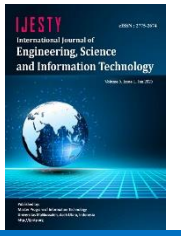
International Journal of Engineering, Science and Information Technology
Volume 5 No. 1 (2025) pp. 350-362
ISSN 2775-2674 (online)
Website: http://ijesty.org/index.php/ijesty
DOI: https://doi.org/10.52088/ijesty.v5i1.747
Research Paper, Short Communication, Review, Technical Paper



Application of Fuzzy C-Means and Borda in Clustering Crime–Prone Areas and Predicting Crime Rates Using Long Short-Term Memory in Northern Aceh Regency

Syahrul Andika Lubis*, Munirul Ula, Sujacka Retno

Department of Informatics, Faculty of Engineering, Universitas Malikussaleh, Aceh, Indonesia

**Corresponding author Email: syahrul.200170129@mhs.unimal.ac.id*

The manuscript was received on 10 May 2024, revised on 18 August 2024, and accepted on 3 January 2025, date of publication 16 January 2025

Abstract

North Aceh is a district with diverse geographical conditions, ranging from vast lowland areas in the north stretching from west to east to mountainous regions in the south. The average altitude in North Aceh is 125 meters. The district covers an area of 2,694.66 km² with a population of 614,640 people in 2022. The issue of crime in North Aceh District has caused significant discomfort among the community. According to data from the Central Bureau of Statistics (BPS) of Aceh Province, the number of criminal cases increased from 6,651 in 2022 to 10,137 in 2023. Using the Fuzzy C-Means clustering method, the data was grouped into three clusters: cluster 1 represents safe areas, cluster 2 represents moderately vulnerable areas, and cluster 3 represents susceptible areas. For ranking using the Borda method, the Dewantara Police Sector ranked first for the physical aspect, while the Muara Batu Police Sector ranked first for the item aspect. As for predictions using the LSTM model, almost all subdistricts achieved MAPE values below 20%, indicating that the LSTM model is quite effective in predicting crime-prone areas. For example, Baktiya District recorded a MAPE value of 15.85% for the physical aspect, while Simpang Keramat District achieved the best result for the item aspect with a MAPE value of 0.00%. However, in Syamtalira Bayu District, the item aspect reached a MAPE value of 20.07%. Although the MAPE value for the item aspect in Syamtalira Bayu is relatively high, it is still considered acceptable as it remains below 50%.

Keywords: *North Aceh Crime, Fuzzy C-Means, Borda, Long Short-Term Memory, MAPE.*

1. Introduction

North Aceh is a district with various geographical conditions, ranging from vast lowland areas in the north that stretch from west to east to mountainous regions in the south. The average altitude in North Aceh is 125 meters. North Aceh District has an area of 2,694.66 km² with a population of 614,640 people in 2022. The issue of crime in North Aceh Regency is also enough to make people uneasy. According to the Central Statistics Agency (BPS) of Aceh Province, there was an increase in the number of criminal cases from 2022, which amounted to 6,651 cases, to 10,137 cases in 2023.

Crime or criminality refers to all actions and behaviors that cause economic and psychological harm contrary to the applicable laws in Indonesia and social and religious norms. There are several types of crimes or criminals, such as crimes against life, physical crimes, crimes against decency, etc.

The notion of criminality can be understood through the evolution of criminal concepts that are still developing. In the legal system, criminality is seen as an act contrary to the law's provisions and is threatened with criminal sanctions for violators. Actions that are not legally regulated to be prohibited cannot be considered crimes, although society may consider them as inappropriate behavior [1].

A Geographic Information System (GIS) is a computer-based information system designed to work with data containing location-related information. Geographic Information Systems function to collect, analyze, examine, integrate, manage, process, and display spatial data related to the shape or condition of the earth's surface. Considering these main elements, it is clear that Geographic Information Systems are a type of information system that focuses on “Geographic Information.” A Geographic Information System is software for input, storage, manipulation, visualization, and output of geographic data and its associated attributes [2].

2. Literature Review

2.1. Previous Research

According to the research by Dinata et al., (2022), "Clustering the Spread of ISPA Disease Using the Fuzzy C-Means Algorithm in Aceh Utara," This research classifies the data into 3 Cluster 1 is the cluster with the lowest level or the least spread of ARI disease. Cluster 2 is the medium level, and Cluster 3 is the cluster with the highest level. Based on the data obtained in this study, the most common type of ARI disease suffered by residents of North Aceh is pneumonia, with 414 cases or 64% of the total data[3].

According to the research by Ula et al. (2022), "Fuzzy C-Means with Borda Algorithm in Cluster Determination System for Food Prone Areas in Aceh Utara," This study concludes that the application of the Fuzzy C-Means method successfully classifies food insecure areas based on six sub-criteria into three clusters: moderately prone, prone, and highly prone. Validation with results close to 1 indicates a high level of accuracy of the Fuzzy C-Means method. Through voting using the Borda algorithm on highly vulnerable clusters, the Sawang sub-district was selected for food availability, the Syamtalira Aron sub-district for food affordability, and the Lapang sub-district for food utilization[4].

According to the research by Pardosi et al. (2020), "Clustering Job Vacancy Data Based on Fuzzy C-Means" Research Results: Based on the discussion in the previous chapters, several things can be concluded in this study:

1. Fuzzy C-Means can provide accurate job recommendations with a low error rate.
2. From the test results, it can be concluded that the accuracy of the recommendation results will increase as the number of applications and interest received for each job increases.

The results of testing the difference in click-to-rate (CTR) predictions show that CTR predictions for most jobs have satisfactory accuracy [5].

According to the research by Rahayu et al. (2023), "Impelementation of Longs Short Term Memory (LSTM) Algorithm for Predicting Stock Price Movements of LQ45 Index (Case Study: BBCA Stock Price)," The implementation of the LSTM method in this study resulted in several significant findings. First, using 14 data points from January 1, 2020, to January 14, 2020, and historical BBCA stock data, training data was taken from January 1, 2020, to January 13, 2020, including Open, Close, High, and Low data. During the process, weights and biases are used with the values W_f=[0.2, 0.3, 0.4, 0.1] and b_f=0.5 on the normalized data. After going through stages such as input gate, dot product, cell gate, output gate, hidden gate, and sigmoid activation, the RMSE value is obtained at 6.658%. Comparison between predictions and actual data results in a MAPE value of 5.4772%. In addition, the comparison between test data and prediction results shows a difference of -27.4242%, where the test data on January 14, 2020, is 0.824, while the prediction shows a value of 0.598. Furthermore, in the implementation of the developed application using the latest historical data from May 25, 2023, to June 23, 2023, the MAPE results for the closing price were 4.8747%, the opening price was 4.3686%, the highest cost was 5.8048%, and the lowest price was 3.7763%. LSTM predictions show an error rate below 6% for BBCA stock price predictions[6].

2.2. Data Mining

Data mining is extracting information from large data sets using various techniques and algorithms. The goal is to gain valuable insights from the data. Large companies often use data mining to optimize their business operations by leveraging information discovered from data analysis[7].

The stages of data mining are as follows:

1. Data Selection
   Data selection from a set of operational data needs to be done. The data selected for the data mining process is then stored in a separate file from the operational database.
2. Preprocessing/cleaning
   Before the data mining process is carried out, it is necessary to clean the selected data. The cleaning process includes removing duplicate data and correcting errors in the data.
3. Transformation
   Coding is a transformation step on the data that has been selected so that the data can be used effectively in the data mining process.
4. Data Mining
   Data mining is finding interesting patterns or information in data that has been selected using various techniques or methods.
5. Evaluation/Interpretation
   The results of information patterns generated from the data mining process need to be presented in a format that related parties can easily understand. This interpretation stage involves evaluating whether the patterns or information found are by pre-existing facts or hypotheses

2.3. Clustering

In the Big Indonesian Dictionary (KBBI), grouping is the process or act of grouping or categorizing something based on specific traits or characteristics. It involves organizing or classifying objects, ideas, or individuals according to the properties or attributes possessed by each. This categorization helps understand, analyze, and communicate about the subject or object under study.

Clustering is the process of grouping similar data into different groups or dividing a data set into subsets so that the data in each subgroup has valuable meaning. In this clustering, each cluster consists of objects that are similar to each other and different from objects in other clusters [8].

2.4. Fuzzy C-Means

Fuzzy C-Means is a data clustering method that uses membership degrees to determine the cluster of each data. Determination of data membership in a cluster in Fuzzy C-Means depends on its membership degree. Fuzzy C-Means is a supervised data clustering algorithm where the number of clusters is determined based on the desired output [9]. The steps to complete Fuzzy C-Means are as follows:

1. Input data is to be clustered X in the form of an n x m matrix (n = number of data samples, m = input variables for each data). Xij = ith sample data (i = 1,2,3,...,n), jth input variable (j = 1,2,...,m).

2. Determine the number of clusters (c), rank (w), Maximum iterations (maxIter), Least expected error (ε), Initial objective function (P0 = 0), and Initial iterations (t = 1).
3. Generate random numbers μik, as elements of the initial partition matrix U with I = 1,2,...,n and k =1,2,3,...,m. Calculate the degree of membership:

$$\mu ik = \frac{\mu ik}{Qi} \quad (1)$$

4. Calculate the kth cluster center: Vkj with k=1,2,...,c and j=1,2,...,m, where Xij is the fuzzy variable used and w is the weight.

$$V_{kj} = \frac{\sum_{i=1}^{n}((\mu ik)^{w}) * X_{ij}}{\sum_{i=1}^{n}((\mu ik)^{w})} \quad (2)$$

5. Calculate the objective function at the tth iteration:

$$P_t = \sum_{i=1}^{n}\sum_{k=1}^{c}\left[\sum_{j=1}^{m}(X_{ij} - X_{kj})^2\right](\mu ik)^{w} \quad (3)$$

6. Calculate the change of the partition matrix:

$$\mu ik = \frac{\left[\sum_{j=1}^{m}(X_{ij} - X_{kj})^2\right]^{\frac{-1}{w-1}}}{\sum_{k=1}^{c}\left[\sum_{j=1}^{m}(X_{ij} - X_{kj})^2\right]^{\frac{-1}{w-1}}} \quad (4)$$

7. Stop condition if : (Pt - Pt-1 < ε) or (t > MaxIter) then stop. If: t = t+1, repeat step - 4 (equation 2)

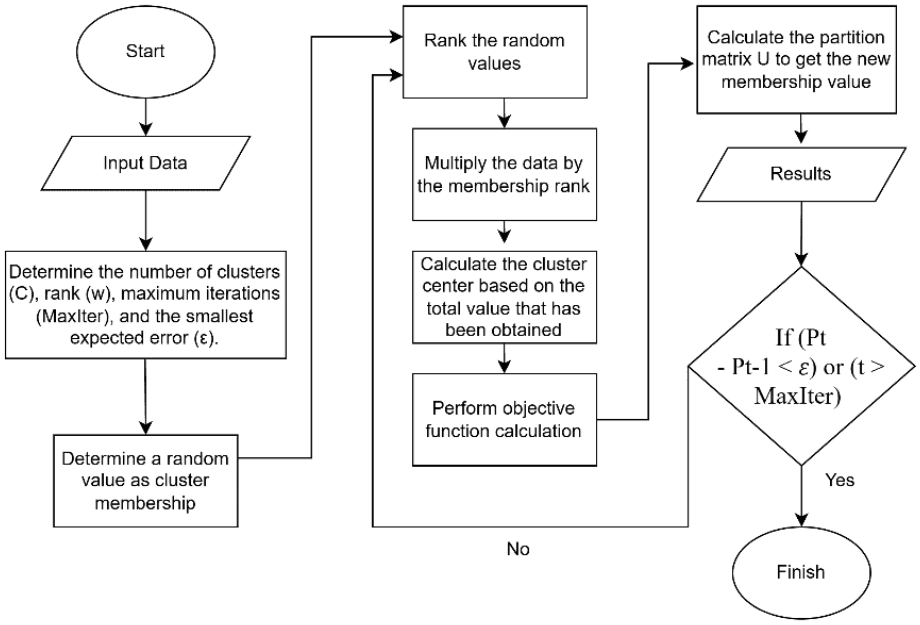


Fig 1. Flowchart Fuzzy C-Means

2.5. Borda

Jean-Charles de Borda invented the Borda method in the 18th century. It is a technique used to rank decisions based on preferences. The Borda method is used in group decision-making to rank candidates based on the preferences of each decision-maker [10].

The Borda method is a ranking that incorporates preferences by assigning weights to each option. This technique is often used in various situations, such as elections, product evaluation, and decision analysis. The Borda method assesses each alternative by assigning points based on its ranking in each criterion or dimension evaluated. The higher the rank of an alternative in a criterion, the more points are given to it. The advantage of the Borda method is that it is simple and easy to understand. There are also steps in calculating the Borda method, which are as follows:

1. Determine the values and criteria for each alternative
2. Determine the number of points for each alternative
3. Sorting the number of points from the most significant (ranking)
4. Selecting the alternative.

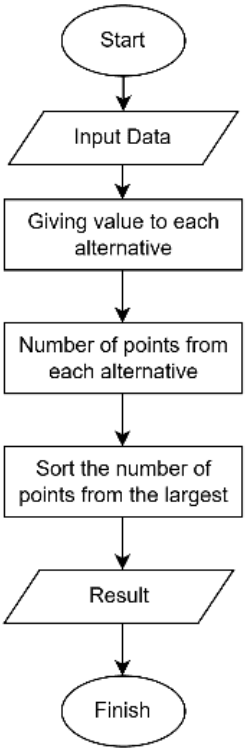


Fig 2. Flowchart Borda

2.6. Long Short Term Memory

Long Short-Term Memory (LSTM) is an artificial neural network used in machine learning to process and predict sequential data with long temporal dependencies. Long Short-Term Memory is a variant of Recurrent Neural Network (RNN).

Long Short-Term Memory (LSTM) is an algorithm renowned for its strong ability to build prediction models. LSTM is an extension of Recurrent Neural Network (RNN), a method designed to process sequential data. RNN has problems with vanishing and exploding gradients; LSTM is specifically designed to overcome these problems. The LSTM architecture consists of three main layers: the input, hidden, and output layers. In the hidden layer, there is a memory cell composed of three gates: input gate, forget gate, and output gate. [11].

1. Forget Gate

   Forget gate is a sigmoid layer that receives input in the form of output from time t-1 and input at time t, then combines them and applies a sigmoid activation function. Because it uses a sigmoid function, the output of this gate is between 0 and 1. If ft = 0, the previous state will be forgotten; if ft = 1, the previous state will be retained. The formula for ft is:

   $f_t = \sigma(W_{xf} \cdot xt + W_{hf} \cdot h_{t-1} + b_f)$ ..... (5)

   Description:

ft	: forgate gate activation at time t.
Wxf and Whf	: weight matrix for input.
xt	: input at time t.
ht-1	: hidden state value at time t-1.
bf	: bias forgate state.
σ	: sigmoid activation

2. Input Gate

   The input gate functions to receive the output of the previous step and the new input, forwarding it through the sigmoid layer. This gate returns a value of 0 or 1. The input gate formula is as follows:

   $it = \sigma(W_{xi} \cdot xt + W_{hi} \cdot h_{t-1} + b_i)$ ..... (6)

   Description:

it	: input gate activation at time t.
Wxi and Whi	: weight matrix for input.
ht-1	: hidden state value at time t-1.
Xt	: input at time t.
bi	: input gate bias.
σ	: sigmoid activation function

3. Candidate Score

   $\tilde{C}_t = \tanh(W_{xc} \cdot X_t + W_{hc} \cdot h_{t-1} + b_c)$ ..... (7)

   Description:

Ct	: candidate value at time t.

Wxc and Whc : weight matrix for input.
Xt : input at time t.
ht-1 : hidden state value at time t-1.
bc : candidate value bias.
tanh : tanh activation function.

4. Update Cell State

$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t$ ..... (8)

Description:
Ct: cell state at time t.
ft : forgate gate activation at time t.
Ct-1 : cell state at time t-1.
C̃t : candidate value at time t.

5. Output Gate

The output gate regulates how much information from the cell state is passed to the output and works together with other gates to produce a new cell state (ht). The steps that need to be taken to calculate the output gate value include calculating the input gate, forget gate, and cell gate values. Here is the formula for the output gate:

$o_t = \sigma(W_{xo} \cdot xt + W_{ho} \cdot h_{t-1} + b_o)$ ..... (9)

Description:
Ot : gate output activation at time t.
Wxo and Who : weight matrix for input.
xt : input at time t.
ht-1 : hidden state value at time t.
bo : let the gate output.
σ : sigmoid activation function.

6. Hidden State

Hidden state is the representation of information carried from one-time step to the next in LSTM. The formula of the hidden state is as follows:

$ht = Ot \cdot \tanh(Ct)$ ..... (10)

Description:
Ot : gate output at time t
Ct : cell state at time t
tanh : tanh activation function

2.7. Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) is the average value of the absolute percentage error. MAPE measures the relative error rate between predicted and actual values, expressed as a percentage (Agrippina & Pamuji, 2024). The formula for MAPE is as follows:

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \frac{|y - yi|}{y} \times 100\% \quad (11)$$

Description:
y : Actual data
yi : Forecasting data in period i
n : Number of data

The value of the MAPE percentage is considered good if the result is less than 20%, as explained in the following range table:

Table 1. MAPE Range

MAPE Range	Description
<10%	Very Accurate Accuracy Level
10 – 20%	Good Accuracy Level
20 – 50%	Fair/Acceptable Accuracy Level
>50%	Inaccurate Accuracy Level

3. Research Methods

The research conducted by the author focuses on the stages involved in classifying, ranking, and predicting crime in North Aceh. The author uses the research stages as guidelines to ensure that the research results remain following the predetermined objectives and do not deviate from the desired goals.

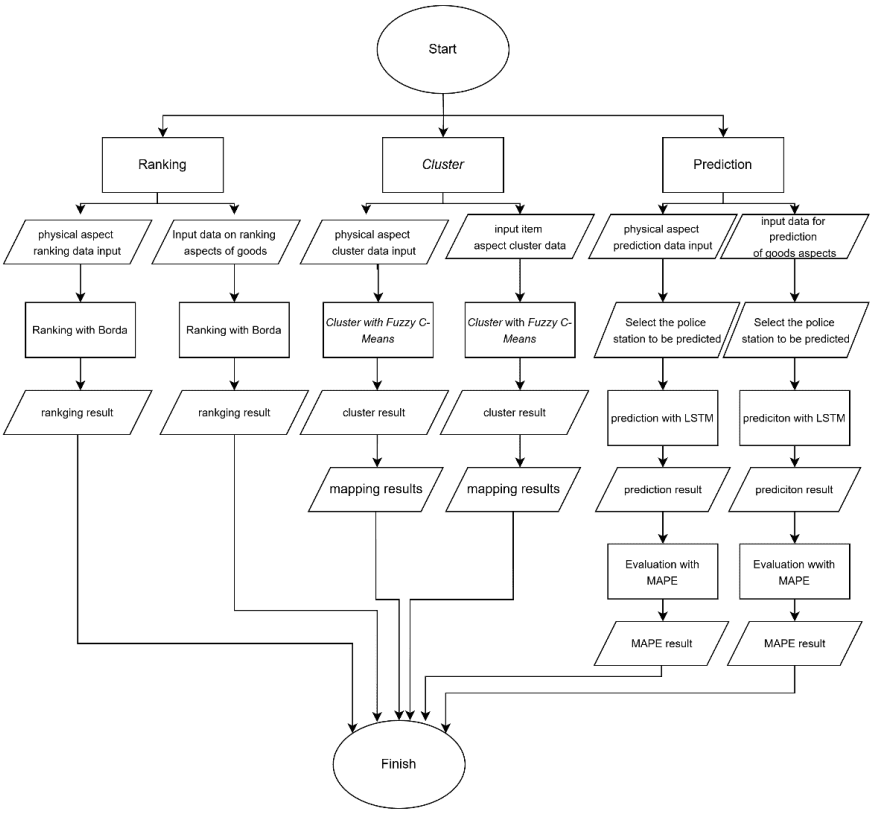


Fig 3. Flowchart System

4. Result and Discussions

4.1. Dataset

The dataset used for clusters and predictions is data on physical aspects of crime and criminality aspects of goods for 2020-2024 in North Aceh Regency obtained from North Aceh Police and Lhokseumawe Police. Meanwhile, data for ranking is used from questionnaires distributed to the people of North Aceh.

Table 2. Physical Aspect Cluster Data

No	Polsek Name	2020	2021	2022	2023	2024	Number of Villages
1	Baktiya	0	7	5	6	5	57
2	Baktiya Barat	0	1	0	2	0	26
3	Cot Girek	0	1	0	2	0	24
4	Langkahan	0	3	2	1	0	23
…	…	…	…	…	…	…	…
19	Sawang	7	7	2	3	4	39
20	Simpang Keramat	1	2	0	0	0	16
21	Syamtalira Bayu	3	2	8	4	0	49
22	Muara Batu	3	3	0	1	1	24

Table 3. Goods Aspect Cluster Data

No	Polsek Name	2020	2021	2022	2023	2024	Number of Villages
1	Baktiya	5	6	9	6	4	57
2	Baktiya Barat	0	1	0	0	0	26
3	Cot Girek	0	6	5	11	11	24
4	Langkahan	3	6	3	2	3	23
…	…	…	…	…	…	…	…

19	Sawang	2	0	1	2	2	39
20	Simpang Keramat	3	0	0	0	0	16
21	Syamtalira Bayu	3	7	5	4	5	49
22	Muara Batu	12	11	16	14	4	24

Table 4. Physical Aspect Ranking Data

No.	Name or Initials	Ranking of Police Stations					
		Baktiya	**Baktiya Barat**	**Cot Girek**	**Langkahan**	**Lhoksukon**	**Matang Kuli**
1	Icut	14	18	19	13	15	12
2	Via	2	3	4	17	5	22
3	G	19	17	18	15	11	20
4	Asmaul Husna	8	2	19	9	11	7
...	…		...			...	...
47	Cut zahra	8	15	7	5	4	16
48	Azzahra Humaira	11	13	10	12	6	5
49	Lindawati	8	11	7	20	9	19
50	M Syuhada	15	7	14	8	13	6

Table 4. Physical Aspect Ranking Data (Continued)

No.	Name or Initials	Ranking of Police Stations					
		Nibong	**Paya Bakong**	**Seunuddon**	**Syamtalira Aron**	**Tanah Luas**	**Dewantara**
1	Icut	16	5	4	6	17	17
2	Via	6	12	11	7	21	21
3	G	21	22	14	13	3	3
4	Asmaul Husna	18	10	3	12	1	1
...	…		...			...	...
47	Cut zahra	9	17	3	19	22	22
48	Azzahra Humaira	14	16	2	4	9	9
49	Lindawati	10	18	22	4	5	5
50	M Syuhada	22	9	5	10	17	17

Table 4. Physical Aspect Ranking Data (Continued)

No.	Name or Initials	Ranking of Police Stations					
		Kuta Makmur	**Meurah Mulia**	**Nisam**	**Samudera**	**Sawang**	**Simpang Keramat**
1	Icut	8	7	2	9	3	22
2	Via	13	14	8	19	15	16
3	G	9	8	12	7	6	5
4	Asmaul Husna	14	16	5	15	17	20
...	…		...			...	...
47	Cut zahra	18	20	2	12	14	21
48	Azzahra Humaira	17	21	18	7	8	20
49	Lindawati	12	21	16	2	13	15
50	M Syuhada	12	18	11	19	3	21

Table 4. Physical Aspect Ranking Data (Continued)

No.	Name or Initials	Ranking of Police Stations	
		Syamtalira Bayu	Muara Batu
1	Icut	10	11
2	Via	18	9
3	G	4	2
4	Asmaul Husna	21	22
...	…	...	...
47	Cut zahra	13	6
48	Azzahra Humaira	19	1
49	Lindawati	14	1
50	M Syuhada	20	2

Table 4. Item Aspect Ranking Data

No.	Name or Initials	Ranking of Police Stations					
		Baktiya	Baktiya Barat	Cot Girek	Langkahan	Lhoksukon	Matang Kuli
1	Icut	10	14	11	7	6	21
2	Via	16	22	14	21	2	13
3	G	6	22	7	9	10	13
4	Asmaul Husna	19	18	22	20	14	6
...	…		...			...	...
47	Cut zahra	11	13	8	12	3	7
48	Azzahra Humaira	10	8	20	9	13	19
49	Lindawati	11	10	12	6	9	5
50	M Syuhada	11	12	21	5	6	14

Table 5. Item Aspect Ranking Data

No.	Name or Initials	Ranking of Police Stations					
		Nibong	Paya Bakong	Seunuddon	Syamtalira Aron	Tanah Luas	Dewantara
1	Icut	12	8	5	19	4	9
2	Via	15	4	5	17	6	7
3	G	20	19	12	8	14	4
4	Asmaul Husna	13	17	21	5	7	12
...	…		...			...	...
47	Cut zahra	21	14	6	15	16	18
48	Azzahra Humaira	12	11	4	16	3	7
49	Lindawati	19	18	7	4	13	21
50	M Syuhada	16	13	15	9	4	17

Table 5. Item Aspect Ranking Data (Continued)

No.	Name or Initials	Ranking of Police Stations					
		Kuta Makmur	Meurah Mulia	Nisam	Samudera	Sawang	Simpang Keramat
1	Icut	22	15	3	17	18	20
2	Via	3	19	10	1	8	20
3	G	18	11	21	4	16	15
4	Asmaul Husna	8	3	1	9	2	10
...	…		...			...	...
47	Cut zahra	17	19	4	2	9	20

48	Azzahra Humaira	15	21	22	6	5	14
49	Lindawati	17	15	14	2	20	22
50	M Syuhada	18	19	8	10	2	20

Table 5. Item Aspect Ranking Data (Continued)

No.	Name or Initials	Ranking of Police Stations	
		Syamtalira Bayu	Muara Batu
1	Icut	16	1
2	Via	9	11
3	G	5	2
4	Asmaul Husna	16	11
...	…	...	...
47	Cut zahra	10	1
48	Azzahra Humaira	17	2
49	Lindawati	16	1
50	M Syuhada	22	3

Table 6. Physical Aspect Prediction Data

No	Police Stations	Month-Year	Total
1	Baktiya	43831	0
2	Baktiya	43862	0
3	Baktiya	43891	0
4	Baktiya	43922	0
...	...	...	...
1317	Tanah Pasir	Sep-24	0
1318	Tanah Pasir	Oct-24	0
1319	Tanah Pasir	Nov-24	0
1320	Tanah Pasir	Dec-24	0

Table 7. Items Aspect Prediction Data

No	Police Stations	Month-Year	Total
1	Baktiya	Jan-20	0
2	Baktiya	Feb-20	1
3	Baktiya	Mar-20	0
4	Baktiya	Apr-20	0
...	…	...	...
1317	Tanah Pasir	Sep-24	0
1318	Tanah Pasir	Oct-24	0
1319	Tanah Pasir	Nov-24	1
1320	Tanah Pasir	Dec-24	0

4.2. Clustering Result

In this study, the authors used the Fuzzy C-Means method to group crime-prone areas in North Aceh Regency. The variables X1 = Total cases in 2020, X2 = Total cases in 2021, X3 = Total cases in 2022, X4 = Total cases in 2023, X5 = Total cases in 2024, and X6 = number of villages in the jurisdiction of each police station. The following are the final results of the Fuzzy C-Means calculation.

Table 8. Physical Aspect Cluster Results

No	Police Stations	Cluster
1	Baktiya	1
2	Baktiya Barat	3
3	Cot Girek	3
4	Langkahan	3
...	…	…
19	Sawang	2
20	Simpang Keramat	3
21	Syamtalira Bayu	2
22	Muara Batu	3

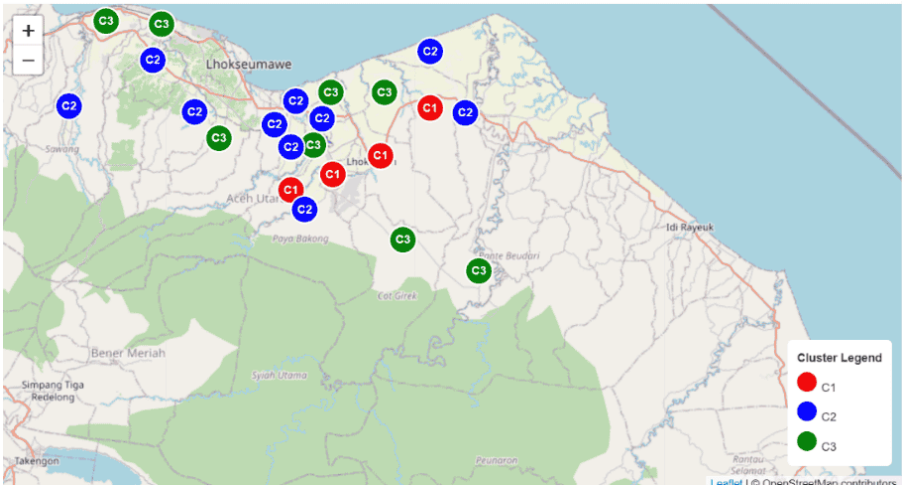


Fig 4. Physical Aspect Mapping Result

Table 9. Items Aspect Cluster Results

No	Police Stations	Cluster
1	Baktiya	2
2	Baktiya Barat	1
3	Cot Girek	1
4	Langkahan	1
...	…	…
19	Sawang	3
20	Simpang Keramat	1
21	Syamtalira Bayu	3
22	Muara Batu	1

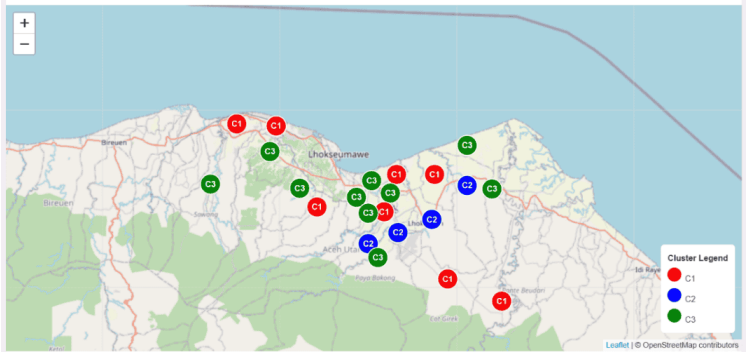


Fig 5. Items Aspect Mapping Result

4.2. Rangking Result

The Borda score is calculated based on the rank given by each respondent, where the formula n - rank is used to determine the Borda score. The following is a manual calculation for the Borda algorithm.

Table 10. Physical Aspect Rangking Results

No	Polsek/Respondent	1	2	3	…	48	49	50	Total	Rangking
1	Baktiya	8	20	3	…	11	14	7	522	10
2	Baktiya Barat	4	19	5	…	9	11	15	508	12
3	Cot Girek	3	18	4	…	12	15	8	452	18
4	Langkahan	9	5	7	…	10	2	14	462	15
…	…	…	…	…	…	…	…	…	…	…
19	Sawang	19	7	16	…	14	9	19	564	5
20	Simpang Keramat	0	6	17	…	2	7	1	363	22
21	Syamtalira Bayu	12	4	18	…	3	8	2	403	21
22	Muara Batu	11	13	20	…	21	21	20	739	2

Table 10. Items Aspect Rangking Results

No	Polsek/Respondent	1	2	3	…	48	49	50	Total	Rangking
1	Baktiya	12	6	16	…	12	11	11	468	16
2	Baktiya Barat	8	0	0	…	14	12	10	476	14
3	Cot Girek	11	8	15	…	2	10	1	407	21
4	Langkahan	15	1	13	…	13	16	17	501	7
…	…	…	…	…	…	…	…	…	…	…
19	Sawang	4	14	6	…	17	2	20	479	13
20	Simpang Keramat	2	2	7	…	8	0	2	399	22
21	Syamtalira Bayu	6	13	17	…	5	6	0	409	20
22	Muara Batu	21	11	20	…	20	21	19	920	1

4.3. Prediction Result

The prediction dataset consists of crime data related to physical aspects and items, collected monthly from 2020 to 2024. The data was obtained from the North Aceh Police Department and the Lhokseumawe Police Department. Below are the data used for the prediction.

Table 10. Physical Aspect Prediction Results Baktiya

No	Name of Police Station	Month - Year	Prediction Result
1	Baktiya	Jan-25	0,62
2	Baktiya	Feb-25	0,6
3	Baktiya	Mar-25	0,58
4	Baktiya	Apr-25	0,75
5	Baktiya	May-25	0,83

Table 10. Items Aspect Prediction Results Baktiya

No	Name of Police Station	Month - Year	Prediction Result
1	Baktiya	Jan-25	0,57
2	Baktiya	Feb-25	0,58
3	Baktiya	Mar-25	0,58
4	Baktiya	Apr-25	0,57
5	Baktiya	May-25	0,57

4.3. Evaluation Prediction Result

The prediction evaluation results using MAPE aim to determine how well the method is able to produce the best predictions. The comparison results of plantation crops are shown in the following table.

Table 11. Evaluation of Prediction Results for Physical Aspect MAPE

No	Name of Police Station	MAPE (%)
1	Baktiya	15,85

2	Baktiya Barat	4,43
3	Cot Girek	4,37
4	Langkahan	7,67
…	…	…
19	Sawang	15,99
20	Simpang Keramat	3,10
21	Syamtalira Bayu	11,13
22	Muara Batu	8,45

Table 12. Evaluation of Prediction Results for Item Aspect MAPE

No	Name of Police Station	MAPE (%)
1	Baktiya	15,74
2	Baktiya Barat	1,60
3	Cot Girek	16,17
4	Langkahan	15,98
…	…	…
19	Sawang	7,17
20	Simpang Keramat	0,00
21	Syamtalira Bayu	20,07
22	Muara Batu	14,22

Based on the overall prediction research results, almost all districts achieved MAPE values of less than 20% for crimes related to physical and item aspects. For example, Baktiya District recorded a MAPE value of 15.85% for the physical aspect. However, in Syamtalira Bayu District, the item aspect reached a MAPE value of 20.07%. Although the MAPE value for the item aspect in Syamtalira Bayu is relatively high, it is still considered acceptable as it remains below 50%.

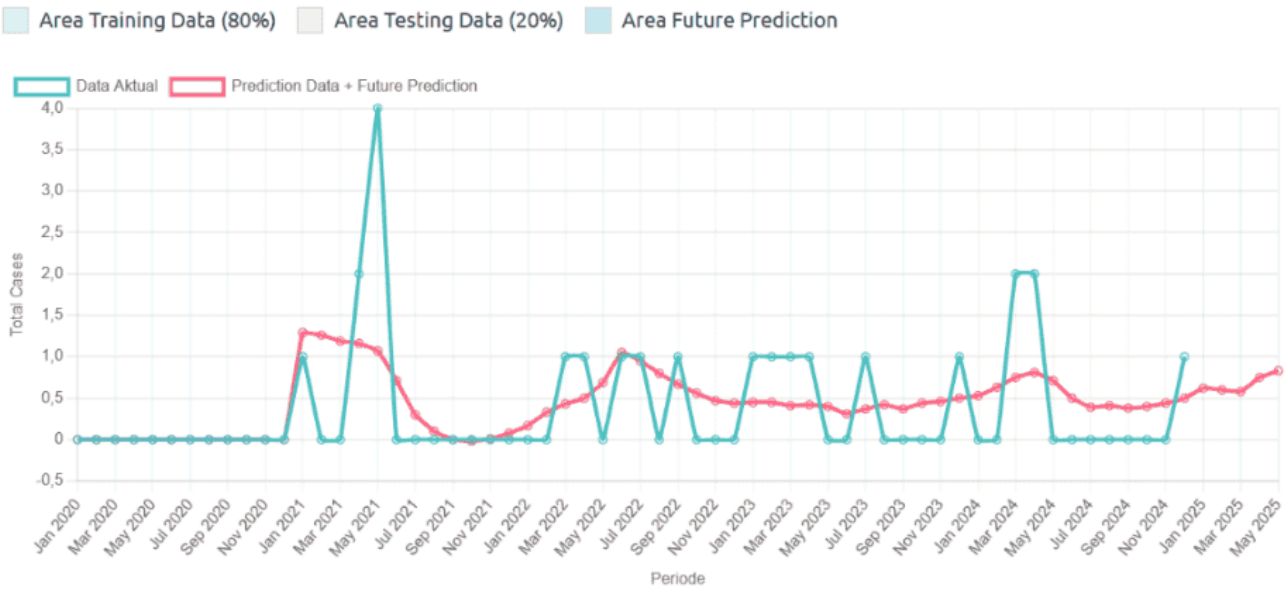


Fig 6. Prediction Graph of Baktiya Physical Aspect

5. Conclusion

Based on the research that has been conducted, the author draws the following conclusions:

1. For clustering with Fuzzy C-Means, the data is grouped into 3 clusters, where cluster 1 represents safe areas, cluster 2 represents moderately vulnerable areas, and cluster 3 represents vulnerable areas.
2. For ranking using the Borda method, the Dewantara Police Sector ranked first for the physical aspect, while the Muara Batu Police Sector ranked first for the item aspect.
3. As for predictions using LSTM, almost all districts achieved MAPE values below 20%, indicating that the LSTM model is quite effective for predicting crime-prone areas.

References

[1]M. Dulkiah, *Sosiologi Kriminal*. 2020.

[2]M. Jannah *dkk.*, "Implementation of Geographic Information System for Tourist Locations and Lodging Services in Lhokseumawe City Based on Android," *Int. J. Eng. Sci. Inf. Technol.*, vol. 2, no. 4, hal. 39–47, 2022, doi: 10.52088/ijesty.v2i4.320.

[3]R. K. Dinata, B. Bustami, S. Retno, dan ..., "Clustering the Spread of ISPA Disease Using the Fuzzy C-Means Algorithm in Aceh Utara," *Int. ...*, vol. 1, no. 2, hal. 21–30, 2022, [Daring]. Tersedia pada: http://ejournal.enlightenlearner.com/index.php/ejournal/article/view/16

[4]M. Ula, M. Ula, D. Yulisda, dan S. Susanti, "Fuzzy C-Means with Borda Algorithm in Cluster Determination System for Food Prone Areas in Aceh Utara," *Ilk. J. Ilm.*, vol. 15, no. 1, hal. 21–31, 2023, doi: 10.33096/ilkom.v15i1.1481.21-31.

[5]D. L. Pardosi dan I. D. Siagian, "Klasterisasi Data Lowongan Pekerjaan Berdasarkan Fuzzy," *J. Ilmu Komput. dan Sist. Inf.*, vol. 3, no. 2, hal. 27–31, 2020.

[6]C. Rahayu, D. Abdullah, dan Z. Yunizar, "Implementation of Long Short Term Memory (Lstm) Algorithm for Predicting Stock Price Movements of Lq45 Index (Case Study …," *Bull. Eng. Sci. ...*, vol. 1, no. 2, hal. 50–59, 2023, [Daring]. Tersedia pada: https://bestijournal.org/index.php/go/article/view/6

[7]S. Asyuti dan A. A. Setyawan, "Data Mining Dalam Penggunaan Presensi Karyawan Denga Cluster Means," *J. Ilm. Sains Teknol. Dan Inf.*, vol. 1, no. 1, hal. 01–10, 2023, [Daring]. Tersedia pada: https://jurnal.alimspublishing.co.id/index.php/JITI/article/download/6/6

[8]Bomanthara dan R. M. Nasrul Halim, "Implementasi Metode K-Means Clustering UntukMenentukan Tingkat Peminatan Konsumen Terhadap TypeSmartphone Di Tio Cell Dan Tio Cell 2," *JUPITER J. Penelit. Ilmu dan Teknol. Komput.*, vol. 15, no. 1a, hal. 49–60, 2023, [Daring]. Tersedia pada: https://jurnal.polsri.ac.id/index.php/jupiter/article/view/5191

[9]N. Ulinnuha, "Provincial Clustering in Indonesia Based on Plantation Production Using Fuzzy C-Means," *ITSMART J. Ilm. Teknol. dan Inf.*, vol. 9, no. 1, hal. 8–12, 2020.

[10]E. Satriani, Ilhamsyah, dan R. P. Sari, "Sistem Pendukung Keputusan Kelompok Penerima Zakat Program Pendidikan Dengan Metode Saw Dan Borda," *J. Komput. dan Apl.*, vol. 07, no. 02, hal. 71–81, 2019.

[11]Moch Farryz Rizkilloh dan Sri Widiyanesti, "Prediksi Harga Cryptocurrency Menggunakan Algoritma Long Short Term Memory (LSTM)," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 1, hal. 25–31, 2022, doi: 10.29207/resti.v6i1.3630.