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

Exploring Clustering Algorithms for Parkinson's Disease Data: A Comparative Analysis

Dip Das[1] and Adepu Rajesh[2]
Assistant Professor, Computer Science and Engineering[1]
Associate Professor, Computer Science and Engineering[2]
Guru Nanak Institute of Technology, Hyderabad, India
hellodip2@gmail.com[1] and adeprajesh@gmail.com[2]

Abstract: *Clustering, an essential analytical approach utilized in data mining, encompasses the act of grouping alike data items into clusters. It is crucial to note that the clustering outcome is significantly impacted by the employed clustering algorithm. This research paper presents a thorough analysis of various clustering algorithms, such as k-means, hierarchical, and DB-scan clustering algorithms, among others, while simultaneously scrutinizing their strengths and limitations. Within each algorithm type, the computation of the distance between data objects and cluster cantres is executed in every iteration, which inevitably poses a challenge to the efficiency of clustering. This paper provides an extensive summary of the fundamental techniques and highlights the associated challenges with clustering algorithms, such as recall, precision, and f-measure, to produce superior outcomes under diverse circumstances. The paper concludes with a discussion of the results obtained from a high-dimensional dataset of Parkinson's disease.*

Keywords: Data mining, Clustering algorithm, k-means, Parkinson's disease

REFERENCES

[1]. "Universityofberkley,"http://www.sims.berkeley.edu/research/projects/how- much-info2003/.
[2]. N. Ye, The Hand Book of Data Mining. Mahwah, New Jersey: Lawrence ErlbaumAssociates, 2003.
[3]. J. Han and M. Kamber, Data Mining: Concepts and Techniques, 2nd ed. Morgan Kaufmann, Elsevier, 2006.
[4]. R. Dubes and A. Jain, Algorithms for Clustering Data. Prentice Hall, 1988.
[5]. R. Duda, P. Hart, and D. Stork, Pattern Classification, 2nd ed. New York: JohnWiley & Sons, 2001.
[6]. S. S. Stevens, "On the theory of the scales of measurement science," Science, vol.103, no. , pp. 677–680, June 1946.
[7]. E. Sungur, "Overview of multivariate statistical data analysis," http://www.mrs.umn.edu/ sungurea/multivariatestatistics/overview.html.
[8]. T. Joachims, "Transductive inference for text classification using support vector machines," in Proc. 16th International Conference on Machine Learning. MorganKaufmann, 1999, pp. 200–209.
[9]. A. Zhang, Advanced analysis of gene expression microarray data: Science, Engineering, and Biology Informatics, Singapore, 596224:World Scientific Publishing Co. Pte .Ltd., 2006, vol. 3.
[10]. S. Bandyopadhyay, U. Maulik, and J.T.L.Wang, Analysis of biological data: as oft comp uting approach: Science, Engineering, and Biology Informatics. Singapore, 596224:World Scientific Publishing Co. Pte. Ltd.,2007, vol. 3.
[11]. I. S. Kohane, A. T. Kho, and A. J. Butte, Microarrays for an Integrative Genomics. Massachusetts, London, England: MIT Press Cambridge, 2003.
[12]. J. T. L. Wang, M. J. Zaki, H. T. T. Toivonen, and D. Shasha, Data Mining in Bioinformatics. London: Springer-Verlag, 2005.
[13]. J. MacQueen. Some methods for classification and analysis of multivariate observations. Proceeding of the 5th Symposium on Mathematical Statistics and Probability. 1: 281–297. 1967.

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]. J. F. Curry, R. J. Thompson, Jr. Patterns of behavioural disturbance in developmentally disabled and psychiatrically referred children:A cluster analytic approach. Journal of Pediatric Psychology. 10:151-167. 1985.
[15]. S. C. Johnson. Hierarchical clustering schemes. Psychometrika. 32: 241– 254. 1967.
[16]. P. Sneath, R. Sokal. Numerical taxonomy. W. H. Freeman, San Francisco. 1973.
[17]. https://archive.ics.uci.edu/ml/datasets.php