

Design and Implementation The Learning of Classification Aromatherapy made from Indonesian Spices using K-Nearest Neighbor (KNN)

Maulia Wijiyanti Hidayah^{*}, Muhammad Ashar, I Made Wirawan

Department of Electrical Engineering, State University of Malang Malang, Indonesia

**Email: mauliahidayah21@gmail.com*

Abstract

Indonesia is country that has various plants which have many benefits for human. There are more than 31 types of medicinal plants as one of material that needed by industry as traditional medicine and spices. Traditional spiceal medicine comes from spiceal plants that used from Indonesian spices. This spices can produce aromatherapy include essential oils. Aromatherapy can help in maintaning the healthy of human body. This is necessary that aromatherapy can be classfified using K-Nearest Neighbor to classify aromatherapy from Indonesian spices. The accuracy result which shown by K-Nearest Neighbor algorithm is 97.5%, the accuracy data testing using confusion matrix which will be followed by front end and back end testing that shown the valid result for application design and valid using weka application with an accuracy result of 97.5%. This research will produce product such as android application that can be accessed by android users.

Keywords: aromatherapy, spices, K-Nearest Neighbor.

I. INTRODUCTION

Indonesia is country that has various plants with many benefits. There are 31 species of medicinal plants as raw material of traditional medicine (spiceal medicine), industrial non spices and spices. This spice plant can produce essential oils. The native Indonesian spices plants that can be used as essential oils are clove, cajuput, cinnamon, nutmeg [1]. According to Ministry of Trade of the Republic of Indonesia, essential oils are most widely used as fragrances, medicine and to add the taste of food. In addition essential oils is one of compound of aromatherapy products [2].

Aromatherapy is consisted of “aroma” that means fragrant and “therapy” that means treatment. So aromatherapy can be interpreted as complementary therapy that is used to maintain healthy body from certain diseases [3]. In the beginning aromatherapy aims to improve the mood, change the cognitive area, but now aromatherapy is used as an additional medication. Aromatherapy has a wide range of products such as medicinal oil, soaps and perfumes. All of these products contain different spices that spice will be classified appropriately in order to produce aromatherapy products.

Classification is the process of finding model that can distinguish class of data and concepts in order to be used in predicting the class of object whose class label is unknown [4]. Classification can also be interpreted as a work of assessing the data object, then inserted into certain class of several classes which available. Classification can be conducted in order to determine in which class of data object that want to be known [5]. To get the right kind of aromatherapy, spices ingredients are used can be classified.

Aromatherapy classification made from Indonesian spices has not been studied for now, and therefore to overcome this problem would be used technology that can classify aromatherapy products made from spices. The right technology to classify aromatherapy products can use the method of K-Nearest Neighbor (KNN). This study aims to classify aromatherapy products made from spices using K-Nearest Neighbor method.

II. METHODOLOGY

A. Research Design

Waterfall model is used to design an aromatherapy classification system. Waterfall development model can be used for system which is generic, it can be identified all of the requirements needed from the beginning until the system requirements gathering system built according to research topics are built up to the system under test. [6]. Waterfall which will be used are shown on Figure 1.

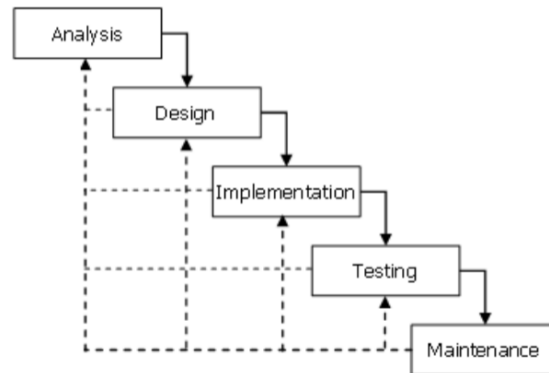


Figure 1. Waterfall Model [7]

(1) Analysis is an analysis process to the needs of the overall system as service system needed, constraints, and system specification details. (2) Design is detailed design process of the system by allocating the overall system requirements. (3) Implementation is process of implementing the design by using the program, write the program using language program required. (4) Testing or experiments is stage of system testing to determine the performance of the system is in accordance with the requirements needed. (5) Maintenance treatment is the last stage system, that is maintaining the system, so that the system remains running in accordance with the requirements needed.

B. Data Research

Aromatherapy of data research made from Indonesian spices is downloaded from the official data website from the Ministry of Agriculture of Republic Indonesia which is managed by the Council of Essential Indonesia <http://www.atsiri-indonesia.com/data-atsiri.php>

C. Application Design

In this study, there are several design system include Usecase Diagram and Data Flow Diagram.

1) Usecase Diagram

Here is the Use Case diagram which explain the application used by the user is the application of Aromatherapy Classification Made from Indonesian Spices using K-Nearest Neighbor algorithm.

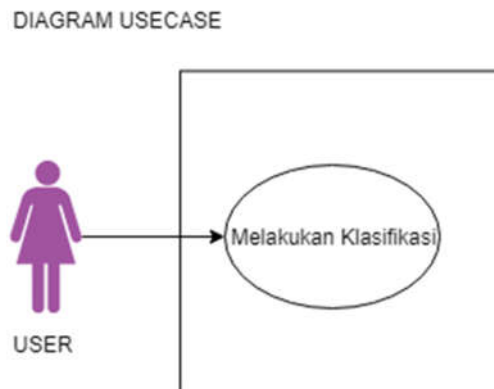


Figure 2. Usecase Diagram

2) Data Flow Diagram level

Data Flow Diagrams which is used to determine the flow of application data of Spices classification, in this study will be described on the context diagram and data flow diagram 0 and data flow diagram level 1.

- Context Diagram

The relationship between processes and external entities described on a simple diagram as shown on Figure 3.

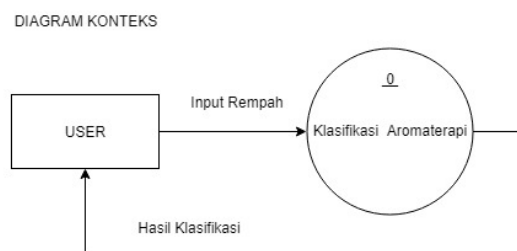


Figure 3. Context Diagram

- Data Flow Diagrams

Flow data that has been made in context diagram will be described to provide detailed data flow so that the data flow can be seen clearly. In aromatherapy classification system has two main processes for member registration process and classification process. Here are the details of the system which can be seen in the Data Flow Diagram Level 1 as shown on Figure 4.

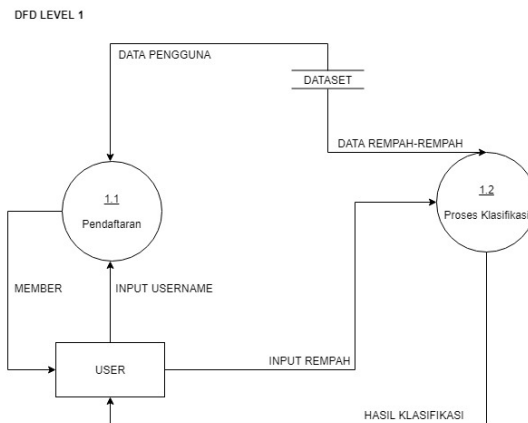


Figure 4. Data flow diagrams level 1

D. K-Nearest Neighbor

K-Nearest Neighbor algorithm which includes in the group instance-based learning. With lazy learning techniques sehingga K-Nearest Neighbor has the advantages of being able to classify large amounts of data, and tough in the training data that has a lot of noise [8]. Here is the process of K-Nearest Neighbor algorithm as shown on Figure 5.

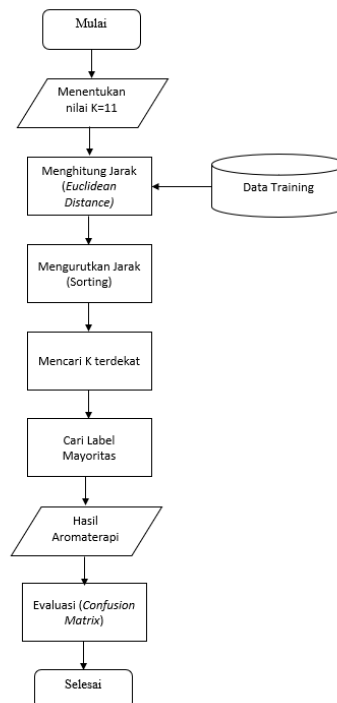


Figure 5. Classification Process

- 1) Determine the parameter of $K = 11$
- 2) Calculate the distance between data which will be evaluated with all training data
- 3) Sort range formed (ascending)
- 4) Determining the closest distance to the order of K
- 5) Pair the corresponding class
- 6) Finding the number of classes from the nearest neighbor and set the class as class data which will be evaluated.

III. RESULT AND DISCUSSION

Tests were conducted to determine differences in various K-Neighbor value that may affect the results of algorithm method K-Nearest Neighbor (KNN) performance on the data. K-Neighbors values used are 5,7,11,13,17,19,23,29,31,33.

Table 1.

Table of K Value Test

No.	K Value	Accuracy
1	5	97.5%
2	7	97.5%
3	11	97.5%
4	13	93.75%
5	17	93.75%
6	19	93.75%
7	23	91.25%
8	29	91.25%
9	31	91.25%
10	33	73.75%

Then the researcher will use the highest K value and the highest accuracy result is the K value = 11 with an accuracy result of 97.5%.

According to the results of aromatherapy classification using K-Nearest Neighbor (KNN) algorithm can be calculated the value of accuracy and error rate using the confusion matrix. Aromatherapy which yielded has 3 types of categories then the confusion matrix used is confusion matrix 3x3. Here is the result of the confusion matrix that can be seen in Table II.

Table 2.

Result of Confussion Matrix

		<i>Predict class</i>		
		Medicinal oil	Perfume	Soap
<i>Actual class</i>	Medicinal oil	20	0	0
	Perfume	1	43	1
	Soap	0	0	15

Table 2. shows the confusion matrix, according to that can be seen that the number of medicinal oil is 20 data, perfume is 45 data and soap is 15 data. The table above can be used to measure the performance of K-Nearest Neighbor method using equation (2) for accuracy and equation (3) for the error rate, as follows.

$$Accuracy = \frac{20+43+15}{80} \times 100\% = 97,5\% \quad (2)$$

$$Error\ rate = \frac{1+0+1}{80} \times 100\% = 2,5\% \quad (3)$$

Application design testing using testing front end and back end. Testing front end is one form of testing to see application design. By using the new data validation of 100 users who install the application data show that every aspects on user questionnaire that includes aspects of software, visual communication, and presentation of information have been met and got valid predicate. So the application has demonstrated the functionality and interface applications can comply with the specifications and procedures that have been designed.

The next test using testing back end. This test serves to equalize whether computational algorithm in the application system is working properly. At the back end testing processes using auxiliary tools called Weka. Weka has proven to be a reliable tool in computing algorithms. Data processes using the tools weka as shown on Figure 6.

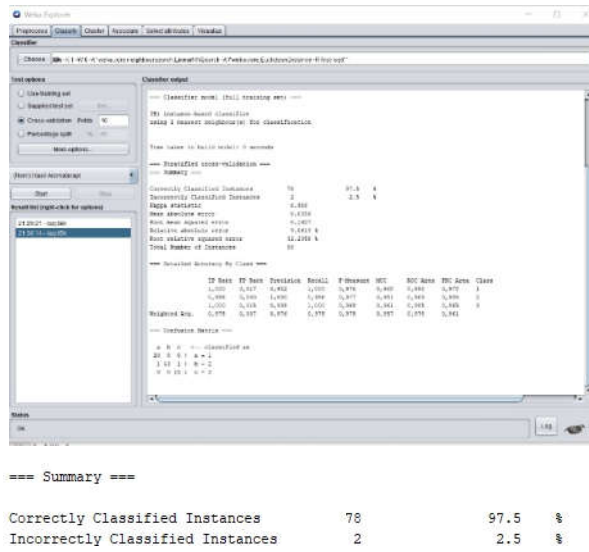


Figure 6. Testing data using application weka

Information obtained according to the experiment result on application of accuracy weka shown the accuracy up to 97.5%

In creating the application process, will be displayed the scheme of application design, before displaying the scheme of application design will be shown the flowchart, as in Figure 7.

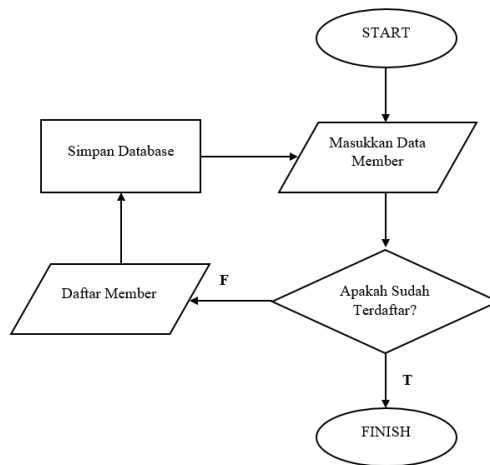
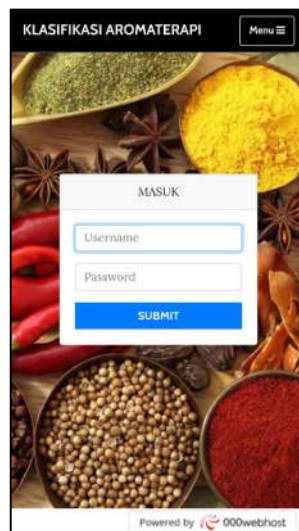


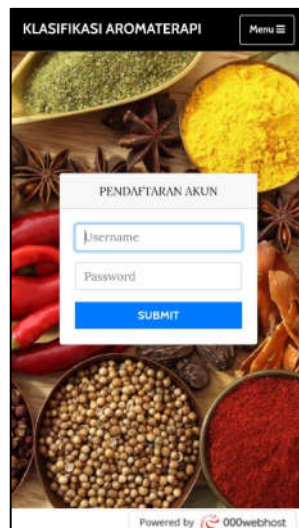
Figure 7. Flowchart Member

1. Start
2. Enter member data
3. If already registered, if yes then finished if not will go to the list members.
4. Store the new member data into database.
5. Repeat step to enter member data
6. Operate application system

1) *Login Interface*



2) *Account Registration Interface*



The following scheme design demonstrates the involved classification in the Figure 8.

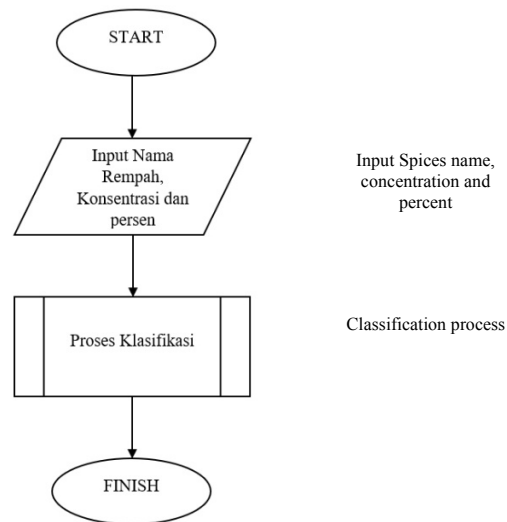


Figure 8. Flowchart classification menu

1. Start
2. User entering data: spices name, concentration and percent to be used
3. The classification process uses the K-Nearest Neighbor application.
4. In the process of classification will be determined value of $K = 11$, looking for euclidean distance, sort and search results that is closest to being searched.
5. Result found

3) *Display menu classification interface*

The screenshot shows the 'KLASIFIKASI AROMATERAPI' interface. At the top, there is a black header with the text 'KLASIFIKASI AROMATERAPI' and a 'Menu' button. Below the header, the title 'KLASIFIKASI AROMATERAPI BERBAHAN REMPAH-REMPAH INDONESIA' is displayed. The form contains two dropdown menus: 'Nama Rempah' with the value 'Adas' and 'Jenis' with the value 'Bubuk'.

Konsentrasi

PROSES



4) Process Classification Results Interface

KLASIFIKASI AROMATERAPI Menu

Nama rempah : cengkeh
Jenis : Cair
Persentase : 17%
Hasil klasifikasi : sabun
Akurasi : 97.5 %

CONFUSION MATRIX :

20	0	0
1	43	1
0	0	15

Maaila Wijayanti Hidayah
SI Teknik Informatika 2014, Universitas Negeri Malang

Powered by 000webhost

IV. CONCLUSION

According to the research result that have been conducted, can be concluded the following conclusions:

- a. K-Nearest Neighbor method can be used to classify aromatherapy made from Indonesian spices. Testing K neighbors which have been tested as many as 10 types of K neighbors on K-Nearest Neighbor algorithm that has been conducted to get the best K value is 11. Then researcher used the highest K value and the highest accuracy result with K value = 11 is 97.5%.
- b. Accuracy Results calculated using confusion matrix, shows accuracy of 97.5% and error rate of 2.5%
- c. Testing results of front end and back end design shows application design and system running includes into valid category, and testing using credible tool “weka” shows the same accuracy result that is 97.5%.

References

- [1] Ayu. F C, Susilawati. Y, “*Aromaterapi Asli Indonesia Sebagai Alternatif Pengobatan*”, Farmaka, Volume 4 No 3. Suplemen 1, 2013.
- [2] Pribadi. E R., “*Pasokan dan Permintaan Tanaman Obat Indonesia Serat Arah Penelitian dan Pengembangannya*”, ISSN: 1412-8004, 2012.
- [3] Ali. B, Al-wabel. N, Shams et al. S, “*Essential oil used in aromatherapy : A systemic review, Elsevier (Singapore)*” Pte Ltd, 2015.
- [4] Leidiyana. H. “*Penerapan Algoritma K-Nearest Neighbor Untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor.*” Jurnal Penelitian Ilmu Komputer, System Embedded & Logic Volume 1 Nomor 1, ISSN : 65-76, 2013.
- [5] Putri. Riyan E, Suparti, Rahmawati. Rita, “*Perbandingan Metode Klasifikasi Naïve Bayes Dan K-Nearest Neighbor Pada Analisis Data Status Kerja Di Kabupaten Demak Tahun 2012*”, Jurnal Gaussian Volume 3 Nomor 4 , ISSN : 2339-2541, 2014.
- [6] Susanto. R, Andriana. A. D., “*Perbandingan model waterfall dan prototyping untuk pengembangan sistem informasi*”, Majalah Ilmiah UNIKOM, Vol. 14 No. 1, 2014.
- [7] Bassil. Y, “*A Simulation Model For The Waterfall Software Development Life.*” International journal of Engineering & Technology (iJET), 2(5), 2012.

- [8] Ginting. S.L.B, Zarman. Z, Darmawan. A, “*Teknik Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma K-Nearest Neighborhood*”, Jurnal Teknik Komputer Unikom Vol 3 No. 2, 2014.