Data Clustering Algorithms for Grouping Liver Cirrhosis Medical Records to Support Treatment

ICTU Student 153
Thai Nguyen University of Information and Communication Technology
Thai Nguyen University, Thai Nguyen, Vietnam
student153@ictu.edu.vn

***Abstract*—This thesis focuses on the setup and application of data clustering algorithms, specifically the C-Means algorithm, to the problem of grouping liver disease cases for treatment support. The author describes the process of setting up and implementing these algorithms in a Visual Basic 6.0 application, highlighting the use of Fuzzy C-Means and other clustering techniques. The work covers the basics of data mining and clustering, and its application in the medical field, particularly for liver disease. The thesis concludes with limitations and future research directions, emphasizing the need for further optimization and practical application of the algorithms.**

***Index Terms*—Data mining, Clustering, C-Means, Fuzzy C-Means, Liver disease, Visual Basic**

I. Introduction

The thesis introduces the field of data mining and clustering, explaining the concepts and techniques involved. It discusses the importance of data mining in extracting useful knowledge from large databases and its applications in various fields, including medicine. The author sets up and applies the C-Means algorithm to group liver disease cases to support treatment. The source also distinguishes the completed demonstration from a clinical deployment: its application had not yet been used in practice, and the available input data were limited [1]. Published books on the methods and subject of this project provide context for this English summary [4].

II. Methods

The methods section details the setup of the C-Means algorithm in a Visual Basic 6.0 application. The author describes the implementation of the algorithm and its application to the problem of grouping liver disease cases. The thesis also covers the use of other clustering techniques such as Fuzzy C-Means and their implementation. The author explains the use of various controls and forms in Visual Basic 6.0, including labels, text boxes, and command buttons, to create a user-friendly interface for input and output. The implementation includes a series of steps to initialize the algorithm, process the input data, and display the results [1].

III. Results

The results include the successful setup and application of the C-Means algorithm and other clustering techniques. The author describes the process of setting up the application and the interface used for input and output. However, the thesis notes that the application has not been fully optimized and the input data is limited. The author demonstrates the application by grouping liver disease cases based on specific patient attributes such as age, gender, and medical history. The clustering results are displayed in a user-friendly interface, allowing for easy interpretation of the data [1].

IV. Discussion

The discussion section highlights the limitations of the current work, such as the lack of practical application and optimization of the algorithms. The author suggests further research into clustering algorithms and the development of a more complete application for practical use. The author also mentions the need for more comprehensive and diverse datasets to improve the accuracy and reliability of the clustering results. The limitations include the limited scope of the application and the need for more extensive testing and validation [1].

V. Conclusion

The thesis reports a simulated clustering algorithm and an application of C-Means to group liver disease cases. It also states that the application had not been deployed in practice, that the algorithm had not been optimized, and that the available input data were limited. The proposed next steps are deeper study of clustering methods and a more complete program suitable for practical use. These limits prevent the summary from claiming any measured improvement in clinical decisions or treatment outcomes [1].

Supervisor: ICTU Teacher 96, Thai Nguyen University of Information and Communication Technology.

References

[1] ICTU Student 153, "Data Clustering Algorithms for Grouping Liver Cirrhosis Medical Records to Support Treatment," internal undergraduate thesis (original in Vietnamese), Thai Nguyen University of Information and Communication Technology, 2010.
[2] Đinh Xuân Lâm, Sách Visual Basic cho sinh viên và kỹ thuật viên Khoa Công nghệ thông tin, NXB Thống kê
[3] Nguyễn thị Kiều Duyên, Lập trình Visual Basic, NXB Trẻ
[4] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, Introduction to Algorithms, 2nd ed. MIT Press, 2001. ISBN 9780262032933.