

Sistem Pakar Diagnosa Penyakit Kulit Anjing Dengan Algoritma Pohon Keputusan dan Perhitungan Gini

Jonathan Chrisnanda Santoso¹, Hendra Bunyamin²

Jurusan Teknik Informatika, Universitas Kristen Maranatha

Jl. Surya Sumantri No. 65 Bandung

¹santosojonathan95@gmail.com

²hendra.bunyamin@it.maranatha.edu

Abstract — An expert system is a computer system that has separable components for storing knowledge and replicating the knowledge to solve a problems. Likewise, the expert system can be utilized to help a particular field solve problems; moreover, the expert system still maintains related parties for working together in the field. Specifically, an expert system is a software package or computer program intended to provide pieces of advice and aids in solving problems in certain fields, such as science, engineering, mathematics, medicine, and education. A specific application of expert systems is to perform a diagnosis on pets. This article attempts to elaborate an implementation and analysis of an expert system that help users determine the illness suffered by pets, particularly dogs. With the help of the decision tree algorithm, the system can produce a diagnosis of symptoms of skin diseases. Dataset of various illness is obtained directly from a veterinarian. The dataset used in this study were 105 instances consisting of 11 illnesses and 11 symptoms. Additionally, a decision tree is trained by using scikit-learn library. Having constructed the decision tree, analysis is performed by taking 10 cases that were given directly by a vet. In addition to the analysis, a questionnaire is distributed to respondents who have a dog. The survey shows about 8 people (25.8%) answered strongly agreed, and 10 (32.3%) responded agreed that the expert system does help explain the illness suffered by their dogs. Furthermore, 10 cases provide by the vet can be determined by the system with an accuracy of 80%. Finally, results of cross validation performed 4 times show that the average prediction of the system is 67.6%.

Keywords— decision tree, expert systems, skin diseases of the dog, scikit-learn.

I. PENDAHULUAN

Sistem pakar merupakan sistem komputer yang dicirikan oleh fakta bahwa perbedaan yang dibuat antara bagian masalah pada pengetahuan yang digambarkan dalam suatu program dan bagian untuk meniru pengetahuan tersebut untuk memecahkan masalah adalah dengan menggunakan data yang ada [1]. Wujud sistem pakar adalah paket perangkat lunak atau paket program komputer yang ditujukan sebagai penyedia nasihat dan sarana bantu dalam memecahkan masalah di bidang-bidang spesialisasi tertentu seperti sains, perekayasaan, matematika, kedokteran, dan pendidikan. Salah satu aplikasi sistem pakar adalah untuk melakukan diagnosis. Di dalam konsep diagnosis telah tercakup pula konsep prognosinya. Prognosis adalah kemungkinan akhir suatu penyakit berdasarkan pengetahuan umum. Proses diagnosis bukan hanya sekedar mengidentifikasi jenis dan karakteristiknya, serta latar belakang dari suatu kelemahan penyakit tertentu melainkan mengimplikasikan suatu upaya untuk meramalkan kemungkinan dan menyarankan tindakan pemecahannya. Kemampuan sistem pakar dalam mendiagnosis gejala penyakit dan memprediksi penyakit tidak akan 100% sama dengan diagnosis seorang dokter; masih banyak hal yang tidak pasti atau tidak konsisten untuk diselidiki sehingga prediksi sistem pakar dapat menyebabkan kemungkinan kesalahan dalam diagnosis.

Rumusan masalah penelitian ini adalah sebagai berikut:

Bagaimana mengimplementasikan sebuah sistem pakar berbasis web yang dapat membantu mendiagnosa gejala penyakit dan memprediksi penyakit kulit anjing?

Tujuan dilakukannya penelitian ini adalah sebagai berikut:

Mengimplementasikan sebuah sistem pakar berbasis web yang dapat membantu mendiagnosis gejala penyakit dan memprediksi penyakit kulit pada anjing.

II. TINJAUAN PUSTAKA

A. *Rekayasa Perangkat Lunak*

Rekayasa perangkat lunak adalah sebuah disiplin yang mengadopsi pendekatan rekayasa seperti metodologi, proses, alat, standar, metode organisasi, metode manajemen, dan sistem jaminan kualitas [2]. Dalam aplikasinya, rekayasa perangkat lunak memiliki 5 (lima) aspek penting, yaitu :

1. *MAINTAINABILITY* (MAMPU DIRAWAT)
2. *DEPENDABILITY* (MAMPU BERKEMBANG)
3. *ROBUST* (MENGIKUTI KEINGINAN PENGGUNA)
4. *EFFECTIVE* (EFEKTIF) DAN *EFFICIENT* (EFISIEN)
5. *USABILITY* (MEMENUHI KEBUTUHAN)

B. *Sistem Pakar*

Sistem pakar adalah sistem yang berusaha mengadopsi pengetahuan manusia ke komputer yang dirancang untuk memodelkan kemampuan menyelesaikan masalah seperti layaknya seorang pakar dalam bidang tersebut [3]. Dengan menggunakan sistem pakar, orang awam dapat menyelesaikan masalahnya atau membantu menyelesaikan masalahnya atau mencari suatu informasi yang diperoleh dari bantuan para ahli dibidangnya. Komponen-komponen yang terdapat dalam struktur sistem pakar sebagai berikut:

1. ANTARMUKA PENGGUNA (USER INTERFACE)
2. BASIS PENGETAHUAN
3. AKUISISI PENGETAHUAN
4. MESIN INFERENSI
5. FASILITAS PENJELASAN
6. PERBAIKAN PENGETAHUAN

C. *Scikit Learn*

Scikit-learn adalah modul Python yang menyediakan berbagai macam algoritma machine learning. Scikit-learn menggunakan *task-oriented interface* yang konsisten sehingga dapat dengan mudah dilakukan perbandingan antara metode [3]. Scikit-learn tersedia dalam bentuk library python.

D. *Flask*

Suatu web *application framework* yang terdapat pada bahasa pemrograman *Python*. *Flask* digunakan sebagai server yang melakukan perhitungan dan klasifikasi data uji terhadap dataset [4]. *Flask* digunakan untuk mempersingkat dan mempermudah pengembangan Web *Application Flask*.

E. *Pohon Keputusan*

Metode ini merupakan salah satu metode yang ada pada teknik klasifikasi dalam data mining. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Pohon keputusan juga berguna untuk mengeksplorasi data, menemukan hubungan tersembunyi antara sejumlah calon variabel input dengan sebuah variabel target [5].

Dalam pembuatan pohon keputusan digunakan gini untuk mencari *parent* dan *child* pada pohon keputusan.

$$Gini(t) = 1 - \sum_{i=0}^{c-1} [p(i|t)]^2 \quad (1)$$

$$\Delta = I(\text{perent}) - \sum_{j=1}^k \frac{N(v_j)}{N} I(v_j) \quad (2)$$

F. Cross Validation

Cross-validation adalah teknik di mana kita melatih model kita menggunakan subset dari data-set dan kemudian mengevaluasi menggunakan subset pelengkap dari data-set. Hal utama dibalik *cross-validation* adalah untuk membagi data, sekali atau beberapa kali, untuk memperkirakan resiko dari masing-masing algoritma: Bagian dari data (sampel pelatihan) digunakan untuk melatih setiap algoritma, dan bagian yang tersisa (sampel validasi) digunakan untuk memperkirakan risiko algoritma [6].

Fold	Data	Subset
Fold 1	Training Testing	$S_2, S_3, S_4, S_5, S_6, S_7, S_8, S_9, S_{10}$ S_1
Fold 2	Training Testing	$S_1, S_3, S_4, S_5, S_6, S_7, S_8, S_9, S_{10}$ S_2
Fold 3	Training Testing	$S_1, S_2, S_4, S_5, S_6, S_7, S_8, S_9, S_{10}$ S_3
Fold 4	Training Testing	$S_1, S_2, S_3, S_5, S_6, S_7, S_8, S_9, S_{10}$ S_4
Fold 5	Training Testing	$S_1, S_2, S_3, S_4, S_6, S_7, S_8, S_9, S_{10}$ S_5
Fold 6	Training Testing	$S_1, S_2, S_3, S_4, S_5, S_7, S_8, S_9, S_{10}$ S_6
Fold 7	Training Testing	$S_1, S_2, S_3, S_4, S_5, S_6, S_8, S_9, S_{10}$ S_7
Fold 8	Training Testing	$S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_9, S_{10}$ S_8
Fold 9	Training Testing	$S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_{10}$ S_9
Fold 10	Training Testing	$S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8, S_9$ S_{10}

Gambar 1. Skema Cross validation

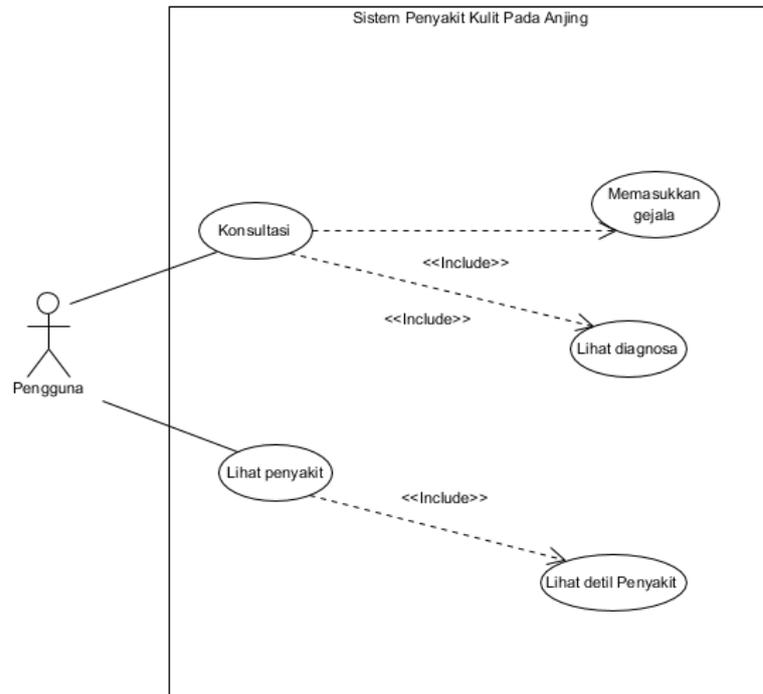
Pada gambar 1 merupakan contoh skenario uji *cross-validation* dengan jumlah 10-fold dan pengujian yang dilakukan sebanyak 10 kali. Pengujian pertama pada fold pertama menggunakan data yang terdiri dari data subset ke-2 sampai subset ke-9 dan sisanya akan digunakan sebagai data testing pada subset pertama. Proses testing dan training akan dilakukan sampai pengujian ke-10.

III. ANALISIS DAN RANCANGAN

A. Metode Penelitian

Metode penelitian yang dipakai adalah kuantitatif. Pada penelitian menggunakan metode kuantitatif ini akan memanfaatkan pengguna grafik, diagram dan lainnya. Data penelitian yang dilakukan pada paper IPB digunakan untuk membantu penelitian ini. Pembuatan model pohon keputusan pada penelitian ini akan menggunakan data yang didapat dari paper IPB [7]. Pohon keputusan yang digunakan akan dibuat menggunakan *scikit-learn*.

B. Use Case



Gambar 2. Use case dari Sistem Penyakit kulit pada anjing

Berikut merupakan penjelasan Gambar 2. Pengguna dapat melakukan konsultasi dan melihat penyakit. Pada konsultasi pengguna akan diminta memasukkan gejala. Pada use case Lihat Penyakit, pengguna dapat melihat detail penyakit dari setiap penyakit.

C. Dataset

Dataset memiliki peran penting dalam penelitian ini, karena untuk membantu dalam pengembangan sistem pakar yang dipakai. Dataset dalam penelitian ini adalah dataset yang didapat dari Joanna Fransisca (2006) [7]. Paper ini direkomendasikan oleh dokter atau pakar secara langsung. Pada paper ini dataset yang dipakai berjumlah seratus lima (105) dengan jumlah sebelas (11) penyakit dan sebelas (11) gejala. Dataset yang dibangun merupakan kumpulan gejala dan penyakit.

Gambar 3 adalah model data yang didapat pada paper. Pembacaan data pada gambar tersebut yaitu pada penyakit atopy ditemukan 4 gejala alopecia, pada penyakit caplak ditemukan 1 gejala bintik.

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	-	-
dermodectosis	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	-	-	-

Gambar 3. Data Pada Paper[7]

	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

Gambar 4. Data yang digunakan pada *scikit-learn*

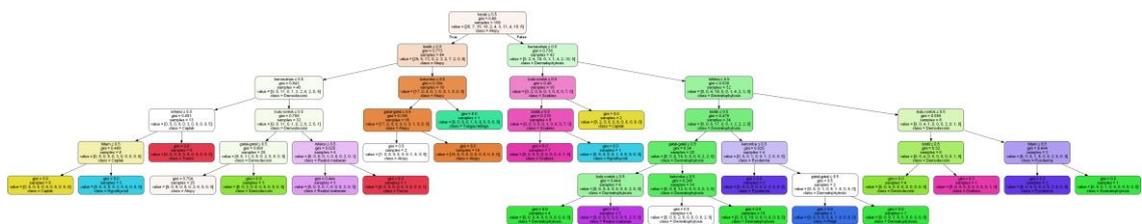
Gambar 4 merupakan data yang telah diubah dan yang digunakan pada *scikit-learn*. Data yang diperoleh setelah dilakukan perubahan adalah 105 data.

D. Pohon Keputusan (Decision Tree)

Gambar 5 menjelaskan bagaimana cara menghasilkan gambar pohon keputusan seperti gambar 6 dengan menggunakan *scikit-learn*. Hasil pohon keputusan yang muncul berisikan hasil perhitungan gini secara otomatis seperti contoh pada root pohon keputusan tersebut mendapatkan gini yaitu **0,89**.

```
import graphviz
dot_data = tree.export_graphviz(clf, out_file=None, feature_names=features,
class_names=Penyakit, filled=True, rounded=True, special_characters=True)
graph = graphviz.Source(dot_data)
graph
```

Gambar 5. Code Untuk Menghasilkan Pohon Keputusan

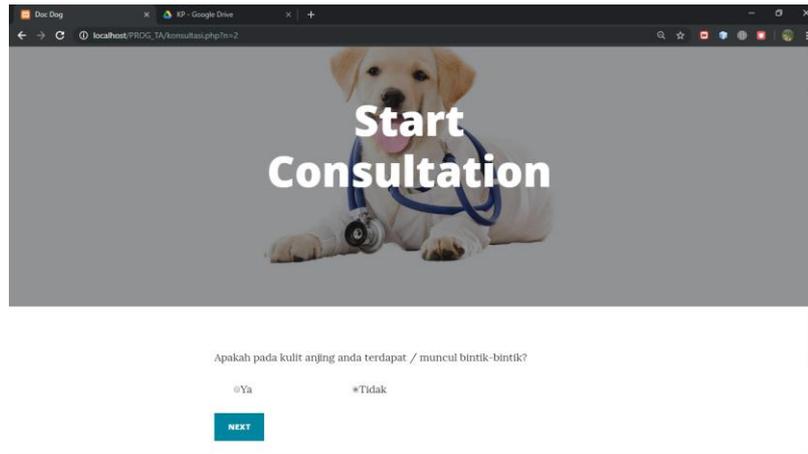


Gambar 6. Pohon Keputusan

IV. IMPLEMENTASI

A. Implementasi UI a. Home

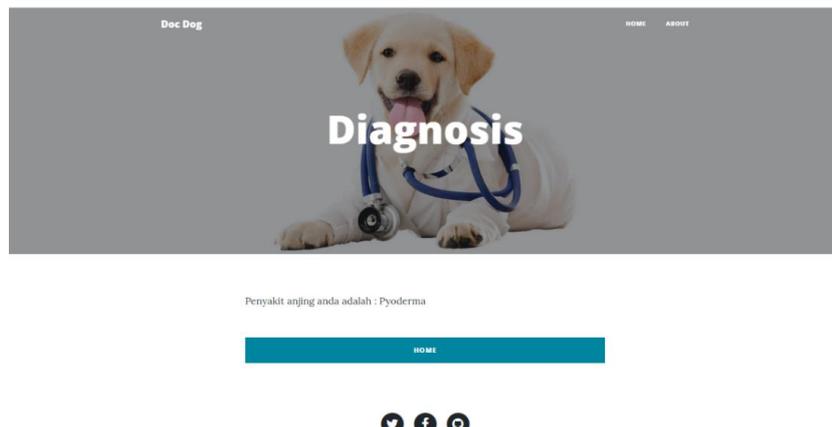
Gambar 7 Menunjukkan hasil implementasi untuk tampilan pada konsultasi. Pada halaman ini pengguna akan diminta untuk mengisi pertanyaan yang diberikan program. Total pertanyaan yang dijawab oleh pengguna adalah 11 pertanyaan. Ketika pertanyaan selesai dijawab maka program akan memunculkan hasil dari konsultasi.



Gambar 7. Konsultasi

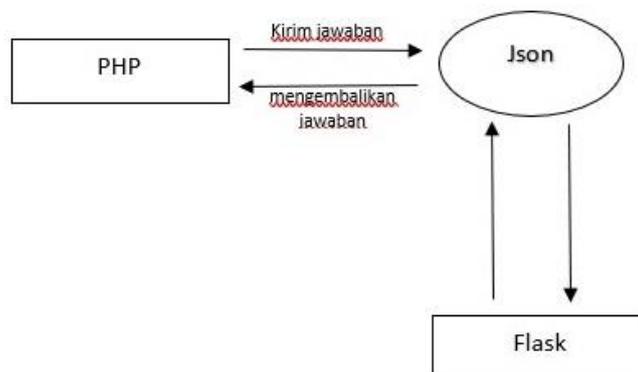
b. Hasil diagnose

Gambar 8 menunjukkan hasil implementasi untuk tampilan pada hasil diagnosis. Pada halaman ini pengguna akan menerima hasil dari diagnosis setelah melakukan konsultasi. Pada halaman ini akan memunculkan nama penyakit yang terdapat pada anjing dan penanganan awal.



Gambar 8. Hasil Konsultasi

B. Implementasi program



Gambar 9. Proses Program Bekerja

Gambar 9 merupakan proses program bekerja. Pada php akan mendapatkan jawaban yang telah diberikan oleh pengguna ketika konsultasi. Setelah jawaban didapat akan dikirim kepada json. Json akan mengirim jawaban berupa array pada flask. Jawaban yang diterima pada flask akan dites pada data training. Flask akan mendapatkan jawaban dari data test tersebut dan akan mengirim kembali pada json untuk ditampilkan pada web.

V. PENGUJIAN

A. Usability Testing

Langkah pertama yang dilakukan dalam usability testing adalah memberikan sejumlah tugas yang sudah dipersiapkan sebelumnya kepada responden saat berinteraksi dengan sistem yang diuji. Pada pengujian ini program dicoba kepada pengguna yang dikhususkan memiliki anjing.

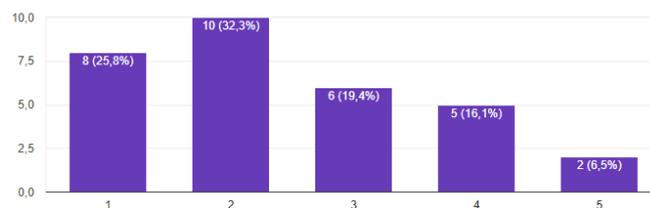
Berikut ini adalah tugas-tugas yang diberikan kepada pengguna :

1. Pengguna membuka web yang telah diberikan dan memperhatikan tampilan web.
2. Pengguna mengakses dan mencoba menu informasi dan melakukan konsultasi.
3. Pengguna melakukan konsultasi dengan menjawab pertanyaan yang terdapat pada web

B. Hasil Pengujian dengan Kuisisioner

6. Dalam penggunaannya web tersebut apakah membantu mempermudah dalam mengetahui hasil diagnosa dari penyakit kulit anjing anda?

31 tanggapan



Gambar 10. Hasil Survey

Gambar 10 merupakan hasil persentase dari pertanyaan “dalam penggunaannya web tersebut apakah membantu mempermudah dalam mengetahui hasil diagnosis dari penyakit kulit anjing anda?” dari 31 response, 8 orang (25,8%) menjawab sangat setuju, 10 orang (32,3%) menjawab setuju, 6 orang (19,4%) menjawab cukup setuju, 5 orang (16,1%) menjawab kurang setuju, 2 orang (6,5%) menjawab tidak setuju.

C. Analisis

Pada analisis berikutnya, sepuluh kasus penyakit kulit anjing yang telah didiskusikan dengan dokter secara langsung diambil untuk dianalisis. Kasus yang benar akan dinilai satu (1) sedangkan kasus yang salah akan bernilai nol (0). Dari 10 kasus yang diambil, program memprediksi 10 kasus tersebut dengan delapan benar dan dua salah.

TABLE 1
 HASIL UJI 10 KASUS

No	Hasil diagnosa	Benar / Salah
1	Dermatophytosis	Benar
2	Tungau telinga	Benar
3	Atopy	Benar
4	Hypothyroid	Benar
5	Reaksi makanan	Benar
6	Caplak	Benar

7	Demodecosis	Benar
8	Dermatophytosis	Benar

Berdasarkan tabel 1 dari 10 kasus, nilai akurasi yang dihasilkan adalah seperti berikut :

$$Akurasi = \frac{\sum match}{Semua\ data} \times 100\% \quad (3)$$

$$akurasi = \frac{8}{10} \times 100\% \quad (4)$$

$$akurasi = 80\% \quad (5)$$

Jadi dapat disimpulkan bahwa akurasi sistem pakar berdasarkan 10 kasus yang diuji adalah 80% yang menunjukkan bahwa program ini berfungsi baik.

D. Cross-validation

Pada pengujian yang dilakukan dengan 25 kali mendapatkan hasil seperti gambar 11 dan gambar 12. Hasil rata-rata yang didapat dari cross-validation pada gambar 5.14 yaitu 0,743.

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

Gambar 11. Cross Validation Sebanyak 25

```
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
```

Gambar 12. Hasil Cross Validation Sebanyak 25

Dengan hasil yang telah didapat pada cross validation maka dapat disimpulkan bahwa rata-rata adalah seperti pada Tabel 2. Rata-rata yang didapat dihitung dari hasil perhitungan penjumlahan dari setiap cross validation yang dilakukan dengan jumlah cross validation yang dilakukan dan menghasilkan rata-rata sebesar 0,676.

TABLE 2
HASIL RATA-RATA CROSS VALIDATION

<i>k-fold cross validation</i>	Akurasi
k=10	0,643
k=15	0,642
k=20	0,698
K=25	0,743

Rata-rata	0,676
-----------	-------

VI. SIMPULAN DAN SARAN

A. Simpulan

Berdasarkan penyebaran survei pada 31 responden yang memiliki anjing dan pengujian dengan 10 kasus yang didapat dari pakar dapat disimpulkan pengimplementasian sebuah sistem pakar berbasis web ini dapat membantu dalam mendiagnosa gejala penyakit dan memprediksi penyakit kulit pada anjing. Dengan dataset yang digunakan sebanyak 105 yang terdiri dari 11 gejala dan 11 penyakit. Penelitian dan program sistem pakar ini mendapatkan hasil yang baik dengan data yang diperoleh dari survei yang telah dilakukan dapat diambil sebuah kesimpulan bahwa program yang dikembangkan dapat membantu responden dalam mengetahui penyakit kulit yang dialami oleh anjing pengguna. Melalui salah satu survei membuktikan sekitar 8 orang (25,8%) menjawab sangat setuju dan 10 orang (32,3%) menjawab setuju. Pada analisis yang dilakukan dengan 10 kasus yang didapat dari pakar atau dokter mendapatkan akurasi sebesar 80%. Pada *cross-validation* yang dilakukan sebanyak 4 kali mendapatkan rata-rata sebesar 6,76%.

B. Saran

Berdasarkan pada kesimpulan maka ada beberapa pengembangan yang dapat dilakukan di masa yang akan datang, yaitu antara lain:

1. Program dapat dikembangkan menjadi aplikasi mobile. Hal ini akan membantu pengguna karena umumnya pengguna lebih sering menggunakan aplikasi mobile daripada menggunakan aplikasi web.
2. Program tidak hanya menangani penyakit kulit anjing saja tetapi berkembang dengan menangani penyakit anjing lainnya.
3. Data yang digunakan bisa berkembang dengan menggunakan data berupa gambar penyakit. Dengan data berupa gambar diharapkan aplikasi dapat dikembangkan dengan dapat mendeteksi penyakit kulit yang terdapat pada anjing lebih akurat.

DAFTAR PUSTAKA

- [1] P. and C. Hetty, "Aplikasi Web Sistem Pakar untuk Diagnosa Penyakit Gizi," vol. 1, p. 2, 2018.
- [2] C. J. Kalung, A. S. M. Lumenta and F. D. Kambey, "SISTEM PAKAR PENGENALAN DAN PENANGANAN AWAL PENYAKIT PADA ANJING DI MANADO".
- [3] F. Pedegrosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blodel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderpias, A. Passos, D. Cournapeau, M. Brucher and E. Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825-2830, 2011.
- [4] M. F. Ponighzwa Rizkanda, R. Sarno and D. Sunaryo, "Rancangan Bangun Aplikasi MusicMoo dengan Metode MIR (Music Information Retrieval) pada Modul Fingerprint dan Song Recommendation," *Jurnal Teknik ITS*, vol. 6, p. A239, 2017.
- [5] F. Slamet and B. D. Meilani Achmad, "Klasifikasi Data Karyawan untuk Menentukan Jadwal Kerja Menggunakan Metode Decision Tree," *Jurnal IPTEK*, vol. 16, no. 1, p. 18, 2012.
- [6] S. Arlot and A. Calisse, "A Survey of cross-validation procedures for model selection," *Statistic survey*, vol. 4, no. 1935-7516, p. 42, 2010.
- [7] J. Francisca, "Profil Penyakit Kulit pada Pasien Anjing: Studi Kasus di Klinik Dokter Hewan Praktek Bersama Drh. Cucu K. Sajuthi Dkk Ruko Green Garden Jakarta," p. 38, 2006.