

Design and Build A Customer-Finding Application For Leko Restaurant Using The K-Means Algorithm

Mohamad Yusuf¹, Muhaimin Hasanudin*², Ifan Prihandi³

^{1,2,3}Universitas Mercu Buana, Indonesia

mhd.yusuf@mercubuana.ac.id¹, muhaimin.hasanudin@mercubuana.ac.id²,
ifan.prihandi@mercubuana.ac.id³

Correspondence by email : muhaimin.hasanudin@mercubuana.ac.id

Abstract

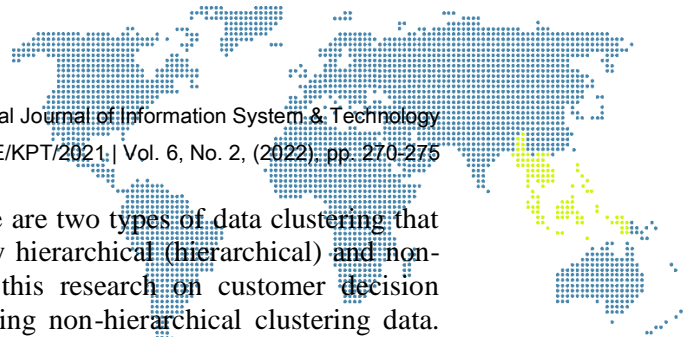
Warung Makan Leko is one of the restaurants in the Jakarta area that offers local cuisine, a diverse menu and delivery orders via phone order. Customers are one source of income for Warung Makan Leko. The amount of competition makes Warung Makan Leko have difficulty in retaining loyal customers. For this reason, further analysis is needed to find out who these potential customers are. Then an application was developed to classify customer data using the K-Means (clustering) algorithm. The data used as an example in this study is the sales transaction data of Warung Makan Leko. Run the process to calculate the total sales to customers and the number of transactions with customers to classify customer data. The K-Means clustering method tries to group the existing data into groups. Data in groups have the same properties. Customer data is grouped into two clusters, no and implicit. Each cluster is then classified based on the prioritized criteria. The cluster with the highest centroid value is the cluster that is rewarded, and the cluster with the lowest centroid value is the non-rewarded cluster. The results of this process form clusters, which are used for advice and consideration to determine sales strategy, namely to reward customers who rank higher in the cluster.

Keywords: restaurants, customers, clustering, algorithm K-Means,

1. Introduction

The customer is one of the important positions in the development of business strategy, the customer is also one of the sources of profit in the company. This requires a good understanding of the customer. A good understanding of customers can be used by companies to retain customers which must be maintained [1,10]. The problem that is often faced is the difficulty in determining the right customer quickly, this can result in the company losing customers that should be retained. Customer segmentation is a popular method used to select the right customers or customers to start a promotion. By segmenting customers based on their behavior[13], we can better target their actions. Such as giving loyalty that can meet customer expectations so that customers will stay because they are satisfied with the services provided. The process of determining customers manually by inputting customer data one by one and then sorting customer data often causes several problems, including taking a long time and high accuracy. In addition, the computational process of determining the customer is also one of the problems. So we need a system that can assist in the decision-making process anyone [2,11] recommended customers receive awards based on predetermined criteria. The solution is to create an application system by making more use of computer technology as a tool in supporting operational activities in the business field that makes it easier for users to process data or information quickly, precisely and accurately so that work effectiveness and efficiency is achieved [6,13].

However, analyzing large amounts of customer data requires a lot of energy and time. So it is necessary to apply an algorithm, the application of a clustering algorithm for



determining customers based on existing data. There are two types of data clustering that are often used in the data grouping process, namely hierarchical (hierarchical) and non-hierarchical (non-hierarchical) data clustering. In this research on customer decision support systems, the authors focus more on grouping non-hierarchical clustering data. Clustering algorithms on non-hierarchies with the partition approach are of several kinds, including Fuzzy C-Means, K-Means and so on. The author uses a clustering algorithm, namely the K-Means Clustering algorithm. The K-Means algorithm is a non-hierarchical data clustering method that partitions data into one or more clusters/groups, so that data with the same characteristics are grouped into the same cluster and data with different characteristics are grouped into other groups. [7,8,13]. This algorithm will help users determine customers faster. The K-Means Clustering Algorithm, apart from being able to cut the time for determining potential customers, can also make precise and accurate determinations [14].

2. Research Methodology

The method that can be used is based on the experience of previous research that the K-means method can be used to classify user satisfaction data. Furthermore, the results of the clustering are used as labels which will later be processed using classification techniques. The results of the classification will be evaluated for its performance for consideration for the next step.

The flow of this research is the first to analyze and survey the research site. Furthermore, the collection of supporting data used for research. After the data is collected, data preprocessing is carried out so that the data can be processed. After the data is normal, the next step is the application of the K-Means method. The last is to evaluate the research results.

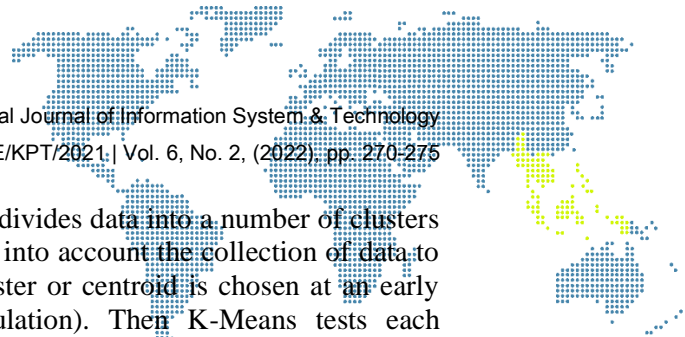
3. Result and Discussion

3.1. Analysis

The development of technology is growing day by day. Advances in technology have led to many web-based applications that can display various kinds of digital information. These web-based applications do not rule out the possibility that the information provided is similar. Customer determination application to find out which customers are not based on existing transactions with data that has been stored in the database. The current system is still manual. The current system analysis is carried out to find out the problems caused by the current system. These problems can be used to analyze user needs for the new system. Then, after the problems and system requirements are analyzed properly, the results of the analysis can be used as a basis for designing a new system, so that a new system can be produced that can meet user needs.

3.1.1. K-Means Clustering Algorithm Analysis

In the analysis of clustering to determine the previous customer, when the data grouping process requires a clustering algorithm for grouping data according to the characters inputted from the user with the transaction data information database. For this reason, the following is an analysis of the Clustering search using the K-Means Clustering Algorithm in data grouping. In an article published by Elsevier "Data clustering: 50 years beyond K-means" [7,13] it was revealed that organizing data into a cluster is the most basic model for understanding and learning. Cluster analysis is a formal study for grouping, or clustering objects according to the characteristics that are measured based on their resemblance to one another. Clustering is grouping using unsupervised learning techniques where no training is needed on the method or in other words, there is no learning phase and does not use labeling for each group. In the same journal, it is also disclosed about the use of K-means as an algorithm which has been reliable since the past and is even currently quite reliable for the clustering process. The K-Means Clustering



algorithm is a distance-based clustering method that divides data into a number of clusters where the clustering process is carried out by taking into account the collection of data to be grouped. In this algorithm, the center of the cluster or centroid is chosen at an early stage randomly from a collection of data (population). Then K-Means tests each component in the data population and marks the component to one of the previously defined centroids based on the minimum distance between the components (data) and each centroid. The centroid position will be recalculated until all data components are grouped into each centroid and finally a new centroid position will be formed. This iteration will continue until convergent conditions are created

3.1.2. Requirements Analysis

To build a system, it is necessary to do a system requirements analysis stage. Analysis of system requirements can be grouped into two parts, namely: functional requirements and non-functional requirements.

a) Functional Requirements

System functional requirements are activities and services that must be owned by a system in the form of input, process, output, and data storage (3,11,15)

The Functional Requirements needed are:

- 1) The user enters input data in the form of a transaction
- 2) The system performs clustering or grouping of customer data based on the transactions of each customer that has just been updated
- 3) The system must be able to determine the sales strategy, namely providing rewards for customers with the top cluster position.

b. Non-Functional Requirements

The non-functional requirements of the system are the requirements of what the system must do. Some of the non-functional requirements that must be met by the designed system are as follows:

- 1) Performance, The system to be built must be able to show the results of the customer determination process
- 2) Easy to learn and use, The system to be built must be simple and easy to learn by the user.
- 3) Documentation, The system to be built can store word suggestions added by the user (user) into the database
- 4) Save cost, The system to be built does not require additional devices in the execution process

3.1.3. How the K-Means Algorithm Works

The way the algorithm works in conducting customer searches is as follows:

- a) Define the number of K clusters.
- b) Initialize K cluster centers (centroids) as seed points (initial cluster prototype). This centroid can be obtained randomly or selected from the first K data objects.
- c) For each data component, calculate and mark the distance (distance) to the initial centroid then enter the data to the centroid closest to the distance 2
- d) Calculate and change the centroid of each cluster as the average of all members of the cluster.
- e) Check all data again and put each data closest to the new centroid. If the members of each cluster do not change (converge), then the step stops and if it still changes, returns to step 2.

The following is a flowchart of the K-Means Clustering algorithm on customer determination.

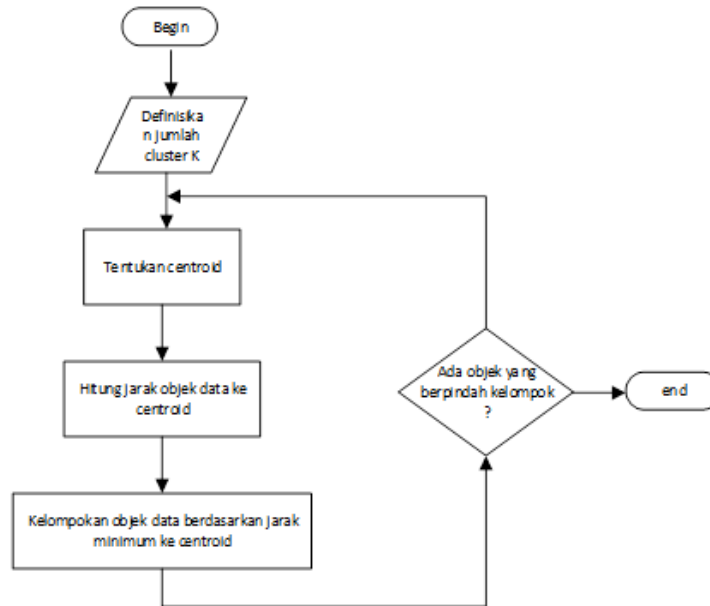


Figure 1. K-Means Clustering Algorithm on customer search

3.2. Design

3.2.1. Design of customer search system

At Warung Makan Leko, a web-based customer search application system will be created. The design of this application was built to make it easier for owners to quickly and precisely obtain customer information, simply by opening the website, information about who got customers can be seen properly and clearly. The design of the Warung Makan Leko website application that was built is object oriented (object-oriented) using Unified Modeling Language (UML) as a modeling language.

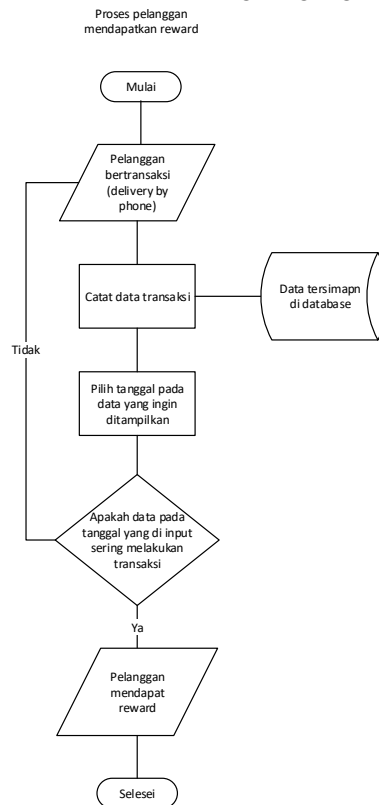
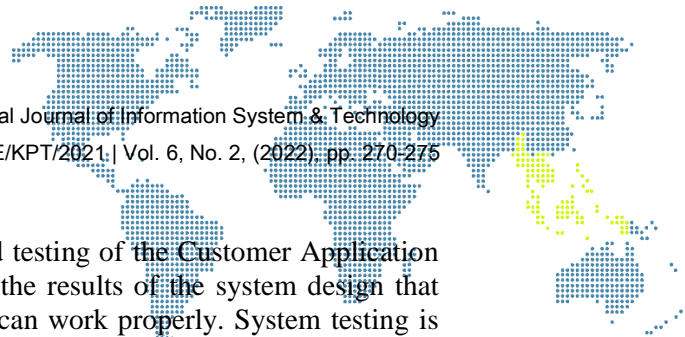


Figure 2. Customer Process Get reward



3.2.2. Implementation

At this stage, it discusses the implementation and testing of the Customer Application System. This implementation is a way of realizing the results of the system design that has been done so as to produce an application that can work properly. System testing is done using Mozilla Firefox, Google Chrome as a web browser, the database engine used is MySQL which works on the Apache web server. The following is an application to add a food menu, which can be seen in the image below:

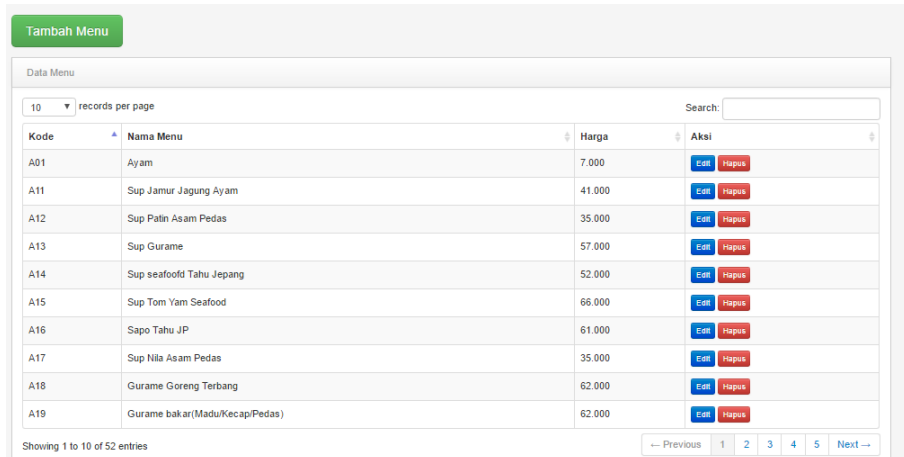


Figure 3. add food menu

3.2.3. Test Results

Table 1. Table of Program Test Results

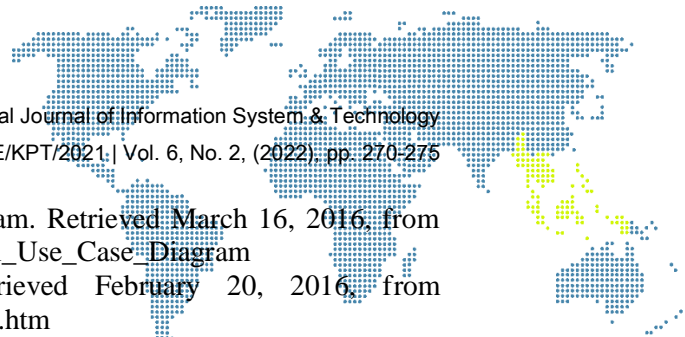
Transaction Data	Types
Data Year	January 2016 to June 2016
Total Transactions	600 Records
Number of Customers for Transactions	± 650 Records
Clusters	2
Clustering Process	± 10 Minutes
Iterations	10 terations

4. Conclusion

Some conclusions that can be drawn from this research are the K-Means Clustering algorithm can be used to group customer data as a support for customer determination decisions. Customer data is grouped into two clusters (and not). Then each cluster is classified based on which criteria are prioritized, namely one of the criteria for the number of transactions and the nominal number of transactions. The cluster with the largest value at the final centroid is a cluster that is recommended to receive a reward, while the cluster with the smallest value is a cluster that is not entitled to receive a reward. Applied the application with the concept of decision makers using the K-Means Algorithm to group customer data to find out data that has potential by calculating the number of transactions to classify customer data. customer with the top cluster position.

References

- [1] Setyaningsih, P. W. (2021). Sistem Informasi Penjualan Kopi Khas Lampung Berbasis Website. INTEK: Jurnal Informatika dan Teknologi Informasi, 4(1), 39-44.
- [2] Rahayu, S. (2016). Analisa Dan Perancangan Sistem Pemilihan Ketua Himpunan Mahasiswa Studi Kasus: Program Studi Sistem Informasi Universitas Mercu Buana. Jurnal Ilmiah Fifo, 8(1), 85-96.



- [3] Christine. (2016). Mengenal Use Case Diagram. Retrieved March 16, 2016, from https://www.academia.edu/5295802/Mengenal_Use_Case_Diagram
- [4] etunas. (2012). Pengertian MySQL. Retrieved February 20, 2016, from <http://www.etunas.com/web/pengertian-mysql.htm>
- [5] Fadul, F. (2016). Apa Itu Bootstrap? bagaimana Memulai Belajar Bootstrap Untuk Pemula. Retrieved March 17, 2016, from <http://dul.web.id/bootstrap/3/tuts-tips/belajar-bootstrap-untuk-pemula.php>
- [6] Hasanudin, M., Kristiadi, D. P., Roihan, A., & Haris, H. (2020). Rancang Bangun Sistem Informasi Jasa Proyek (SiJasPro) berbasis Mobile. *IT Journal Research and Development*, 4(2).
- [7] fadli. (2016). K-Means, Custering . Retrieved June 14,2014 , from - <https://fadlikadn.wordpress.com/2013/06/14/tahap-tahap-k-means-clustering/>
- [8] Siburian, D., Andani, S. R., & Sari, I. P. (2022). Implementation of K-Means Algorithm for Clustering Books Borrowing in School Libraries. *JOMLAI: Journal of Machine Learning and Artificial Intelligence*, 1(2), 115-124.
- [9] Virantika, E., Kusnawi, K., & Ipmawati, J. (2022). Evaluasi Hasil Pengujian Tingkat Clusterisasi Penerapan Metode K-Means Dalam Menentukan Tingkat Penyebaran Covid-19 di Indonesia. *Jurnal Media Informatika Budidarma*, 6(3), 1657-1666.
- [10] Novianto, R. (2019). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menganalisa Bisnis Perusahaan Asuransi. *Jatisi (Jurnal Teknik Informatika Dan Sistem Informasi)*, 6(1), 85-95.
- [11] Yulisopianti. (2015). Perbedaan Use Case, Activity Diagram, Sequence Diagram. Retrieved March 16, 2016, from <http://yulisopianti.staff.telkomuniversity.ac.id/2015/05/02/perbedaan-use-case-activity-diagram-sequence-diagram/>
- [12] Mulyana, Y., Fadlil, A., & Riadi, I. (2022). Classification of Organizational Data Using the K-Means Algorithm. *Komputasi: Jurnal Ilmiah Ilmu Komputer dan Matematika*, 19(2), 122-136.
- [13] Zou, F., & Li, R. Construction of Student Innovation and Entrepreneurship Experience System Integrating K-Means Clustering Algorithm. *Mathematical Problems in Engineering*, 2022.
- [14] Sumadikarta.2014.Peneraan Algoritma K-Means:Studi Kasus PT MEga Arvia Utama. *Jakatta:Jurna Teknik Informatika*. Vol. 1, No.1:12-22.
- [15] Teorey, T. (2006). The Unified Modeling Language (UML), 3, 33–51. <http://doi.org/10.1016/B978-012685352-0/50003-3>