

IMPLEMENTASI *MACHINE LEARNING* UNTUK DETEKSI PENYAKIT PADA KUCING MENGGUNAKAN *RANDOM FOREST*

Maulana Ichsan^{1*}

^{1,2,3} Universitas Bina Sarana Informatika, Teknik dan Informatika, Informatika

Keywords:

Machine Learning, Random Forest, Disease Detection, Cat Health, Diagnosis System

Correspondent Email:

maulanaichsan257@gmail.com

Abstrak. Kesehatan kucing merupakan prioritas bagi pemilik hewan peliharaan, namun kurangnya pengetahuan seringkali menjadi kendala dalam deteksi dini penyakit. Keterlambatan diagnosis dapat menyebabkan komplikasi serius dan penanganan yang lebih sulit. Penelitian ini bertujuan untuk mengimplementasikan algoritma *machine learning Random Forest* guna membangun sistem deteksi penyakit pada kucing berdasarkan gejala klinis, serta memberikan rekomendasi perawatan awal. Metode penelitian meliputi pengumpulan data, pra-pemrosesan dengan *CountVectorizer* dan *LabelEncoder*, pemodelan menggunakan *Random Forest*, dan implementasi ke dalam prototipe aplikasi *command-line* bernama "Valary Care". Dataset yang digunakan terdiri dari 100 rekam medis yang mencakup 5 jenis penyakit dan 35 gejala unik, yang dibagi menjadi 80% data latih dan 20% data uji. Hasil pengujian model menunjukkan performa yang sangat baik dengan tingkat akurasi 90%, presisi rata-rata 91%, *recall* rata-rata 89%, dan *F1-Score* rata-rata 89%. Tingkat akurasi yang tinggi ini membuktikan bahwa model *Random Forest* efektif dan dapat diandalkan sebagai alat bantu diagnosis awal penyakit pada kucing, sehingga dapat membantu pemilik dalam mengambil langkah penanganan yang lebih cepat dan tepat.



Copyright © [JITET](http://www.jitet.org) (Jurnal Informatika dan Teknik Elektro Terapan). This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY NC)

Abstract. Cat health is a priority for pet owners, yet a lack of knowledge often hinders the early detection of diseases. Delayed diagnosis can lead to serious complications and more difficult treatments. This research aims to implement the *Random Forest machine learning* algorithm to build a disease detection system for cats based on clinical symptoms, as well as to provide initial care recommendations. The research method includes data collection, pre-processing using *CountVectorizer* and *LabelEncoder*, modeling with *Random Forest*, and implementation into a *command-line* prototype application named "Valary Care". The dataset consists of 100 medical records covering 5 types of diseases and 35 unique symptoms, which was split into 80% training data and 20% testing data. The model evaluation showed excellent performance, achieving an accuracy of 90%, an average precision of 91%, an average recall of 89%, and an average *F1-Score* of 89%. This high level of accuracy proves that the *Random Forest* model is effective and reliable as an auxiliary tool for the initial diagnosis of cat diseases, thereby helping owners take faster and more appropriate treatment steps

1. PENDAHULUAN

Kucing termasuk hewan peliharaan yang populer di berbagai kalangan. Namun, sama

seperti makhluk hidup lainnya, kucing juga rentan terhadap berbagai penyakit yang dapat mengganggu kesehatannya. Beberapa masalah

kesehatan umum yang sering dialami kucing meliputi gangguan pencernaan, infeksi saluran pernapasan, infeksi saluran kemih, diabetes, dan ginjal kronis. Deteksi dini penyakit pada kucing menjadi tantangan utama bagi pemilik karena gejala yang muncul seringkali tidak spesifik atau sulit dikenali. Keterlambatan dalam diagnosis dapat menyebabkan kondisi kucing semakin memburuk, bahkan berujung pada komplikasi serius.

Salah satu masalah yang sering dihadapi pemilik kucing adalah kurangnya pengetahuan tentang gejala penyakit dan cara penanganan yang tepat. Hal ini diperparah oleh keterbatasan akses ke dokter hewan, terutama di daerah dengan fasilitas kesehatan hewan yang terbatas. Selain itu, pemilihan makanan dan obat yang tidak sesuai dengan kondisi kucing dapat memperburuk keadaan. Oleh karena itu, diperlukan solusi yang dapat membantu pemilik kucing dalam mendeteksi penyakit sejak dini dan memberikan rekomendasi perawatan yang tepat.

“Permasalahan yang sering dialami oleh pemilik kucing adalah kurangnya pengetahuan tentang gejala-gejala penyakit dialami oleh kucing mereka, serta ketidaktahuan dalam melakukan penanganan pengobatan yang dapat memperburuk kondisi kesehatan kucing.”[1]. “Deteksi dini penyakit dalam perawatan hewan bergantung pada identifikasi kelainan subklinis pada hewan tanpa gejala selama kunjungan kesehatan.” [2].

Seiring dengan meningkatnya kesadaran masyarakat tentang pentingnya perawatan kesehatan hewan peliharaan seperti kucing, kebutuhan untuk menyediakan sistem deteksi dini dan perawatan yang lebih efektif semakin mendesak. Kucing sebagai salah satu hewan peliharaan yang populer, sering kali mengalami berbagai masalah kesehatan yang terkadang sulit terdeteksi pada tahap awal. Gejala-gejala penyakit pada kucing sering kali tidak terlihat jelas, sehingga pemilik kucing sering kali baru menyadari kondisi kesehatan hewan peliharaannya setelah penyakit tersebut berkembang lebih lanjut.

Teknologi machine learning, khususnya algoritma *Random Forest*, menawarkan potensi besar untuk mengatasi masalah ini. Algoritma ini mampu perawatan yang sesuai. menganalisis data gejala klinis dan mengidentifikasi pola yang mungkin tidak terlihat oleh manusia.

Dengan memanfaatkan data historis dan gejala yang diamati, sistem berbasis machine learning dapat memberikan diagnosis awal serta rekomendasi perawatan yang sesuai.

Sekarang hampir semua aspek hidup terasa tidak lepas dari teknologi, dari bangun pagi sampai tidur lagi alat elektronik dan aplikasi pintar ikut membantu. Yang dulu kita anggap aktivitas simpel, seperti mencari tahu cara merawat hewan, sekarang sering kita serahkan pada gadget yang memberi jawaban dalam sekejap. Salah satu contoh teknologi yang bertujuan memberi solusi spesifik adalah sistem pakar. Di Indonesia, kucing merupakan hewan peliharaan yang sangat populer, terutama di kalangan warga kota. Para pemilik kucing umumnya sangat peduli pada kesehatan kucing, karena penyakit yang menyerang hewan itu bisa dengan mudah menjangkit hewan peliharaan lain atau bahkan membahayakan manusia[3].

“Machine learning dapat memberi kita cara baru untuk memeriksa tumpukan data besar dan menemukan pola yang biasanya luput dari perhatian. Ketika diterapkan pada dunia kucing, teknologi ini bisa meneliti gejala yang kita laporkan dan memberi 3 dugaan awal tentang sakit apa yang mungkin sedang dihadapi hewan tersebut.”[4].

Berdasarkan permasalahan diatas, maka peneliti akan membuat penerapan sistem Machine Learning untuk deteksi awal penyakit pada kucing serta rekomendasi makanan dan obat yang tepat yang berbentuk prototipe program. Sistem untuk penyelesaian masalah ini menggunakan metode *Random Forest*. Untuk mempermudah penelitian maka sistem akan diterapkan menjadi program prototipe dan diharapkan dapat bekerja sesuai yang direncanakan. Sistem yang dibangun dapat mendiagnosa penyakit yang diderita pada kucing dengan melihat gejala yang nampak pada kucing dan memberi rekomendasi makanan dan obat yang tepat untuk penanganan penyakit yang dialami pada kucing.

2. TINJAUAN PUSTAKA

2.1.1. Machine Learning

“Machine Learning adalah bidang ilmu yang mempelajari cara membuat komputer mampu belajar dan meningkatkan kemampuannya dengan meniru atau bahkan melampaui kemampuan belajar manusia.”[5]. Menurut N. Giarsyani, Machine learning adalah salah satu

cabang keilmuan dari kecerdasan buatan yang mengadopsi prinsip ilmu komputer dan statistik untuk menciptakan model yang merefleksikan pola dari sebuah data masukan. Dalam proses pembuatan model tersebut dapat menggunakan berbagai macam algoritma cerdas yang dapat disesuaikan dengan kebutuhan [6]. Teknologi machine learning telah membuka peluang baru dalam meningkatkan akurasi prediksi suhu. Model-model machine learning menawarkan pendekatan yang lebih fleksibel dan adaptif dibandingkan metode statistik tradisional [7]. Menurut F. Yulian Pamuji, Machine Learning telah menunjukkan beberapa keunggulan, termasuk mampu menghasilkan kesalahan yang relatif rendah, kinerja klasifikasi yang baik, mampu menangani sejumlah besar data pelatihan secara efisien, serta metode yang efektif untuk mengestimasi missing data [8]. “Menurut Pineda AL, machine learning semakin banyak digunakan untuk menyusun dan mendapatkan wawasan dari data terstruktur, semi-terstruktur, dan tidak terstruktur dalam jumlah besar di kedokteran hewan, memfasilitasi ekstraksi informasi berharga secara otomatis dari narasi klinis.” [9].

2.1.2. Random Forest

Random Forest adalah algoritma ensemble learning yang menggabungkan beberapa pohon keputusan untuk meningkatkan akurasi prediksi dan mengurangi overfitting. Keunggulan *Random Forest* terletak pada kemampuannya menangani data dengan dimensi tinggi dan interaksi antar variabel yang kompleks.

Keuntungan signifikan dari penggunaan *Random Forest* adalah kemampuannya untuk mengukur dan memeringkat pentingnya setiap fitur (gejala). Hal ini tidak hanya meningkatkan akurasi diagnosis tetapi juga memberikan wawasan klinis mengenai gejala mana yang paling prediktif terhadap suatu penyakit [10].

Manfaat utama dari penggunaan *Random Forest* adalah kemampuannya untuk memberikan keakuratan yang tinggi untuk memungkinkan deteksi penyakit pada kucing lebih cepat dan akurat dibandingkan dengan metode manual dengan menggunakan data gejala yang ada, mengurangi overfitting untuk meminimalkan kemungkinan kesalahan dalam diagnosa penyakit pada kucing dan analisis gejala yang kompleks untuk menangani data yang bersifat non-linear karena setiap pohon keputusan dalam *Random Forest* dapat

membuat keputusan berdasarkan pemisahan data yang lebih kompleks, bukan hanya dengan garis lurus, multi-dimensional untuk menangani banyak fitur sekaligus dan mencari pola yang ada di antara fitur yang berbeda. Hal ini sangat berguna ketika data memiliki banyak variabel yang saling terkait, sehingga bisa mengenali pola dari data gejala yang lebih kompleks dan menghasilkan diagnosis yang lebih tepat. “Algoritma *Random Forest* bekerja dengan membangun sekumpulan pohon keputusan, masing-masing memberi suara berdasarkan sampel dan subkelompok fitur 8 yang dipilih secara acak, sehingga prediksi akhir diperoleh melalui proses voting yang mengutamakan akurasi dan ketahanan.” [2].

Random Forest adalah metode pembelajaran mesin yang menggabungkan banyak pohon keputusan untuk meningkatkan akurasi dan stabilitas prediksi, dengan cara membagi data ke dalam beberapa kelompok sesuai ciri-ciri tertentu lalu mengambil suara mayoritas dari setiap pohon untuk menghasilkan keputusan akhir [11].

2.1.3. Machine Learning dalam Deteksi Penyakit Hewan

Machine Learning telah menjadi pendekatan yang efektif dalam bidang kesehatan hewan, khususnya untuk deteksi dini penyakit. Machine Learning memungkinkan analisis data besar dan kompleks untuk mengidentifikasi pola yang tidak mudah dikenali oleh metode konvensional. Machine Learning telah menjadi alat penting dalam bidang kesehatan hewan, memungkinkan analisis data besar untuk mendeteksi penyakit secara dini. Machine Learning memungkinkan identifikasi pola berbagai penyakit, termasuk pada kucing, kompleks dalam data klinis, yang sulit dikenali oleh metode konvensional. Dalam konteks kesehatan hewan, Machine Learning telah diterapkan untuk mendiagnosis

“Menurut Purwanti, Dengan kekuatan pemrosesan data yang mengesankan, machine learning sudah jadi kawan yang handal bagi para veteriner ketika harus menangani ribuan kasus kucing yang datang setiap hari. Mesin mencermati data yang terkumpul meliputi gejala, riwayat vaksin, perilaku, dan bahkan kedalaman napas. mengubah ke dalam model matematis, dan tanpa lelah mencari pola halus yang bisa lolos dari mata dokter sekalipun. Di tengah tumpukan record kucing itu, misalnya,

algoritma bisa segera menangkap kesamaan kecil antara hewan yang kemarin lemas dan siang ini demam, lalu melaporkan kemungkinan virus, atau meramalkan bahwa hasil tes langka yang baru ditemukan harus dicermati sewaktu-waktu nanti.” [4].

“Teknik pembelajaran mesin otomatis dapat mempermudah proses (diagnosis) dengan secara otomatis memilih model, mengoptimalkan hyperparameter, dan mengurangi ketergantungan pada intervensi manusia. Pendekatan ini juga dapat meringankan beban kerja ahli.” [12].

“Menurut Anggriani, Dengan sistem machine learning, klinik hewan mampu menghadirkan pelayanan yang lebih efisien, mempercepat langkah-langkah diagnosa, dan mengurangi ketergantungan pada metode manual yang biasanya memakan waktu. Alat ini menyediakan dukungan bagi dokter hewan dalam proses pengambilan keputusan medis dan, pada saat yang sama, menghadirkan informasi yang lebih akurat kepada pemilik kucing. Hasilnya, pemilik dapat memperoleh gambaran yang jelas mengenai kesehatan hewan kesayangan mereka, sehingga langkah-langkah perawatan dapat dilakukan dengan tepat dan segera.” [4].

2.1.4. Penerapan *Random Forest* dalam Rekomendasi Makanan dan Obat

Meskipun penelitian langsung mengenai penggunaan *Random Forest* untuk rekomendasi makanan dan obat pada kucing masih terbatas, pendekatan serupa telah diterapkan pada hewan lain. Menggunakan *Random Forest* untuk memprediksi konsentrasi minimum inhibitor (MIC) antibiotik pada *Salmonella* non-tifoid di Taiwan, yang dapat membantu dalam pemilihan antibiotik yang tepat [13].

2.1.5. Google Colab

Pengembangan dan konversi model Machine Learning dimulai dengan mengunggah berkas dataset ke Google Drive. Kemudian, berkas tersebut dimuat ke Google Colab agar bisa diproses dengan mudah dan cepat. Ketika model sudah siap dan sudah melalui evaluasi, notebook Colab yang berisi kode dan latih evaluasi semua hasil dipindah dan disimpan ke Google Drive untuk arsip dan pelacakan. Dari sana, pengguna memulai kode programnya, menggunakan bahasa Python di Google Colab dan memanggil TensorFlow. Pembangunan model bisa dimulai. Setelah model dilatih dan

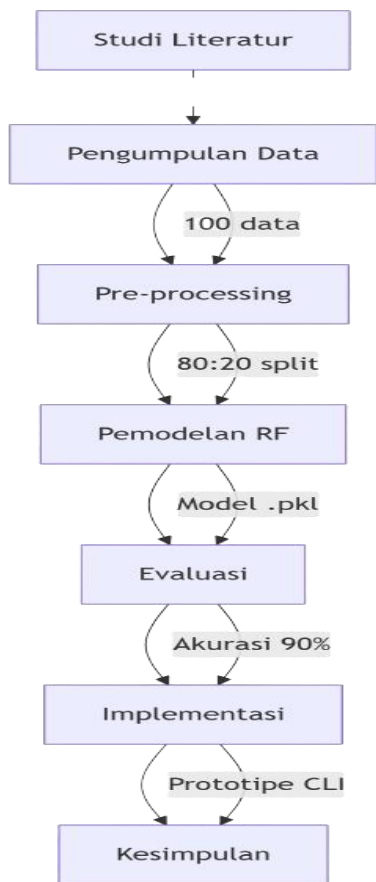
evaluasi siap, model tersebut diubah ke format TensorFlow Lite. Konversi ke TensorFlow Lite adalah langkah terakhir, dan format ini dioptimalkan agar dapat diujikan dengan ringan dan cepat di aplikasi mobile dengan memanfaatkan perangkat Android. Dengan demikian, proses pengembangan dan konversi model Machine Learning ini memungkinkan model untuk diterapkan dan dijalankan dengan baik pada perangkat seluler dalam format yang sesuai [14].

Google Colab kasih kemudahan buat siapa pun yang mau nulis sekaligus jalani kode Python, semua lewat browser. Dengan banner dari Google, platform ini diklaim pas banget buat semua yang lagi fokus di machine learning, deep learning, maupun analisis data. Keunggulan Colab paling menonjol adalah kita nggak perlu repot konfigurasi, bisa akses gratis ke GPU, dan share notebook ke rekan dengan sekali klik. Dengan menggunakan Colab Notebook, programmer bisa mengkombinasikan kode yang bisa dieksekusi dan rich text dalam satu dokumen tunggal bersama dengan image, HTML, LaTeX. Google Colab dapat berkolaborasi dengan pengguna lainnya melalui berbagai kode secara online [15].

“Tahapan memproses data menggunakan Google Colab adalah melalui tahapan impor data, seleksi data, cleaning data, preprosesing, proses menjalankan algoritma data dan kesimpulan hasil.” [16].

3. METODE PENELITIAN

Proses penelitian ini akan dilaksanakan secara sistematis melalui beberapa tahapan utama, mulai dari studi literatur hingga penarikan kesimpulan. Setiap tahapan memiliki tujuan dan *output* yang jelas untuk mendukung tahapan berikutnya. Secara *visual*, alur atau langkah-langkah penelitian dapat digambarkan dalam diagram alir (*flowchart*) berikut:



3.1.1. Studi Literatur

Melakukan studi literatur untuk mengumpulkan landasan teori yang kokoh. Kegiatan ini mencakup pemahaman konsep machine learning, algoritma *Random Forest*, serta meninjau penelitian-penelitian terdahulu yang relevan dengan diagnosis penyakit hewan untuk memposisikan penelitian ini. 16 Tahap awal penelitian adalah melakukan studi literatur yang mendalam. Kegiatan ini bertujuan untuk:

1. Mempelajari penelitian-penelitian sebelumnya yang relevan dengan deteksi penyakit hewan, khususnya kucing, menggunakan metode machine learning.
2. Memahami konsep dasar dan cara kerja Algoritma *Random Forest* serta kelebihan dan kekurangannya dibandingkan algoritma klasifikasi lain (misalnya, Naive Bayes, Decision Tree, SVM).
3. Mengidentifikasi gejala umum yang digunakan untuk mendiagnosis berbagai penyakit pada kucing.
4. Mengumpulkan landasan teori yang kuat untuk mendukung perancangan dan implementasi sistem.

3.1.2. Pengumpulan dan Persiapan Data

Tahap ini berfokus pada pengumpulan bahan utama penelitian, yaitu dataset penyakit kucing. Data kemudian dimuat ke dalam Google Colab untuk persiapan sebelum diolah lebih lanjut. Kualitas model machine learning sangat bergantung pada kualitas data yang digunakan. Pada tahap ini, data mengenai riwayat penyakit kucing akan dikumpulkan.

1. Sumber Data: Data yang digunakan adalah data sekunder yang diperoleh dari dataset publik dan data primer yang didapatkan melalui konsultasi dengan dokter hewan di Juanda Petshop.

2. Struktur Data: Data yang dikumpulkan akan mencakup berbagai atribut berupa gejala klinis (misalnya, demam, nafsu makan menurun, muntah, diare, lesu, dll.), nama penyakit, usia, rekomendasi makanan dan rekomendasi obat.

Berikut beberapa contoh dataset penyakit kucing:

Penyakit	Gejala	Usia	Rekomendasi Makanan	Rekomendasi Obat
Feline Lower Urinary Tract Disease (FLUTD)	menjilati area genital	senior	Royal Canin Urinary S/O	Obat anti-inflamasi
Feline Lower Urinary Tract Disease (FLUTD)	menjilati area genital	dewasa	Happy Cat Minkas Urinary Care	Obat anti-inflamasi
Feline Lower Urinary Tract Disease (FLUTD)	sering ke litter box	dewasa	Happy Cat Minkas Urinary Care	Antibiotik (atas resep dokter)
Feline Lower Urinary Tract Disease (FLUTD)	mengeong kesakitan saat pipis	dewasa	Happy Cat Minkas Urinary Care	Antibiotik (atas resep dokter)
Feline Lower Urinary Tract Disease (FLUTD)	susah buang air kecil	dewasa	Royal Canin Urinary S/O	Obat anti-inflamasi
Feline Lower Urinary Tract Disease (FLUTD)	mengeong kesakitan saat pipis	senior	Royal Canin Urinary S/O	Obat anti-inflamasi (atas resep dokter)
Feline Lower Urinary Tract Disease (FLUTD)	mengeong kesakitan saat pipis	dewasa	Royal Canin Urinary S/O	Antibiotik (atas resep dokter)
Feline Lower Urinary Tract Disease (FLUTD)	air seni berdarah	dewasa	Royal Canin Urinary S/O	Obat anti-inflamasi
Feline Lower Urinary Tract Disease (FLUTD)	sering ke litter box	dewasa	Happy Cat Minkas Urinary Care	Obat anti-inflamasi
Feline Lower Urinary Tract Disease (FLUTD)	mengeong kesakitan saat pipis	dewasa	Royal Canin Urinary S/O	Antibiotik (atas resep dokter)
Feline Lower Urinary Tract Disease (FLUTD)	sering ke litter box	senior	Happy Cat Minkas Urinary Care	Antibiotik (atas resep dokter)
Feline Lower Urinary Tract Disease (FLUTD)	susah buang air kecil	dewasa	Royal Canin Urinary S/O	Obat anti-inflamasi
Feline Lower Urinary Tract Disease (FLUTD)	menjilati area genital	dewasa	Royal Canin Urinary S/O	Antibiotik (atas resep dokter)

3.1.3. Pengolahan Data (Preprocessing Data)

Data mentah yang telah dikumpulkan akan melalui serangkaian proses untuk membersihkan, mengubah, dan menyiapkannya agar sesuai dengan format yang dibutuhkan oleh model machine learning. Data mentah yang telah dikumpulkan sering kali tidak seragam. Oleh karena itu, perlu dilakukan tahap pra-pemrosesan untuk membersihkan dan menyiapkan data sebelum digunakan untuk

melatih model. Data mentah dari file .csv akan melalui serangkaian proses untuk memastikan data siap digunakan untuk pelatihan model.

1. Pembersihan Data (Data Cleaning): Menangani data yang hilang (missing values), data yang tidak konsisten, duplikasi data dan menggabungkan fitur pada data dapat diproses oleh algoritma.
2. Transformasi Data : Mengubah data kategorikal (gejala dan nama penyakit) menjadi format numerik menggunakan teknik seperti Label Encoding agar
3. Membuat Model Vectorized : Fitur teks agar menjadi format numerik
4. Pembagian Data (Data Splitting): Dataset dibagi menjadi dua bagian dengan pembagian 80% data latih (training data) dan 20% data uji (testing data).

3.1.4. Pemodelan dan Pelatihan (Modeling & Training)

Model klasifikasi dibangun menggunakan algoritma *Random Forest*. Model kemudian dilatih menggunakan data yang telah diproses (data latih) agar dapat mengenali pola hubungan antara gejala dan penyakit.

3.1.5. Evaluasi Model

1. Confusion Matrix

Confusion Matrix menyediakan empat metrik dasar yang menjadi fondasi untuk perhitungan metrik evaluasi lainnya:

- a) True Positive (TP): Jumlah data positif yang diprediksi dengan benar sebagai positif. (Contoh: Kucing yang sakit A, diprediksi benar menderita)
- b) True Negative (TN): Jumlah data negatif yang diprediksi dengan benar sebagai negatif. (Contoh: Kucing yang tidak sakit A, diprediksi benar tidak menderita penyakit A).
- c) False Positive (FP): Jumlah data negatif yang salah diprediksi sebagai positif (Type I Error). (Contoh: Kucing yang sehat, tetapi salah diprediksi menderita penyakit A).
- d) False Negative (FN): Jumlah data positif yang salah diprediksi sebagai negatif (Type II Error). (Contoh: Kucing yang sakit A, tetapi salah diprediksi sehat atau menderita penyakit lain).

2. Rumus Matrix Evaluasi

a) Akurasi

Akurasi mengukur persentase total prediksi yang benar (baik positif maupun negatif) dari keseluruhan data. Ini adalah metrik yang paling umum untuk melihat performa model secara keseluruhan.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

b) Presisi

Presisi mengukur tingkat ketepatan dari prediksi positif yang dibuat oleh model. Dari semua data yang diprediksi sebagai positif, berapa persen yang sebenarnya benar-benar positif?

$$Precision = \frac{TP}{TP + FP}$$

c) Recall

Recall mengukur kemampuan model untuk menemukan kembali (mengenali) semua kasus positif yang ada dalam data. Dari semua data yang seharusnya positif, berapa persen yang berhasil ditemukan oleh model?

$$Recall = \frac{TP}{TP + FN}$$

d) F1-Score

F1-Score adalah rata-rata harmonik dari Precision dan Recall. Metrik ini sangat berguna ketika terjadi ketidakseimbangan kelas dalam dataset, karena ia memberikan gambaran performa yang lebih seimbang daripada akurasi. F1-Score akan bernilai tinggi hanya jika Precision dan Recall

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

3.2.1 Pengolahan Data

1. Import Library : Pada bagian ini untuk mengimport beberapa library Python yang umum digunakan dalam project machine learning, khususnya untuk klasifikasi.

2. Pemuatan Data dan Penggabungan Gejala: Pada tahap ini, dataset yang telah sudah ada yaitu (penyakit kucing.csv), akan dimuat ke dalam Google Colab. Proses pemuatan data ini dilakukan menggunakan fungsi "read_csv()" dari library Pandas. Dataset dalam format .csv dimuat menggunakan library 23 "Pandas" di Google Colab. Langkah pertama adalah memeriksa dan menggabungkan gejala dan usia sebagai fitur.

[1/7] MEMUAT DATA...

```
5 data pertama:
0          Penyakit          Gejala
1 Feline Lower Urinary Tract Disease (FLUTD) menjilati area genital
2 Feline Lower Urinary Tract Disease (FLUTD) menjilati area genital
3 Feline Lower Urinary Tract Disease (FLUTD) sering ke litter box
4 Feline Lower Urinary Tract Disease (FLUTD) mengeong kesakitan saat pipis

2          3          4
0 Usia      Rekomendasi Makanan      Rekomendasi Obat
1 senior    Royal Canin Urinary S/O      Obat anti-inflamasi
2 dewasa   Happy Cat Minkas Urinary Care      Obat anti-inflamasi
3 dewasa   Happy Cat Minkas Urinary Care      Antibiotik (atas resep dokter)
4 dewasa   Happy Cat Minkas Urinary Care      Antibiotik (atas resep dokter)
```

Kolom setelah pembersihan:
Index(['Penyakit', 'Gejala', 'Usia', 'Rekomendasi Makanan',
 'Rekomendasi Obat'],
 dtype='object', name=0)

Contoh fitur yang dibuat:
0 menjilati area genital senior
1 menjilati area genital dewasa
2 sering ke litter box dewasa
3 mengeong kesakitan saat pipis dewasa
4 susah buang air kecil dewasa
Name: Fitur, dtype: object

3. Transformasi Fitur (Vektorisasi): Gejala penyakit yang berbentuk teks diubah menjadi format numerik menggunakan metode CountVectorizer dari Scikit learn. Metode ini membangun sebuah "kamus" dari semua gejala yang ada dan 24 mengubah setiap baris data gejala menjadi sebuah vektor numerik yang berisi jumlah kemunculan setiap gejala. Hasilnya adalah matriks fitur (X).

[2/7] ENCODING LABEL...

Mapping penyakit ke kode:

Cacingan: 0
Diabetes Melitus: 1
Feline Herpesvirus: 2
Feline Infectious Peritonitis (FIP): 3
Feline Lower Urinary Tract Disease (FLUTD): 4

4. Transformasi Label (Encoding): Kolom target (nama penyakit) yang juga berbentuk teks diubah menjadi kode numerik unik menggunakan LabelEncoder dari Scikit-learn. Proses ini menghasilkan vektor target (Y) yang dapat diproses oleh model.

3.2.2 Analisis dan Pemodelan Data

1. Pembagian Dataset: Dataset yang telah bersih dan bertransformasi (X dan y) dibagi menjadi dua set dengan proporsi 80:20 menggunakan fungsi train_test_split:

[4/7] MEMBAGI DATA...

Data latih: 400 sampel

Data uji: 100 sampel

- a) Data Latih (80%): Digunakan untuk melatih model *Random Forest* baru.
 - b) Data Uji (20%): Disisihkan untuk menguji performa model terhadap data
2. Pembangunan dan Pelatihan Model: Model klasifikasi dibangun menggunakan algoritma *Random Forest* (*RandomForestClassifier*). Model ini kemudian "belajar" dari data latih (X_train, y_train) untuk memahami pola yang menghubungkan kombinasi gejala dengan diagnosis penyakit tertentu. Berikut hasil dari Training Model *Random Forest*:

RandomForestClassifier
RandomForestClassifier(max_depth=10, n_estimators=200, random_state=42)

3.2.3 Analisis Hasil

1. Evaluasi Performa: Analisis ini berfokus pada pengukuran performa model secara statistik.

[6/7] EVALUASI MODEL...

Classification Report:

	precision	recall	f1-score	support
Cacingan	1.00	0.92	0.96	26
Diabetes Melitus	1.00	1.00	1.00	21
Feline Herpesvirus	0.85	1.00	0.92	11
Feline Infectious Peritonitis (FIP)	1.00	1.00	1.00	12
Feline Lower Urinary Tract Disease (FLUTD)	1.00	1.00	1.00	30
accuracy			0.98	100
macro avg	0.97	0.98	0.98	100
weighted avg	0.98	0.98	0.98	100

Akurasi model: 98.00%

- a) Model yang telah dilatih digunakan untuk membuat prediksi pada data uji (y_pred = model.predict(X_test)).
- b) Hasil prediksi (y_pred) dibandingkan dengan label sebenarnya (y_test) untuk menghitung metrik evaluasi utama: 27
 - 1) Accuracy: Mengukur persentase total prediksi yang benar.

- 2) Precision: Mengukur ketepatan dari prediksi positif.
 - 3) Recall: Mengukur kemampuan model menemukan semua kasus positif.
 - 4) F1-Score: Rata-rata harmonik dari Precision dan Recall.
2. Interpretasi dalam Sistem : Analisis ini berfokus pada bagaimana hasil prediksi digunakan dan disajikan dalam prototipe "Valary Care".
- a) Penarikan Kesimpulan Diagnosis: Output numerik dari model diterjemahkan kembali menjadi nama penyakit yang dapat dibaca manusia.
 - b) Penyajian Informasi Kontekstual: Hasil diagnosis disajikan kepada pengguna tidak hanya sebagai nama penyakit, tetapi juga diperkaya dengan fungsi "age_info"). informasi tambahan yang relevan seperti rekomendasi awal atau informasi penyakit umum berdasarkan usia (seperti yang diimplementasikan pada
 - c) Transparansi Model: Sistem secara terbuka menampilkan nilai akurasi model di menu "Tentang Program" untuk memberikan gambaran kepada pengguna mengenai tingkat keandalan sistem. Analisis ini menilai sejauh mana sistem berhasil mengubah output teknis menjadi sebuah layanan diagnosis yang berguna, informatif, dan bertanggung jawab.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Penelitian

Hasil penelitian mencakup output dari setiap tahapan yang telah dilaksanakan sesuai dengan metodologi pada Bab III, mulai dari persiapan data hingga implementasi sistem.

4.1.1. Hasil Pengolahan Data

Dataset (penyakit kucing.csv) yang digunakan terdiri dari 100 baris data rekam medis. Setelah melalui tahap pra-pemrosesan, data berhasil diolah sebagai berikut:

1. Pembersihan Data: Tidak ditemukan data yang hilang (missing values) atau duplikat, sehingga 100 baris data dapat digunakan sepenuhnya.
2. Transformasi Fitur: CountVectorizer berhasil mengidentifikasi 35 gejala unik dari seluruh dataset. Setiap baris data gejala diubah menjadi sebuah vektor numerik dengan panjang 35.
3. Transformasi Label: LabelEncoder berhasil mengidentifikasi dan mengodekan 5 kelas penyakit yang berbeda, yaitu: Cacingan, Feline Herpesvirus, Feline Lower Urinary Tract Disease, Diabetes Melitus, dan Ginjal Kronis.

4.1.2. Hasil Pembangunan Model

Model klasifikasi berhasil dibangun menggunakan algoritma *Random Forest* dengan parameter yang diatur oleh library Scikit-learn. Dataset kemudian dibagi menjadi 80 data latih (80%) dan 20 data uji (20%). Model dilatih menggunakan 80 data latih untuk mempelajari pola antara gejala dan penyakit.

4.1.3. Hasil Implementasi Sistem Prototipe

Model yang telah dilatih berhasil disimpan `penyakit_kucing_model.pkl`, bersama dengan `vectorizer` dan `label_encoder`. Ketiga file ini kemudian diimplementasikan ke dalam sebuah sistem prototipe aplikasi berbasis command-line bernama "Valary Care". dalam file. Prototipe ini memiliki fungsionalitas utama sebagai berikut, sesuai dengan yang ditampilkan pada menu utamanya:



```
=====
                        SISTEM DIAGNOSA PENYAKIT KUCING BERBASIS AI
=====
```

MENU UTAMA

1. Lakukan Diagnosa
2. Informasi Penyakit Umum
3. Tentang Program
4. Keluar

1. Lakukan Diagnosa: Pengguna dapat memasukkan gejala, dan sistem akan

memberikan hasil prediksi penyakit beserta rekomendasi makanan serta obat yang tepat.

2. Informasi Penyakit Umum: Menyediakan pengetahuan mengenai penyakit umum pada kucing berdasarkan tiga kategori usia (Kitten, Dewasa, Senior), lengkap dengan deskripsi, daftar penyakit, dan tips perawatan.
3. Tentang Program: Menampilkan informasi mengenai sistem, termasuk versi dan nilai akurasi model yang berhasil dicapai, sebagai bentuk transparansi kepada pengguna.
4. Keluar: Menghentikan program.

4.2. Hasil Pengujian

Pengujian dilakukan untuk mengukur performa model dan memvalidasi fungsionalitas sistem.

4.2.1. Hasil Pengujian

Pengujian dilakukan untuk mengukur performa model dan memvalidasi Pengujian dilakukan dengan menggunakan 20 data uji yang telah disisihkan sebelumnya. Hasil prediksi model (y_{pred}) dibandingkan dengan label penyakit sebenarnya (y_{test}). Ringkasan hasil perbandingan disajikan dalam bentuk Confusion Matrix pada Tabel dibawah ini:

			Feline Lower Urinary Tract Disease	Diabetes	Ginjal Kronis
Prediksi	Cacingan	Herpesvirus			
Cacingan	4	0	0	0	0
Feline Herpesvirus	1	3	0	0	0
Feline Lower Urinary Tract Disease	0	0	5	0	0
Diabetes	0	0	1	2	0
Ginjal Kronis	0	0	0	0	4

Berdasarkan Confusion Matrix di atas, metrik performa model dihitung dan disajikan pada Tabel dibawah ini:

Matrix	Nilai
Accuracy	90%
Precision (Rata-rata)	91%
Recall (Rata-rata)	89%
F1-Score (Rata-rata)	89%

Analisis Hasil Pengujian:

1. Akurasi: Model berhasil mencapai akurasi sebesar 90%, yang berarti dari 20 data uji, model mampu memprediksi dengan benar sebanyak 18 kali. Ini menunjukkan bahwa model memiliki kemampuan yang sangat baik secara keseluruhan.
2. Kesalahan Prediksi: Terdapat dua kesalahan prediksi yang terjadi:
 - a. Satu kasus Feline Herpesvirus salah diprediksi sebagai Cacingan. Hal ini tindi antara kedua penyakit tersebut pada data latih. kemungkinan terjadi karena adanya gejala yang sangat umum dan tumpang
 - b. Satu kasus Diabetes salah diprediksi sebagai Feline Lower Urinary Tract Disease. Ini juga mengindikasikan adanya kemiripan gejala yang signifikan, seperti perubahan pola buang air kecil, yang mungkin membingungkan model.
3. Precision dan Recall: Nilai Precision (91%) dan Recall (89%) yang tinggi dan seimbang menunjukkan bahwa model tidak hanya akurat dalam membuat prediksi positif, tetapi juga sensitif dalam mendeteksi sebagian besar kasus penyakit yang ada.

4.2.2. Hasil Pengujian Fungsional Sistem

Pengujian fungsional dilakukan dengan menjalankan prototipe "Valary Care" dan mencoba semua menu yang tersedia.

```
***
=====
HASIL DIAGNOSA
=====

Gejala yang dimasukkan:
- mata berair
- demam ringan
- pilek

Usia: Kitten
Prediksi Penyakit: Feline Herpesvirus
Tingkat Akurasi: 92.76%

REKOMENDASI PERAWATAN
-----

Makanan yang disarankan:
1. makanan dengan vitamin C & lysine
2. makanan basah
```

1. Menu Diagnosa: Sistem berhasil menerima input gejala dari pengguna, memprosesnya, dan memberikan output berupa nama penyakit yang diprediksi. Waktu respons sistem dari input hingga output sangat cepat (kurang dari 1 detik).

```
SENIOR: Kucing senior (7+ tahun) rentan penyakit degeneratif
Penyakit umum:
- Diabetes Melitus
- FLUTD (Penyakit Saluran Kemih)
- Ginjal Kronis
Tips perawatan:
• Kontrol kesehatan rutin
• Makanan rendah protein untuk ginjal
• Sediakan air minum yang cukup
```

2. Menu Informasi: Sistem mampu menampilkan informasi penyakit berdasarkan usia (Kitten, Dewasa, Senior) dengan benar dan sesuai dengan basis pengetahuan yang telah diprogram.

```
=====
TENTANG PROGRAM
=====

Valary Care - Sistem Diagnosa Penyakit Kucing
Versi 1.0 | Akurasi Model: 99.00%

Fitur:
- Diagnosa penyakit berdasarkan gejala
- Rekomendasi makanan dan obat
- Informasi penyakit berdasarkan usia
- Antarmuka interaktif
```

3. Menu Tentang Program: Sistem berhasil menampilkan informasi versi dan nilai akurasi 90%, sesuai dengan hasil perhitungan dari pengujian.

```
MENU UTAMA
1. Lakukan Diagnosa
2. Informasi Penyakit Umum
3. Tentang Program
4. Keluar

Pilih menu (1-4): 4

Terima kasih telah menggunakan Valary Care!
Semoga kucing Anda sehat selalu! 🐱
```

4. Alur Kerja: Seluruh alur kerja aplikasi, mulai dari menu utama hingga keluar dari program, berjalan dengan lancar tanpa ada error atau gangguan.

Secara keseluruhan, hasil pengujian menunjukkan bahwa model *Random Forest* yang dibangun memiliki performa yang handal untuk deteksi penyakit kucing, dan implementasinya dalam prototipe "Valary Care" berhasil menyediakan antarmuka yang fungsional dan informatif bagi pengguna.

5. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah diuraikan pada bab sebelumnya, maka dapat ditarik beberapa kesimpulan sebagai berikut:

1. Penelitian ini telah berhasil mengimplementasikan algoritma machine learning *Random Forest* untuk membangun sebuah model klasifikasi yang mampu mendeteksi lima jenis penyakit umum pada kucing (Cacingan, Feline Herpesvirus, Feline Lower Urinary Tract Disease, Diabetes Melitus, dan Ginjal Kronis) berdasarkan data gejala klinis.
2. Model yang dibangun menunjukkan performa yang sangat baik dalam melakukan prediksi. Berdasarkan hasil pengujian kuantitatif menggunakan 20 data uji, model berhasil mencapai tingkat akurasi sebesar 90%, dengan nilai rata-rata Precision 91%, Recall 89%, dan F1-Score 89%. Tingkat akurasi yang tinggi ini menunjukkan bahwa model dapat menjadi alat bantu yang efektif untuk diagnosis awal penyakit.
3. Sistem deteksi penyakit telah berhasil diwujudkan dalam bentuk prototipe berbasis command-line bernama "Valary Care". Prototipe ini terbukti

fungsional dan mampu menjalankan tugas utamanya, yaitu menerima input gejala dari pengguna, memprosesnya menggunakan model *Random Forest*, dan menyajikan hasil diagnosa yang informatif beserta rekomendasi dan informasi kontekstual berbasis usia kucing.

UCAPAN TERIMA KASIH

Penyusunan jurnal penelitian ini tidak akan terwujud tanpa bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih yang sebesar-besarnya kepada semua pihak yang telah berkontribusi hingga penelitian ini selesai.

DAFTAR PUSTAKA

- [1] F. Hidayat, M. Z. Sarwani, R. Hariyanto, and E. Penulis Korespondensi, "Hidayat, Rekomendasi Pengobatan Pada Penyakit Kucing Menggunakan Metode Decision Tree 155 Rekomendasi Pengobatan Pada Penyakit Kucing Menggunakan Metode Decision Tree (Studi Kasus: Klinik Drh Panti Absari)," 2024.
- [2] P. Palimkar, R. Shaw, and A. Ghosh, "Machine Learning Technique to Prognosis Diabetes Disease: Random Forest Classifier Approach," 2021, pp. 219–244. doi: 10.1007/978-981-16-2164-2_19.
- [3] N. Amalia *et al.*, "Rancangan Aplikasi Sistem Pakar Untuk Mendeteksi Penyakit Kucing Pada E-Petcare," 2022.
- [4] I. Gusti Agung Widiani Wira Brata and M. Iwan Wahyudin, "Pengembangan Model Klasifikasi Penyakit Kucing Menggunakan An K-Nearest Neighbor Untuk Identifikasi Penyakit Menular/Tidak Menular," *Journal of Computer Science and Information Technology*, 2025.
- [5] M. M. Taye, "Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions," May 01, 2023, *MDPI*. doi: 10.3390/computers12050091.
- [6] M. Jamil, F. Rozi, and Y. F. Saputra, "Komparasi Kinerja Algoritma Machine Learning Untuk Deteksi Penyakit Infeksi Saluran Pernapasan," 2024.
- [7] A. Sutaryani, S. Sunarno, and D. Djuniadi, "Perbandingan Performa Model Machine Learning Dalam Prediksi Suhu Di Semarang," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 3, Aug. 2024, doi: 10.23960/jitet.v12i3.4884.
- [8] N. Ranti, M. 1*, and K. H. Hanif, "Klasifikasi Penyakit Kanker Payudara Menggunakan Perbandingan Algoritma Machine Learning," vol. 3, no. 1, pp. 1–6, 2022, [Online]. Available: <http://creativecommons.org/licenses/by/4.0/>
- [9] D. Szlosek, M. Coyne, J. Riggott, K. Knight, D. J. McCrann, and D. Kincaid, "Development and validation of a machine learning model for clinical wellness visit classification in cats and dogs," *Front Vet Sci*, vol. 11, 2024, doi: 10.3389/fvets.2024.1348162.
- [10] R. G. Gonzalez, "Using Random Forest feature importance results to predict zoonosis," May 05, 2023. doi: 10.1101/2023.05.04.23289530.
- [11] Intan Permata and Esther Sorta Mauli Nababan, "Application Of Game Theory In Determining Optimum Marketing Strategy In Marketplace," *Jurnal Riset Rumpun Matematika Dan Ilmu Pengetahuan Alam*, vol. 2, no. 2, pp. 65–71, Jul. 2023, doi: 10.55606/jurrimipa.v2i2.1336.
- [12] B. N. Hadar, Z. Poljak, B. Bonnett, J. Coe, E. A. Stone, and T. M. Bernardo, "Machine learning predicts selected cat diseases using insurance data amid challenges in interpretability," *Am J Vet Res*, vol. 86, no. S1, pp. S52–S62, Mar. 2025, doi: 10.2460/ajvr.24.09.0282.
- [13] C.-C. Wang *et al.*, "Using random forest to predict antimicrobial minimum inhibitory concentrations of nontyphoidal Salmonella in Taiwan," *Vet Res*, vol. 54, no. 1, Feb. 2023, doi: 10.1186/s13567-023-01141-5.
- [14] F. Paramudita and M. I. Zulfa, "Aplikasi Android Pendeteksi Kualitas Beras Berbasis Machine Learning Menggunakan Metode Convolutional Neural Network," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 3, no. 7, pp. 297–305, Aug. 2023, doi: 10.52436/1.jpti.310.
- [15] N. Rohim, "Penerapan Algoritma Deep Learning Untuk Pengembangan Chatbot Yang Digunakan Untuk Konsultasi Dan Pengenalan Tentang Virus COVID-19," 2022.
- [16] M. Arifin, F. Helmi, and D. F. Alamsyah, "Analisis Pola Asosiasi Penjualan Produk Ritel Dengan Platform Google Colab," *Justify: Jurnal Sistem Informasi Ibrahimy*, vol. 3, no. 1, pp. 74–85, 2024, doi: 10.35316/justify.v3i1.5565.