

Application of K-Means algorithm to predict consumer interest according to the season on place reservation and food online software

J S Pasaribu^{1*}

¹Department of Informatics Engineering, Politeknik Piksi Ganesha, Jl. Gatot Subroto no. 301, Bandung, Indonesia

*johni_0106@yahoo.com

Abstract. The background of what this report discussed is the problem on the food menu due to the unsold food menu causes that make the restaurant loose their money because of the unused food menu. In this case K-Means can be used to cluster consumer interest where the interest will be adjusted with the season. K-Means algorithm will be used in software to find consumer interest from existing sales data, in this case can be said as data mining where the data will be made and sorted by customer's interest like summer or rainy season. The determination of these problems based on the season due to consumer interest that will be more easily to calculated when we following the season in Indonesia where Indonesia itself is a tropical country which has a diverse diet. With the grouping of data from sales data restaurant can determine the menu of food and beverages in accordance with consumer interest where the menu will be adjusted to the season in Indonesia that is summer and rainy season. And the results of data processing will be implemented on online reservation software and food reservations previously made in previous research entitled Online Food Place and Food Packages In Food Garden Miko Mall. This is intended to facilitate the restaurant to determining in the purchase settings of feed ingredients. If there is a lot of menus sold then raw materials will be reproduced and if there is a food menu that is less enthusiastic or the sale is spelled out less then the raw material will be reduced.

1. Introduction

Data mining is a technique of processing large amounts of data for grouping. This technique is used in the Knowledge Discovery process in Database (KDD). Clustering is one of the data mining methods that is non-directive (unsupervised) and a method for finding and grouping data that has similar characteristics between one data and another data [1]. According to the category of compactness, grouping is divided into two, namely complete and partial. If all data can be joined together, all data can be said to be compact into one group [2]. In this clustering there are several grouping algorithms to easily group data. One of them is the K-Means algorithm, which is a group analysis method that leads to N partitions of observational objects into K groups, where each object is observed by a group of data with the closest mean (mean) [2]. Grouping is generally applied to group documents or objects that are not neatly arranged and do not match the arrangement in its place. But the function of grouping is not just grouping documents or objects. K-Means algorithm will be used in software to find consumer interest from existing sales data.



Many companies today are updating their marketing systems to sell their products. Various methods are carried out such as conducting questionnaires, advertisements, and even predicting consumer interest. This is done so that the products sold can be adjusted and not too deviated from the wishes of consumers. There are various ways to predict consumer interest, one example is using a tool in the form of software. Using software is an easy way to predict because the system of the software that calculates, calculates and classifies products. Where one of them is software using algorithms and data mining. The algorithm will determine decision making regarding the interests of consumers based on sales data.

Software that uses algorithms is expected to help the company to make decisions about customer interest so that the product can be in accordance with market desires in this case consumers. Likewise, restaurants and restaurants will require predictions about consumer interest to adjust food and beverage menus. The many types of food and drinks that are in restaurants or restaurants that sometimes have unsold or less interested menus will certainly harm the restaurant or restaurant. So in this case K-Means algorithm can be used to cluster consumer interests where the interest will be adjusted to the season. K-Means algorithm will be used in software to find consumer interest from existing sales data, in this case it can be said as data mining where the data will be created and sorted according to customer interests such as summer or rainy season. K-means clustering algorithm is one method of non-hierarchical clustering data that groups data in one or more clusters / groups. Overall data that has the same characteristics are grouped in one cluster / group and data that has different characteristics are grouped with other clusters / groups so that data in one cluster / group has a small level of variation [3]. And the determination of the problem is based on the season because consumer interest will be easier to calculate following the seasons in Indonesia where Indonesia itself is a tropical country where there is a diverse diet.

With the grouping of data from sales data it is expected that restaurants or restaurants can determine the menu of food and drinks that are in accordance with consumer interests where the food menu will be adjusted to the seasons in Indonesia, namely summer and rainy season. And the results of processing the data will be implemented in online food ordering and food software that was previously made in a previous study entitled Online Ordering and Food Software Food Garden at Miko Mall. This research is an additional feature in the software where the software changes the application of the ordering method which is generally done in a conventional manner where prospective customers must come to the Food Garden to order online with no need to come customers can already book a place through the website. This additional feature is in the form of a suggested product menu, such as hot goods or items that are sold a lot or items that are trending like in an online shop. Another additional feature is the annual report to determine the food menu that is sold in the summer or dry season and the rainy season and what food items are not sold in the rainy and summer seasons based on data that has been processed by K-means algorithm. This is intended to facilitate the restaurant in determining the arrangement of raw material purchases. If there is a menu that is sold a lot, the raw material will be reproduced and if there is a food menu that is less interested or the sales are spelled out less or less then the raw material will be reduced.

Based on the background above, the problem can be identified as follows:

- Difficulty in determining the type of food and drink that is in accordance with consumer interest, making the existing menu unsold and prepared food or beverage ingredients in this case causing harm to the restaurant or restaurant.
- Consumers feel confused in choosing a food or drink menu.

2. Literature Review

As a material consideration and to enrich the literature in this study, the authors put forth some previous studies that have been examined as references in the analysis of this study. Nurul Rohmawati, Sofi Defiyanti, and Mohamad Jajuli [4] made a research entitled K-Means Algorithm in Classification of Scholarship Application Students. Data grouping scholarship applicants for Student Learning Assistance (BBM) grouped into 3 categories are entitled to receive, considered and not eligible to receive the scholarship. Grouping into 3 groups is useful to facilitate in determining scholarship recipients. K-means

algorithm is an algorithm of clustering technique based partitions. This technique can categorize student data scholarship applicants. The purpose of this research is to determine the algorithms for performance measurement, and measurement in view of the results of the cluster by calculating the value of purity (purity measure) of each - each cluster is generated. The data used in this research is data of students who apply for a scholarship to the School of Computer Science UNSIKA (Universitas Singaperbangsa Karawang) many as 36 students. The data will be converted into 3 datasets with different formats, ie attribute data codification in part, attributes and attribute the overall codification of the original data. Then it can be concluded that the k-means algorithm is more suitable for use in datasets with formatting attributes that codified a whole.

Johan Oscar Ong [5] made a research with the title of K-Means Clustering Algorithm Implementation to Determine Marketing Strategies President University. In this research, he start from calculation of data set of students who have graduated from President University using K-Means Clustering Algorithm, namely by classifying the data of students into several clusters based on the characteristics of this data in order to discover the information hidden from the data set of student who have graduated from President University. The attribute data that is used in this study is hometown, major and GPA. The purpose of this study is to help the President University's marketing department in predicting promotion strategies undertaken in the cities in Indonesia. Information gained in this study can be used as a references in determining the proper strategy for marketing team in their promotion activities in the cities in Indonesia so that the campaign will be more effective and efficient.

Green F Mandias, Green A Sandag, Susi Susanti, and Haryanto Reza Musak [6] made a research entitled Application of K-Means Algorithm for Academic Achievement Analysis of Faculty of Computer Science Universitas Klabat. This research was conducted to find out the achievement of computer science faculty students who are at level 4 which has 52 active students by utilizing data mining method. Based on data from computer science faculty students, this research is done to find out how many students have academic achievement in databases, networking and programming using data mining algorithm. This research uses K-Means algorithm in analyzing academic achievement of computer science faculty student at Universitas Klabat.

3. Analysis and Design

The software that will be created is a Website-Based Information System Application as an Implementation of Public Information Disclosure at Diskominfo Kabupaten Bandung. The functional requirements of the software are:

Table 1. Functional Needs.

No	Functional Needs	The people can do
1	The system enters the home page in the software as supervisor, and cashier	The supervisor and cashier log in to be able to run features in the ordering application
2	The system creates and view place reservation	Buyer orders place
3	The system manages place reservations: edit and delete place reservation data	Cashiers can create and view place reservation that occurs every month
4	The system manages upload proof of payment	Buyers upload proof of payment
No	Functional Needs	The people can do

5	The system manages reservations	Cashiers can search and change or delete data on the booking form that was previously made
6	The system knows the customer's payment	Supervisors can search and change or delete data on the payment form that was previously created
7	The system makes and see food packages	Supervisors can make food packages in the food garden to order the place
8	The system makes and see food menu	Supervisors can make food menus in the food garden to order the place
9	The system manages the food menu report	Supervisors can download reports on the calculation of food menus and beverage menus sold and unsold
10	The system manages order report	Supervisors can download order reports on the food menus and beverage menus sold and unsold
11	The system manages K-Means criteria	Supervisors can make criteria to determine the classification of reports to get a food determination report
12	The system imports sales data	Supervisors can import sales data that will be used as supporting calculators and processing of K-Means Clustering

3.1. Usecase Diagram

The use case diagram is used to find out what functions are in a system and who has the right to use these functions. The following contains a use case diagram for a Application of K-Means Algorithm to Predict Consumer Interest According to the Season on Place Reservation and Food Online.

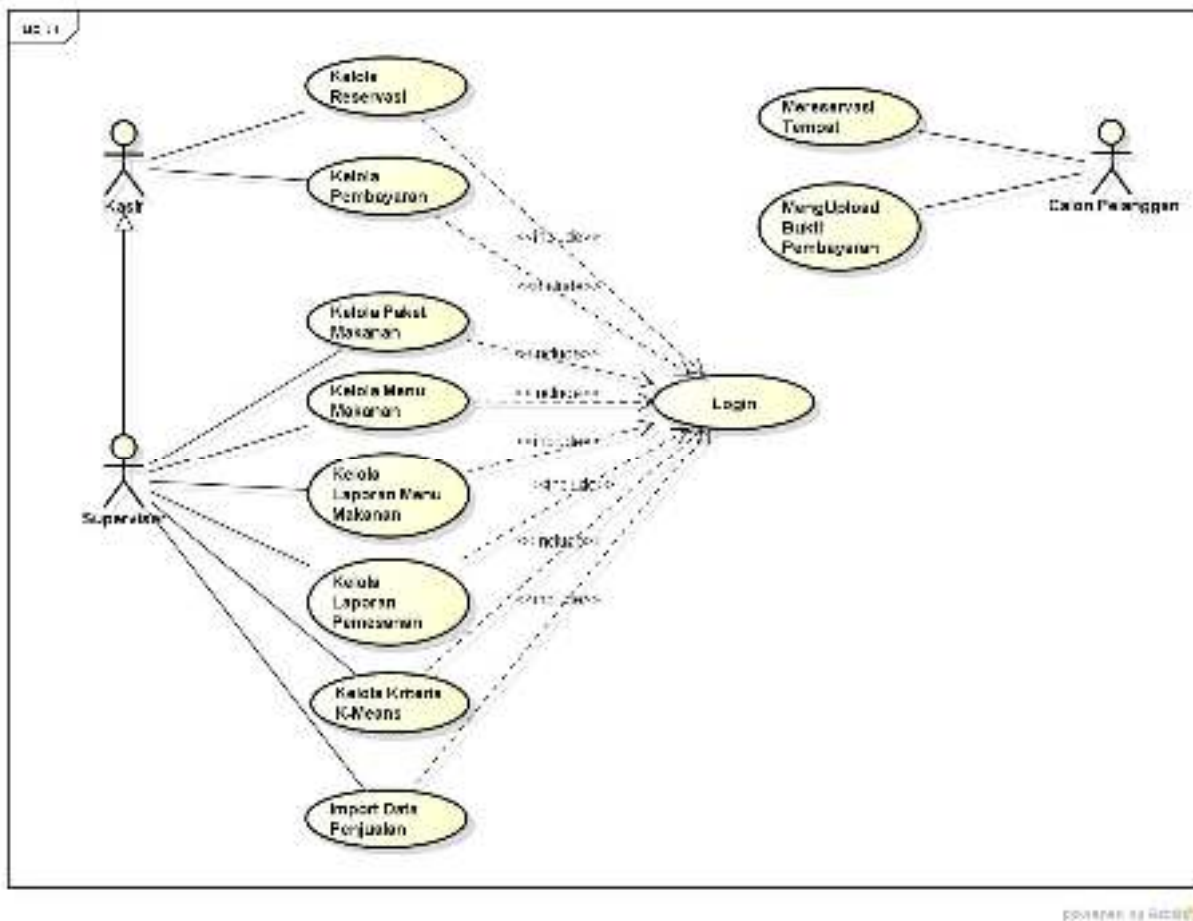


Figure 1. Usecase Diagram

3.2. Usecase Diagram

Class Diagrams is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among objects.

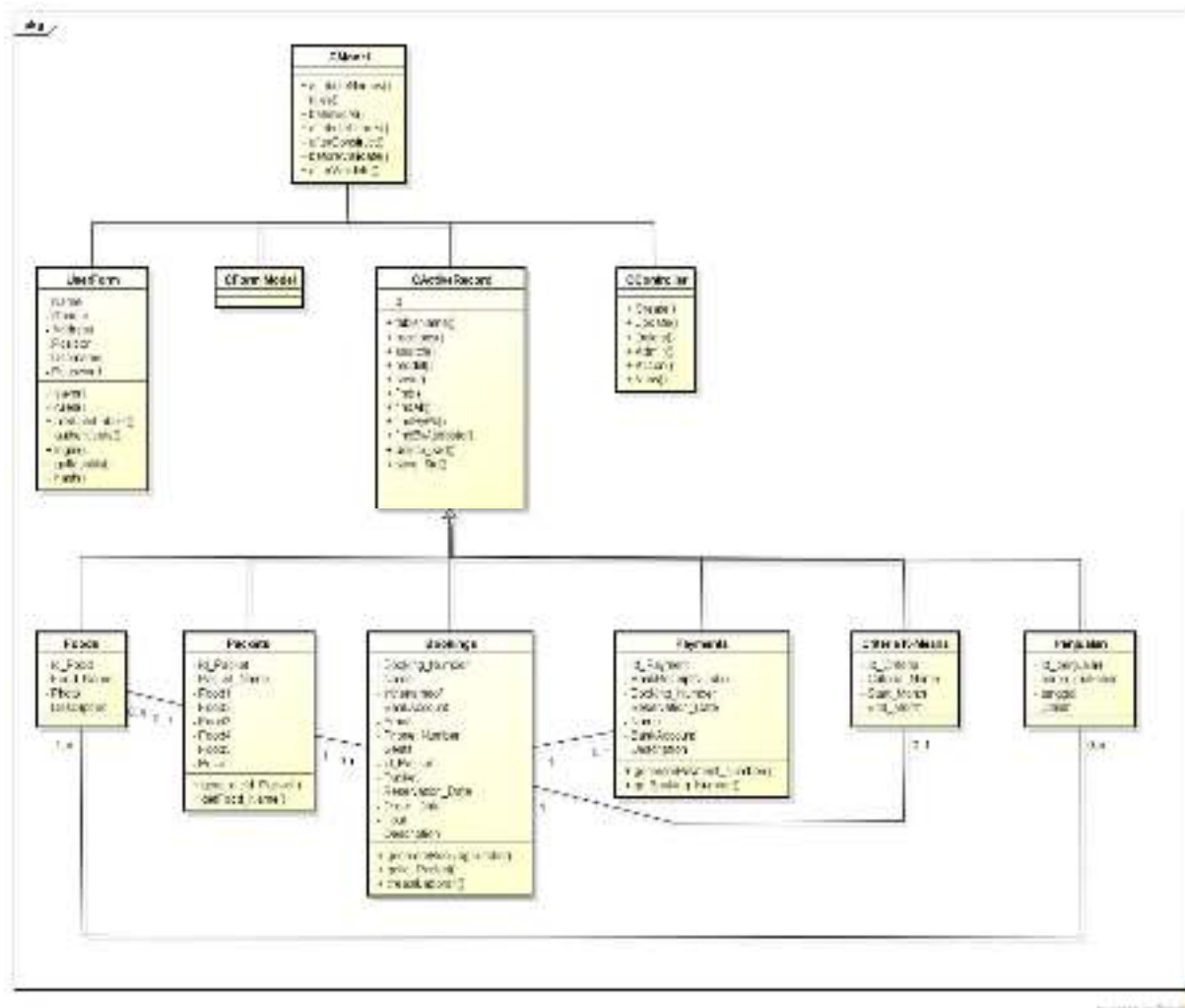


Figure 2. Class Diagram

4. Interface Implementation

The following will be given some interface implementations.



Figure 3. Screen dialog the main menu of customers in the rainy season



Figure 4. Screen dialog for main menu for customers in summer



Figure 5. Screen dialog K-Means food menu

5. Conclusion and Suggestion

Based on the description of the discussion in the previous sections, some conclusions are drawn which include:

- The implementation of the K-Means Clustering algorithm on the food menu based on sales data can be done and the results obtained in accordance with what is expected in the form of a report to be considered by top management to determine the food menu according to the season. The

Application of K-Means Algorithm to Predict Consumer Interest According to the Season on Place Reservation and Food Online is in accordance with the criteria possessed by K-Means clustering. Where the food menu data is in demand and the food menu that is not in demand can be found by K-Means and this can help the restaurant not to suffer losses due to unused food ingredients.

- The results of the calculation of the K-Means algorithm can be used as a suggested product because the data taken is obtained from the first cluster, namely the cluster whose criteria are products that are sold a lot or in this case are products that are purchased by consumers. It also explains that what is explained in Chapter I in the problem identification sub-chapter in the second point has been successfully implemented, and the results of the K-Means can also be a selling power to show or introduce the best food menu that is owned by the restaurant. In this case it can be concluded that the data obtained from the K-means calculation can be used as a recommendation menu for customers so that customers can find out the menu recommendations from the restaurant.

Based on the research that has been done and based on the results that have been made from the research conducted, then suggestions are made that can be accepted and useful. Therefore, it is expected that future researchers will develop this system for the better and more useful. The software development suggestions:

- Criteria for determining centroid points can be made into input or not static.
- Reports are made in more detail where the results of the K-Means calculation can be grouped again based on the criteria of food sold, such as Cow Soup has the criteria of Warm, Food, Berkuah, Spicy, it can help restaurants to create new menus based on criteria from foods that are sold a lot.

References

- [1] Yudi A 2007 K-Means – Penerapan, Permasalahan dan Metode Terkait *Jurnal Sistem dan Informatika* pp. 47-60
- [2] Mehmed K 2003 *Data Mining Concepts, Models, Methods, and Algorithms*. (Hoboken: John Wiley & Sons)
- [3] Budi S 2007 *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis* (Yogyakarta: Graha Ilmu)
- [4] Nurul R W, Sofi D and Mohamad J 2015 Implementasi Algoritma K-Means Dalam Pengklasteran Mahasiswa Pelamar Beasiswa *JITTER Jurnal Ilmiah Teknologi Informasi Terapan* **1** pp 62-68
- [5] Johan O O 2013 Implementasi Algoritma K-Means Clustering Untuk Menentukan Strategi Marketing President University *Jurnal Ilmiah Teknik Industri* **12** pp 10-20
- [6] Green F M, Green A S, Susi S and Haryanto R M 2017 Penerapan Algoritma K-Means Untuk Analisis Prestasi Akademik Mahasiswa Fakultas Ilmu Komputer Universitas Klabat *Cogito Smart Journal* **3** pp 230-239
- [7] Fauziah N, Muhammad Z and Benny B N 2017 Penerapan Algoritma K-Means Pada Siswa Baru Sekolah Menengah Kejuruan Untuk Clustering Jurusan *InfoTekJar (Jurnal Nasional Informatika dan Teknologi Jaringan)* **1** pp 100-105
- [8] Jhonsen 2004 *Web Designer untuk Pemula* (Jakarta: Elex Media Komputindo)
- [9] Muhammad S and Rosa A S 2013 *Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek* (Bandung: Informatika)
- [10] Puspitosari H A 2011 *Pemrograman Web Database dengan PHP & MySQL* (Yogyakarta: Scripta)
- [11] Roger S P 2010 *Rekayasa Perangkat Lunak* (Yogyakarta: Penerbit Andi)
- [12] Rulianto K 2010 *PHP dan MySQL untuk orang Awam* (Palembang: Maxikom)

- [13] Saputra A 2011 *Trik dan Solusi Jitu Pemrograman PHP* (Jakarta: PT Elex Media Komputindo)
- [14] Wixom B H, Dennis A and Tegarden D 2005 *System Analysis and Design with UML version 2.0* (Virginia: Wiley-Interscience)