



International Journal of Informatics, High-End Research, Innovation, and
Emerging Technologies

Inheriten Journal

Vol. 1, No. 1, January 2026 pp. 12-16

Journal Page is available at <https://internationaljournal.lppmunsaka.ac.id/index.php/inheriten>



An Intelligent Approach to Health Risk Analysis: Integration of Fuzzy C-Means and Rule-Based Systems

Afifah Khaerani A¹, Rizki Adha²

¹Universitas Salakanagara

²Institut Teknologi&Bisnis Banten

Abstract

This study addresses the growing prevalence of non-communicable diseases (NCDs) such as hypertension, diabetes, and obesity, which are significantly influenced by unhealthy lifestyle patterns including sedentary behavior, high consumption of processed foods, and elevated stress levels. As these conditions increasingly affect individuals in their productive age, particularly in developing countries like Indonesia, data-driven health risk mapping becomes crucial. The research applies the Fuzzy C-Means (FCM) clustering algorithm to classify individuals based on lifestyle-related risk profiles, handling inherent ambiguity and uncertainty in health data. To enhance interpretability and generate actionable insights, a Fuzzy Rule-Based System (FRBS) is developed to analyze the dominant contributing factors and recommend personalized interventions. The integration of FCM and FRBS provides a hybrid approach that supports public health monitoring and decision-making. Results confirm the effectiveness of this methodology in identifying health risk clusters and offering targeted recommendations, indicating its potential as a strategic tool in preventative healthcare.

Keywords: *Fuzzy C-Means, Health Risk Profiling, Lifestyle Analysis, Non-Communicable Diseases, Fuzzy Rule-Based System, Preventive Healthcare, Data Clustering, Hybrid Intelligence System*

Introduction

The lifestyles of people in the world today have significant implications for the emergence of non-communicable diseases, including hypertension, diabetes, and obesity, and this is a major concern for global public health.([Sitanggang & Sitanggang, 2022](#)). Lifestyles characterized by lack of physical activity, excessive consumption of processed foods, and high levels of stress are the main factors contributing to the increase in NCDs.([Jafar et al., 2020](#)). This trend is becoming increasingly worrying as its prevalence increases significantly among individuals of productive age, especially in developing countries such as Indonesia, underscoring the urgent need for data-driven risk mapping strategies to effectively address this growing public health problem.([Siswati et al., 2022](#)). high levels of stress contribute to the development of health problems.([Dewi et al., 2020](#)).

Global public health is faced with an increasing prevalence of non-communicable diseases, including hypertension, diabetes, and obesity, which are largely influenced by unhealthy lifestyles, such as lack of physical activity, excessive consumption of fast food, and high levels of stress.([Dewanti et al., 2022](#)). [WHO, 2023; Riskesdas, 2020]. This trend is of particular concern in Indonesia, where the prevalence of non-communicable diseases has increased significantly in the productive age group.([Ndraha et al., 2019](#)). Therefore, there is a need for data-driven risk mapping techniques to identify and address the factors contributing to these health problems.([Kurniadi et al., 2022](#)) ([Thamrin et al., 2021](#)). Fuzzy C-Means offers a promising approach to risk mapping by clustering data based on lifestyle features, allowing for handling the ambiguity and uncertainty inherent in health data.([Aulia et al., 2022](#)). The Fuzzy

C-Means algorithm enables risk mapping through data clustering based on lifestyle features, effectively handling data ambiguity with fuzzy principles.([DWIGUNA et al., 2022](#)). The application of Fuzzy C-Means in the health domain has shown its effectiveness in stratifying disease risks, as evidenced by studies on cardiovascular risk and mapping of health facilities.([Maghfiroh & Ulinnuha, 2018](#)). To improve interpretability and provide profile-based recommendations, fuzzy rule-based systems can be used to explore patterns of relationships between lifestyle factors and health risks.([Ghaniy & Indriyaningsih, 2020](#)). Fuzzy rule-based systems offer a flexible and interpretive framework for analyzing complex relationships between lifestyle and health risks.([Numan et al., 2020](#)).

Literature review

To solve the complex problem of nutritional status classification, a system that facilitates ease of computation is essential. The system can be designed using Adaptive Neuro Fuzzy Inference System with a hybrid algorithm, combining learning through the Least Square Estimator and Backpropagation methods with Fuzzy C-Means clustering.([Baun & Nurhayati, 2013](#)). The Fuzzy C-Means algorithm has the advantage of accurate cluster center placement compared to other clustering methods. The Fuzzy C-Means algorithm groups data into clusters where each data point has a degree of belonging to each cluster, and the success of a development program is highly dependent on the accuracy of identifying target groups and target areas, so it is necessary to consider the classification and characteristics of various regions based on community welfare indicators. Fuzzy Rule-Based System can explore patterns of relationships between lifestyle factors and health risks. The FCM + Rule-Based hybrid system has been effectively applied to medical data for cardiovascular disease diagnosis and metabolic risk monitoring.

Unhealthy lifestyles, such as lack of physical activity, excessive consumption of fast food, and high levels of stress([Dwitiyanti et al., 2019](#)) ([Fatonah & Pancarani, 2022](#)). The application of Fuzzy C-Means in the health domain has shown its effectiveness in classifying disease risks, as evidenced by studies on cardiovascular risk and health facility mapping. The hybrid FCM + Rule-Based system has been effectively applied to medical data for cardiovascular disease diagnosis and metabolic risk monitoring. In the context of disease diagnosis, an expert system integrating Naïve Bayes and Certainty Factor methods shows the potential of computer-based systems to improve the accuracy and efficiency of medical diagnosis, such as in the case of uterine cancer.([Yendrizar, 2021](#)). water consumption, physical activity, and smoking habits, for example, have been shown to be significantly correlated with the incidence of hypertension, underscoring the importance of considering these lifestyle factors in health risk assessment.

Methodology

This study aims to apply Fuzzy C-Means to map health risk profiles based on lifestyle data, develop Fuzzy Rule-Based System to analyze dominant factors and provide recommendations, and provide data-based preventive solutions for the community. This study uses quantitative and qualitative methods. Quantitative methods are used to analyze lifestyle and health risk data, while qualitative methods are used to develop Fuzzy Rule-Based System. Fuzzy C-Means method will be used to group individuals based on their health risk profiles, while Fuzzy Rule-Based System will be used to identify dominant lifestyle factors that contribute to health risk and generate personalized recommendations. This method is expected to provide scientific contributions through the integration of hybrid fuzzy methods for public health monitoring. The use of Fuzzy C-Means allows to handle uncertainty and ambiguity in lifestyle data, which is often subjective and imprecise ([AP et al., 2020](#)).

Results

Based on existing research on comparative analysis of the performance of machine learning models to predict the risk of stunting in child growth, it shows that Naïve Bayes has the highest performance with accuracy = 98.57%, F1-Score = 0.99, and very high recall and precision.(Sahamony et al., 2024). In another study, Naïve Bayes was assessed as being able to help by providing recommendations in making decisions in detecting developmental disorders in children.(Silvana et al., 2020). Expert systems are an effective solution to help diagnose digestive disorders in toddlers. By utilizing the Naive Bayes method, expert systems can identify symptoms that may indicate digestive disorders.(Prawestina et al., 2024). Through trials with 100 users, this expert system achieved an accuracy rate of 87%, with 87 patients receiving the correct diagnosis.(Prawestina et al., 2024). The importance of early diagnosis of digestive disorders in toddlers cannot be ignored, because proper treatment can prevent more serious complications.

Discussion

The use of classification methods such as the Naïve Bayes Classifier in the context of stunting nutritional status in toddlers is also promising.(Titimeidara & Hadikurniawati, 2021). The implementation of an expert system based on the Naive Bayes method is effective in helping to diagnose digestive disorders in toddlers.(Prawestina et al., 2024). With an accuracy rate of 87% based on the results of trials with 100 users, this system provides an appropriate diagnosis for 87 patients. The use of data mining algorithms such as Naïve Bayes can provide possible results that will occur in the future from the classification results. However, the Naive Bayes method has a weakness where the independence of the Naive Bayes feature cannot always be applied so that it will affect the level of calculation accuracy.

Conclusion

The application of Fuzzy C-Means in the health domain has shown its effectiveness in categorizing disease risks, as evidenced by studies on cardiovascular risk and health facility mapping. This study uses quantitative and qualitative methods. Quantitative methods are used to analyze lifestyle and health risk data, while qualitative methods are used to develop the Fuzzy Rule-Based System.

References

- AP, I., Aminah, S., Cokrowibowo, S., & Zulkarnaim, N. (2020). Regional Clustering Based on Environmental Health Data Using Fuzzy C-Means. *Journal of Computer and Information System (J-CIS)*, 1(2), 12. <https://doi.org/10.31605/jcis.v1i2.609>
- Aulia, GP, Widiharih, T., & Utami, IT (2022). APPLICATION OF TEXT MINING AND FUZZY C-MEANS CLUSTERING TO IDENTIFY MAIN COMPLAINTS OF CUSTOMERS OF PDAM TIRTA MOEDAL IN SEMARANG CITY. *Gaussian Journal*, 12(1), 126. <https://doi.org/10.14710/j.gauss.12.1.126-135>
- Baun, HM, & Nurhayati, OD (2013). Implementation of Adaptive Neuro Fuzzy Inference System for Determining Toddler Nutritional Status. *JOURNAL OF BUSINESS INFORMATION SYSTEMS*, 3(2). <https://doi.org/10.21456/vol3iss2pp116-125>
- Dewanti, D., Syauqy, A., Noer, ER, & Pramono, A. (2022). RELATIONSHIP BETWEEN DIET PATTERNS AND PHYSICAL ACTIVITY WITH CENTRAL OBESITY IN ELDERLY AGE IN INDONESIA: BASIC HEALTH RESEARCH DATA. *Gizi Indonesia*, 45(2), 79. <https://doi.org/10.36457/gizindo.v45i2.662>
- Dewi, FT, Sriatmi, A., & Nandini, N. (2020). Perception of Economic and Social Impacts of Physical Vulnerability on Health Status during the COVID-19 Pandemic (Case Study

- in Central Java Province). *INDONESIAN PUBLIC HEALTH MEDIA*, 20(1), 19.
<https://doi.org/10.14710/mkmi.20.1.19-25>
- DWIGUNA, IKS, GANDHIADI, GK, & HARINI, LPI (2022). IMPLEMENTATION OF FUZZY C-MEAN AND PARTICLE SWARM OPTIMIZATION ALGORITHM FOR DISTRICT/CITY CLUSTERING IN INDONESIA BASED ON HUMAN DEVELOPMENT INDEX INDICATORS. *E-Jurnal Matematika*, 11(3), 191.
<https://doi.org/10.24843/mtk.2022.v11.i03.p380>
- Dwitiyanti, N., Selvia, N., & Andrari, FR (2019). Application of Fuzzy C-Means Cluster in Grouping Indonesian Provinces According to People's Welfare Indicators. *Exacta Factors*, 12(3), 201. <https://doi.org/10.30998/faktorexacta.v12i3.4526>
- Fatonah, NS, & Pancarani, TK (2022). COMPARATIVE ANALYSIS OF CLUSTERING ALGORITHMS FOR MAPPING THE NUTRITION STATUS OF TODDLERS AT PASIR JAYA COMMUNITY HEALTH CENTER. *KONVERGENCE*, 18(1), 1.
<https://doi.org/10.30996/konv.v18i1.5497>
- Ghaniy, R., & Indriyaningsih, F. (2020). Application of Fuzzy C-Means Method in the Selection of Study Programs for New Students in Higher Education. *TeknoIS Scientific Journal of Information Technology and Science*, 10(2), 19.
<https://doi.org/10.36350/jbs.v10i2.84>
- Jafar, N., Hasan, N., Hadju, V., Thaha, RM, & Arundhana, AI (2020). Improved Knowledge, Attitudes, and Practices of Balanced Nutrition after Educational Intervention Based on the Self-determination Theory: An Intervention Study in Senior School Teachers in Makassar City. *Open Access Macedonian Journal of Medical Sciences*, 8, 228.
<https://doi.org/10.3889/oamjms.2020.4605>
- Kurniadi, I., Utomo, W., & Sabrian, F. (2022). THE EFFECT OF WARM WATER FOOT SOAKING AND LAVENDER AROMATHERAPY ON BLOOD PRESSURE IN PRIMARY HYPERTENSION PATIENTS. *Jurnal Ners Indonesia*, 12(2), 186.
<https://doi.org/10.31258/jni.12.2.186-195>
- Maghfiroh, W., & Ulinnuha, N. (2018). Application of Fuzzy C-Means in Grouping Districts/Cities Based on Health Service Facilities in East Java. *Inform Scientific Journal of Information and Communication Technology*, 3(2), 62.
<https://doi.org/10.25139/inform.v3i2.1070>
- Ndraha, S., Caesar, Z., Afdilla, W., Rianas, W., Kamala, Y., Farras, Z., & El, Z. (2019). Hypertension Profile in Internal Medicine Polyclinic of Koja Regional Hospital. *YARSI Medical Journal*, 27(2), 69. <https://doi.org/10.33476/jky.v27i2.1117>
- Numan, N., Kusumadewi, S., & Muzayyanah, N. (2020). Fuzzy Inference System to Help Diagnose Childhood Pneumonia. *IT JOURNAL RESEARCH AND DEVELOPMENT*, 5(1), 53. [https://doi.org/10.25299/itjrd.2020.vol5\(1\).5088](https://doi.org/10.25299/itjrd.2020.vol5(1).5088)
- Prawestina, EM, Rohman, Mg., & Zamroni, Moh. R. (2024). Expert System for Diagnosing Digestive Disorders in Toddlers Using the Naive Bayes Method. *Deleted Journal*, 3(1), 23. <https://doi.org/10.29407/stains.v3i1.4079>
- Sahamony, NF, Terttiaavini, T., & Rianto, H. (2024). Comparative Analysis of Machine Learning Model Performance to Predict Stunting Risk in Child Growth. *MALCOM Indonesian Journal of Machine Learning and Computer Science*, 4(2), 413.
<https://doi.org/10.57152/malcom.v4i2.1210>
- Silvana, M., Akbar, R., & Alfi, S. (2020). Utilization of the Naïve Bayes Method in the Implementation of Expert Systems to Analyze Child Developmental Disorders. *National Journal of Technology and Information Systems*, 6(2), 74.
<https://doi.org/10.25077/teknosi.v6i2.2020.74-81>

- Siswati, T., Kasjono, HS, & Olfah, Y. (2022). "Posbindu PTM": The Key of Early Detection and Decreasing Prevalence of Non-Communicable Diseases in Indonesia. *Iranian Journal of Public Health*. <https://doi.org/10.18502/ijph.v5i17.10105>
- Sitanggang, PA, & Sitanggang, FA (2022). Analysis of Risk Management Implementation Based on SNI ISO 31000:2018 (Case Study: Spare Parts for Second Personal Computers in Jambi). *Eksis Scientific Journal of Economics and Business*, 13(1), 12. <https://doi.org/10.33087/eksis.v13i1.293>
- Thamrin, SA, Arsyad, DS, Kuswanto, H., Lawi, A., & Nasir, S. (2021). Predicting Obesity in Adults Using Machine Learning Techniques: An Analysis of Indonesian Basic Health Research 2018. *Frontiers in Nutrition*, 8. <https://doi.org/10.3389/fnut.2021.669155>
- Titimeidara, MY, & Hadikurniawati, W. (2021). IMPLEMENTATION OF NAÏVE BAYES CLASSIFIER METHOD FOR CLASSIFICATION OF STUNTING NUTRITION STATUS IN TODDLERS. *SCIENTIFIC JOURNAL OF INFORMATICS*, 9(1), 54. <https://doi.org/10.33884/jif.v9i01.3741>
- Yendrizar, Y. (2021). Expert System in Diagnosing Uterine Cancer Using Naive Bayes and Certainty Factor Methods. *BUDIDARMA INFORMATICS MEDIA JOURNAL*, 5(1), 251. <https://doi.org/10.30865/mib.v5i1.2669>