

Analysis of the K-Nearest Neighbor Algorithm to Determine the Prediction of Tofu Production

Munif^{1*}, Mustain², Kurnia Yahya³

^{1,2} Teknik Informatika, Fakultas Teknik, Universitas Islam Lamongan, Indonesia

³SMK Muhammadiyah 1 Lamongan

^{1*}munif@unisla.ac.id, ²mustain@unisla.ac.id, ³yahya@unisla.ac.id

Article history:

Received 12 December 2022
Revised 29 December 2022
Accepted 29 December 2022
Available online 29 December 2022

Keywords:

K-Nearest Neighbor,
Application,
Tofu Production,
Prediction of Production
Amount.

Abstract

The tofu factory in Kedungpring does not yet have a prediction system to estimate the number of tofu that will be predicted for the next month. As a result, companies cannot meet market demand in a timely and appropriate amount. Therefore, it is necessary to make a prediction system to determine the amount of tofu production in Kedungpring District. In this research, the application of K-Nearest Neighbor Algorithm Analysis Application to determine the prediction of the number of Web-Based production to make it easy to predict the number of tofu production. The system functional test results show that all features in the application are able to run properly and functionally. Testing the accuracy of the prediction system K-Nearest Neighbor algorithm to determine the prediction of the number of web-based tofu production that can produces a MAPE of 0.68%.



This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2021 by author.

I. INTRODUCTION

The At this time companies engaged in the industrial sector are faced with a problem, namely the increasing competition in production, this requires companies to plan or determine the amount of production so that they can meet market demand in a timely manner and in the right amount. Consumer demand that changes every day and even every month can be a serious problem for the company. Companies sometimes cannot meet market demand because demand increases and there is no supply [1][2]. Companies can also experience losses due to decreased market demand so that production in the Kedungpring sub-district cannot be marketed and increases storage and warehousing costs, companies need a method that can optimize the amount of production every day or even every month, to avoid these problems In the tofu factory in the Kedungpring sub-district there is a problem where in producing tofu there is no way to predict how much will be produced in the following month, so a writer will develop a prediction system to determine the amount of tofu production[3][4].

^{1*} Corresponding author

One manner in which can be used is forecast where forecast can be specified, specifically a procedure of methodically estimating something that's probably to occur in the future based upon previous and present information that's owned, so that the mistake (distinction in between something that happened and the anticipated outcome) can be reduced [5].

From this explanation, the K-Nearest Neighbord method is used to predict the amount of tofu production in the Kedungpring sub-district based on the web with the hope that later it will make it easier for tofu entrepreneurs to produce tofu and be able to meet market needs [8].

II. METHODS

The K-Nearest Neighbor algorithm (k-NN or KNN) is a technique for categorizing objects based upon learning information that's closest to the item [8][9].

The point whose new classification is predicted is included in the classification of many of these points.

$$distance = \sqrt{\sum_{i=1}^n (X_{training}^i - X_{testing})^2}$$

$X_{training}^i$: training data i

$X_{testing}$: data testing

I: the Ith record (row) of the table

N: amount of training data

Functional requirements contain requirements that include what processes will be carried out by the system [10][11]. The functional requirements of the prediction system for the amount of tofu production have several features that will be included in the system to be made, which include the following:

1. Login menu
2. Menu Admin (user)
3. Production Data Input Menu
4. Menu Predict the amount of production
5. Accuracy Menu

Non-functional requirements are requirements outside the system, which in this requirement include requirements when making and running applications [3][12]. Hardware Requirements (Hardware). In this requirement required some hardware (Hardware). When creating the application:

1. CPU Specifications: Processor Core i3, Minimum 4GB of RAM
2. Monitors
3. Mouse and keyboard
4. Smartphones

Software is needed to be able to process data. The web-based information system for the amount of tofu production in Kedungpring District is made using the following software:

1. Operating system (Windows 7, 8, 10)
2. Sublime Text used for the application script that was created

3. Google Chrome as the media used in running the application
4. Xampp as a web server maker that is used in the process of building and debugging applications.

III. RESULTS AND DISCUSSIONS

A. System Test Results Data

In the experimental results, it can be discussed the outcomes of the application of the program thoroughly in accordance with the design that is made, an application program is made based on the aim of providing convenience to the user and carrying out a service process, therefore a trial is carried out on the program that's made, whether the program this can function and can accomplish the objective as expected or otherwise. From the trial results it can be seen that this system really produces according to the desired goals, and a program must go through the refinement stage before it can be operated:

B. System Functionality Test Results

The speculative information explains the outcomes of the program application thoroughly in accordance with the design that is made. The stages carried out in the application trial use the Black-Box testing method. Black-box testing tries to find errors such as functional errors and application display errors. Black-box testing can be used to test conventional applications and object-oriented applications. It can be seen that all the features being tested which include the login menu, production data page, prediction calculation page, accuracy calculation page and logout menu are all running well and functioning, as it should be.

C. System Design

System design is the stage of planning and making a design by bringing together several separate components into a unified whole to clarify the shape of a system

1. Flowcharts

Flowchart or flow chart is a type that represents a process algorithm. Where in the flowchart will explain all the web flows.

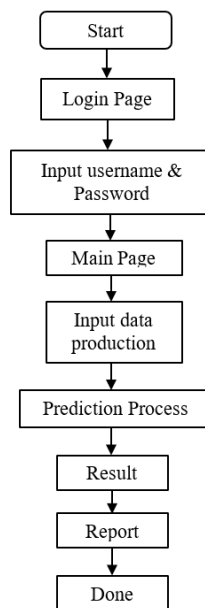


Figure 1. Flowchart System

In Figure 1, The flowchart explains the running of the system, if we are going to predict a tofu

production in the coming month we first log in first then enter the next data in the process to be predicted after that the results come out and the report is complete.

2. Data Flow Diagrams (DFD)

Data flow diagrams are made to describe where the data comes from and where the data will go out of the system, where the data is stored, what process produces the data and the interaction between the stored data and the process flow for predicting the amount of production. The following is the DFD (data flow diagram) stage.

a. Context Diagram

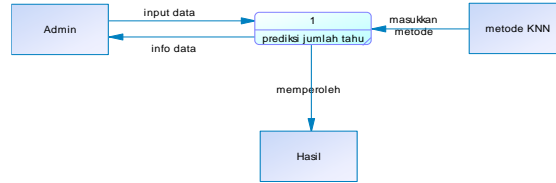


Figure 2. Context Diagram

Figure 2 From the image above the admin enters data that will be predicted and processed by the system to generate predictions for the following month.

b. Data Flow Diagram Level 0

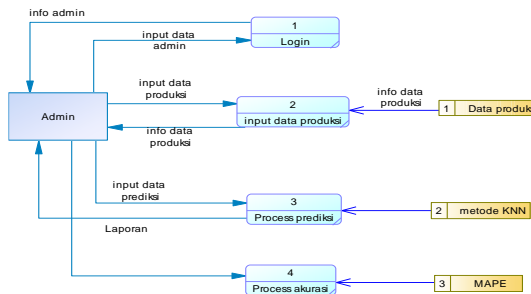


Figure 3. Data Flow Diagram Level 0

Figure 3 From the Data Flow Diagram above, the system flow can be explained as follows:

- a) Process 1: Admin performs the Login Process as an Administrator
- b) Process 2: Admin inputs production data, Admin can also edit and delete production data.
- c) Process 3: Admin performs the process of calculating the amount of tofu production.

c. Data Flow Diagram Level 1

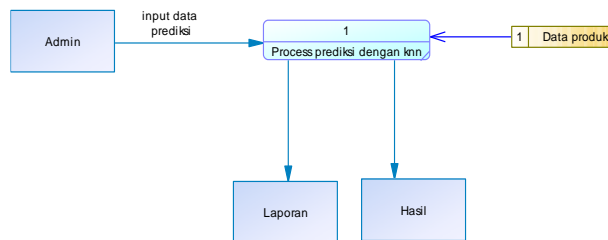


Figure 4. Data Flow Diagram Level 1

In the picture above, it explains about the admin entering predictive data and processing it using the KNN method after that the data enters temporary storage after which results and reports appear.

3. Conceptual Data Model (CDM)

CDM is used to describe in detail the CDM database structure consisting of objects that are not implemented directly into the actual database.

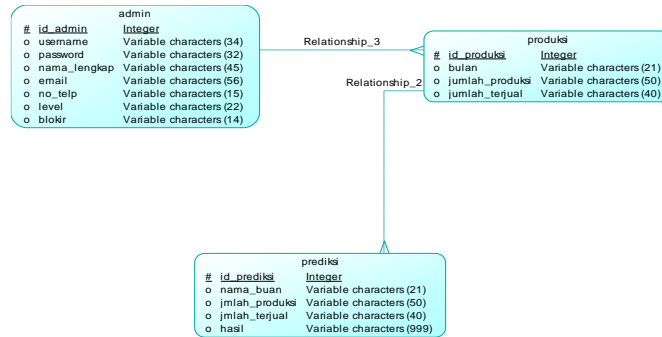


Figure 5. Conceptual Data Model (CDM)

The CDM above is the CDM of the prediction application for the amount of tofu production which explains that the admin can input production data then the data is predicted to determine the amount of production in the following month.

4. Physical Data Model (PDM)

PDM is an image of a running CDM showing the correct data storage structure in the database that is actually used.

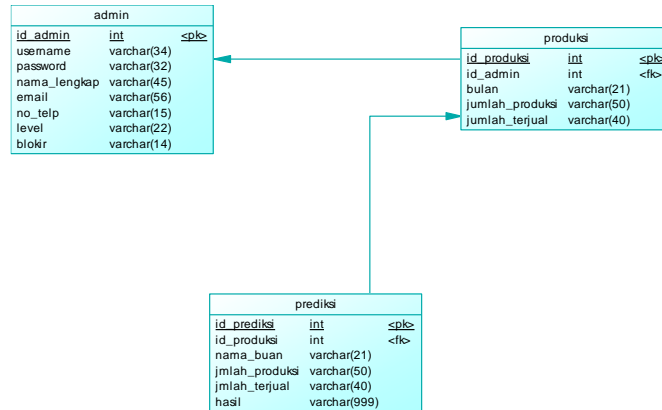


Figure 6. Physical Data Model (PDM)

In the PDM image above, there are 4 tables showing the generet results from the Conceptual Data Model that was created previously.

D. Discussion

With testing we can find out whether there is significant knowledge about this system. This section is part of the discussion along with the things that happen when the process is in progress.

1. Admin Page

The appearance of the admin page is the page that appears when the admin has successfully logged in and on that page there are several menus that can be accessed by users after logging in.

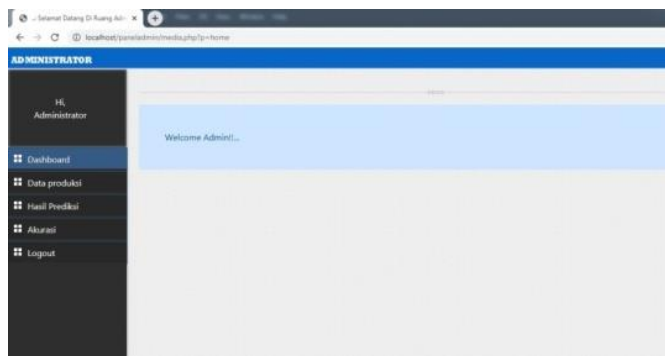


Figure 6. Admin page

2. Production Data Pages

The production data page display is a page that contains production data that has been input, the data of which was taken from January to December in 2019. On this page, users can enter, edit and delete data.

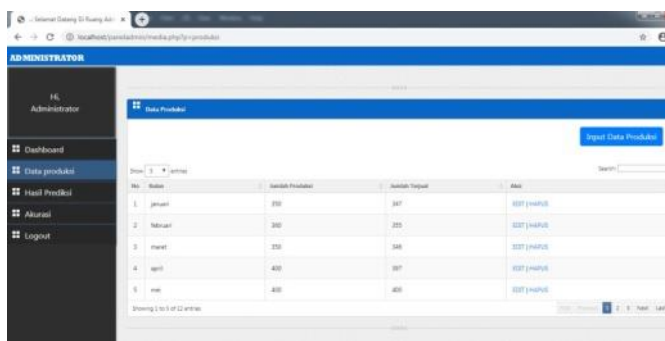


Figure 7. Production Data Pages

3. Add Production Data Pages

The display of the added production data page is a page that contains orders to enter the month, Production Amount, Sold Amount then there is a save button where the button can save the data that has been inputted, then the data is automatically saved into the database.

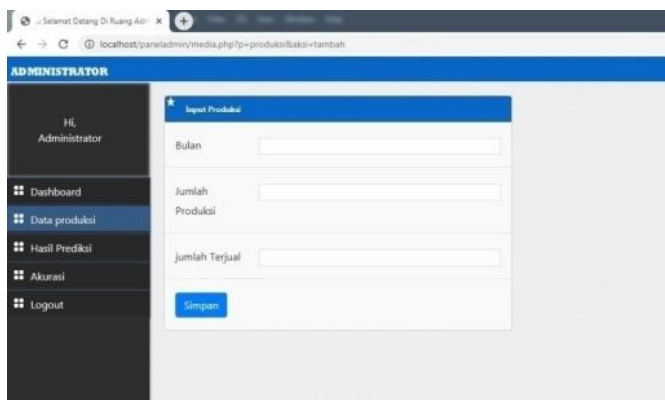


Figure 8. Add Production Data Page

4. Prediction Calculation Pages

On the prediction calculation page there is an order to enter the month, production amount, amount sold then there is a calculate button where the calculate button can calculate how much production will be in the coming months and the results will appear.



Figure 9. Prediction Calculation Page

5. Accuracy Page Discussion

The accuracy of the prediction system analysis of the K-Nearest Neighbor Algorithm for Predicting the Amount of Tofu Production (Case Study: Kedungpring District) Web-based is classified as high with a MAPE of 0.68 or with an accuracy level of 99.32%, this accuracy value has an average accuracy, in previous research that is equal to 86.2%.

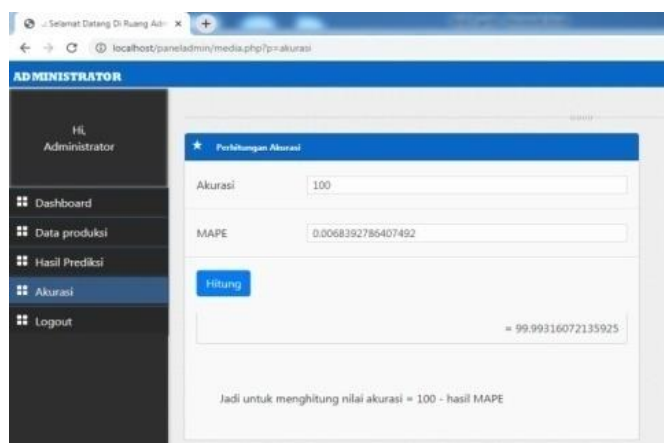


Figure 10. Page Accuracy

IV. CONCLUSIONS AND RECOMMENDATIONS

Based on implementation and testing of the prediction system for the amount of tofu production in Kedungpring District, it can be concluded as follows: Web-Based Tofu Production Prediction Application (Case Study: Kedungpring District) with the K-Nearest Neighbor method has been successfully made and can run well. Web-based Tofu Production Prediction Application (Case Study: Kedungpring District) using the K-Nearest Neighbor method produces a MAPE of 0.68% and an accuracy rate of 99.32%. In the results of making the K-Nearest Neighbor Algorithm Analysis application to determine the Predicted Amount of Tofu Production (Case Study: Kedungpring District) Web-based can make it easier to predict the amount of Tofu production in Kedungpring District.

V. REFERENCES

- [1] D. Satriawan, H. Fitriyah, and A. S. Budi, "Sistem Klasifikasi Tahu Putih Murni dan Tahu Putih Mengandung Formalin Menggunakan Metode K-Nearest Neighbor," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 10, pp. 10287–10293, 2019.

- [2] R. K. Niswatin, R. Wulanningrum, and U. Syaidah, "Sistem Prediksi Nilai Ipk Mahasiswa Menggunakan Metode K-Nearest Neighbor," *J. Maklumatika*, vol. 3, no. 1, pp. 56–66, 2016.
- [3] A. J. T, D. Yanosma, and K. Anggriani, "Implementasi Metode K-Nearest Neighbor (Knn) Dan Simple Additive Weighting (Saw) Dalam Pengambilan Keputusan Seleksi Penerimaan Anggota Paskibraka," *Pseudocode*, vol. 3, no. 2, pp. 98–112, 2017, doi: 10.33369/pseudocode.3.2.98-112.
- [4] M. R. Fauzi, R. A. Pratama, P. Laksono, and P. Eosina, "Penerapan Big Data Menggunakan Algoritma Multi-Label K-Nearest Neighbor dalam Analisis Sentimen Konsumen UMKM Sektor Kuliner," *Krea-TIF*, vol. 9, no. 1, p. 9, 2021, doi: 10.32832/kreatif.v9i1.3587.
- [5] Y. Yahya and W. Puspita Hidayanti, "Penerapan Algoritma K-Nearest Neighbor Untuk Klasifikasi Efektivitas Penjualan Vape (Rokok Elektrik) pada 'Lombok Vape On,'" *Infotek J. Inform. dan Teknol.*, vol. 3, no. 2, pp. 104–114, 2020, doi: 10.29408/jit.v3i2.2279.
- [6] M. M. Baharuddin, H. Azis, and T. Hasanuddin, "Analisis Performa Metode K-Nearest Neighbor Untuk Identifikasi Jenis Kaca," *Ilk. J. Ilm.*, vol. 11, no. 3, pp. 269–274, 2019, doi: 10.33096/ilkom.v11i3.489.269-274.
- [7] D. Yunita, "Perbandingan Algoritma K-Nearest Neighbor dan Decision Tree untuk Penentuan Risiko Kredit Kepemilikan Mobil," *J. Inform. Univ. Pamulang*, vol. 2, no. 2, p. 103, 2017, doi: 10.32493/informatika.v2i2.1512.
- [8] P. Putra, A. M. H Pardede, and S. Syahputra, "Analisis Metode K-Nearest Neighbour (Knn) Dalam Klasifikasi Data Iris Bunga," *J. Tek. Inform. Kaputama*, vol. 6, no. 1, pp. 297–305, 2022.
- [9] S. Ernawati and R. Wati, "Penerapan Algoritma K-Nearest Neighbors Pada Analisis Sentimen Review Agen Travel," *J. Khatulistiwa Inform.*, vol. 6, no. 1, pp. 64–69, 2018.
- [10] L. Farokhah, "Implementasi K-Nearest Neighbor untuk Klasifikasi Bunga Dengan Ekstraksi Fitur Warna RGB," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 6, p. 1129, 2020, doi: 10.25126/jtiik.2020722608.
- [11] W. Yustanti, "Algoritma K-Nearest Neighbour untuk Memprediksi Harga Jual Tanah," *J. Mat. Stat. dan komputasi*, vol. 9, no. 1, pp. 57–68, 2012.
- [12] S. Sumarlin, "Implementasi Algoritma K-Nearest Neighbor Sebagai Pendukung Keputusan Klasifikasi Penerima Beasiswa PPA dan BBM," *J. Sist. Inf. Bisnis*, vol. 5, no. 1, pp. 52–62, 2015, doi: 10.21456/vol5iss1pp52-62.