ISSN: 2477-5126
e-ISSN: 2548-9356

EXPERT SYSTEM FOR IDENTIFICATION OF SKIN DISEASE IN HUMANS USING NAIVE BAYES CLASSIFIER METHOD ON WEB

Dinar Nugroho Pratomo[1], Diyah Utami Kusumaning Putri [2]
[1]Jurusan Teknologi Elektro dan Informatika, Sekolah Vokasi, Universitas Gadjah Mada, Yogyakarta
[2]Jurusan Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Gadjah Mada, Yogyakarta
[1]Sekip Unit 1 Catur Tunggal, Yogyakarta, 55281, Indonesia
[2]Sekip Utara Bulaksumur, Yogyakarta, 55281, Indonesia
email: [1]dinar.nugroho.p@ugm.ac.id, [2]diyah.utami.k@ugm.ac.id

***Abstract*—Expert System Identification of Skin Disease in Human is an application that adopts expert knowledge, in this case, a dermatologist, to identify skin diseases in humans found in Indonesia. This expert system is constructed using a naive bayes classifier. To make the system interface more dynamic and easy to access by anyone, this system is made based on the web. Through this web application, users can consult with a system like an expert to know the skin disease that occurs to the human and find the correct treatment to solve the problems.**

***Keywords*—expert system, skin disease, naive bayes classifier, web.**

I. INTRODUCTION

The skin is the largest organ of the human body composed of the outermost and covering the entire surface of the body . Because of its outermost location, the skin first receives stimuli such as touch stimuli, pain, and bad influences from the outside. Functions of the skin include protecting the body surface, maintaining body temperature, and removing certain impurities. Skin is also important for the production of vitamin D by the body that comes from ultraviolet light. Given the terms of skin as a protective organ of the body in it`, then the skin should be maintained health. Apart from being a body armor, the skin also has an aesthetic value. Clean and well-groomed skin will look beautiful.

Disorders of the skin often occur due to various factors, among others are climate, living environment, unhealthy living habits, allergies, and others. Knowledge of skin disease is needed to solve the skin disease problem quickly and precisely. Although sometimes people consider it trivial, skin disorders can be very dangerous if wrong in the treatment. For that treatment should not be done carelessly.

*) **penulis korespondensi**: Dinar Nugroho Pratomo
Email: dinar.nugroho.p@ugm.ac.id

The development of information technology is currently very fast development and also plays an important role in various human activities. Almost all areas utilize information technology to help complete human work. Similarly in the medical world, knowledge-based technology, facts and reasoning to help solve problems in various disciplines have been developed. An expert system is a system that can capture and use the concept of choice used by one or several experts in a particular field. It can collect and store it from someone or several people perish in a knowledge base and build a criminal system that is very important in solving the problem. Until now there have been some results of the development of expert systems in various fields according to one's expertise, for example in the fields of education, agriculture, animal husbandry, and in the field of medicine.

The goal of expert systems is not to replace the role of humans, but a variety of material human knowledge which is presented in the form so that it can be in use by many people[7]. Some previous research on expert system give good result to finish problems that use complex data in the case of prediction or diagnosis, for example, an expert system based on certainty factors for early detection of red chili disease[8] and internal diseases[9] and naive bayes for detection of corn disease[10-11]. the naïve bayes method also used in diagnosis of diseases of the diabetes mellitus using 140 data obtained an accuracy rate of 100%, disease in cats using Naïve Bayes and Certainty Factor method with 25 data tests obtained 80% accuracy rate, disease in goats using naïve bayes and certainty Factor obtained accuracy rate of 86.80% [12-14].

Based on the problems presented above, then by utilizing the advancement of computer technology created an expert system of skin diseases in humans using naive bayes classifier. This system is expected to provide knowledge and solutions to skin problems humans are often encountered. In addition, to provide easy and free access, the system is built on the web. A good website is a website that is dynamic and user friendly or easily understood by users of any circle. For that web-based system will be built using MySQL database and CodeIgniter Framework.

II. RELATED WORK

Various previous studies expert systems, one of which research conducted by Utami [4] who examine the skin disease diagnosis experts on children using web-based forward chaining. Another study related expert system on the skin of children developed by Ulya, [6] using certain factors method. Based on these two studies it is concluded that each disease has a unique phenomenon that is in the options in the expert system search engine.

Subsequent research-related examine expert systems with the same method but using different test samples of animal skin resulted in that forward chaining and certain factor methods can only process uncertainties from two data, so the combination of these two methods is often linked by Mahardika [3]. However, over time, forward chaining methods with certain factors begin to be abandoned because of the lack of complexity of knowledge from an expert system because of the more complex knowledge available in an expert system it is necessary to perform several times of data processing for each selected phenomenon, not just comparing two data just like certain factor method [5].

The development of related system systems skin is done using image image-based and image processing. Expert systems can detect disease by photographing the symptoms and the system will detect the disease. However, this system is very dependent on the detection of symptoms that are photographed.

Based on these problems developed an expert system using the naïve bayes method. In this method the system will calculate each selected symptom of each option that appears and then compare it with each disease, then the system will provide output results of detection of several diseases that match the symptoms entered. Therefore, this study will try to analyze the expert system using the naïve bayes classifier method to detect a skin disease suffered by humans.

III. RESEARCH METHODS

Naïve Bayes Classifier is a simple probabilistic classifier based on the Bayes theorem[15] Bayes's theorem is combined with "Naïve" meaning every attribute/variable is independent. Naïve Bayes Classifier can be trained efficiently in supervised learning. The advantage of classification is that it requires only a small amount of training data to estimate the parameters (means and variance of variables) required for classification. Since independent variables are assumed, only variations of the variables for each class should be determined, not the entire covariance matrix. [2]

In the process, the Naïve Bayes Classifier assumes whether or not a feature in a class is related to the presence or absence of other features in the same class. At the time of classification, the bayes approach will produce the highest probability category label (VMAP) with attribute inputs a1, a2, a3, ... an [1]

$$V_{MAP} = argmax_{v_j \epsilon V}\, P(a_1 a_2 \ldots . a_n) \quad (1)$$

V_{MAP} = Highest probability.

$a_1 a_2 \ldots . a_n$ = Atribut (Inputan)

Bayes Theorem Declares:

$$P(B|A) = \frac{P(A|B)\, P(B)}{P(\mathrm{A})} \quad (2)$$

P (B | A) = Probability B if known state of the disease type A.
P (A | B) = Probability of evidence A if known hypothesis B
P (B) = Probability of hypothesis B regardless of any evidence.
P (A) = Probability of evidence of disease A.
Using Bayes's theorem, equation (3) can be written as follows:

$$V_{MAP} = argmax_{v_j \epsilon V} \frac{P(a_1 a_2 \ldots . a_n | v_j) P(v_j)}{P(a_1 a_2 \ldots . a_n)} \quad (3)$$

V_{MAP} = Highest Probability.
P (v_j) = Probability of skin disease type
P (a_1 a_2 ... .a_n | v_j) P (v_j) = Probability of attributes (input) if known state of vj.
P (a_1 a_2 ... .a_n | v_j) P = Probability of attributes (input)

Because the value of P (a1, a2, ... an | vj) is constant for all vj so that this equation can be write:

$$V_{MAP} = argmax_{v_j \epsilon V}\, P(a_1 a_2 \ldots . a_n | v_j) P(v_j) \quad (4)$$

V_{MAP} = Highest Probability.
P (v_j) = Probability type of skin disease j
P (a_1 a_2 ... .a_n | v_j) P (v_j) = Probability of attributes (input) if known state of v_j

To count $P(a_1 a_2 \ldots . a_n | v_j) P(v_j)$ more difficult because of the number of symptoms $P(a_1 a_2 \ldots . a_n | v_j) P(v_j)$ can be very large. This is because the number of symptoms is equal to the sum of all combinations of symptoms multiplied by the number of categories.
Calculation Naïve bayes classifier is:
Calculate P(a_i|v_j) by the formula:

$$P(a_i | v_j) = \frac{n_c + m_p}{n + m} \quad (5)$$

n_c = number of records in data learning v = v_j and a = a_i
p = 1/ number types of *class* / disease
m = number of parameters / symptoms
n = number of records in the learning data v = vj / each class
Equation (5) is solved by the following calculation:

1. Calculating value n_c for every *class*
2. Calculating value of *P* (a_i|v_j) and calculating value of *P* (v_j)
   $$V_{MAP} = argmax_{v_j \epsilon V}\, P(v_j) \Pi_i P(a_i | v_j) \quad (6)$$
   $$P(a_i | v_j) = \frac{n_c + m_p}{n + m}$$
3. Calculating value of P(a_i|v_j) x P(v_j) for every v
4. Determining the classification result v that has the greatest multiplication results
   $$V_{MAP} = argmax_{v_j \epsilon V}\, P(v_j) \Pi_i P(a_i | v_j)$$

ISSN: 2477-5126
e-ISSN: 2548-9356

IV. RESULT AND ANALYSIS

Based on the analysis of knowledge obtained from dr.Iswani, Sp.KK, books and journals related to human skin disease, there are 32 skin diseases with symptoms that amount to 88 symptoms. After collecting and formulating the knowledge that will be entered into the expert system, then determine the representation of knowledge using the form of production rules (production rule) using naive bayes classifier method.

There are several models of knowledge representation that can be used to make expert systems effective. In this study, the model of production rules is used. Production rules are declarative so that the representation of knowledge in this form is easily understood, in accordance with the way the human mind in solving a problem. Before forming the rules need to present the knowledge that has been successfully obtained.

In presenting a knowledge, needed supporting information so that knowledge can be presented appropriately. In the development of SP-PKPM, information needed is information about the symptoms and skin diseases that occur in humans. Such information is obtained from books, the Internet, and the knowledge of an expert. Once information is obtained, the information is organized to match what is required. The information can be seen in table 1 and table 2.

TABLE I. SYSMTOMS TABLE

#	Symptoms
G1	Body shivering
G2	Feeling weak
G3	Body feels hot
G4	Lumps on the legs
G5	Lumps on the buttocks
G6	A lump in the groin
G7	Bumps on the face
G8	Large patches in the middle and around small
G9	Brown spots on the skin
G10	Black spots on the soles of the feet / hands
G11	Red spots on the skin
G12	Fever
G13	Erosion on the sidelines of a finger
G14	Eruption in the face
G15	Fissures on the toes
G16	Nail damage
G17	Nails dingy, weathered and brittle
G18	There is hollow in the palms of the hands and feet
G19	Slightly tanned skin
G20	Blackish skin
G21	Skin redness
G22	The outer skin is brownish yellow
G23	Blistered skin
G24	Flawless skin
G25	Skin feels soft
G26	Typical lesions
G27	The wound is festering
G28	Oval or oval macula
G29	Find eggs (nits) ‘tuma’ / lice on the hair
G30	Feeling nauseous
G31	Feeling a bad smell
G32	Feeling itchy
G33	Feeling Pain
G34	Feel the heat on the skin
G35	Feel the pain on the chin / beard
G36	A lump appeared on the scar
G37	Red nodules
G38	Muscle ache
G39	Joint pain
G40	Nerve Pain
G41	Papules and pustules are penetrated by hair
G42	Swelling of the skin
G43	Skin inflammation
G44	Skin surface feels rough
G45	Hair loss
G46	Skin rashes are firm and scaly
G47	Headache
G48	Sore throat
G49	Often feel tingling
G50	Skuama roughly resembles mica
G51	It's hard to move the limbs
G52	There was a meandering lesion
G53	Looks macula erythematous
G54	There is a lump filled with clear fluid
G55	There is a dark brown bump
G56	There is a blackish brown skin lump
G57	There is a red leather lump
G58	There is a lump on the skin
G59	There is a lump like cauliflower
G60	There are brown spots
G61	There are red spots
G62	There are white patches
G63	There are reddish spots
G64	There is an ulcer on the skin
G65	There is a big bull on the skin
G66	There is a yellowish liquid
G67	There is a crustaceous impetigo
G68	There are blackheads
G69	There is crust
G70	There is a 'skin injury' on the skin
G71	There are minor injuries to the skin
G72	There is a macula on the head
G73	There is a hyperpigmented macula in the chest
G74	There is a hyperpigmented macula on the back
G75	There is a node
G76	There is a papule
G77	There are papules and pustules
G78	There are papules in the belt area
G79	There are papules in the armpit area
G80	There are vesicles on the face
G81	Did not feel the itch
G82	Tuma / lice attached to the skin around the base of the hair
G83	Typical vesicles break easily
G84	Grouped vesicles
G85	Vesicles at the corners of the mouth
G86	Vesicles (bumps) on the face
G87	Skin color is different from the surrounding skin
G88	Hair is attached to each other

TABLE II. DISEASE TABLE

#	Disease
1	Tinea Versikolor
2	Tinea Nigra Palmaris
3	Tinea Kapitis
4	Tinea barbae & sikosis barbae
5	Tinea pedis
6	Tinea Unguium
7	Tinea Kruris
8	Kromomikosis
9	Pioderma

ISSN: 2477-5126
e-ISSN: 2548-9356

10	Impetigo bulosa
11	Folikulitis
12	Furunkel
13	Erisipelas
14	Veruka vulgaris
15	Veruka plana
16	Keratolisis
17	Herpes zoster
18	Herpes simpleks
19	Varisela
20	Psoriasis
21	Pitiriasis rosea
22	Uritikaria pigmentosa
23	Kusta
24	Reaksi kusta
25	Skabies
26	Pedikulosis kapitalis
27	Pedikulosis korporis
28	Pedikulosis pubis (phthiriasis pubis)
29	Creeping eruption
30	Acne
31	Pemfigoid bulosa
32	Keloid

Each disease has several symptoms, and so each symptom can be associated with several diseases. The relationship between illness and symptoms is illustrated in table 3.

TABLE III. TABLE RELATION

#	Disease	Symptoms
1	Tinea Versikolor	G60, G62, G87, G32
2	Tinea Nigra Palmaris	G33, G10, G32
3	Tinea Kapitis	G45, G32, G33, G72
4	Tinea barbae & sikosis barbae	G35, G63, G27, G76
5	Tinea pedis	G13, G15, G31,
6	Tinea Unguium	G16, G17
7	Tinea Kruris	G46, G32
8	Kromomikosis	G4, G5, G59
9	Pioderma	G22, G69, G70, G67
10	Impetigo bulosa	G23, G66
11	Folikulitis	G11, G41, G43
12	Furunkel	G33, G37, G58,64
13	Erisipelas	G3, G21, G24, G43
14	Veruka vulgaris	G44, G55
15	Veruka plana	G9, G20, G25
16	Keratolisis	G18, G19
17	Herpes zoster	G34, G40, G84, G21, G32
18	Herpes simpleks	G12, G38, G71, G85
19	Varisela	G12, G80, G83
20	Psoriasis	G26, G50, G53, G33
21	Pitiriasis rosea	G53, G8, G28, G2, G47, G48
22	Uritikaria pigmentosa	G42, G73, G74, G32
23	Kusta	G61, G62, G81, G49, G51, G45
24	Reaksi kusta	G12, G1, G30, G39, G75
25	Skabies	G77, G5, G6
26	Pedikulosis kapitalis	G29, G82, G32, G88
27	Pedikulosis korporis	G78, G79, G32
28	Pedikulosis pubis (phthiriasis pubis)	G29, G82, G32
29	Creeping eruption	G52, G32, G33
30	Acne	G7, G14, G68, G32
31	Pemfigoid bulosa	G32, G65, G54
32	Keloid	G36, G56, G57, G32

After the data obtained symptoms and diseases and has made relations, then the next data can be used in expert systems method Naïve Bayes Classifier. Calculation Naïve bayes classifier is:

$$P(a_i|v_j) = \frac{n_c+\text{mp}}{n+m}$$

n_c = number of records in data learning v = v_j and a = a_i
p = 1/ number of class types / disease
m = number of parameters / sysmptom
n = number of records in data learning v = v_j / every *class*

So for the variable value p is 1/32 = 0.03125 and the variable value m is 88 in accordance with the number of existing symptoms. Then for the value of variable n will always be worth 1 because the calculation using one by one of the symptoms entered. As for the variable nc is worth 1 if vj or symptoms are input in accordance with the data relation for the calculation of each disease, but if vj or symptoms that entered did not exist in relation with the disease then the value of the variable is 0. After all probability for each selected phenomenon in a disease has been calculated. The next step is to do the following calculations:

$$V_{MAP} = argmax_{v_j \epsilon V}\, P(a_1 a_2 \ldots . a_n|v_j) P(v_j)$$

So after calculating the value of symptoms in each disease, the value of each symptom P (vj) is multiplied by the value of the disease P (vj) to obtain the value of V_{MAP} for each disease. This calculation is done continuously until all diseases have V_{MAP} value. After the calculation is complete, the largest value of the classification results for each value of V_{MAP} was obtained. The value is the result of disease identification in accordance with the symptoms that have been entered.

Fig. 1. Home Page

The home page is the first page the user sees when starting using an expert skin disease identification system application in humans. Implementation of the main page made berdadsarkan designing the main page. Here is the main page interface implementation that contains the home menu, disease list,

ISSN: 2477-5126
e-ISSN: 2548-9356

identification, faq and login. Figure 1 is the main page implementation view.

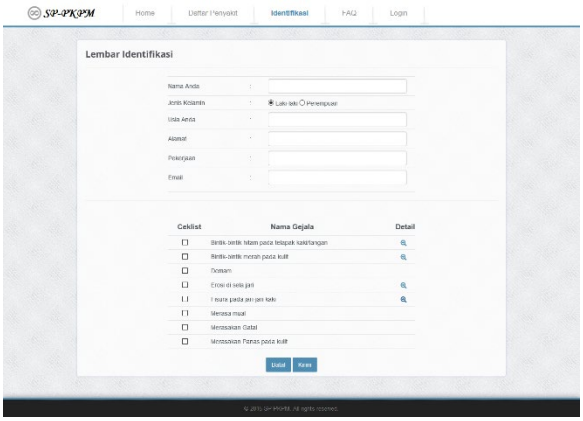

Fig. 2. Identification Page

The identification page is the main page in an expert skin disease identification system in humans. This page contains the form of self data and the question of knowing the symptoms experienced. This page when processed will produce the result of identification as in figure2.

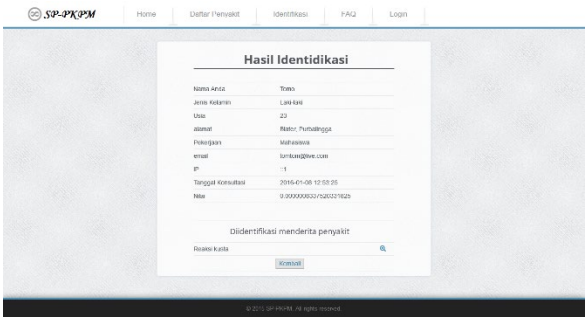

Fig. 3. The result of identification

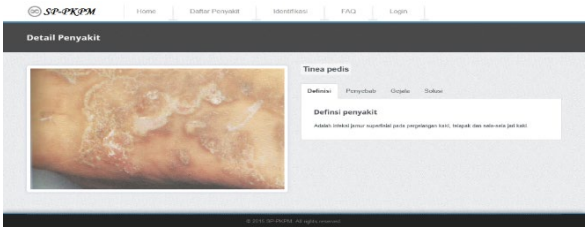

Fig. 4. The information of disease

The identification page is the result page of the identification process in Figure 3. This page can produce a skin disease solution suffered and clearly describe the information about the skin disease by selecting the detail button. The disease detail page will be displayed as shown in Figure 4.

V. CONCLUSIONS

Based on the discussion that has been described in this thesis report, it can be concluded that the method of naive bayes classifier is able to be applied in expert system, and with this expert system can be made easier for everyone to know the common skin diseases.

VI. FUTURE WORK

Suggestion for further research of Expert Skin Identification Expert System on Human that is Further development, can be expanded scope of system function, which originally only to detect skin disease is also expected to detect genital skin disease.

The addition of features for mobile apps so as to enhance the convenience of system users. Combining Genetic Algorithms to gain weighting of each symptom.

REFERENCES

[1] Aribowo, T. 2009. Aplikasi Inferensi Bayes pada Data Mining terutama Pattern Recognition. Departemen Teknik Informatika, Institut Teknologi Bandung.

[2] Aribowo, T. 2009. Aplikasi Inferensi Bayes pada Data Mining terutama Pattern Recognition. Departement of informatic engineering, Institut Teknologi Bandung.

[3] Mahardika, AP. 2010 “Sistem Pakar Pendeteksi Penyakit Kulit Pada Manusia Dengan Metode Forward Chaining”.

[4] Puteri, UH. 2014. “ Sistem Pakar Identifikasi Penyakit Kulit Pada Kucing Menggunakan Metode Forward Chaining Berbasis Web”.

[5] Setiawan, W. 2014 “Sistem Pakar Diagnosis Penyakit Mata Menggunakan NaÏve Bayes Classifier”.

[6] Ulya, D. 2007. “Sistem Pakar Diagnosa Penyakit Kulit Pada Anak Menggunakan Certainty Factor”.

[7] Prayoga, ND., Hidayat N., & Dewi RK. 2017. Sistem Diagnosis Penyakit Hati Menggunakan Metode Naïve Bayes. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 2 (8): 2666-2671.

[8] Fahrul, A. Hernandha, EW. Indah, FA. 2017. Expert System With Certainty Factor For Early Diagnosis Of Red Chili Peppers Diseases. Journal of Applied Intelligent System,Vol.2, No. 2 pp. 52 – 66

[9] Munandar, Suherman and Sumiat. 2012. The Use of Certainty Factor with Multiple Rules for Diagnosing Internal Disease. International Journal of Application or Innovation in Engineering & Management, Vol. 1, No. 1, pp.58-64

[10] Syarief, M. Prastiti, N. Setiawan, W. 2017. Comparison of Naïve Bayes and Certainty Factor Method for Corn Disease Expert System: Case in Bangkalan, Indonesia. International Journal of Engineering Research and Application, Vol. 7, No. 11, pp.30-34

[11] Syarief, M. Mukminin, A. Prastiti, N. Setiawan, W. 2017. Penerapan Metode Naïve Bayes Classifier Untuk Deteksi Penyakit Pada Tanaman Jagung. NERO. Vol. 3 No. 1, pp 61- 68

[12] Ferdiansyah, W. R., Muflikhah, L., & Adinugroho. 2018. Sistem Pakar Diagnosis Penyakit Pada Kucing Menngunakan Metode Naive Bayes dan Certainty Factor. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 2(2): 451-458.

[13] Suprayogi, AA., Hidayat, N., & Fanani, L. 2018. Sistem Pakar Diagnosis Penyakit Kucing Menggunakan Metode Naïve Bayes – Certainty Factor Berbasis Android. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 2(2): 650-658

[14] Insani, MI. Alamsyah. Putra, AT. 2018. Implementation of Expert System for Diabetes Diseases using Naïve Bayes and Certainty Factor Methods. Scientific Journal of Informatics, 2 (5): 185-193.

ISSN: 2477-5126
e-ISSN: 2548-9356

[15] Bestari, D. P. B., Saptono, R., & Anggrainingsih, R. (2018). Academic Articles Classification Using Naive Bayes Classifier (Nbc) Method. Jurnal Ilmiah Teknologi Dan Informasi, 7(2), 74–81.