JUSIKOM PRIMA (Journal of Information Systems and Computer Science Prima)
Vol. 7 No. 1, August 2023 E-ISSN : 2580-2879

K-NEAREST NEIGHBOR (KNN) ANALYSIS FOR CLOTHING SALES CLASSIFICATION BASED ON MATERIALS USED

Christian Frederic Lumbantobing[1], Regina Tukashi Siregar[2], Muhammad Alfathan Ridho[3], Muhammad Fikri Akbar Zuhdi[4], Jepri Banjarnahor*[5]
[1.2.3.4]Indonesian Prima University
Jl. Sampul No.4, Sei Putih Bar., Kec. Medan Petisah, Medan City, North Sumatra
E-mail : * jepribanjarnahor@unprimdn.ac.id

ABSTRACT-Abstract Classification is a method to see the behavior and characteristics of certain groups. The K-nearest neighbor method is a learning algorithm for classifying new data based on the K-Nearest Neighbor majority class. The primary purpose of this algorithm is to classify new objects based on attributes and training samples. In today's digital era, competition in the business world is getting more challenging and overgrowing, especially regarding online marketing systems. Every market driver must always pay attention to the needs and desires of consumer satisfaction when buying products from online stores. However, the problem that consumers often complain about is the use of clothing size charts in online stores that do not match the consumer's body size. This study aims to reduce consumers' frustration buying clothes online so that products do not have to be returned via the Internet. Based on these problems, these conditions must be improved by selecting clothes to achieve optimal customer satisfaction. This application was built using the K-Nearest Neighbor (KNN) method and Profile Matching to help you determine the most suitable clothes for your consumer size.

Keywords: KNN, Classification, data, Internet.

1. INTRODUCTION

In any business, sales are significant for the industry to keep going and generate revenue to make various products to sell. Each company competes in product quality and quality so that sales continue to grow as expected. Businesses generally use product sales forecasts, or forecasts, to determine which products will outsell in the future.

Sales and advertising systems on the Internet are overgrowing. The company uses web-based technology as a business strategy to provide products to all consumers without being limited by space and time. The online sales system has become a cheap and affordable means of sales promotion for large small, and medium businesses. To maintain the quality of the fabrics that will be sold at Transmart stores, it is necessary to have a system for making sales based on the raw materials used.

Currently, many companies need to pay more attention to quality in doing business but are only concerned with profits; this type of business will cause a decrease in our business turnover, and the fatal effect is that the company we run can close. So with that, there needs to be a system to solve the problem above.

Previous related work carried out by research introduces the use of the KNN algorithm. This last article only explained the basic idea of the data mining extraction process with KNN, while the articles in this study process the extraction of furniture sales data mining to predict furniture sales results. Investigated the KNN method and its modified version to overcome the weaknesses of the KNN method to make it more efficient. This previous research did not develop a KNN application program, unlike the research in this article, which created an intelligent application program [7,8]. Predict sales of furniture items from a furniture company as a case study. This previous research indicates sales with the association method of company sales transactions [9]. Unlike the research in this article predicting furniture sales using the KNN method. In addition, the research in this article builds a website-based application program that states its advantages compared to previous research [10,12]. Analyzes data mining for predicting the future of work in a city using the KNN algorithm. The difference with this article is that the expected data mining and intelligent application systems that were developed in this previous article used the Python programming language [13,14], and the application system that was built was not website-based. In contrast, this article is website-based, which was created with the PHP programming language. From a review of several previous related works,

2. RESEARCH METHOD

2.1 Types of Research

Quantitative research systematically investigates a phenomenon by collecting data that can be measured using statistical, mathematical, or computational techniques.

This research aims to develop and use the K-NN method, theories, and hypotheses related to natural phenomena.

This type of research groups data by data mining with a more extensive database.

The methods contained in this type of quantitative research are:

- **Experimental Method**

This research method aims to examine the effect of a particular treatment on the symptoms of a specific group compared to other groups that use different medicines.

- **Survey Method**

Research is conducted on large and small populations. Still, the studied data are from samples taken from these populations to find relative occurrences, distribution, and relationships between sociological and psychological variables.

This type of quantitative research is a method of quantitative data mining and objective statistics through scientific calculations derived from data samples that are asked to answer several questions to determine the type of fabric to be increased inventory.

In this type of research, the instrument that is often used is a questionnaire or questionnaire. The contents of the discussion are also more on the surface and not in-depth, with more general problems having broad areas with complex levels of variation.

2.2 Work Procedures

The analysis process includes an analysis of criteria and an analysis of the selection process. It determines the type of fabric that is feasible to increase inventory using the k-Nearest Neighbor (K-NN) method and system modeling.

Figure 2.1 Research Framework

Information:

1. Lecturer Data Input
   This section is part of inputting the data of the goods to be processed
2. Input criteria
   This section will process inputting the criteria to get better results.
3. K-NN Method process
   In this section, we will work with the method that will be used; in this case, we will use the K-NN method.
4. Display the Test Result
   At this stage, it will test the system that has been made and whether it is feasible to implement.
5. Selection Result
   Selection results are the results of selection from a system made and executed by the K-NN method.

3. ANALYSIS AND DESIGN

3.1 Analysis

The analysis process includes an analysis of criteria and an analysis of the selection process for online sales, which can be selected according to the type of fabric to be used. This research has made a research flowchart so that the process of completing this system can be carried out according to the stage that has been determined. Can be seen in the image below.

Figure 3.1 System Flowchart

3.1.1 Criteria Analysis

The criteria used in the process of selecting the best fabric are as follows:

1. Budget, namely the budget owned by the customer for sales classification with criteria.

Table 3.1 Baby Canvas Criteria

Dari	Sampai	Nilai
13,000,000	15,000,000	1
15,000,001	20,000,000	2
20,000,001	25,000,000	3
25,000,001	30,000,000	4
30,000,001	35,000,000	5
35,000,001	50,000,000	6

2. Transmission, namely the transmission of the desired motorbike. The criteria used can be seen in Table 3.2.

Table 3.2 Denim Criteria

Transmisi	Nilai
Tebal	1
Tipis	9

3. Fuel consumption, namely the fuel oil consumption profile of the selected motorcycle. The criteria used can be seen in Table 3.3.

Table 3.3. Cotton Consumption Criteria

Katun	Nilai
Bebulu	1
Kasar	2
Tipis	3
Tebal	4

3.2 Result in Data From Program

In this study, the criteria that will be used in the classification system with the K-NN method have been determined, which can be seen in the image below:

Figure 3.2 Display Criteria

3.3 Data Sets Used

This system uses a dataset to guide researchers in making the system taken from research sites which can be seen in the image below.

Figure 3.3 Data Set Display

3.4 Calculations With Lift Each Criterion

So that this research can be carried out according to plan, the researcher will determine the results of the selection based on the K-NN calculation, which can be seen in the image below**.**

Figure 3.4 Image of known data

3.5 Data Normalization

After inputting the data, as shown in Figure 3.3, there will be a dataset normalization data table, as shown in the image below.

Figure 3.5 Display of Dataset Normalization

3.6 Nearest Data

After the data is normalized using the K-NN method, the next process is compiling the data with the nearest data.

Figure 3.6 Nearest data

3.7 Results of the above assessment

After carrying out the testing and normalization stages, the selection results are as follows.

Figure 3.7 Selection Results

4. CONCLUSIONS AND RECOMMENDATIONS

4.1 Conclusion

The conclusions of this study are as follows.

1. This system can determine the type of fabric that will be feasible to add stock so that consumer demand can be fulfilled
2. This system has succeeded in minimizing the company's work in determining the inventory of fabric types every month.

4.2 Suggestions

This research has been completed by the steps that have been determined, but always done research has drawbacks such as this system is still lacking in determining insufficient data with the type of fabric to be used, so future research is expected to be able to

find accurate data and use other methods such as fuzzy.

BIBLIOGRAPHY

[1] Alfani, A., Rozi, F., & Sukmana, F. (2021). Unilever Product Sales Prediction Using the K-Nearest Neighbor Method. JIPI (Scientific Journal of Informatics Research and Learning).

[2] Amalia Rizki, Y. (2018). Application of Data Mining for Predicting Best Selling Electronic Products Sales Using the K-Nearest Neighbor Method (Case Study: PT. Bintang Multi Sarana Palembang). Retrieved from http://eprints.radenfatah.ac.id/3302/

[3] Annur, H. (2018). Classification of the Poor Using the Naïve Bayes Method. ILKOM Scientific Journal, 10, 6. Retrieved from http://103.226.139.203/index.php/ILKOM/article/view/303

[4] Bode, A. (2017). K-Nearest Neighbor With Feature Selection Using Backward Elimination To Predict Arabica Coffee Commodity Prices. ILKOM Scientific Journal, 9(2), 188–195. https://doi.org/10.33096/ilkom.v9i2.139.188-195

[5] Cholil, S., Handayani, T., & Tria, A. (2021). Implementation of the K-Nearest Neighbor (KNN) Classification Algorithm for Scholarship Recipient Selection Classification. IJCIT (Indonesian Journal on Computer and Information Technology), 6(2), 118–127.

[6] Hakim, LAR, Rizal, AA, & Ratnasari, D. (2019). K-Nearest Neighbor (K-NN) Based Student Graduation Prediction Application. JTIM : Journal of Information Technology and Multimedia, 1(1), 30–36. https://doi.org/10.35746/jtim.v1i1.11 [3] Jauhari, nurdin, Edge Linking Detection and Comparison of the 3 Methods,http://ahtovicblogs.blog.ug.ac.id/?p=80, March 17, 2011 20.25

[7] Harun, R., Pelangi, K., & Yuliyanti, L. (2020). Application of Mining Data to Determine Daily Rain Potential Using the K Nearest Neighbor (KNN) Algorithm. MISSION (Journal of Information Management & Information Systems). Retrieved from http://ejournal.stmiklombok.ac.id/index.php/misi/article/view/125/84

[8] Hermawan, F., & Agung, H. (2017). Implementation of the K-Nearest Neighbor Method in the Sales Data Application of PT. Multitek Mitra Sejati.

[9] Wahyuningsih, S., & Utari, DR (2018). Comparison of the K-Nearest Neighbor, Naive Bayes and Decision Tree Methods for Predicting Creditworthiness. Information Systems National Conference 2018 STMIK Atma Luhur Pangkalpinang, 8 – 9 March 2018, 619–623.

[10] Lauwrenza, X., Fitriyah, H., & Syauqi, D. (2021). Design of a Shirt Size Classification System based on Body Size with the Arduino-based K-Nearest Neighbor Method. Journal of Information Technology Development and Computer Science, 6(1), 74-81.

[11] Nikmatun, IA, & Waspada, I. (2019). Implementation of Data Mining for Classification of Student Study Period Using the K-Nearest Neighbor Algorithm. Symmetric: Journal of Mechanical Engineering, Electrical and Computer Science, 10(2),421-432.

[12] Rizaldi, R., Kurniawati, A., & Angkoso, CV (2018). Implementation of the Euclidean distance method for clothing size recommendations in virtual dressing room applications. Journal of Information Technology and Computer Science, 5(2), 129-138.

[13] Yolanda, Ike and Hasanul Fahmi, 2021, Application of Data Mining for Predicting Sales of Best Selling Bread Products at PT. Nippon Indosari Corpindo Tbk Using the K-Nearest Neighbor Method, Journal of Computer Science and Information Systems. E-ISSN: 2723-6129. pp. 9-15.

[14] Widaningsih, Sri, 2019, Comparison of Data Mining Methods for Predicting Value and Graduation Time of Informatics Engineering Study Program Students with the C4.5 Algorithm, Naïve Bayes, K-NN and SVM, Incentive Tekno Journal. ISSN(p):1907-4964. Vol.13, No. 1 (pp. 16-25). [4]Jose, Stephane, Why Should I Care About SQL Server,http://blog.iweb.com/en/2010/06/why-should-i-care-about-sql-server/4772.html, March 17, 2011 19.30

[15] Umam, Khaerul and Muhammad Hilman Fakhriza, 2021, Analysis of Best Selling Products Using the K-

Means Clustering Method at PT. Sukanda Jaya, Journal of Informatics. E-ISSN: 2722-2713. p. 8-15.

[16] Meliala, Meilida Dina, Penda Hasugian, 2020, Comparison of the K-Nearest Neighbor Algorithm with a Decision Tree in Predicting Pet Food Sales at Petshop Dore Vet Clinic, Information Technology Journal. ISSN: 1907-2430. Vol. XV (pp. 35-39).

[17] Pradnyana, Aditra, Gede, Agus Aan Jiwa Permana, 2017, Comparison of K-Means and Hybrid K-Means KNN Algorithms for Class Division of Student Lectures, National Innovative Research Seminar. ISBN: 978-602-6428-11-0 (pp. 941-949).