

Naïve Bayes Approach for Expert System Design of Children Skin Identification Based on Android

Hartatik¹, A Purnomo¹, R Hartono¹, H Munawaroh²

¹ Informatics Engineering Department, Faculty of Mathematics and Natural Sciences, Sebelas Maret University, Surakarta, Indonesia

² Chemistry Department, Master Program Graduate School, Sebelas Maret University, Surakarta, Indonesia

E-mail: hartatik.uns@gmail.com

Abstract. The development of technology gives some benefits to each person that we can use it properly and correctly. Technology has helped humans in every way. Such as the excess task of an expert in providing information or answers to a problem. Thus problem that often occurs is skin disease that affecting on child. That because the skin of children still vulnerable to the environment. The application was developed using the naïve Bayes algorithm. Through this application, users can consult with a system like an expert to know the symptoms that occur to the child and find the correct treatment to solve the problems.

1. Introduction

Growth period becomes very important period for the development of the human body. In childhood, the body is still quite vulnerable to both food and the environment. The children's phase is an interesting phase between the other phases. At the stage of development of children, they begin to have a much greater curiosity. Know a lot and want a lot of things too. In this phase most of the human phase is very keen to be noticed (pampered). Rapid physical growth also occurs in childhood. For example, feet that can grow faster, to see the child's height is rapidly increasing. For that, the child's health should get more attention on the growth period, especially on the skin of children.

The development of information technology is currently very fast development and also plays an important role in various human activities [1]. 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 [2].

The technology is expected to assist people in solving problems encountered, such as about decision support system or expert system about skin disease in children. [3]. In previous research has been about the development of decision support system among others has been done by some researchers, including decision support system for news classification [4], decision support system for nutritional status [5]. In addition to the development of applications about decision support systems, some research that can help people in solving problems is the development of expert systems Prediction System for Heart Disease [6], expert systems for Prediction of Different Dermatological [7], Expert System for Diagnosis of Tropical Infectious Diseases [8], as well as an Expert System to Diagnose Corn Plant Diseases [9], and Expert System for Mineral Identification [10].



2. Experimental

In this study the authors use Naïve Bayes Classification algorithm the main focus in determining the classification of Throat Disease [11]. The meaning of the Naïve Bayes Classification algorithm is to use the probability approach to produce a classification of skin disease types.

Steps of work in this study can be seen in Figure 1:

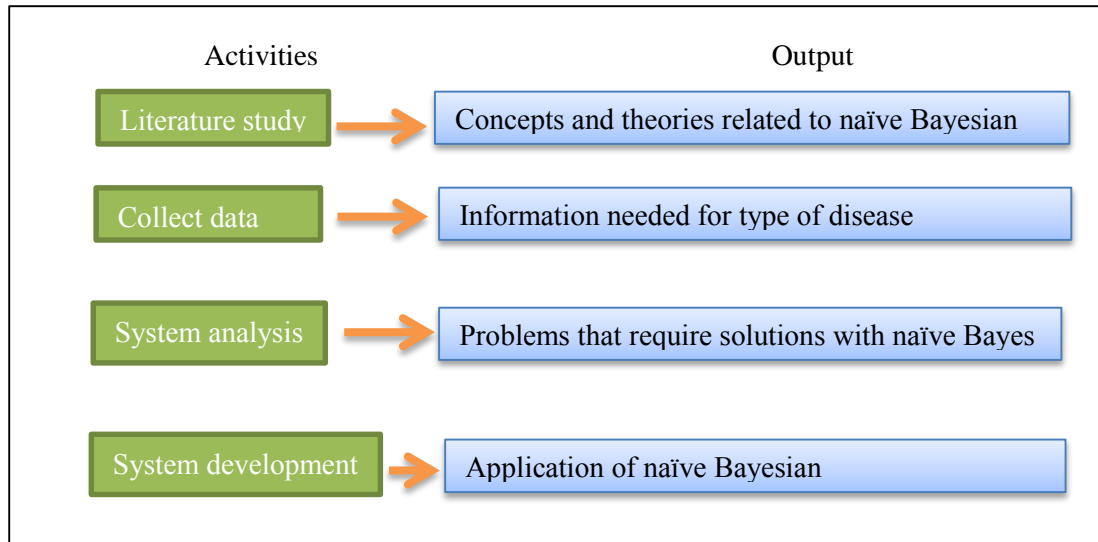


Figure 1. Research flow for detect naïve Bayes

To support good results, the object of research is children aged 1-12 years. Conducted interviews of children with complaints of skin diseases. Furthermore, based on the results of interviews formed categories of diseases and symptoms. This will be the basis for decision making. Sampling is done by Purposive sampling.

3. Result and Discussion

Naïve Bayes Classifier is a simple probability classifier based on Bayes theorem. Bayes's theorem is combined with "Naïve" which means every attributes / variables are independent (independent). Naïve Bayes Classifier can be trained efficiently in supervised learning.

In this research causes of skin diseases

1. Poor self-hygiene
2. Viruses
3. Bacteria
4. Allergic Reactions
5. Low immune system

Bayes method as a method for problem solving in decision support system to detect the type of skin disease, especially for children aged 1-12 years. In the process of forward chaining of the inference engine for the determination of appropriate healthy lifestyle categories, it is possible that the final diagnostic results have several categories selected. To select one of several selected categories requires management of uncertainty. This uncertainty will be calculated using the Bayes probability theorem, with equation 1.

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (1)$$

Where Variable X represents the class, while the variable $X_1 \dots X_n$ represents the characteristics of the instructions needed to classify the types of skin diseases. Then the formula explains that the probability of entering a sample of certain characteristics in class H (Posterior) is the probability of occurrence of class H (before the entry of the sample, often called prior), multiplied by the probability of occurrence of sample characteristics in class H (also called likelihood) with the probability of occurrence of sample characteristics globally (also called evidence). Therefore, the above formula can also be written simply as follows:

$$\text{Posterior} = \text{prior} \times \text{likelihood} \times \text{evidence} \quad (2)$$

Evidence values are always fixed for each class on a single sample. The value of the posterior will then be compared with the posterior values of the other classes to determine to which class a sample will be classified. A further description of the Bayes formula is made by elaborating $(H | X_1, \dots, X_n)$ using the multiplication rule as follows:

$$\begin{aligned} (H|X_1, \dots, X_n) &= (H)P(X_1, \dots, X_n | H) \\ &= (H)P(X_1 | H)P(X_2, \dots, X_n | H, X_1) \\ &= P(H)P(X_1 | H)P(X_2 | H, X_1)P(X_3, \dots, X_n | H, X_1, X_2) \\ &= (H)P(X_1 | H)P(X_2 | H, X_1)P(X_3 | H, X_1, X_2)P(X_4, \dots, X_n | H, X_1, X_2, X_3) \\ &= P(H)P(X_1 | H)P(X_2 | H, X_1)P(X_3 | H, X_1, X_2) \dots P(X_n | H, X_1, X_2, X_3, \dots, X_{n-1}) \end{aligned} \quad (3)$$

It can be seen that the results of such descriptions lead to more and more complex factors that affect the probability value, which is almost impossible to analyze one by one. As a result, the calculation becomes difficult to do. Here is the very high (naive) assumption of independence, that each clue (X_1, X_2, \dots, X_n) is independent of each other. With these assumptions, then apply a similarity as follows:

$$P(X_i | X_j) = P(X_i \cap X_j)P(X_j) = P(X_i)P(X_j)P(X_j) = P(X_i) \quad (4)$$

For $i \neq j$, so

$$(X_i | H, X_j) = (X_i | H) \quad (5)$$

The above equation is a model of the Naive Bayes theorem which will then be used in the classification process. For classification with continuous data used Gauss Density formula:

$$(X_i = x_i | Y = y_j) = \frac{1}{\sqrt{2\pi}\sigma_{ij}} e^{\frac{-(x_i - \mu_j)^2}{2\sigma_{ij}^2}} \quad (6)$$

Where:

P: Opportunities

X_i : Attribute to i

x_i : The attribute value to i

Y: Classes searched

y_i : Sub class Y is searched

μ : mean, states the mean of all attributes

σ : Standard deviation, declares a variant of all attributes.

Furthermore the naïve Bayes method is used to determine the type of disease, with the design as follows:

Table 1. Type of disease

Number	Code	Desease
1	D1	measles
2	D2	eczema
3	D3	water fleas
4	D4	Panu
5	D5	Boils

Based on the type of disease, then known the symptoms of each disease meaning. For example, measles, there are some symptoms as shown in figure 2.

Table 2. Symptoms of the disease

S11	Hot sores
S12	Limp
S13	blisters rupture and wound
S14	blister and yellow discharge
S15	Wet

Based on the symptoms of the disease it fits with all types of diseases, and will be searched for compatibility with the highest level of compatibility. Thus, a symptom with a high titled fit is expressed as a possible skin disease suffered, can be seen in Table 3

Table 3. Table decision

	D1	D2	D3	D4	D5
S1	√	√		√	
S2		√			√
S3			√		
S4				√	
S5					√
S11	√				
S12	√				
S13					
S14					
S15					
S21		√		√	
S22		√			
S23		√			

4. Conclusion

Simulations for the application of naïve Bayesian methods for the determination of further skin disease types to be developed in android applications, so that the community can use to detect the type of skin disease experienced.

References

- [1] Jogiyanto H M 2005 *Analisis dan Desain Sistem Informasi: Pendekatan Terstruktur Teori dan Praktek Aplikasi Bisnis* (Yogyakarta: Andi Offset)
- [2] Durkin J 1994 *Expert System Design and Development* (New Jersey: Prentice Hall International Inc)
- [3] Efraim T 1995 *Decision Support Systems and Expert System* (New Jersey: Prentice Hall International Inc)
- [4] Wibisono Y 2005 *Klasifikasi Berita Berbahasa Indonesia menggunakan Naïve Bayes Classifier, Seminar Nasional Matematika 2005* (Bandung: Universitas Pendidikan Indonesia)
- [5] Nur Anggraeni Diana Rahmawati and Firli Irhamni 2012 *Jurnal Sarjana Teknik Informatika* **1(1)** 85 – 92
- [6] Pattekari S A, Parveen A 2012 *Int. J. Adv. Comp. Math. Sci.* **3(3)** 290-294
- [7] Manjusha K K, Sankaranarayanan K, Seena P 2014 *Int. J. Adv. Res. Comp. Sci. Software Eng.* **4(1)** 864-868
- [8] Diema Hernyka Satyareni 2011 *Jurnal Teknologi* **1(2)** 61-67
- [9] Munanda, Edi, Nanang P 2013 *Jurnal Litek* **10(2)** 113-117
- [10] I O Folorunso, O C Abikoye, R G Jimoh and K S Raji 2012 *JETCIS* **3(2)** 205-210
- [11] Ting S L, Ip W H, Tsang A H C 2011 *IJSEA* **5(3)** 37-46