Jurnal Mantik, 5 (4) (2022) 2682-2687

Published by:Institute of Computer Science (IOCS)

Jurnal Mantik

Journal homepage: www.iocscience.org/ejournal/index.php/mantik/index

Decision Tree Algorithm and Gini Calculation for Puppies Skin Disease Diagnosis

Khairun Annisa[1], Arpan[2]

[1,2]Computer System, Sains and Teknologi, Universitas Pembangunan Panca Budi Medan,
Jl. Gatot Subroto km. 4,5 Medan, 20122, Indonesia

Email: khairunannisa2000@gmail.com,

A R T I C L E I N F O

Article history:

Received: Jan 10, 2022
Revised: Jan 30, 2022
Accepted: Feb 14, 2022

Keywords:

Expert Systems,
Skin Diseases of the Puppies,
Scikit-learn.
Decision Tree

A B S T R A C T

To put it another way, an expert system is a computer system that can store and reproduce knowledge in order to solve a problem. It is also possible to employ expert systems to assist a specific field with problem-solving, and the expert system serves to keep the necessary parties working together. Software or computer programs that provide guidance and assistance in specialized subjects, such as science, engineering, mathematics, medicine, and education can be referred to as expert systems. The expert system can be used to perform diagnostics on pets. This article aims to discuss the implementation and study of an expert system that assists users in determining the ailment that pets, particularly dogs, are suffering from. The system can generate a symptom diagnostic of a skin illness using a decision-tree algorithm. Veterinarians provided data sets on a variety of disorders. There were 105 cases of 11 diseases and 11 symptoms in the dataset used in this study. The scikit-learn library was also used to train decision trees. Following the creation of a decision tree, a case study analysis was conducted using ten real-world examples provided by the veterinarian. Participants who own dogs received questionnaires in addition to those used for analysis. In the study, 25.8% of respondents said they strongly agreed with the expert system, and 32.3% said they agreed that the expert system helped explain the ailment their dog had. A further 80 percent of the instances provided by vets can be accurately determined by the system. When four independent tests were run, the mean system prediction was found to be 67.6 percent.

1. Introduction

An expert system is a computer system that distinguishes between the problem part of the knowledge described in a program and the part that replicates that information in order to solve the problem using existing data [1] [2] [3]. An expert system is a software package or computer program package that provides advise and assists in the solution of issues in certain fields such as science, engineering, mathematics, medicine, and education [4] [5] [6]. Making a diagnostic is one of the expert system's applications. The term "diagnosis" has come to incorporate the term "prognosis." The probability of a disease's end based on general information is called prognosis [7] [8] [9]. The diagnosis procedure entails not only determining the type, characteristics, and background of a certain illness weakness, but also attempting to anticipate the likelihood and provide a cure [10]. An expert system's capacity to diagnose disease symptoms and anticipate disease will not be 100 percent identical to a doctor's diagnosis [11] [12]; there are still numerous unknown or inconsistent factors to study, therefore expert system predictions may lead to diagnostic errors.

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

2. Method

The research method used is quantitative. In research using this quantitative method, users of graphs, diagrams and others will be used [13]. The research data conducted on the IPB paper is used to assist this research. Making the decision tree model in this study will use data obtained from the IPB paper [14]. The decision tree used will be created using scikit-learn.

2.1 Use Case

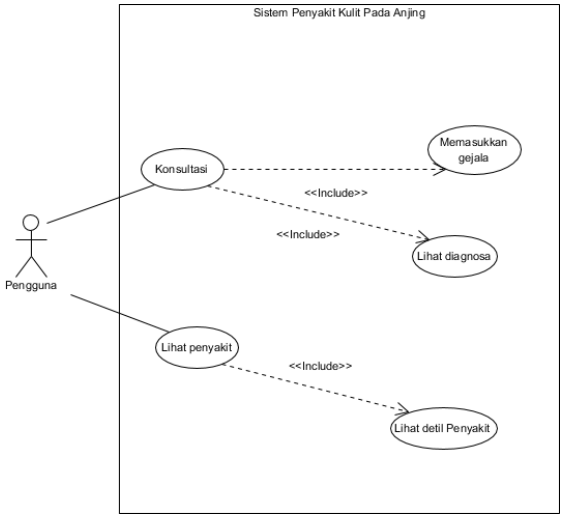

Figure 1. Use case of Skin Diseases System in Puppiess

The following is an explanation of Figure 2. Users can consult and view diseases. At the consultation the user will be asked to enter symptoms. In the View Disease use case, the user can view the disease details of each disease.

2.2 Dataset

The dataset has an important role in this research, because it helps in the development of the expert system used. The dataset in this study is a dataset obtained from Joanna Fransisca (2006) [15]. This paper is recommended by doctors or experts directly. In this paper, the dataset used is one hundred and five (105) with eleven (11) diseases and eleven (11) symptoms. The dataset that is built is a collection of symptoms and diseases.

Figure 2 is the data model obtained in the paper. Reading the data in the image, there are 4 symptoms of alopecia found in atopy, and 1 symptom in ticks.

diagnosa	alopecia (botak/ pitak)	bintik (papula)	bintil (nodula)	bulu kusam	bulu rontok	eritema (kemerahan)	hitam	infeksi (bengkak/ radang/ abses)	kerak	ketombe	pruritus (gatal/ garuk-garuk)
atopy	4	17	-	-	4	26	3	-	-	-	16
caplak	1	1	-	1	7	-	-	-	2	-	-
demodecosis	5	1	-	-	-	15	1	5	4	1	7
dermatophytosis	8	2	2	4	10	18	5	1	18	12	14
tungau telinga	1	1	-	-	2	2	-	-	-	1	1
hypothyroid	-	1	-	-	-	-	4	-	1	-	-
pinjal	-	1	-	-	1	3	-	-	1	1	-
pyoderma	1	5	1	-	1	11	-	2	4	1	-
reaksi makanan	1	-	-	-	4	4	-	-	2	1	-
scabies	1	1	1	-	-	3	-	1	10	1	4
tumor	-	-	-	-	2	1	1	6	-	-	-

Figure 2. Data on Paper

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

	kebotakan	bintik	bintil	bulu kusam	bulu rontok	kemerahan	hitam	infeksi	kerak	ketombe	gatal-gatal	penyakit
0	1	1	0	0	1	1	1	0	0	0	1	1
1	1	1	0	0	1	1	1	0	0	0	1	1
2	1	1	0	0	1	1	1	0	0	0	1	1
3	1	1	0	0	1	1	0	0	0	0	1	1
4	0	1	0	0	0	1	0	0	0	0	1	1
5	0	1	0	0	0	1	0	0	0	0	1	1
6	0	1	0	0	0	1	0	0	0	0	1	1
7	0	1	0	0	0	1	0	0	0	0	1	1
8	0	1	0	0	0	1	0	0	0	0	1	1
9	0	1	0	0	0	1	0	0	0	0	1	1
10	0	1	0	0	0	1	0	0	0	0	1	1
11	0	1	0	0	0	1	0	0	0	0	1	1
12	0	1	0	0	0	1	0	0	0	0	1	1
13	0	1	0	0	0	1	0	0	0	0	1	1
14	0	1	0	0	0	1	0	0	0	0	1	1
15	0	1	0	0	0	1	0	0	0	0	1	1
16	0	1	0	0	0	1	0	0	0	0	0	1
17	0	0	0	0	0	1	0	0	0	0	0	1
18	0	0	0	0	0	1	0	0	0	0	0	1
19	0	0	0	0	0	1	0	0	0	0	0	1
20	0	0	0	0	0	1	0	0	0	0	0	1

Figure 3. Data used in scikit-learn

Figure 3 is the data that has been modified and used in scikit-learn. The data obtained after the changes were made were 105 data.

2.3 Decision Tree

Figure 2 explains how to generate a decision tree like Figure 6 using scikit-learn. The result of the decision tree that appears contains the results of the Gini calculation automatically, for example, the root of the decision tree gets a Gini of 0.89.

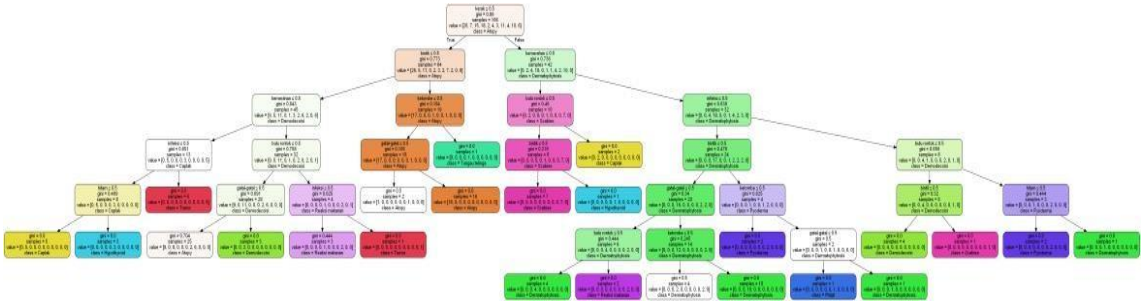

Figure 4. Code To Generate Decision Tree

3. Result and Discussion

3.1 Implementation UI

Figure 5 Shows the implementation results for the display at the consultation. On this page the user will be asked to fill in the questions given by the program. The total questions answered by the users were 11 questions. When the questions are answered, the program will display the results of the consultation.

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

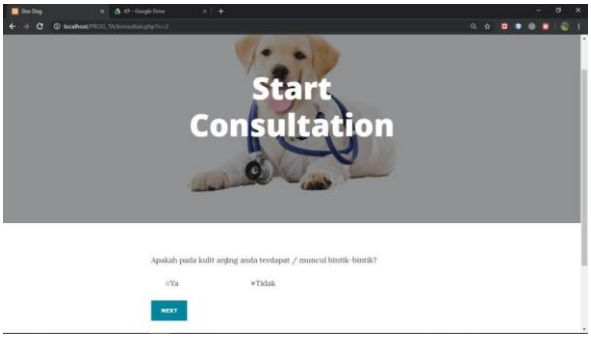

Figure 5. Consultation

Figure 6 shows the implementation results for the display on the diagnosis results. On this page the user will receive the results of the diagnosis after conducting a consultation. On this page will display the name of the disease found in Puppiess and initial treatment.

Figure 6. Consultation result

3.2 Program Implementation

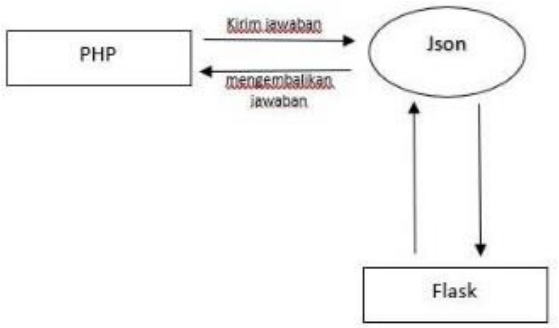

Figure 7. Work Program Process

Figure7 is a working program process. In php will get the answers that have been given by the user during the consultation. After the answer is obtained it will be sent to json. Json will send the answer as an array on flask. Answers received on the flask will be tested on the training data. Flask will get the answer from the test data and will send it back to json to display on the web.

3.3 Usability Testing

The first step in usability testing is to give respondents a number of tasks that have been prepared in advance when interacting with the system being tested. In this test, the program is tested on users who are devoted to owning a Puppies.

Following are the tasks assigned to the user:

a. The user opens the web that has been given and pays attention to the appearance of the web.
b. Users access and try the information menu and conduct consultations.

c. Users conduct consultations by answering questions contained on the web

3.4 Test Result with Questionnaire

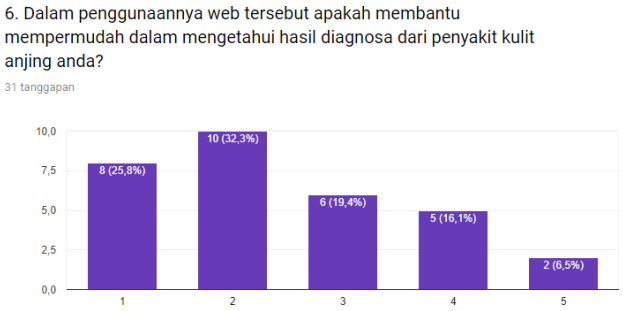

Figure 8. Survey Result

Figure 8 is the percentage result of the question "Does the web use help make it easier to find out the results of the diagnosis of your Puppies's skin disease?" of 31 responses, 8 people (25.8%) answered strongly agree, 10 people (32.3%) answered agree, 6 people (19.4%) answered quite agree, 5 people (16.1%) answered disagree , 2 people (6.5%) answered disagree.

3.5 Analysis

In the next analysis, ten cases of Puppies skin disease that had been discussed with the doctor directly were taken for analysis. A correct case will be scored as one (1) while an incorrect case will be scored zero (0). Of the 10 cases taken, the program predicts those 10 cases with eight true and two false.

TABLE 1.
CASE TEST RESULTS

No	Diagnosis Result	correct / incorrect
1	Dermatophytosis	correct
2	Tungau telinga	Correct
3	Atopy	correct
4	Hypothyroid	correct
5	Reaksi makanan	correct
6	Caplak	correct
7	Demodecosis	correct
8	Dermatophytosis	correct

Based on table 1 of 10 cases, the resulting accuracy values are asture follows:

$$Akurasi = \frac{\sum match}{Semua\ data} \times 100\%$$

$$akurasi = \frac{8}{10} \times 100\%$$

$$akurasi = 80\%$$

So it can be concluded that the accuracy of the expert system based on 10 tested cases is 80% which indicates that this program is functioning well.

3.6 Cross-Validation

In the tests carried out with 25 times the results are as shown in Figure 9 and Figure 10. The average result obtained from the cross-validation in Figure 5.14 is 0.743.

```
In [21]: cv_result = cross_val_score(clf, X, Y, cv=25)
```

Figure 9. Cross Validation A total of 25

Accredited "Rank 4"(Sinta 4), DIKTI, No. 36/E/KPT/2019, December 13th 2019.

```
In [22]: cv_result

Out[22]: array([0.33333333, 0.45454545, 0.7       , 0.88888889, 0.71428571,
                0.85714286, 0.83333333, 0.6       , 0.6       , 0.6       ,
                0.5       , 0.66666667, 0.33333333, 0.33333333, 0.66666667,
                0.5       , 1.        , 1.        , 1.        , 1.        ,
                1.        , 1.        , 1.        , 1.        , 1.        ])

In [23]: cv_result.mean()

Out[23]: 0.7432611832611833
```

Figure 10. Cross Validation Results of 25

With the results that have been obtained on cross validation, it can be concluded that the average is as in Figure 2. The average obtained is calculated from the results of the sum calculation of each cross validation carried out with the number of cross validations carried out and produces an average of 0.676 .

4. Conclusion

Based on the survey spread to 31 respondents who own Puppiess and testing with 10 cases obtained from experts, it can be concluded that the implementation of a web-based expert system can assist in diagnosing disease symptoms and predicting skin diseases in Puppiess. The dataset used is 105 consisting of 11 symptoms and 11 diseases. This research and expert system program got good results with the data obtained from the surveys that have been carried out, it can be concluded that the developed program can assist respondents in knowing the skin diseases experienced by the user's Puppies. Through one survey, it was proven that about 8 people (25.8%) answered strongly agree and 10 people (32.3%) answered agree. In the analysis carried out with 10 cases obtained from experts or doctors, an accuracy of 80% was obtained. The cross-validation carried out 4 times got an average of 6.76%.

5. References

[1] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," J. Mach. Learn. Res., vol. 12, pp. 2825–2830, 2011.
[2] J. S. P. Tyoso, Sistem Informasi Manajemen. Deepublish, 2016.
[3] H. Pratiwi, Buku Ajar: Sistem Pakar. STMIK Widya Cipta Dharma, 2019.
[4] I. Sukma and M. Petrus, "Sistem Pakar Penyakit Kucing Menggunakan Metode Forward Chaining Berbasis Web," Simtek J. Sist. Inf. dan Tek. Komput., vol. 5, no. 1, pp. 52–58, 2020.
[5] M. Lika Ifa Murdiansari, "SISTEM PAKAR UNTUK MENENTUKAN KELAYAKAN SEKOLAH MELAKSANAKAN PEMBELAJARAN TATAP MUKA (PTM) DIMASA PANDEMI COVID-19 DENGAN METODE CASE BASED REASONING." Universitas Muhammadiyah Ponorogo, 2021.
[6] Y. Kuryanti, "Sistem Pakar Diagnosa Keluhan Kesehatan pada Orang Dewasa Berbasis Web dengan Metode Backward Chaining."
[7] J. Durkin, "Expert System," Handb. Appl. Expert Syst., vol. 4, p. 15, 4AD.
[8] A. S. Sitio, "Sistem Pakar Untuk Mendiagnosa Penyakit Jantung Menggunakan Metode Dempster Shafer," J. Inform. Pelita Nusant., vol. 3, no. 1, 2018.
[9] R. Rosnelly, Sistem Pakar: Konsep dan Teori. Penerbit Andi, 2012.
[10] P. DOYOK, "STUDI KASUS SISWA YANG BERJUDI KELAS XI SEKOLAH MENENGAH ATAS NEGERI 1 SUNGAI BETUNG KABUPATEN BENGKAYANG." IKIP PGRI PONTIANAK, 2020.
[11] H. T. Sihotang, "Perancangan Aplikasi Sistem Pakar Diagnosa Diabetes Dengan Metode Bayes," J. Mantik Penusa, vol. 1, no. 1, 2017.
[12] E. D. S. Mulyani and N. N. F. SM, "Aplikasi Pakar Diagnosa Penyakit Gigi Menggunakan Metode Forward Chaining Berbasis Mobile," E-Proceedings KNS&I STIKOM Bali, pp. 119–124, 2017.
[13] D. Haryono, "Pengendalian kualitas produksi dengan model grafik kontrol p pada PT. Asera Tirta Posidonia," J. Varian, vol. 1, no. 1, pp. 27–34, 2017.
[14] P. Gunarso, M. Kanninen, and H. Priyadi, "MODEL DINAMIKA SISTEM UNTUK PENGEMBANGAN," 2003.
[15] J. C. Santoso and H. Bunyamin, "Sistem Pakar Diagnosa Penyakit Kulit Anjing Dengan Algoritma Pohon Keputusan dan Perhitungan Gini," J. Strateg. Maranatha, vol. 1, no. 2, pp. 336–344, 2019.