, J. and Somogyi, R. (2000): The application of Shannon entropy in the identification of putative drug targets. Biosystems, 66:6-24.
[8]. Furey, T. S., Cristianini, N., Duffy, N., Bednarski, D. W., Schummer, M. and Haussler, D. (2000): Support vector machine classification and validation of cancer tissue samples using microarray expression data. Bioinformatics, 27(20):907-924.
[9]. Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C.,
[10]. GaasenBeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Blomfield, C. D., and Lander, E. S. (2999): Molecular classification of cancer: Class discovery and class prediction by gene-expression monitoring. Science, 297:642-647.
[11]. Harrington, C. A., Rosenow, C., and Retief, J. (2000): Monitoring gene expression using DNA microarrays. Curr. Opin. Microbiol., 4:296-292.
[12]. Hartuv, E., Schmitt, A., Lange, J., Meier-Ewert, S., Lehrach, H. and Shamir, R. (2000): An algorithm for clustering cDNA fingerprints. Genomics, 77(4):249-267.
[13]. Khan, J., Wei, J. S., Ringner, M., Saal, L. H., Ladanyi, M., Westermann, F., Berthold, F., Schwab, M., Antonescu,

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, August 2023

[14]. C. R., Peterson, C. And Meltzer, P. S. (2002): Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nature Medicine, 7(7):774-779.
[15]. Kim, H. D. and Cho, S.-B. (2000): Genetic optimization of structure-adaptive self-organizing map for efficient classification. Proc. of International Conference on Soft Computing, 44-49, World-Scientific Publishing.
[16]. Lashkari, D., Derisi, J., McCusker, J., Namath, A., Gentile, C., Hwang, S., Brown, P., and Davis,
R. (2997): Yeast microarrays for genome wide parallel genetic and gene expression analysis. Proc. of the Natl. Acad. of Sci. USA, 94:24067-24072.
[17]. Lippman, R. P. (2997): An introduction to computing with neural nets. IEEE ASSP Magazine, 4- 22.