Application for Monitoring the Prices of Basic Foods in Traditional Markets by the Department of Industry and Trade Using the Machine Learning Algorithm

Lidya Wati[1] , Niky Hardinata[2] , Muhammad Ridho Nosa[3] , Khairus Suhada[4]

{lidyawati@polbeng.ac.id[1], nikyhardinata@polbeng.ac.id[2] , ridhonosa@polbeng.ac.id [3] , khairussuhada@polbeng.ac.id[4]}

Informatics Engineering Department, State Polytechnic of Bengkalis [1, 2, 3, 4.]

Abstract. The trade sector at the Bengkalis Regency Trade and Industry Service and also the general public need reference information on basic food prices that are reliable and easily accessible via the web because so far the public has often speculated about food prices due to a lack of accurate information about developments in prices prevailing at that time. This problem is often exploited by unscrupulous traders, especially traders in traditional markets, by raising prices inappropriately. This study aims to build a basic food price data system application to support the work of supervising and monitoring prices in traditional markets in 11 sub-districts in the Bengkalis Regency. The research data used is the price of necessities in traditional markets from 2019-2023. The research system design utilizes *the Laravel* and *Ionic frameworks* with the Model View and Control (MVC) concept. In the early stages of research, the system was built with 2 access interfaces to the servers, namely through public access web applications (Front-End), and operator/administrator access (Back-End). This study applies the machine learning method with *clustering* or grouping techniques using the K-Means algorithm on food price data. The results of the grouping of basic food prices, it is divided into 3 *clusters*, namely rising prices, falling prices, and fixed prices. Prices increased in 3 sub-districts (Bengkalis, Mandau, Bathin Solapan), prices remained constant in 7 sub-districts (Siak Kecil, Rupat Utara, Bukit Batu, Laksmana, Talang Mandau) and prices fell in 1 sub-district (Pinggir). Cluster evaluation uses the Silhouette Coefficient and Dunn Index, with the results of optimal k = 2 with a value of 0.55 for silhouette and optimal k = 4 with a value of 0.67 for Dunn Index.

Keywords: Sembako, Machine Learning, K-Mean, Cluster

1 Introduction

One of the products that we can find in the market is food products, especially the nine basic ingredients (SEMBAKO). Groceries are the basic needs of our society, and are the main commodity as a reference for people's welfare, but unfortunately at the middle-class level of society often happen speculation price material food because lack of information accurate information about price developments prevailing at that time. This problem is often exploited by unscrupulous traders, especially traders in traditional markets by raising prices that are not

ABEC 2023, September 21, Bengkalis, Indonesia

DOI 10.4108/eai.21-9-2023.2342904

appropriate. One of the media that is often used to obtain information is the internet, the information needed can be obtained quickly and easily. In 2022-2023 internet users in Indonesia will reach 215.63 million. This number increased by 2.67% compared to the previous period of 210.03 million users, based on the results of a survey by the Indonesian Internet Service Providers Association (APJII). The number of internet users is equivalent to 78.19% of Indonesia's total population of 275.77 million people. In 2018, internet penetration in the country reached 64.8% and increased to 73.7% in 2019-2020. It can be concluded that the trend of internet penetration in Indonesia is increasing from year to year.

Application nor system information for food commodity prices by Rahmayuni et al (2020), based application web For monitoring the price of groceries, Jauansyah (2018), prediction of basic food prices by Siti Mujilahwati and Suci Nur Fauziah (2018), system information prices of food and groceries, April Lia Hananto and Bayu Priyatna (2017), to a realtime food price information system for inflation control, Rahman and Sry Wahyuni (2018) can help the public to obtain basic food information and assist the government in monitoring basic food prices.
The system development model uses a prototype development model because it has advantages, namely, faster system completion, minimizing errors due to participation between developers and owners. Web-based applications use prototype methods such as employee payroll systems (Fridayanti, et al. 2021), office document filing systems (Kurniati, 2021) and management information systems (Kustanto GEA, 2021).

Machine Learning (ML) or machine learning is an approach in Artificial Intelligence (AI) (Russell, 2016). This research focuses on one of the machine learning algorithms, namely unsupervised learning. Unsupervised learning algorithm that analyzes unlabeled data. Unsupervised learning is often called cluster because there is no need for labels in the data set and the results do not identify examples in predetermined classes (Thupae, et al. 2018). *Clustering* is a step in grouping data mining based on data similarity. This study uses the K-Mean *clustering* method to classify the data using the simple K-Mean method. The K-Means algorithm is a grouping method based on the shortest distance, where this shortest distance is used to divide data into a *cluster (*Siburian, T., et all, 2019).

Some applications that apply *machine learning*: are school mapping, by Aprilia, et all (2022), application on application nutritional status of toddlers, by Sulastri, et all (2021), application to the paper review application, by Roihan, et all (2020), shows that the application of *the machine learning algorithm* has Worked with Good as well as capable produce application And system information Which as needed.

From the results of the identification, it is felt that an application is needed that can provide information on basic food prices, statistics on basic food prices per community, statistics on basic food prices per market, monitoring sub-district prices in the form of price increases, price decreases and average prices using machine learning algorithms. So that it can help the public to find out information on basic food prices more easily and can be done online, it can also help the Department of Industry and Trade to monitor and control basic food prices in traditional markets in Bengkalis Regency.

2 Research Methods

The stages of the research use the software development method, namely the prototype model,

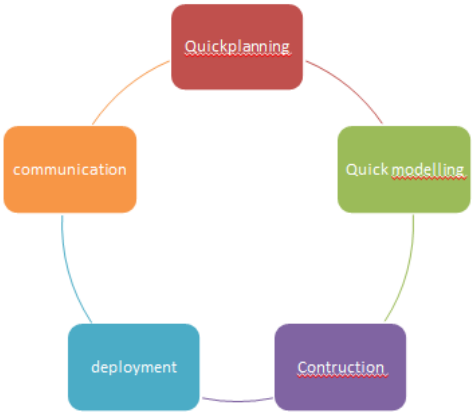

Fig. 1. Prototyping Development Model

Stages done in planning according to the method used that is:

a. Communications

Communicating with the Department of Industry and Trade as data collection material to get an overview of the system to be made, and conducting literature studies through journals and books.

List of questions asked during an interview with the Department of Industry and Trade:

1 What are groceries and what are the types of groceries?
2 How many markets are there in the district? Bengkalis, and mention it!
3 Which markets often experience increases in the price of staple foods, and what causes them?
4 Which markets often have high prices?
5 Which markets are often fixed or moderately priced?
6 Which markets often drop/lower prices?
7 The price of groceries for each market from the 1st week of February – 2nd week of May 2023?
8 What types of groceries often go up/high prices?
9 What types of groceries often have fixed prices?
10 What types of groceries do the prices drop/lower frequently?
11 What are the criteria factors that cause the increase and decrease in the price of basic foodstuffs, for example, stock, number of purchases, number of sales, market location, and unit price?
12 What is the data based on the criteria in no 4?

Based on interviews conducted with the Department of Industry and Trade, the available market data are 11 markets in all sub-districts in Bangkalis Regency, a list of basic food prices

from 2019-June 2023 with 15 basic commodities, namely rice, soybeans, flour, cooking oil, butter, chili, onions, fish, chicken, meat, eggs, tofu, tempeh, salt, and milk.

b. Quick Planning

At this stage, the researcher makes a business plan, and timeline based on the system requirements that have been drawn at the communication stage.

Table 1 Research schedule.

No	Activity	Month					
		4	5	6	7	8	9
1	Communication						
2	Quick Planning _						
3	Model Design						
4	Construction						
5	Submission						
6	Reports, Articles						

c. Quick Modeling (Model design)

At this stage, the proposed system flow design is carried out. The system design uses the Unified Modeling Language model, namely use case diagrams. Input design, output design in the form of an application mockup. Then data processing and data analysis is carried out using *machine learning* algorithms which will later be applied to the system at the construction stage.

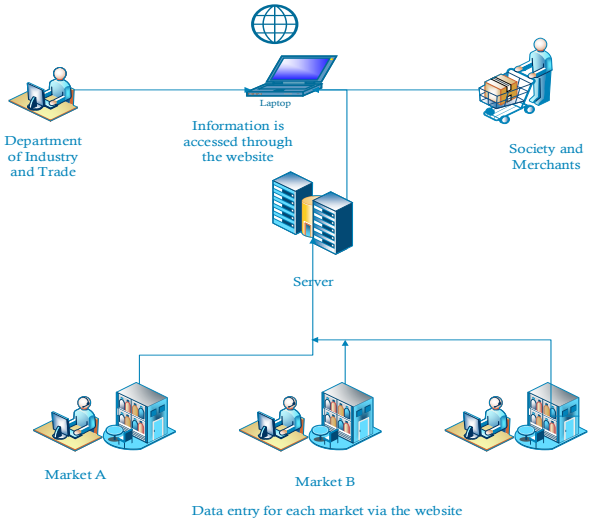

Fig. 2. Proposed system flow.

Application monitoring price market is done by Service Industry And Trading Kab. Bengkalis. Public And traders Can see the price of groceries based on the input made _ by existing operators at each market in each district. Party Department of Industry and Trade will monitor And supervise the price of groceries by monitoring the map price existing market _ For every subdistrict that experienced an increase And decline. Party Department of Industry and Trade Also Can see statistics price of commodity goods, and statistics on the price market as well see a chart monitoring the price market so that expected Can control the price of groceries that are in each market.

Following use case diagram design with three actors viz party department of industry and trade, society and market admin For every district,

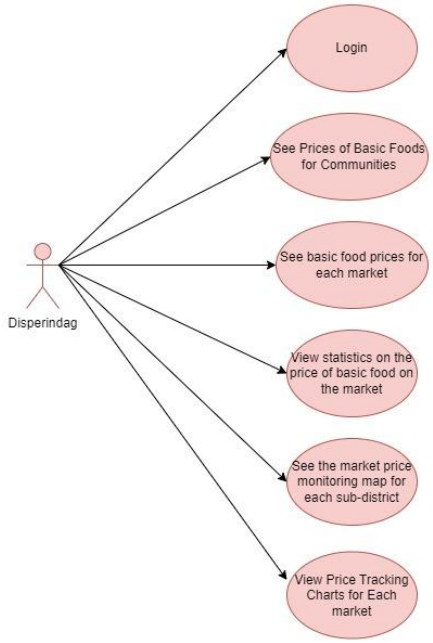

Fig. 3. Use Case diagram for party department of industry and trade

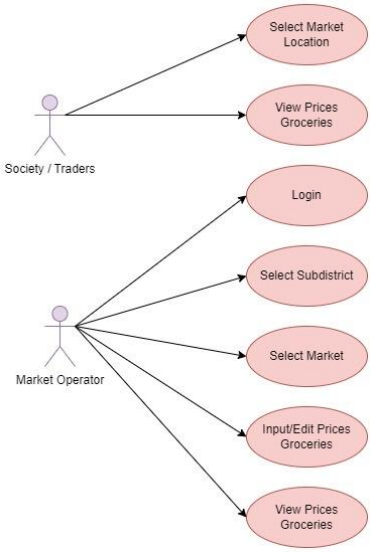

fig. 4. Use case diagram for public and market admin

Following Database design used, database modeling using class diagrams, such as pictures following

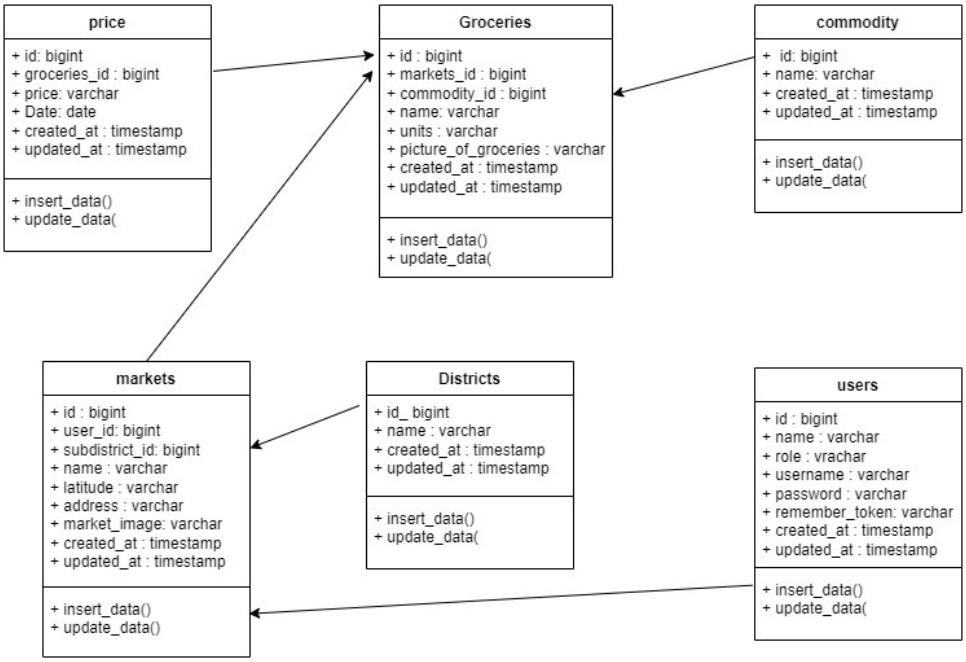

Fig. 5. Class Diagrams

In the study, The K-Mean algorithm is used For group price commodity material food for districts in Kab. Bengkalis to three groups: price up, price still, And price down based on the distance closest. Following K-Mean stages according to Novita A, and Seta AB in (2021) namely,

1. Determine the number of clusters to be used, namely 3 clusters (price up, price stay, and price down)
2. Determine centroid point (i.e. point initial cluster center) i.e. price go on obtained from price highest (max), price still obtained of the average (avg), and price down obtained from price lowest (min). Example testing on commodity rice child virgin period January 2023.
   Price up (max) = 16,000
   Price fixed (avg) = 13.930
   Price Down (min) = 12,000
3. Determine the distance nearest every data point with a centroid point using equality Eucledian Distance: 2, where = Centroid point and = Data point/price. As seen in the picture following

ANAK DARA RICE JANUARY 2023 PERIOD				
ITERASI 1				
Basic foods / District	Average price / month	RISING PRICE	FIXED PRICE	PRICES GO DOWN
		16.000,00	13.930,56	12.000,00
District 1	16.000,00	0,00	2.069,44	4.000,00
District 2	13.000,00	3.000,00	930,56	1.000,00
District 3	13.000,00	3.000,00	930,56	1.000,00
District 4	14.000,00	2.000,00	69,44	2.000,00
District 5	15.000,00	1.000,00	1.069,44	3.000,00
District 6	14.000,00	2.000,00	69,44	2.000,00
District 7	15.375,00	625,00	1.444,44	3.375,00
District 8	12.000,00	4.000,00	1.930,56	0,00
District 9	13.000,00	3.000,00	930,56	1.000,00

Fig. 6. Result data calculation Eucledian distance on iteration 1

4. Done grouping the distance of every data on the nearest centroid data point For each cluster. From Figure 6 above Group prices go on to districts 1, 5, 7. Group price is still in Districts 2, 3, 4, 6, and 9 meanwhile group price is down on District 8.
5. For each cluster specify a new centroid point with the count average (mean) of each data in the cluster, with equation: =, where = the number of data in the cluster, and = the number of distance values in each cluster. Centroid calculation for iteration next with,
   Price Up = (16,000 + 15,000 + 15,375)/3 = 15,458.33
   Price fixed = (13,000 + 13,000 + 14,000 + 14,000 + 13,000)/5 = 13,400
   Price Down = 12,000
6. Repeat steps 3-5, if member cluster No changed then the iteration process is done. Figure 7 follows the results grouping the data For iterations 1 and 2 The same that is Group price goes on in districts 1, 5, 7. Group price is still in districts 2, 3, 4, 6, and 9 meanwhile group price is down in District 8. Then the process is complete.

ITERASI 2				
Basic foods / District	Average price / month	RISING PRICE	FIXED PRICE	PRICES GO DOWN
		15.458,33	13.400,00	12.000,00
District 1	16.000,00	541,67	2.600,00	4.000,00
District 2	13.000,00	2.458,33	400,00	1.000,00
District 3	13.000,00	2.458,33	400,00	1.000,00
District 4	14.000,00	1.458,33	600,00	2.000,00
District 5	15.000,00	458,33	1.600,00	3.000,00
District 6	14.000,00	1.458,33	600,00	2.000,00
District 7	15.375,00	83,33	1.975,00	3.375,00
District 8	12.000,00	3.458,33	1.400,00	0,00
District 9	13.000,00	2.458,33	400,00	1.000,00

Fig. 7. Result data calculation Eucledian distance on iteration 2

d. Construction

The stage that is carried out is making the system through coding using the XML programming language, PHP, and MySql as the database with the Laravel framework. And then testing using the technique:

1. White Box Testing

White box testing is testing a system program from start to finish and data flow, whether or not it is by the expected research.

2. BlackBox testing

Black box testing is system application testing that involves users (community, traders, operators, and the Department of Industry and Trade) which aims to find out the deficiencies in the application system that has been built.

e. Deployment

After carrying out further testing the application is implemented and the Department of Industry and Trade will discuss the system repair and maintenance process.

3. Results and Discussions

The results achieved in this study are in the form of a web-based application system for monitoring the prices of groceries in traditional markets that can be used by the industry and trade department. uses a machine learning algorithm, which can be accessed on the www.sembakobengkalis.com page.

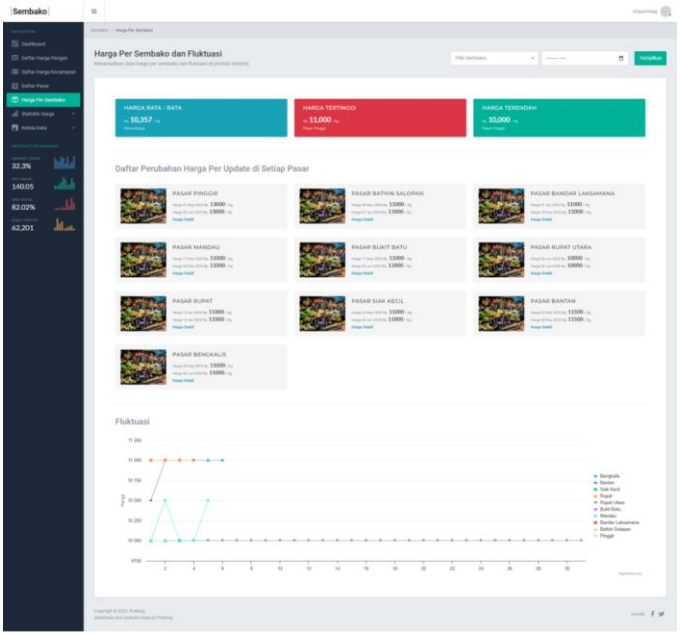

Fig. 8. Price list of foodstuffs

Tests were also carried out on the algorithm used, namely the K-Mean algorithm contained in machine learning, as shown in Figure 9, namely processing using Excel, and 10, namely application to the system. Based on testing the algorithm for the rice commodity for the January 2023 period, there were 2 iterations with 3 clustering, namely prices rising in 3 sub-districts (Bengkalis, Mandau, Bathin Solapan), fixed prices in 7 sub-districts (Siak Kecil, Rupat Utara, Bukit Batu, Laksmana, Talang Mandau) and prices fell in 1 district (Pinggir).

ANAK DARA RICE JANUARY 2023 PERIOD				
ITERASI 1				
Basic foods / District	Average price / month	RISING PRICE	FIXED PRICE	PRICES GO DOWN
		16.000,00	13.930,56	12.000,00
District 1	16.000,00	0,00	2.069,44	4.000,00
District 2	13.000,00	3.000,00	930,56	1.000,00
District 3	13.000,00	3.000,00	930,56	1.000,00
District 4	14.000,00	2.000,00	69,44	2.000,00
District 5	15.000,00	1.000,00	1.069,44	3.000,00
District 6	14.000,00	2.000,00	69,44	2.000,00
District 7	15.375,00	625,00	1.444,44	3.375,00
District 8	12.000,00	4.000,00	1.930,56	0,00
District 9	13.000,00	3.000,00	930,56	1.000,00
ITERASI 2				
Basic foods / District	Average price / month	RISING PRICE	FIXED PRICE	PRICES GO DOWN
		15.458,33	13.400,00	12.000,00
District 1	16.000,00	541,67	2.600,00	4.000,00
District 2	13.000,00	2.458,33	400,00	1.000,00
District 3	13.000,00	2.458,33	400,00	1.000,00
District 4	14.000,00	1.458,33	600,00	2.000,00
District 5	15.000,00	458,33	1.600,00	3.000,00
District 6	14.000,00	1.458,33	600,00	2.000,00
District 7	15.375,00	83,33	1.975,00	3.375,00
District 8	12.000,00	3.458,33	1.400,00	0,00
District 9	13.000,00	2.458,33	400,00	1.000,00

Fig. 9. Processing of the K-Mean algorithm using Microsoft Excel

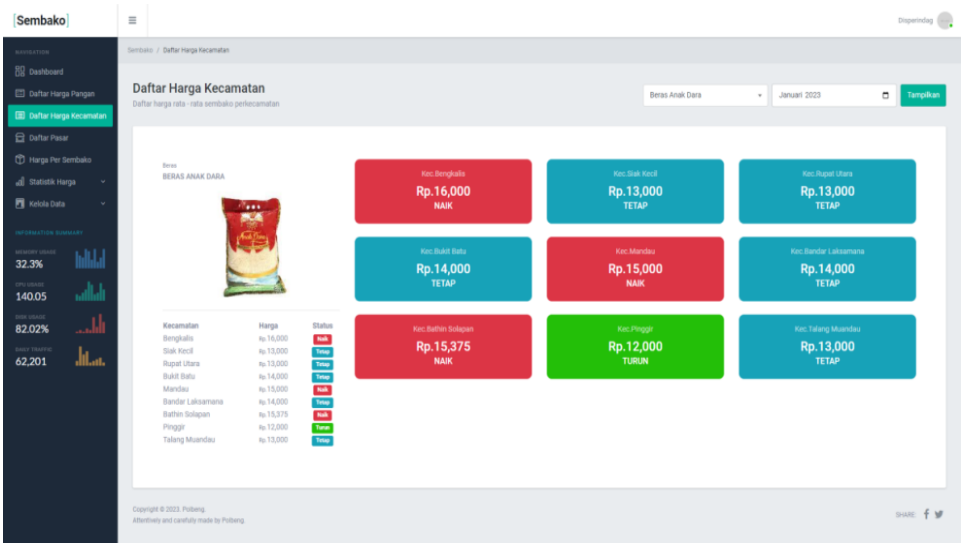

Fig. 10. Page of implementing the K-mean algorithm on the system

Validation Cluster results using the method *Silhouette coefficient and* Dunn Index.

1. *Silhouette coefficients*
   Determination *k* is optimal at method *Dunn Index* looks on from processing use the following Python language

```
#evaluasi jumlah cluster untuk K=2-5
from sklearn.metrics import silhouette_score
for i in range (2,6):
    kmeans = KMeans(n_clusters=i)
    kmeans.fit(x)

    score = silhouette_score(x, kmeans.labels_)
    print("silhoutte score untuk K-Means dengan K = ", i, "adalah ", score)
```

```
silhoutte score untuk K-Means dengan K =  2 adalah  0.5554663423362192
silhoutte score untuk K-Means dengan K =  3 adalah  0.46139304143642645
silhoutte score untuk K-Means dengan K =  4 adalah  0.46112700541701346
silhoutte score untuk K-Means dengan K =  5 adalah  0.45770278150568944
```

Fig. 11. Evaluation of the number of clusters (k) using the *Silhouette coefficient method*

2. *Dunn* Index
   Determination *k* is optimal at method *Dunn Index* visible from processing use the following Python language

```
#evaluasi jumlah cluster untuk K=2-5
from sklearn.metrics import davies_bouldin_score
for i in range (2,6):
    kmeans = KMeans(n_clusters=i)
    kmeans.fit(x)

    scoreDI = davies_bouldin_score(x, kmeans.labels_)
    print("silhoutte score untuk K-Means dengan K = ", i, "adalah ", scoreDI)
```

```
silhoutte score untuk K-Means dengan K =  2 adalah  0.5777500906135428
silhoutte score untuk K-Means dengan K =  3 adalah  0.6168706311664486
silhoutte score untuk K-Means dengan K =  4 adalah  0.6762428343547594
silhoutte score untuk K-Means dengan K =  5 adalah  0.5438563441084899
```

Fig. 12. Evaluation of the Number of Clusters (k) using the Dunn Index method

Based on Pictures 11 and Figure 12 above results from the silhouette has optimal k = 2 with value 0,55. Results from *the Dunn Index* have optimal k = 4 with a mark of 0.67. But based on suggested cluster data collection only 3 clusters.

4. Conclusion

The application of the Machine Learning Algorithm to the application of monitoring the prices of necessities in traditional markets by the Department of Industry and Trade uses a prototype model which consists of five activity frameworks in the implementation of the research, namely communication, quick planning, quick modelling, and construction. (construction), and delivery/ release. The result of this study is that the application of the Machine Learning algorithm, namely K-Mean, can be applied well in the application of monitoring food prices in traditional markets. Based on testing the algorithm for the rice commodity for the January 2023 period, there were 2 iterations with 3 clustering, namely prices rising in 3 sub-districts (Bengkalis, Mandau, Bathin Solapan), fixed prices in 7 sub-districts (Siak Kecil, Rupat Utara, Bukit Batu, Laksmana, Talang Mandau) and prices fell in 1 district (Pinggir). Cluster evaluation uses the Silhouette Coefficient and Dunn Index, with the results of optimal k = 2 with a value of 0.55 for silhouette and optimal k = 4 with a value of 0.67 for Dunn Index.

Acknowledgments

Thank you to those who have helped and contributed to the completion of this research. Hopefully this article is useful for academics and practitioners. This research was fully funded by Bengkalis State Polytechnic

References

[1.] Aprilia US. Ariwibowo B. Cirzun A., (2022), Simulation of School Mapping Design with the Machine Learning Algorithm Method Using Rapid Miner Software, Journal of Al-Azhar Indonesia Science and Technology Series, Vol 7 No 1 January 2022

[2.] April Lia Hananto, et al (2017), Design and Build Product Price Information Applications Food and Staple Foods in the District Market. Karawang, Journal of Computer Science & Information Technology , Vol 2 No:1, April 2017

[3.] Juansyah (2018), E-Monitoring of basic food prices at the Department of Commerce and Industry Musi Banyuasin Regency Industry, Informanika Journal, Volume 4 No.2, July-December 2018
[4.] Kurniati, (2021), Application of the Prototype Method in the Design of a Document Filing System for the Lais District Office, Journal of Software Engineering Ampera Vol 2 No 1 February 2021
[5.] Kustanto GEA. Chernovita AP, (2021), Web-Based Management Information System Design Case Study: PT. Unicorn Intertrnz, JTIIK Journal Vol 8 No 4 August 2021
[6.] Fridayanti EW, Haryanto, Tsabitah T, (2021), Application of the Prototype Method in Designing a Web-Based Employee Payroll Information System (Persis Gawan), Jurnal Paradigma Vol 23 No 2 September 2021
[7.] Rahmayuni I. Sonata Y. Alanda A. Erinda A, (2020), Food Commodity Price Information System for Markets in Padang City, Jitsi Journal Vol 1No 1
[8.] Russel, SJ and Norvig, P. (2016). Artificial intelligence: a modern approach, Malaysia; Pearson Education Limited
[9.] Roihan A, Sunarya PA, Rafika AS. (2020). Utilization of Machine Learning for Various Fields: Review Paper. Journal of IJCIT, pp 75-82
[10.] Siburian, T., Safii, M., Parlina, I., (2019) Application of the K-Mean Clustering Algorithm for Grouping Retail Prices of Rice in Traditional Markets Based on City Region, Proceedings of the National Information Science Research Seminar, September 2019, Pages 927-936
[11.] Siti Mujilahwati, et al (2018), Ooad Modeling for Food Price Prediction Applications Based on Android, Antivirus Journal, Vol. 12 No. May 1st 2018
[12.] Sulastri H, Mubarok H, Iasha SS. (2021). Implementasi Algoritma Mechine Learning untuk Penentuan Cluster Status Gizi Balita. Jurnal Jurti. Vol. 5, No. 2 Agustus 2021
[13.] Thupae, R., Isong, B., Gasela, N., & AbuMahfouz, A. M. (2018). Machine Learning Techniques for Traffic Identification and Classification in SDWSN: A Survey. IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society, 4645–4650.