Journal of Technology Informatics and Engineering (JTIE)
Vol. 5 No. 2 August 2026
E-ISSN: 2961-9068; P-ISSN: 2961-8215, Pages 417-439
DOI: https://doi.org/10.51903/jtie.v5i2.576

Clustering of Infectious and Non-Communicable Disease Distribution in Lhokseumawe City Using the Fuzzy Gustafson-Kessel Method

M. Rahmat Azhari*[1], Eva Darnila[1], Yesy Afrillia[1]
Email: rahmat.210170111@mhs.unimal.ac.id, eva.darnila@unimal.ac.id, yesy.afrillia@unimal.ac.id
[1]*Department of Informatics Engineering, Faculty of Engineering, Universitas Malikussaleh, Lhokseumawe, Indonesia 24353*
*Corresponding Author

Abstract

The high incidence of infectious and non-communicable diseases in Lhokseumawe City requires a data-driven approach for mapping distribution areas. This study designs a disease distribution clustering system using the Fuzzy Gustafson-Kessel (FGK) method and compares three fuzzy membership functions: linear, sigmoid, and Gaussian. FGK excels in handling non-uniform data distributions utilizing an adaptive covariance matrix based on Mahalanobis distance. The data comprised 38,063 cases from the Lhokseumawe City Health Office (2023–2024), aggregated into 4 sub-districts, and processed using z-score standardization. The clustering was implemented in a Python and MySQL web-based system and evaluated using the Xie-Beni validity index. The FGK method successfully grouped the sub-districts into 3 clusters: low, moderate, and high. For infectious diseases (converging at iteration 24), Blang Mangat formed the low cluster, Muara Dua and Muara Satu the moderate, and Banda Sakti the high. For non-communicable diseases (converging at iteration 9), Blang Mangat was low, Muara Satu moderate, while Banda Sakti and Muara Dua were high. The linear membership function proved optimal, yielding the lowest average Xie-Beni Index of 0.1316. In conclusion, the FGK method with a linear membership function effectively maps disease distribution, assisting the local Health Office in targeted policy planning and resource allocation.

Keywords: *Fuzzy Gustafson-Kesse, Clustering, Xie-Beni Index, Disease.*

I. INTRODUCTION

Infectious and non-communicable diseases remain significant public health challenges because of their effects on population well-being, healthcare demand, and the allocation of limited health resources. Infectious diseases can spread through biological agents such as viruses, bacteria, and parasites and continue to contribute substantially to global morbidity and mortality, particularly when newly emerging or previously controlled diseases reappear (Afdal & Humani, 2020; Andika et al., 2020). At the same time, non-communicable diseases (NCDs) develop gradually and often require long-term management. Lifestyle-related factors, including unhealthy diets, insufficient physical activity, and smoking, contribute to conditions such as hypertension, diabetes mellitus, cardiovascular disease, and other chronic disorders (Puspa Sari et al., 2024). The coexistence of these two disease categories creates a need for analytical approaches capable of identifying variations in disease burden across geographic areas.

Lhokseumawe City represents an important context for such analysis. The city consists of four sub-districts, Blang Mangat, Muara Dua, Muara Satu, and Banda Sakti, with different demographic and environmental characteristics (Rizky et al., 2025). Infectious diseases such as tuberculosis, dengue fever, and hepatitis remain relevant public health concerns, while diabetes

Received in May 2026; Revised in June 2026; Accepted in July 2026; Published in August 2026

mellitus, hypertension, and stroke contribute to the burden of non-communicable diseases. Environmental conditions may further influence disease occurrence, particularly for vector-borne diseases, while poor environmental management can increase risks associated with unhealthy living conditions (Rahma et al., 2024; Chaniago et al., 2023). Differences in disease incidence among sub-districts therefore require more than descriptive reporting; they require analytical methods capable of revealing underlying patterns and identifying areas with relatively low, moderate, or high disease burdens.

Data mining provides a computational approach for extracting meaningful patterns, relationships, and structures from large datasets using statistical and artificial intelligence-based techniques (Ardini, 2022; Gumilang & Ariyani, 2023). Within data mining, clustering is an unsupervised learning technique that groups observations according to similarity without requiring predefined class labels. It has been applied to various analytical problems, including regional zoning, pattern recognition, market segmentation, prediction, and spatial analysis (Prayoga et al., 2025). Clustering is particularly relevant for health data because disease distributions are rarely uniform across geographic areas, and identifying groups with similar epidemiological characteristics can provide useful information for data-driven health planning (Nurdin et al., 2024).

However, conventional clustering approaches may have limitations when applied to datasets whose groups differ in shape, orientation, or covariance structure. The Fuzzy Gustafson-Kessel (FGK) method offers an alternative because it extends fuzzy clustering by incorporating an adaptive covariance matrix and Mahalanobis-based distance calculation. This characteristic allows FGK to adjust to the geometric structure of individual clusters rather than assuming that all clusters have a similar spherical form (Gustafson & Kessel, 1978; Saraswati et al., 2024). Previous applications have demonstrated the potential of FGK for grouping heterogeneous datasets. Putri and Rochmawati (2021), for example, applied FGK to social media data and showed that the method could produce measurable clustering performance, while recent research has also applied FGK to regional health-related indicators and characteristics (Saraswati et al., 2024; Dyfa et al., 2025). These studies indicate that adaptive fuzzy clustering has potential for analysing complex regional data.

Nevertheless, an important research gap remains in its application to local disease surveillance. Existing clustering studies have commonly applied methods such as K-Means or other fuzzy clustering techniques, whereas the use of FGK for simultaneously examining the geographical distribution of infectious and non-communicable diseases remains limited. Previous applications also tend to analyse a single category of indicators or a single problem domain, providing limited insight into whether different disease categories produce similar or contrasting regional patterns. In addition, the effect of different fuzzy membership functions on the quality of FGK clustering has received limited attention in regional disease-distribution analysis. This issue is important because the membership function determines the degree to which an observation belongs to a particular cluster and may consequently influence cluster separation and interpretability.

Accordingly, the novelty of this study lies in integrating FGK-based clustering of both infectious and non-communicable disease distributions within a single analytical framework and systematically comparing linear, sigmoid, and Gaussian fuzzy membership functions using the Xie-Beni validity index. Rather than merely producing fixed regional groups, the proposed approach evaluates the degree of cluster membership and identifies the membership formulation that produces the most compact and well-separated clustering structure. The clustering results are further incorporated into a web-based information system with spatial visualization, providing a computational mechanism for transforming health-case data into more interpretable regional information. This contribution is relevant to informatics-based health analytics because it combines adaptive fuzzy clustering, cluster-validity assessment, and digital visualization to support data-driven interpretation of regional disease patterns.

Therefore, this study aims to apply the Fuzzy Gustafson-Kessel method to cluster the distribution of infectious and non-communicable diseases across the sub-districts of Lhokseumawe City, evaluate clustering outcomes using different fuzzy membership functions, and determine the most appropriate membership function based on the Xie-Beni Index. The resulting clusters are expected to provide a more structured representation of regional disease burden and support health authorities in identifying priority areas for monitoring, intervention planning, and resource allocation.

II. LITERATURE REVIEW

Clustering is an important technique in data mining for identifying natural groupings within unlabeled datasets. Unlike supervised learning, which relies on predefined target classes, clustering analyzes similarities and differences among observations to reveal latent structures that may not be immediately visible through descriptive analysis alone. This approach has been widely applied in regional zoning, pattern recognition, segmentation, and spatial analysis because it enables heterogeneous observations to be organized into relatively homogeneous groups (Hidayat et al., 2023; Nurdin et al., 2024; Prayoga et al., 2025). In the context of public health analytics, clustering can provide a useful basis for distinguishing regions with different levels of disease burden and for supporting more targeted interpretation of epidemiological patterns.

Several clustering algorithms have been developed to address different data characteristics. Conventional partition-based methods such as K-Means assign each observation exclusively to a single cluster, whereas fuzzy clustering methods allow an observation to have partial membership in more than one cluster. This characteristic is particularly relevant for health-related regional data because disease profiles may overlap across geographic areas rather than form completely separated groups. Fuzzy approaches therefore offer a more flexible representation of uncertainty and gradual transitions between disease-risk categories (Almeida & Chen, 2025; Kuznetsova & Nkosi, 2025).

The Fuzzy Gustafson-Kessel (FGK) algorithm is an extension of fuzzy clustering that was designed to address limitations associated with fixed distance structures. The method modifies the conventional fuzzy objective function by incorporating a cluster-specific covariance matrix and Mahalanobis-based adaptive distance calculation (Gustafson & Kessel, 1978). Through this mechanism, each cluster can adapt to the local geometric characteristics of the data instead of assuming a uniform spherical shape. This property makes FGK suitable for datasets with unequal variances, different orientations, or non-uniform distributions. Saraswati et al. (2024) emphasized that the use of an adaptive covariance structure enables FGK to provide more flexible grouping when observations exhibit heterogeneous spatial or statistical characteristics.

The suitability of FGK for complex data has also been demonstrated in several application domains. Putri and Rochmawati (2021) applied the FGK algorithm to clustering social media data associated with Go-Jek and Grab Indonesia and reported measurable classification-related performance across accuracy, precision, recall, and F-measure indicators. Although the application domain differs from health analytics, the study illustrates the capability of FGK to handle multidimensional data with overlapping characteristics. In a regional context, Saraswati et al. (2024) used FGK to group districts and cities in Kalimantan based on housing and environmental health indicators, showing the relevance of adaptive fuzzy clustering for geographically aggregated data. More recently, Dyfa et al. (2025) employed FGK to group districts and cities in Central Java based on the ratio of health workers, further demonstrating the potential of the method for regional health-related analysis.

The use of clustering in regional analysis is also supported by studies employing other algorithms. Hidayat et al. (2023), for example, applied K-Means clustering to identify natural-disaster-prone zones in Mandailing Natal, demonstrating how unsupervised methods can support geographic grouping and regional prioritization. Nurdin et al. (2024) similarly used K-Means to cluster capture fisheries products, confirming the broader applicability of clustering to heterogeneous regional datasets. These studies indicate that clustering is valuable for transforming multidimensional regional data into interpretable groups. However, methods based on fixed Euclidean-distance structures may be less adaptive when cluster shapes and covariance patterns differ substantially.

An additional consideration in fuzzy clustering concerns the membership function, which determines the degree to which each observation belongs to a particular cluster. Different membership formulations can produce different degrees of cluster compactness and separation, even when the same underlying data and clustering framework are used. Therefore, evaluating the quality of clustering results is necessary to determine whether a particular membership formulation produces a meaningful partition. One commonly used measure is the Xie-Beni Index, which evaluates the ratio between within-cluster compactness and separation among cluster centers. A lower Xie-Beni value indicates better clustering quality because it reflects relatively compact clusters with stronger separation (Putri & Rochmawati, 2021).

Despite the demonstrated potential of FGK, its application to disease-distribution analysis at the local level remains limited in the literature used in this study. Existing studies tend to apply FGK to environmental indicators, social media data, or regional resource characteristics, while disease-related clustering is often examined through more conventional clustering approaches. In addition, relatively little attention has been given to simultaneously analyzing infectious and non-communicable disease patterns within the same regional framework and evaluating how different fuzzy membership functions affect the resulting cluster validity.

Based on these limitations, this study positions FGK as an adaptive clustering method for identifying spatial patterns in infectious and non-communicable disease distributions across the sub-districts of Lhokseumawe City. The study further extends the analysis by comparing linear, sigmoid, and Gaussian membership functions and evaluating their performance using the Xie-Beni Index. This approach is intended to determine which membership formulation produces the most compact and well-separated clustering structure, while also providing a stronger analytical basis for interpreting regional disease-distribution patterns.

III. RESEARCH METHOD(S)

A. Research Design

This study employed a quantitative computational approach to analyze the spatial distribution patterns of infectious and non-communicable diseases in Lhokseumawe City using the Fuzzy Gustafson-Kessel (FGK) clustering method. The analysis was designed to identify similarities in disease burden among sub-districts and classify them into low, moderate, and high clusters. In addition to implementing the standard FGK procedure, the study compared linear, sigmoid, and Gaussian fuzzy membership functions to determine which formulation produced the most compact and well-separated clusters. The research procedure consisted of data acquisition, aggregation of disease cases by sub-district, data standardization, FGK parameter initialization, iterative clustering, membership-function comparison, cluster validity assessment using the Xie-Beni Index, and visualization of clustering results through a web-based information system. Python was used for computational processing, while MySQL was employed for data storage and management.

B. Study Area and Data Source

The study was conducted in Lhokseumawe City, Aceh Province, Indonesia. The study area comprises four administrative sub-districts: Banda Sakti, Blang Mangat, Muara Dua, and Muara Satu. The details of this information are shown in Table 1 and Table 2. Disease data were obtained from the Lhokseumawe City Health Office for the 2023–2024 period. A total of 38,063 recorded disease cases were analyzed. The dataset was divided into two categories. Infectious diseases consisted of tuberculosis (TB), hepatitis, and dengue fever, whereas non-communicable diseases consisted of diabetes mellitus, hypertension, and stroke.

Table 1. Aggregation Data on Infectious Diseases per District

Subdistrict	TBC	Hepatitis	DBD	Total
Banda Shakti	226	18	37	281
Blang Mangat	45	5	9	59
Muara Dua	138	7	11	156
Muara Satu	102	2	22	126
Total	**511**	**32**	**79**	**622**

The observations were aggregated according to the four sub-districts to enable regional clustering analysis. The original manuscript confirms that these six disease variables and 38,063 cases constitute the analytical dataset. Together, the infectious and non-communicable disease datasets contained 38,063 cases, consisting of 622 infectious disease cases and 37,441 non-communicable disease cases.

Table 2. Aggregation Data on Non-Communicable Diseases per District

Subdistrict	Diabetes Mellitus	Hypertension	Stroke	Total
Banda Shakti	2.839	7.530	250	10.619
Blang Mangat	1.895	6.000	182	8.077
Muara Dua	2.555	7.062	220	9.837
Muara Satu	2.177	6.531	200	8.908
Total	**9.466**	**27.123**	**852**	**37.441**

C. Data Processing

Data standardization was performed before clustering because the disease variables had substantially different numerical ranges. Without standardization, variables with larger values could exert disproportionate influence on distance calculations and subsequently dominate the clustering results. Therefore, each disease variable was transformed using z-score standardization. The standardized value was calculated as Eq. (1).

Journal of Technology Informatics and Engineering (JTIE)
Vol. 5 No. 2 August 2026
E-ISSN: 2961-9068; P-ISSN: 2961-8215, Pages 417-439
Journal Link: https://jtie.stekom.ac.id/index.php/jtie/index

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (1)$$

where z_{ij} represents the standardized value of observation I for variable j, x_{ij} represents the original value, μ_j represents the mean of variable j, and σ_j represents the corresponding standard deviation. The standardization procedure placed all disease variables on comparable scales while maintaining their relative differences across sub-districts. The resulting standardized datasets are presented in Tables 3 and 4.

Table 3. Standardized Infectious Disease Data

Sub-district	Tuberculosis (z)	Hepatitis (z)	Dengue Fever (z)
Banda Sakti	1.495313	1.655212	1.551051
Blang Mangat	-1.259411	-0.496564	-0.966597
Muara Dua	0.156000	-0.165521	-0.786765
Muara Satu	-0.391901	-0.993127	0.202311

The standardized values indicate the relative position of each sub-district for each disease variable rather than the absolute number of cases. These standardized observations were subsequently used as input for FGK clustering. The values above are consistent with the preprocessing results reported in the original manuscript.

Table 4. Standardized Non-Communicable Disease Data

Sub-district	Diabetes Mellitus (z)	Hypertension (z)	Stroke (z)
Banda Sakti	1.314262	1.308037	1.465993
Blang Mangat	-1.311480	-1.363029	-1.228265
Muara Dua	0.524314	0.491005	0.277350
Muara Satu	-0.527096	-0.436012	-0.515079

D. Fuzzy Gustafson-Kessel Clustering

Fuzzy Gustafson-Kessel is an extension of fuzzy clustering that incorporates an adaptive distance norm based on the covariance structure of each cluster. Unlike clustering approaches that assume similar spherical distributions, FGK allows clusters to adapt to different orientations and geometric structures by using a fuzzy covariance matrix and Mahalanobis-based distance measurement (Gustafson & Kessel, 1978; Saraswati et al., 2024). The clustering procedure began by initializing the fuzzy partition matrix, which represents the degree of membership of each observation in each cluster. The membership values for an observation satisfy Eq (2).

$$\sum_{k=1}^{c} u_{ik} = 1 \quad (2)$$

where u_{ik} represents the membership degree of observation i in cluster k, and c denotes the number of clusters. A fixed random seed of 42 was used for initialization to ensure reproducibility of the computational procedure. The centroid of each cluster was calculated as in Eq. (3).

$$v_k^{(t)} = \frac{\sum_{i=1}^{n}(u_{ik}^{(t-1)})^m xi}{\sum_{i=1}^{n}(u_{ik}^{(t-1)})^m}, 1 \leq k \leq c \quad (3)$$

where vk represents the centroid of cluster k, x_i represents observation i, u_{ik} represents its membership degree, and m represents the fuzzy weighting exponent. A key characteristic of FGK is the calculation of a separate fuzzy covariance matrix for each cluster (4).

$$F_k = \frac{\sum_{i=1}^{n}(u_{ik}^{(t-1)})^m (x_i - v_k^{(t)} x_i - v_k^{(t)})^T}{\sum_{i=1}^{n}(u_{ik}^{(t-1)})^m} \quad (4)$$

The covariance matrix captures the internal distribution and geometric characteristics of each cluster. To improve numerical stability, covariance regularization was applied using the parameter γ, following the configuration adopted in the study. The distance between an observation and a cluster centroid was calculated using the Mahalanobis-based adaptive distance in Eq. (5).

$$d_{ik} = \sqrt{(x_i - v_k)^T M_k (x_i - v_k)} \quad (5)$$

where M_k represents the adaptive norm-inducing matrix derived from the covariance structure of cluster k. The membership values were subsequently updated according to the relative distance of each observation from all cluster centroids ias Eq. (6).

$$u_{ik}^{(t)} = \left[\sum_{I=1}^{c}\left(\frac{D_{ikA_k}}{D_{iIA_k}}\right)^{\frac{2}{(m-1)}}\right]^{-1} \quad (6)$$

The centroid, covariance matrix, adaptive distance, and membership matrix were recalculated iteratively until the convergence criterion was satisfied.

E. FGK Parameter Configuration

The FGK configuration was kept consistent for both disease categories so that the resulting clustering patterns could be interpreted under equivalent computational conditions. Three clusters were specified to represent low, moderate, and high disease-distribution levels. These parameters correspond to the configuration reported in the original computational procedure.

Table 5. FGK Parameter Configuration

Parameter	Value	Description
Number of clusters (c)	3	Low, moderate, and high clusters
Fuzzifier (m)	2	Controls the degree of fuzzy membership
Error tolerance (ε)	0.00001	Minimum convergence threshold
Maximum iterations	100	Maximum number of clustering iterations
Regularization parameter (γ)	0.1	Covariance regularization parameter
Eigenvalue threshold (β)	10^{15}	Threshold for covariance eigenvalue ratio
Random seed	42	Ensures reproducible initialization

The iterative procedure was terminated when:

$$max\left|U^{(t)} - U^{(t-1)}\right| U(t) - U(t-1) < \varepsilon$$

with:

Journal of Technology Informatics and Engineering (JTIE)
Vol. 5 No. 2 August 2026
E-ISSN: 2961-9068; P-ISSN: 2961-8215, Pages 417-439
Journal Link: https://jtie.stekom.ac.id/index.php/jtie/index

$$\varepsilon = 0.00001$$

where $U^{(t)}$ represents the membership matrix at iteration t. If the required tolerance was not achieved, the algorithm continued until the maximum limit of 100 iterations.

F. Fuzzy Membership Function Comparison

In addition to the original FGK membership formulation, three alternative membership functions were evaluated to determine whether different transformations of cluster distance affected clustering quality. The functions compared were linear, sigmoid, and Gaussian.

The linear membership function was formulated as:

$$\mu(d) = \frac{1}{d}$$

The sigmoid membership function was formulated as:

$$\mu(d) = \frac{1}{1 + \exp\left(-\frac{1}{d}\right)}$$

The Gaussian membership function was formulated as:

$$\mu(d) = \exp(-\frac{d^2}{2\sigma^2})$$

where represents the distance between an observation and a cluster center, while represents the Gaussian spread parameter. Each membership formulation was evaluated on both infectious and non-communicable disease datasets. The original FGK formulation was retained as a reference to determine whether alternative membership transformations improved cluster compactness and separation. The comparison therefore involved four conditions: original FGK, linear, sigmoid, and Gaussian membership formulations. These are the membership formulations reported and compared in the manuscript.

G. Cluster Validity Assessment

The quality of the resulting clusters was evaluated using the Xie-Beni Index. This index simultaneously considers cluster compactness and separation and is commonly used to assess fuzzy clustering structures. A smaller Xie-Beni value indicates a better clustering solution because observations are more compact within their respective clusters while cluster centers remain relatively well separated (Putri & Rochmawati, 2021). The Xie-Beni Index was calculated as:

$$XB = \frac{\sum_{i=1}^{n} \sum_{k=1}^{c} u_{ik}^{m} \|x_i - v_k\|^2}{n \min_{k \neq 1} \|x_i - v_k\|^2}$$

where n represents the number of observations, c represents the number of clusters, u_{ik}_represents the degree of membership of observation i in cluster k, represents the fuzzifier, x^i represents the observation vector, and represents the corresponding cluster centroid. The Xie-Beni values obtained from the original FGK, linear, sigmoid, and Gaussian membership formulations were compared for both disease categories. The membership function producing the lowest average Xie-Beni value was considered the most appropriate configuration for the dataset.

H. System Implementation and Research Flow

The clustering model was implemented using Python, while MySQL was used to manage disease records and clustering outputs. Python supported the preprocessing, FGK computation, membership calculation, and cluster validation stages, whereas MySQL was used for structured data storage and retrieval. The use of Python and MySQL follows the implementation approach described in the manuscript (Dyfa et al., 2025; Tarina et al., 2025). The overall analytical workflow began with the acquisition of disease data from the Lhokseumawe City Health Office. The records were separated into infectious and non-communicable disease categories and aggregated according to sub-district. Z-score standardization was then applied to minimize the effect of differences in variable scale.

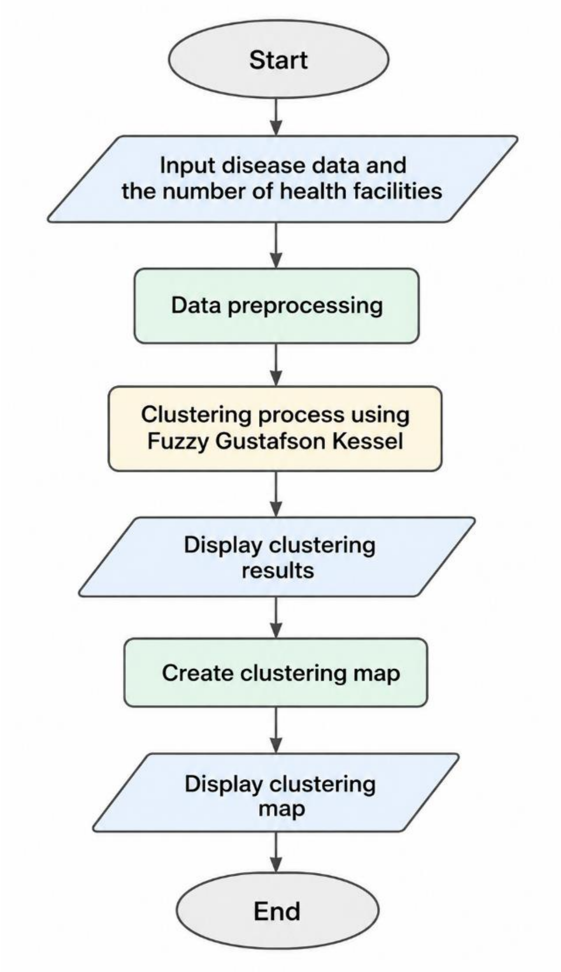

Figure 1. System Scheme

Subsequently, the FGK algorithm was initialized using predefined parameters, followed by iterative estimation of cluster centroids, covariance matrices, Mahalanobis distances, and fuzzy membership degrees. After convergence, each sub-district was assigned to the cluster corresponding to its highest membership degree. The clustering procedure was then evaluated under the original FGK, linear, sigmoid, and Gaussian membership formulations. The Xie-Beni Index was calculated for each configuration to determine the clustering structure with the highest compactness and separation. Finally, the resulting regional clusters were integrated into a web-based information system and presented through map-based visualization to facilitate interpretation of disease-distribution patterns.

I. System Scheme

This system scheme begins with the data input process. Entering data related to infectious and non-infectious diseases, as well as the number of health facilities. The system will perform the clustering process using Fuzzy Gustafson-Kessel. Data preprocessing will then be performed for data cleaning and preparation, such as normalization, feature selection, and handling of empty data. Once the process is complete, the system will present the calculation results along with the data clustering results. Using the obtained clustering results, the system will generate a clustering map. The process ends, and the clustering map will then be displayed.

IV. RESULT/FINDINGS AND DISCUSSION

A. FGK Clustering Performance for Infectious Diseases

The Fuzzy Gustafson-Kessel clustering procedure was first applied to the standardized infectious disease dataset consisting of tuberculosis, hepatitis, and dengue fever cases across the four sub-districts of Lhokseumawe City. The algorithm was configured to generate three clusters representing low, moderate, and high disease-distribution levels. Convergence was evaluated based on the maximum change in the membership matrix between consecutive iterations, with a tolerance value of 0.00001.

Table 6. Convergence Performance for Infectious Disease Clustering

Iteration	Objective Function (J)	Maximum Membership Change	Status
1	0.785434	–	Initial
2	0.430588	0.342455	Continuing
5	0.366359	0.068899	Continuing
23	0.215888	0.000026	Continuing
24	0.215888	0.000009	Converged

As shown in Table 6, the objective function decreased substantially from 0.785434 during the initial iteration to 0.215888 at convergence. The maximum change in membership values also progressively declined and reached 0.000009 at iteration 24, which was below the predefined tolerance of 0.00001. Therefore, the infectious disease clustering process was considered stable after 24 iterations. The reduction in the objective function indicates that observations progressively moved toward a more compact clustering configuration during the iterative optimization process. The stable objective value between iterations 23 and 24 further suggests that additional iterations would have produced negligible changes in cluster membership. The final fuzzy membership values and corresponding cluster assignments are presented in Table 7.

Table 7. Final FGK Clustering Results for Infectious Diseases

Sub-district	μ(C0)	μ(C1)	μ(C2)	Dominant Cluster	Disease Level
Banda Sakti	0.0000	0.0000	1.0000	C2	High
Blang Mangat	0.9999	0.0000	0.0000	C0	Low
Muara Dua	0.0547	0.9348	0.0106	C1	Moderate
Muara Satu	0.0477	0.9436	0.0088	C1	Moderate

Table 7 demonstrates clear differences in infectious disease burden across the four sub-districts. Banda Sakti was assigned to the high cluster (C2) with a membership degree of 1.0000, indicating an exceptionally strong association with the high infectious disease category. This result is consistent with its standardized disease profile, in which tuberculosis, hepatitis, and dengue fever values were consistently higher than those of the other sub-districts. In contrast, Blang Mangat was classified into the low cluster (C0) with a membership degree of approximately 0.9999. The extremely high membership value indicates a clear distinction between Blang Mangat and areas experiencing greater infectious disease burdens.

Muara Dua and Muara Satu were both classified into the moderate cluster (C1), with dominant membership degrees of 0.9348 and 0.9436, respectively. Nevertheless, the membership values were not completely exclusive. Muara Dua retained a membership value of 0.0547 in C0, while Muara Satu retained a value of 0.0477 in C0. These partial memberships illustrate the fuzzy characteristic of FGK, in which an observation may possess different degrees of association with multiple clusters rather than being assigned exclusively to a single group. Overall, the infectious disease clustering formed a clear regional pattern: Banda Sakti represented the high-burden group, Blang Mangat represented the low-burden group, while Muara Dua and Muara Satu represented the moderate-burden group. These assignments correspond to the final membership matrix reported by the FGK procedure.

B. FGK Clustering Performance for Non-Communicable Diseases

The same clustering configuration was subsequently applied to the non-communicable disease dataset consisting of diabetes mellitus, hypertension, and stroke. Compared with the infectious disease dataset, the FGK model required substantially fewer iterations to reach convergence.

Table 8. Convergence Performance for Non-Communicable Disease Clustering

Iteration	Objective Function (J)	Maximum Membership Change	Status
1	0.116060	–	Initial
2	0.067599	0.299768	Continuing
3	0.055198	0.120218	Continuing
8	0.057261	0.000056	Continuing
9	0.057261	0.000006	Converged

As presented in Table 8, the non-communicable disease clustering converged at iteration 9, when the maximum membership change decreased to 0.000006. This value satisfied the predefined convergence tolerance and indicates that stable cluster memberships were achieved substantially earlier than in the infectious disease analysis. The final objective function was 0.057261, considerably lower than the final value of 0.215888 obtained for infectious diseases. Within the context of the present dataset, this indicates a more compact clustering configuration for the non-communicable disease observations. The faster convergence also reflects stronger separation in the standardized profiles of diabetes mellitus, hypertension, and stroke across the four sub-districts. The final membership structure is shown in Table 9.

Table 9. Final FGK Clustering Results for Non-Communicable Diseases

Sub-district	μ(C0)	μ(C1)	μ(C2)	Dominant Cluster	Disease Level
Banda Sakti	0.0013	0.0028	0.9958	C2	High
Blang Mangat	1.0000	0.0000	0.0000	C0	Low
Muara Dua	0.0033	0.0118	0.9850	C2	High
Muara Satu	0.0000	1.0000	0.0000	C1	Moderate

As shown in Table 9, Banda Sakti remained in the high cluster, achieving a dominant membership value of 0.9958 in C2. Blang Mangat also maintained its position in the low cluster, with complete membership in C0. These results show that the relative positions of Banda Sakti and Blang Mangat were highly consistent across both disease categories. A different pattern was observed for Muara Dua. While this sub-district belonged to the moderate cluster for infectious diseases, it shifted to the high cluster for non-communicable diseases, with a membership degree of 0.9850. This finding indicates that the relative non-communicable disease burden in Muara Dua was considerably stronger than its infectious disease burden. Muara Satu, meanwhile, was exclusively classified into the moderate cluster (C1), with a membership degree of 1.0000. The strong memberships observed in Table 9 indicate that the non-communicable disease dataset generated more distinct cluster separation than the infectious disease dataset. The reported final FGK solution confirms high membership certainty for all four sub-districts.

C. Comparison of Infectious and Non-Communicable Disease Clusters

To clarify the differences between the two disease categories, Table 10 summarizes the final regional classifications obtained from FGK. Table 10 reveals both stable and contrasting disease-distribution patterns. Banda Sakti was consistently categorized as a high-burden area, whereas Blang Mangat remained consistently within the low-burden category. These stable assignments suggest substantial differences in disease profiles between these two sub-districts. Muara Satu also exhibited a relatively stable pattern, remaining within the moderate cluster for both disease categories.

Table 10. Comparison of Regional Cluster Assignments

Sub-district	**Infectious Diseases**	**Non-Communicable Diseases**	**Pattern**
Banda Sakti	High	High	Consistently high
Blang Mangat	Low	Low	Consistently low
Muara Dua	Moderate	High	Increased disease-burden category
Muara Satu	Moderate	Moderate	Consistently moderate

In contrast, Muara Dua demonstrated the most notable difference between disease types, changing from the moderate infectious disease cluster to the high non-communicable disease cluster. This distinction is important because it demonstrates that regional health burden cannot be represented adequately by a single overall disease classification. Different disease categories can produce different geographic clustering structures. Consequently, separating infectious and non-communicable diseases provides a more informative representation of local disease-distribution patterns than combining all conditions within a single clustering model.

D. Comparison of Fuzzy Membership Functions

After obtaining the FGK clustering structure, the performance of different fuzzy membership formulations was evaluated using the Xie-Beni Index. Four configurations were compared: the linear, original FGK, Gaussian, and sigmoid membership formulations. Lower Xie-Beni values indicate better cluster compactness and separation.

Table 11. Xie-Beni Index Comparison of Membership Functions

Membership Function	**Infectious Diseases**	**Non-Communicable Diseases**	**Average XB**	**Rank**
Linear	**0.1172**	**0.1459**	**0.1316**	**1**
Original FGK	0.1259	0.1672	0.1466	2
Gaussian	0.2777	0.3491	0.3134	3
Sigmoid	0.7986	0.6602	0.7294	4

As presented in Table 11, the linear membership function produced the lowest Xie-Beni values for both disease categories, with 0.1172 for infectious diseases and 0.1459 for non-communicable diseases. Its average Xie-Beni Index was 0.1316, making it the best-performing membership formulation among those evaluated. The original FGK formulation ranked second, with an average Xie-Beni value of 0.1466. Although its clustering quality was relatively close to the linear formulation, the consistently lower values produced by the linear function indicate better overall compactness and cluster separation for the dataset used in this study.

The Gaussian formulation produced substantially larger Xie-Beni values, resulting in an average of 0.3134. The sigmoid membership function yielded the weakest clustering performance, with an average value of 0.7294. Its Xie-Beni values were several times larger than those obtained using the linear and original FGK formulations. Therefore, the linear membership function was identified as the most appropriate formulation for the disease-distribution dataset, as it generated the lowest overall cluster-validity index. The original manuscript likewise reports an average Xie-Beni value of 0.1316 for the linear function, followed by original FGK, Gaussian, and sigmoid configurations.

E. Overall Clustering Findings

The results collectively demonstrate that FGK successfully differentiated disease-distribution patterns across the four sub-districts of Lhokseumawe City. The algorithm reached stable convergence for both disease categories, although the non-communicable disease dataset converged considerably faster than the infectious disease dataset. For infectious diseases, Banda Sakti was categorized as high, Blang Mangat as low, and Muara Dua and Muara Satu as moderate. For non-communicable diseases, Banda Sakti and Muara Dua were categorized as high, Blang Mangat as low, and Muara Satu as moderate.

The fuzzy membership values further demonstrate that most sub-districts exhibited strong associations with their dominant clusters. However, the partial memberships observed particularly for Muara Dua and Muara Satu in infectious disease clustering illustrate the advantage of fuzzy clustering in representing regional observations that lie between distinct disease-burden categories. Finally, comparison using the Xie-Beni Index demonstrated that the choice of membership formulation influenced cluster quality. Among the evaluated configurations, the linear membership function generated the best overall clustering structure with an average Xie-Beni Index of 0.1316.

F. Discussion

The findings demonstrate that the Fuzzy Gustafson-Kessel (FGK) method was able to distinguish meaningful regional patterns in the distribution of both infectious and non-communicable diseases across the four sub-districts of Lhokseumawe City. As presented in Tables 7 and 9, the resulting clusters showed relatively strong membership certainty, particularly for Banda Sakti and Blang Mangat. Banda Sakti consistently belonged to the high-burden cluster for both disease categories, whereas Blang Mangat consistently formed the low-burden cluster. Muara Satu maintained a moderate classification, while Muara Dua showed a notable shift from the moderate infectious-disease cluster to the high non-communicable-disease cluster. These patterns indicate that regional disease burden is heterogeneous and that different disease categories can generate different clustering structures even within the same administrative area.

The ability of FGK to represent these differences is closely related to its use of an adaptive covariance structure and Mahalanobis distance. Unlike clustering methods that apply the same geometric assumption to all groups, FGK allows individual clusters to adapt to variations in data orientation and covariance (Gustafson & Kessel, 1978). This characteristic is particularly relevant to regional health datasets because disease variables do not necessarily exhibit uniform distributions across geographical areas. Saraswati et al. (2024) similarly applied FGK to group districts and cities based on housing and environmental health indicators and demonstrated the applicability of adaptive fuzzy clustering to heterogeneous regional characteristics. The current results extend this application by showing that FGK can also differentiate local disease-distribution profiles involving both infectious and non-communicable conditions.

The findings are also consistent with the broader use of clustering for regional pattern identification. Hidayat et al. (2023) demonstrated that clustering can be used to categorize geographically distributed observations into meaningful regional groups, while Nurdin et al. (2024) showed that unsupervised clustering is useful for organizing regional datasets according to similarities among multiple variables. Although these studies addressed different application domains, their findings support the general value of clustering as a method for transforming multidimensional regional data into interpretable group structures. In the present study, this transformation was particularly useful because raw disease counts alone would not directly identify how the four sub-districts relate to one another in terms of their combined disease profiles.

The high cluster membership observed for Banda Sakti is consistent with the standardized input data presented in Tables 3 and 4. Banda Sakti recorded positive z-scores across all infectious and non-communicable disease variables, indicating consistently higher relative disease values than the other sub-districts. Conversely, Blang Mangat exhibited negative standardized values across the analyzed variables, corresponding with its assignment to the low cluster. These results indicate that the final FGK clusters were not arbitrary but reflected systematic differences already present in the multidimensional disease profiles. The final membership matrices reported in the analysis confirm these strong regional distinctions.

A particularly important finding concerns Muara Dua. For infectious diseases, Muara Dua was assigned to the moderate cluster with a membership degree of 0.9348, whereas for non-communicable diseases it shifted to the high cluster with a membership value of 0.9850. This difference demonstrates why infectious and non-communicable diseases should not necessarily be aggregated into a single general health-risk indicator. The relative burden of one disease category may differ substantially from another within the same geographic area. From an analytical perspective, separating the two categories therefore provides more detailed information for regional health planning than a single combined clustering result.

Muara Satu also provides an illustration of the value of fuzzy clustering. In the infectious-disease analysis, Muara Satu had a dominant membership value of 0.9436 in the moderate cluster but retained a small membership in the low cluster. Similar partial memberships were observed for Muara Dua. Such patterns are difficult to represent using hard clustering approaches, in which every observation must be assigned completely to one cluster. Fuzzy clustering instead preserves the degree of association between an observation and several clusters. This is relevant for health-related data because boundaries between low, moderate, and high disease burdens are unlikely to be entirely discrete.

This characteristic distinguishes FGK conceptually from conventional K-Means clustering. K-Means has been successfully applied to regional zoning and other data-grouping problems, including disaster-prone regional clustering (Hidayat et al., 2023) and fisheries-based regional grouping (Nurdin et al., 2024). However, K-Means produces hard partitions and relies on centroid-based distance structures. FGK provides an alternative when observations exhibit overlapping characteristics and when the geometric structures of clusters are not expected to be identical. The present study does not empirically claim that FGK outperforms K-Means because a direct algorithmic comparison was not performed; rather, the results demonstrate that FGK provides useful fuzzy membership information and stable regional partitions for the analyzed disease data.

The convergence behavior provides additional information regarding differences between the two disease datasets. As shown in Tables 6 and 8, infectious-disease clustering reached convergence after 24 iterations, whereas non-communicable-disease clustering converged after only nine iterations. The non-communicable disease solution also produced a lower final objective function. This suggests that the diabetes mellitus, hypertension, and stroke profiles formed a more clearly structured grouping across the four sub-districts than the infectious-disease variables. The manuscript results specifically show convergence at iteration 24 with a maximum membership change of 0.000009 for infectious diseases and iteration 9 with a change of 0.000006 for non-communicable diseases.

The difference in convergence should nevertheless be interpreted cautiously. A smaller number of iterations does not automatically indicate that one disease category is epidemiologically less complex than another. Rather, within the current dataset and selected parameter configuration, the non-communicable disease observations reached a stable fuzzy partition more rapidly. The result may be influenced by the relative separation of standardized observations, the small number of regional units, and the covariance structure of the variables. Consequently, convergence speed should be interpreted primarily as a computational characteristic of the observed dataset rather than as a direct epidemiological indicator.

Another central finding of the study is the effect of the membership formulation on cluster validity. As shown in Table 11, the linear membership function produced the lowest Xie-Beni Index for both infectious diseases (0.1172) and non-communicable diseases (0.1459), resulting in an overall average of 0.1316. The original FGK formulation ranked second with an average value of 0.1466, followed by Gaussian at 0.3134 and sigmoid at 0.7294. Because a lower Xie-Beni value represents greater within-cluster compactness relative to between-cluster separation, the linear formulation generated the most favorable clustering structure among the alternatives evaluated in this dataset. The comparative values and ranking are directly supported by the clustering-validation results reported in the manuscript.

This result reinforces the importance of cluster validation rather than if the default membership configuration will necessarily provide the best representation of a particular dataset. Putri and Rochmawati (2021) also demonstrated that FGK performance can be evaluated quantitatively rather than relying solely on visual or descriptive interpretation of clusters. Their application of FGK to social-media data differs substantially from the present health context, but it similarly illustrates the importance of evaluating the quality of fuzzy clustering outputs. In the present study, the Xie-Beni Index provides an objective criterion for determining which membership formulation generated the most compact and separated solution.

The results also complement the application of FGK to regional health-related information reported by Saraswati et al. (2024) and Dyfa et al. (2025). Saraswati et al. employed the method to group regional environmental health indicators, while Dyfa et al. applied FGK to health-worker ratios across districts and cities. Collectively, these applications indicate that FGK can accommodate different forms of geographically aggregated health data. The present study contributes to this line of research by applying FGK directly to disease-distribution variables and by separately modeling infectious and non-communicable disease profiles.

From an informatics perspective, an additional contribution lies in integrating computational clustering with a web-based information system. Data mining is intended not only to identify hidden structures in large or multidimensional datasets but also to transform those structures into information that can support interpretation and decision-making (Ardini, 2022; Gumilang & Ariyani, 2023). In this study, the FGK results were converted into regional disease-level categories and incorporated into a digital visualization environment. This approach provides a more interpretable representation than presenting raw numerical disease records alone. The use of Python and MySQL also enables the analytical procedure to be integrated with a structured data-processing workflow, consistent with contemporary information-system applications for health-related data management (Darnila et al., 2020; Tarina et al., 2025).

The practical implication of these findings is that regional health authorities can use clustering results as an exploratory analytical layer for prioritizing areas that warrant further attention. For example, the consistent classification of Banda Sakti within the high cluster indicates a relatively greater disease burden across the analyzed disease categories, while the transition of

Muara Dua from moderate infectious-disease burden to high non-communicable-disease burden suggests that different intervention priorities may be required for different regions. This type of differentiation can complement routine surveillance and provide a structured basis for allocating monitoring or health-promotion resources.

However, the clustering results should not be interpreted as direct measures of individual disease risk or causal relationships. The model identifies similarities in aggregated disease profiles rather than explaining why a particular sub-district experiences a higher or lower disease burden. Factors such as population size, age distribution, healthcare utilization, environmental exposure, socioeconomic conditions, sanitation, and access to health services were not incorporated into the clustering variables. Environmental health conditions are known to influence disease occurrence and community well-being (Chaniago et al., 2023), while broader behavioral factors are relevant to the development of non-communicable diseases (Puspa Sari et al., 2024). Including such variables in future analyses could provide a more comprehensive representation of regional health conditions.

Another limitation concerns the relatively small number of spatial units. Although the dataset included 38,063 individual disease cases, the clustering analysis was performed on data aggregated into only four sub-districts. Therefore, the effective number of regional observations used by the FGK model was limited. This should be considered when interpreting the high membership certainty observed in several clusters. The current results are appropriate for exploratory regional grouping but should not be generalized beyond Lhokseumawe City without further validation using larger numbers of spatial units.

Future studies could consequently extend the framework to village or *gampong* level, incorporate demographic and environmental attributes, and evaluate the stability of clustering over multiple years. Direct comparisons between FGK and alternative algorithms such as K-Means, Fuzzy C-Means, or other adaptive fuzzy clustering methods would also strengthen evidence regarding algorithm selection. Such comparisons are particularly relevant because previous regional studies have demonstrated the usefulness of K-Means (Hidayat et al., 2023; Nurdin et al., 2024), while FGK applications have shown potential for datasets characterized by heterogeneous covariance structures (Gustafson & Kessel, 1978; Saraswati et al., 2024).

Overall, the results demonstrate that FGK provides a meaningful computational approach for representing heterogeneous disease-distribution patterns in Lhokseumawe City. Its fuzzy membership structure captures both clear and transitional regional characteristics, while the Xie-Beni analysis demonstrates that clustering quality is sensitive to the selected membership formulation. The linear membership function produced the most favorable validity score in the present dataset, and the distinction between infectious and non-communicable disease clusters highlights the importance of disease-specific regional analysis. By combining adaptive fuzzy clustering, quantitative cluster validation, and web-based visualization, the study provides an informatics-oriented framework for transforming routine disease records into more interpretable information for regional health monitoring and planning.

V. CONCLUSION AND RECOMMENDATIONS

Based on the conclusions drawn from the research that has been carried out, the System is a System. The spread of infectious and non-infectious diseases in Lhokseumawe City has been successfully designed, built, and implemented in the form of a web-based application using Python and MySQL, which is equipped with visualization features. Heatmap and mapping of clustering results. Application of the Fuzzy Gustafson-Kessel(FGK) in grouping areas based on disease data has proven effective by producing consistent clustering patterns in both disease categories. In infectious diseases, FGK successfully grouped 4 sub-districts into 3clusterwith convergence at the 24th iteration.Cluster0 (low) contains Blang Mangat District,cluster1 (medium) contains the Muara Dua and Muara Satu Districts, andcluster2 (high) contains Banda Sakti District. For non-communicable diseases, convergence was achieved more quickly in the 9th iteration. Cluster0 (low) contains Blang Mangat District,cluster1 (medium) contains Muara Satu District, andcluster2 (high) contains the Banda Sakti and Muara Dua Districts. Comparison of the three membership functions shows that the linear function is the most optimal, with the lowest average Xie-Beni Index at 0.1316, followed by the functionsigmoid And gaussian. Mark Xie-Beni Index indicates better quality, so the linear function is recommended as the best membership function in the application of the FGK method to disease data in Lhokseumawe City.

REFERENCES

Almeida, S., & Chen, M. (2025). Adaptive Control of Autonomous Mobile Robots Using Fuzzy Logic Based PID Optimization. *Journal of Technology Informatics and Engineering*, *4*(3), 403–416. https://doi.org/10.51903/JTIE.V4I3.454

Andika, F., Safira, A., & Mustina, N. (2020). Education on communicable disease eradication among students at SMA Negeri 5 Banda Aceh. Jurnal Pengabdian Masyarakat (Kesehatan), 2(1). https://www.tbindonesia.or.id

Afdal, M., & Generis Humani, D. (2020). Android-based expert system application for early diagnosis of infectious diseases in toddlers. Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi, 6(1), 55–63. https://doi.org/10.24014/rmsi.v6i1.8979

Ardini. (2022). Grouping of Indonesian archipelago provinces based on economic characteristics and fisheries potential in 2020 using the fuzzy C-means method. Seminar Nasional Official Statistics. https://doi.org/10.34123/semnasoffstat.v2022i1.1335

Chaniago, E., Lubis, A., & Ani, N. (2023). Penyuluhan Menciptakan Lingkungan Hidup yang Bersih dan Sehat di Kabupaten Deli Serdang Kecamatan Batang Kuis Desa Bakaran Batu. Jurnal Deputi: Derma Pengabdian Dosen Peguruan Tinggi, 3(1), 153-156. https://doi.org/10.54123/deputi.v3i1.234

Darnila, E., Al-Kautsar, H., & Iksan, Y. (2020). Analysis of patient disease trends based on medical record data using the C4.5 algorithm. Login: Jurnal Teknologi Komputer, 14(1), 7–12. https://www.neliti.com/publications/326752/analysis-of-patient-disease-trends-based-on-medical-record-data-using-the-c4-5-a

Dyfa, M. A., Haris, A. M., & Imron, A. (2025). Grouping of districts/cities in Central Java based on the ratio of health workers using the fuzzy Gustafson-Kessel method. Prosiding SENADA. https://prosiding-senada.upnjatim.ac.id/index.php/senada/article/view/488

Fahmi, Y., & Sutaji, D. (2025). Design of data flow diagram (DFD) for WhatsApp-based public service system. Journal of Informatics and Applied Electrical Engineering, 13(3S1). https://doi.org/10.23960/jitet.v13i3S1.8083

Fristiohady, A., Trinovitasari, N., Mahmudah, A., & Putri, I. P. (2023). Socialization paradigm healthy living as a promotional and preventive concept of infectious diseases in an effort to support SDGs at State High School 2 Kendari. MJPF. https://jpfi.uho.ac.id/index.php/journal/index

Gumilang, T. A., & Farida Ariyani, P. (2023). Algoritma Apriori Dalam Implementasi Data Mining Untuk Pembuatan Paket Penjualan Di Mesha Petshop. JTRISTE, 12(1), 49-63. https://jurnal.kharisma.ac.id/jtriste/article/view/572

Gustafson D & Kessel W. (1978). "Fuzzy Clustering With A Fuzzy Covariance Matrix," Conference Paper In Proceeding Of The Ieee Conference On Decision And Control. https://doi.org/10.1109/cdc.1978.268028

Hamzah, A., Ramadhan, A. P., Witama, F. Y., & Rosyani, P. (2023). Analysis of Python application in inequality calculation.

Harahap, L. M., Fuadi, W., Rosnita, L., Darnila, E., & Meiyanti, R. (2022). Clustering of superior vegetables using the K-means algorithm. Journal of Informatics Engineering and Information Systems, 8(3). https://doi.org/10.28932/jutisi.v8i3.5277

Hidayat, I., Darnila, E., & Afrillia, Y. (2023). Clustering Zonasi Daerah Rawan Bencana Alam Di Kabupaten Mandailing Natal Menggunakan Algoritma K-Means. G-Tech: Jurnal Teknologi Terapan, 7(3), 1218–1226. https://doi.org/10.33379/gtech.v7i3.2880

Kuznetsova, E., & Nkosi, M. (2025). Adaptive Fuzzy Logic Integration for Optimizing Decision Support Systems under Data Uncertainty. *Journal of Technology Informatics and Engineering*, *4*(3), 463–477. https://doi.org/10.51903/JTIE.V4I3.456

Napitupulu, R. J., Silvia, H., Rhamadan, A. I., & Niqotaini, Z. (2025). Analysis of UPN Veteran Jakarta Library system: Case study of discussion room borrowing using UML. Information Systems Journal (TEKNOFILE), 3, 379–396.

Noviana, R. (N.D.). Pembuatan Aplikasi Penjualan Berbasis Web Monja Store Menggunakan Php Dan Mysql. Jts, 1(2). https://doi.org/10.56127/jts.v1i2.128

Nurdin., Bustami.,Meiyanti, R., & Fahada, A. (2024). Clustering Types Of Capture Fisheries Products Using The K-Means Clustering Algorithm. Journal Of Theoretical And Applied Information Technology, 15(17). https://www.jatit.org/volumes/vol102no17/13vol102no17

Nurdin.,Pradita, C,C.,Fadlisyah. (2023). Sisfo : Jurnal Ilmiah Sistem Informasi Implementasi Data Mining Untuk Menganalisis Kategori Kompetisi Mahasiswa Menggunakan Algoritma Apriori. Sisfo : Jurnal Ilmiah Sistem Informasi, 7(1). https://doi.org/10.29103/sisfo.v7i1.12104

Prayoga, A., Fadlisyah, F., & Aidilof, H. A. K. (2025). Geographic information system for mapping the location of mosques and meunasah in Lhokseumawe City using website-based Haversine algorithm. Rabit: Journal of Technology and Information Systems Univrab, 10(2), 815–825. https://doi.org/10.36341/rabit.v10i2.6381

Puspa Sari, M., Kholika Hamal, D., Annisa Mulyawati, D., Ummu Unais, J., Pratiwi, T., Kesehatan Masyarakat, J., Ilmu-Ilmu Kesehatan, F., Muhammadiyah, U., & Artikel, R. (2024). Sosialisasi Pencegahan Dan Pengendalian Penyakit Tidak Menular. Jurnal Pengabdian Kesehatan Masyarakat, 4(1), 1–7. http://jurnal.uinsu.ac.id/index.php/shihatuna/index

Putri, F. A. A., & Rochmawati, N. (2021). Penerapan algoritma Fuzzy Gustafson-Kessel untuk clustering tweets mention akun Go-Jek dan Grab Indonesia [Application of the Fuzzy Gustafson–Kessel algorithm for clustering tweets mentioning Go-Jek and Grab Indonesia accounts]**.** Journal of Informatics and Computer Science, 2(3), 204–212**.** https://doi.org/10.26740/jinacs.v2n03.p204-212

R. Babuka, P. Van Der Veen, And U. Kaymak. (2002). "Improved Covariance Estimation For Gustafson-Kessel Clustering," In Fuzzy Systems. Fuzz-Ieee'02. Proceedings Of The 2002 Ieee International Conference On, Vol. 2, Pp. 1081 –1085. http://dx.doi.org/10.1109/fuzz.2002.1006654

Rahma, S. N., et al. (n.d.). Achieving TB elimination with the Vibrant Indonesia End Tuberculosis (GIAT) movement. Kementerian Kesehatan Republik Indonesia. https://ayosehat.kemkes.go.id/capai-eliminasi-tbc-dengan-semarak-gerakan-indonesia-akhiri-tuberkulosis-giat

Rizky, M., Fuadi, W., & Maryana. (2025). Application of profile matching method for recommendation of field work practice placement for students. Rabit: Journal of Technology and Information Systems Univrab, 10(2), 752–762. https://doi.org/10.36341/rabit.v10i2.6366

Saraswati, G. S., Nohe, D. A., & Hayati, M. N. (2024). Grouping of districts/cities in Kalimantan Island based on housing and environmental health indicators using the fuzzy Gustafson-Kessel method. Gaussian Journal, 13(1), 156–167. https://doi.org/10.14710/j.gauss.13.1.156-167

Tarina, R., Safwandi, S., & Afrillia, Y. (2025). Implementation of decision support system in blood bag distribution in UDD PMI City of Lhokseumawe. Rabit: Journal of

Technology and Information Systems Univrab, 10(2), 1277–1291. https://doi.org/10.36341/rabit.v10i2.6520