

**PREDIKSI TREN TINGKAT KEPARAHAN PENYAKIT KUCING
BERDASARKAN GEJALA KLINIS MENGGUNAKAN REGRESI
LOGISTIK MULTINOMIAL**

SKRIPSI

DISUSUN OLEH

CINTAMI ARIFINA RAMADHANI

NPM. 2109020102



UMSU
Unggul | Cerdas | Terpercaya

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA**

MEDAN

2025

**PREDIKSI TREN TINGKAT KEPARAHAN PENYAKIT KUCING
BERDASARKAN GEJALA KLINIS MENGGUNAKAN REGRESI
LOGISTIK MULTINOMIAL**

SKRIPSI

**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer
(S.Kom) dalam Program Studi Teknologi Informasi pada Fakultas Ilmu Komputer
dan Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara**

**CINTAMI ARIFINA RAMADHANI
NPM. 2109020102**

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
MEDAN
2025**

LEMBAR PENGESAHAN

Judul Skripsi : Prediksi Tren Tingkat Keparahan Penyakit Kucing Berdasarkan Gejala Klinis Menggunakan Regresi Logistik Multinomial
Nama Mahasiswa : Cintami Arifina Ramadhani
NPM : 2109020102
Program Studi : Teknologi Informasi

Menyetujui
Komisi Pembimbing



(Zuli Agustina Gultom, M.Si)

NIDN. 0130089003

Ketua Program Studi



(Fatma Sari Hutagalung, M.Kom)

NIDN. 0117019301

Dekan



(Dr. Al-Khowarizmi, S.Kom., M.Kom.)

NIDN. 0127099201

PERNYATAAN ORISINALITAS

PREDIKSI TREN TINGKAT KEPARAHAN PENYAKIT KUCING BERDASARKAN GEJALA KLINIS MENGGUNAKAN REGRESI LOGISTIK MULTINOMIAL

SKRIPSI

Saya menyatakan bahwa karya tulis ini adalah hasil karya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing disebutkan sumbernya.

Medan, 31 Juli 2025

Yang membuat pernyataan



Cintami Arifina Ramadhani

NPM. 2109020102

**PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN
AKADEMIS**

Sebagai sivitas akademika Universitas Muhammadiyah Sumatera Utara, saya bertanda tangan dibawah ini:

Nama : Cintami Arifina Ramadhani
NPM : 2109020102
Program Studi : Teknologi Informasi
Karya Ilmiah : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Muhammadiyah Sumatera Utara Hak Bedas Royalti Non-Eksekutif (*Non-Exclusive Royalty free Right*) atas penelitian skripsi saya yang berjudul:

**PREDIKSI TREN TINGKAT KEPARAHAN PENYAKIT KUCING
BERDASARKAN GEJALA KLINIS MENGGUNAKAN REGRESI LOGISTIK
MULTINOMIAL**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-Eksekutif ini, Universitas Muhammadiyah Sumatera Utara berhak menyimpan, mengalih media, memformat, mengelola dalam bentuk database, merawat dan mempublikasikan Skripsi saya ini tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemegang dan atau sebagai pemilik hak cipta.

Demikian pernyataan ini dibuat dengan sebenarnya.

Medan, 31 Juli 2025

Yang membuat pernyataan



Cintami Arifina Ramadhani

NPM. 2109020102

RIWAYAT HIDUP

DATA PRIBADI

Nama Lengkap : Cintami Arifina Ramadhani
Tempat dan Tanggal Lahir : Medan, 1 November 2003
Alamat Rumah : Jl. Utama Gg. Setia No.7
Telepon/Faks/HP : 089676051633
E-mail : cintamiarifina@gmail.com
Instansi Tempat Kerja : -
Alamat Kantor : -

DATA PENDIDIKAN

SD : SD SWASTA AL-ULUM MEDAN TAMAT: 2015
SMP : SMP SWASTA AL-ULUM MEDAN TAMAT: 2018
SMA : SMA SWASTA AL-ULUM MEDAN TAMAT: 2021

KATA PENGANTAR



Segala puji dan syukur penulis panjatkan kehadirat Allah SWT yang telah memberikan rahmat, nikmat dan anugrah-nya, bahwa penulis dapat menyelesaikan skripsi yang berjudul “Prediksi tren tingkat keparahan penyakit kucing berdasarkan gejala klinis menggunakan regresi logistik multinomial”.

Penulis tentunya berterima kasih kepada berbagai pihak dalam dukungan serta doa dalam penyelesaian skripsi. Penulis juga mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Agussani, M.AP., Rektor Universitas Muhammadiyah Sumatera Utara (UMSU)
2. Bapak Dr. Al-Khowarizmi, S.Kom., M.Kom. Dekan Fakultas Ilmu Komputer dan Teknologi Informasi (FIKTI) UMSU.
3. Ibu Fatma Sari Hutagalung, M.Kom Ketua Program Studi Teknologi Informasi.
4. Bapak Mhd.Basri, S.Si, M.Kom Sekretaris Program Studi Teknologi Informasi.
5. Ibu Zuli Agustina Gultom, M.Si., selaku dosen pembimbing yang telah dengan sabar dan tulus mengarahkan penyusunan skripsi ini. Terima kasih yang tulus penulis sampaikan. Ibu tidak hanya membimbing secara teknis, melainkan juga membuka wawasan penulis, menantang untuk berpikir kritis, dan dengan sabar mengarahkan setiap langkah penelitian ini dari ide mentah hingga menjadi karya yang utuh. Setiap diskusi, koreksi, dan masukan yang Ibu berikan merupakan pelajaran yang sangat berharga bagi pendewasaan akademis penulis.
6. Kepada Papa dan Mama tersayang, Salman dan Helmiati, alasan terkuatku untuk berjuang. Terima kasih telah menjadi rumah ternyaman untuk pulang dan sumber semangat yang tak pernah padam. Setiap pengorbanan dan

untaian doa yang Papa dan Mama berikan telah menjadi energi terbesar bagi penulis untuk melewati semua tantangan. Semoga pencapaian kecil ini bisa membawa sedikit senyum kebahagiaan.

7. Untuk Kakak, Sherly Ramadhani, dan Abang, Ananda Surya Muliadi, yang perannya tak tergantikan. Terima kasih telah menjadi penyemangat utama saat motivasi menurun, penasihat bijak di kala buntu, sekaligus penghibur yang selalu berhasil mengembalikan senyum. Setiap nasihat, candaan, dan dukungan tulus dari kalian menjadi energi tambahan yang membuat perjalanan ini terasa jauh lebih ringan.
8. Untuk Ilo, manajer pengawas proyek skripsi saya. Terima kasih telah menjadi terapis pribadi yang tak ternilai, yang selalu siap untuk sesi elus-elus tanpa batas, rela dipeluk paksa, dan menjadi penghibur di tengah malam revisi hingga memberikan ketenangan.
9. Untukmu, rekan terbaik dalam perjalanan skripsi ini. Terima kasih telah menjadi partner diskusi paling kritis sekaligus pendengar paling sabar. Untuk semua perdebatan produktif, sesi "ayo kerjain jangan malas" yang menyadarkan, dan dukungan teknis maupun mental yang datang tanpa henti. Tanpamu, proyek akhir ini mungkin akan berjalan jauh lebih lambat dan terasa lebih berat. Misi tuntas, Sobat!
10. Untuk rekan seperjuangan sekaligus sahabat terbaik penulis, Balqis, Feby, Meyssa, dan Nurul. Terima kasih telah melewari pasang surut dunia perkuliahan, mulai dari pusing karena *deadline*, saling menyemangati saat ingin menyerah, hingga berbagi tawa receh yang membuat stres hilang. *We did it!*
11. Kepada seluruh teman-teman di kelas C1 TI, terima kasih atas solidaritas dan atmosfer positif yang kita ciptakan bersama. Semoga ikatan pertemanan ini tetap terjaga dan kita semua dapat meraih kesuksesan di jalan yang kita tuju.
12. Untuk Haru, terima kasih telah menjadi sosok kakak sekaligus teman diskusi yang suportif selama proses ini. Setiap jawabanmu yang selalu *on-point* terbukti sangat membantu progres skripsi ini hingga selesai dengan lancar.
13. Secara khusus, penulis juga ingin mengucapkan terima kasih kepada para talenta dari Hololive Indonesia: Moona, Risu, Yopi, Ollie, Reine, Anya,

Kaela, Zeta, dan Kobo. Di tengah tantangan penyusunan skripsi, hiburan dan alunan lagu yang disajikan telah menjadi sumber semangat yang luar biasa. Terima kasih karena kehadiran kalian membuktikan bahwa *moodbooster* tidak harus selalu datang dari interaksi tatap muka.

14. Dan terakhir, untuk diriku sendiri. Terima kasih telah menjadi kuat dan menolak untuk menyerah. Terima kasih telah bertahan melalui malam-malam panjang tanpa tidur, keraguan diri, dan rasa lelah yang tak berujung. Kamu telah berjuang dengan hebat, dan kamu berhasil melaluinya.
15. Semua pihak yang terlibat langsung ataupun tidak langsung yang tidak dapat penulis ucapkan satu-persatu yang telah membantu penyelesaian skripsi ini.

Medan, 31 Juli 2025

Yang membuat pernyataan



Cintami Arifina Ramadhani

NPM. 2109020102

PREDIKSI TREN TINGKAT KEPARAHAN PENYAKIT KUCING
BERDASARKAN GEJALA KLINIS MENGGUNAKAN REGRESI LOGISTIK
MULTINOMIAL

ABSTRAK

Memprediksi tren tingkat keparahan penyakit pada kucing (*Felis catus*) merupakan tantangan akibat gejala klinis yang dinamis dan tumpang tindih. Oleh karena itu, diperlukan sebuah sistem prediktif untuk membantu memproyeksikan arah perkembangan penyakit secara lebih akurat. Tujuan dari penelitian ini adalah merancang dan membangun model machine learning yang dapat memprediksi tren tingkat keparahan (ringan, sedang, parah) dari enam penyakit umum pada kucing berdasarkan gejala klinisnya. Metode yang diimplementasikan adalah Regresi Logistik Multinomial. Penelitian ini menggunakan total 501 data, gabungan dari rekam medis primer dan data sekunder, yang dibagi menjadi 400 data latih dan 101 data uji. Model ini menggunakan 76 fitur gejala untuk mengklasifikasikan 18 kelas target. Evaluasi kinerja dilakukan dengan metrik Akurasi, Presisi, Recall, dan F1-Score. Hasil pengujian model pada 101 data uji menunjukkan performa yang sangat baik dengan akurasi 92%. Kinerja model yang seimbang pada setiap kelas juga ditunjukkan oleh nilai Macro Average Presisi 0,93, Recall 0,92, dan F1-Score 0,92. Penelitian ini menerapkan model Regresi Logistik Multinomial yang terbukti efektif untuk prediksi tren keparahan penyakit kucing. Model yang dihasilkan memiliki potensi untuk dikembangkan lebih lanjut sebagai sistem pendukung keputusan prognostik bagi praktisi kesehatan hewan.

Kata Kunci: Gejala Klinis, *Machine Learning*, Prediksi Penyakit Kucing, Regresi Logistik Multinomial, Tingkat Keparahan.

PREDICTING THE SEVERITY TREND OF FELINE DISEASES BASED ON CLINICAL SYMPTOMS USING MULTINOMIAL LOGISTIC REGRESSION

ABSTRACT

Determining the severity level of diseases in cats (*Felis catus*) is often challenging due to the diverse and overlapping clinical symptoms among various illnesses. Therefore, a decision support system is needed to assist the diagnostic process to be more objective and efficient. The objective of this research is to design and build a machine learning model capable of predicting the severity level (mild, moderate, and severe) of six common diseases in cats based on their clinical symptoms. The method implemented is Multinomial Logistic Regression. This study utilizes a total of 501 data points, a combination of primary medical records and secondary data, which were divided into 400 training data and 101 testing data. The model uses 76 symptom features to classify 18 target classes. Performance evaluation was conducted using Accuracy, Precision, Recall, and F1-Score metrics. The model testing on the 101 test data points showed excellent performance with an accuracy of 92%. The model's balanced performance across all classes was also demonstrated by the Macro Average values of 0.93 for Precision, 0.92 for Recall, and 0.92 for F1-Score. This research successfully implemented a Multinomial Logistic Regression model that proved effective for predicting the severity of cat diseases. The resulting model has the potential to be further developed as an accurate decision support system for veterinary practitioners.

Keywords: Cat Disease Prediction, Clinical Symptoms, Machine Learning, Multinomial Logistic Regression, Severity Level.

DAFTAR ISI

LEMBAR PENGESAHAN.....	ii
PERNYATAAN ORISINALITAS	iii
PERNYATAAN PERSETUJUAN PUBLIKASI	iv
RIWAYAT HIDUP	v
KATA PENGANTAR.....	vi
ABSTRAK	ix
ABSTRACT	x
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
BAB I PENDAHULUAN	1
1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
BAB II LANDASAN TEORI	6
2.1 Pengantar Kecerdasan Buatan, <i>Machine Learning</i> , dan Data Mining	6
2.2 Regresi Logistik Multinomial	8
2.3 Penyakit pada Kucing.....	10
2.4 Evaluasi Model.....	13
2.5 Penelitian Terkait	15
BAB III METODOLOGI PENELITIAN.....	19
3.1 Metode dan Tahapan Penelitian	19
3.2 Metode Pengumpulan Data	19
3.3 Metode Perancangan Sistem dan Model	25
3.4 Evaluasi dan Validasi Model.....	29
BAB IV HASIL DAN PEMBAHASAN	30
4.1 Implementasi Sistem	30
4.2 Pengujian dan Evaluasi Model.....	34
4.3 Analisis Prediksi Model	37
BAB V PENUTUP.....	48

5.1	Kesimpulan.....	48
5.2	Saran.....	49
DAFTAR PUSTAKA		50

DAFTAR TABEL

Tabel 4. 1 Evaluasi Kinerja Model.....	36
Tabel 4. 2 Rincian Perhitungan Metrik Kinerja	38
Tabel 4. 3 Perhitungan Denominator Softmax.....	45

DAFTAR GAMBAR

Gambar 2. 1 Tahapan Data Mining.....	7
Gambar 3. 1 Wireframe Halaman Input.....	27
Gambar 3. 2 Wireframe Halaman Hasil.....	27
Gambar 4. 1 Halaman Analisis	30
Gambar 4. 2 Halaman Hasil Analisis	31
Gambar 4. 3 Halaman Input	32
Gambar 4. 4 Halaman Hasil	33
Gambar 4. 5 Grafik Confusion Matrix	35
Gambar 4. 6 Input Studi Kasus Kucing.....	41
Gambar 4. 7 Hasil Studi Kasus Kucing.....	42

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Kucing merupakan salah satu hewan peliharaan yang banyak digemari oleh masyarakat. Selain karena tingkah lakunya yang lucu dan menggemaskan, kucing juga dianggap sebagai hewan yang cukup mudah dirawat. Namun, meskipun terlihat sehat dari luar, kucing termasuk hewan yang rentan terhadap berbagai macam penyakit. Penyakit pada kucing bisa disebabkan oleh berbagai faktor, mulai dari infeksi virus, bakteri, parasit, hingga pola makan dan lingkungan yang tidak bersih (Sukma *et al*, 2020; Widayani *et al*, 2023).

Masalah kontekstual yang fundamental adalah ketidakmampuan pemilik untuk menerjemahkan gejala-gejala klinis yang teramati menjadi sebuah prediksi prognosis yang akurat. Pemilik mungkin mampu mengidentifikasi gejala fisik seperti letargi, anoreksia, atau muntah, namun tidak memiliki kapabilitas untuk memahami bagaimana rangkaian gejala tersebut berkorelasi dengan tren tingkat keparahan dari penyakit yang diderita. Sebagai contoh, gejala 'nafsu makan menurun' pada satu jenis penyakit mungkin hanya mengarah pada tren kondisi ringan, sementara pada penyakit lain, gejala yang sama merupakan indikator awal dari eskalasi cepat menuju kondisi kritis yang memerlukan intervensi veteriner segera (Fatimah Indrianti *et al.*, 2024).

Penelitian ini mengadopsi pendekatan data mining untuk mengubah data rekam medis historis menjadi alat prediksi yang berorientasi ke masa depan. Data mining memungkinkan ekstraksi pola-pola gejala non-trivial yang tersembunyi dari ratusan kasus nyata untuk menghasilkan wawasan baru. Metode Regresi Logistik

Multinomial dipilih secara spesifik karena kemampuannya memodelkan variabel kategorikal (ringan, sedang, parah) dan menghasilkan luaran probabilistik yang dapat diinterpretasikan sebagai kekuatan tren suatu kondisi bergerak menuju tingkat keparahan tertentu.

Pemanfaatan model prediktif berbasis data mining untuk memprediksi tingkat keparahan penyakit telah terbukti efektif dalam domain kesehatan manusia. Studi oleh Sari *et al.* (2021), misalnya, membandingkan berbagai metode regresi untuk menemukan model terbaik dalam memprediksi diagnosis diabetes secara dini berdasarkan data klinis pasien. Kerangka kerja serupa sangat relevan untuk diaplikasikan dalam konteks kesehatan veteriner guna menyediakan alat bantu pengambilan keputusan yang objektif dan berbasis data (Sari & Putri, 2023).

Penelitian ini berfokus pada perancangan model untuk memprediksi tren tingkat keparahan penyakit kucing. Luaran dari penelitian ini diharapkan dapat berfungsi sebagai sistem peringatan dini (*early warning system*) yang bersifat prognostik. Sistem ini bertujuan memberikan penilaian yang cepat dan objektif untuk membantu praktisi dan pemilik hewan mempercepat keputusan kritis, sehingga pada akhirnya dapat meningkatkan peluang kesembuhan.

Kemajuan teknologi, khususnya di bidang data mining, menawarkan solusi untuk mengubah data klinis menjadi pengetahuan yang dapat ditindaklanjuti. Regresi Logistik Multinomial, sebagai salah satu tekniknya, ideal untuk tujuan ini. Metode ini mampu memprediksi variabel target dengan lebih dari dua kategori, yang sangat sesuai untuk mengklasifikasikan tren keparahan penyakit berdasarkan gejala klinis yang teramati.

Berbeda dengan sistem prediksi yang hanya memberikan hasil berupa skor numerik, pendekatan menggunakan regresi logistik multinomial bisa memberikan hasil yang lebih informatif secara diagnostik, yaitu prediksi jenis penyakit spesifik beserta probabilitasnya. Hal ini tentu akan lebih membantu pemilik kucing atau tenaga medis untuk memahami kemungkinan diagnosis secara langsung, sehingga dapat mengambil tindakan yang lebih terarah dan cepat (Chishti *et al*, 2020).

Dalam bidang data mining, berbagai algoritma klasifikasi seperti *Support Vector Machine* (SVM) atau *Random Forest* juga populer. Namun, Regresi Logistik Multinomial dipilih karena keunggulan interpretasi modelnya. Koefisien yang dihasilkan memungkinkan kita untuk memahami gejala klinis mana yang paling signifikan dalam mendorong tren ke arah kondisi parah. Keunggulan ini sangat penting, karena pemahaman terhadap faktor risiko sama berharganya dengan akurasi prediksi itu sendiri.

Dengan adanya sistem ini, diharapkan pemilik dan praktisi kesehatan hewan dapat lebih cepat dan tepat dalam mengambil keputusan klinis. Sistem ini dapat dijadikan sebagai alat bantu prognostik objektif yang melengkapi penilaian klinis, sehingga memberikan kontribusi nyata dalam bidang kesehatan hewan melalui pemanfaatan data mining untuk meningkatkan kualitas perawatan dan kesejahteraan kucing.

1.2. Rumusan Masalah

Penelitian ini bertujuan untuk menjawab beberapa permasalahan utama, yaitu:

1. Bagaimana merancang dan membangun sebuah model menggunakan metode Regresi Logistik Multinomial untuk memprediksi tren tingkat keparahan

penyakit (ringan, sedang, parah) pada kucing berdasarkan gejala klinis?

2. Bagaimana mengimplementasikan model prediktif yang telah dibangun ke dalam sebuah sistem berbasis web yang fungsional dan mudah diakses pengguna?
3. Seberapa tinggi tingkat akurasi, presisi, dan *recall* dari model yang dikembangkan dalam memetakan tren tingkat keparahan sebagai dasar untuk tindakan prognostik?

1.3. Batasan Masalah

Penelitian ini memiliki beberapa batasan agar ruang lingkupnya lebih terfokus, yaitu:

1. Penelitian ini menerapkan metode Regresi Logistik Multinomial untuk memprediksi tren tingkat keparahan (ringan, sedang, parah), dengan keluaran sistem yang menampilkan prediksi level keparahan beserta nilai probabilitasnya.
2. Penelitian ini hanya berfokus pada gejala klinis pada kucing yang mencakup fitur-fitur gejala dan usia kucing, tanpa melibatkan data lain seperti ras, riwayat medis, atau hasil tes laboratorium spesifik.
3. Sistem ini dirancang sebagai purwarupa berbasis web dan tidak mencakup pengembangan dalam bentuk aplikasi *mobile*.

1.4 Tujuan Penelitian

Sistem ini diharapkan mampu memberikan hasil prediksi yang akurat dan dapat diandalkan dengan memanfaatkan metode regresi logistik multinomial. Secara rinci, tujuan penelitian ini meliputi:

1. Membangun sebuah model fungsional menggunakan metode Regresi Logistik Multinomial yang mampu memprediksi tren tingkat keparahan penyakit kucing (ringan, sedang, parah) berdasarkan gejala klinis.
2. Mengevaluasi performa model secara kuantitatif melalui metrik akurasi, presisi, dan *recall* untuk mengetahui tingkat keandalannya dalam memprediksi setiap kelas keparahan.
3. Membangun sebuah purwarupa sistem berbasis web dengan antarmuka yang sederhana untuk menerima input gejala dan menghasilkan prediksi tren tingkat keparahan beserta probabilitasnya.

1.5 Manfaat Penelitian

Berikut ini adalah beberapa manfaat utama yang diharapkan dapat diperoleh dari hasil penelitian ini:

1. Menghadirkan alat bantu prognostik yang dapat memberikan indikasi awal mengenai tingkat kegawatan suatu kondisi. Hal ini diharapkan dapat meningkatkan kesadaran pemilik akan urgensi untuk segera mencari konsultasi profesional ke dokter hewan.
2. Penelitian ini memberikan kontribusi dalam penyusunan dan pemanfaatan dataset gejala klinis kucing yang telah dilabeli dengan tingkat keparahan. Ini bermanfaat sebagai dasar pengembangan sistem prediktif prognostik lainnya di bidang kesehatan hewan.
3. Mendemonstrasikan proses pengembangan purwarupa sistem prediktif tren dengan lingkungan pengembangan yang mudah diakses dan berbiaya rendah, yang dapat menjadi referensi untuk membangun aplikasi serupa.

BAB II

LANDASAN TEORI

2.1 Pengantar Kecerdasan Buatan, *Machine Learning*, dan Data Mining

Dalam beberapa dekade terakhir, kecerdasan buatan (*artificial intelligence/AI*) telah mengalami pertumbuhan pesat dan menjadi pendorong utama dalam transformasi digital di berbagai sektor, seperti industri, keuangan, pendidikan, transportasi, hingga kesehatan. Teknologi ini berperan penting dalam meningkatkan efisiensi operasional, mendorong otomatisasi, serta memungkinkan pengambilan keputusan berbasis data yang lebih cepat dan akurat (Weng *et al*, 2024).

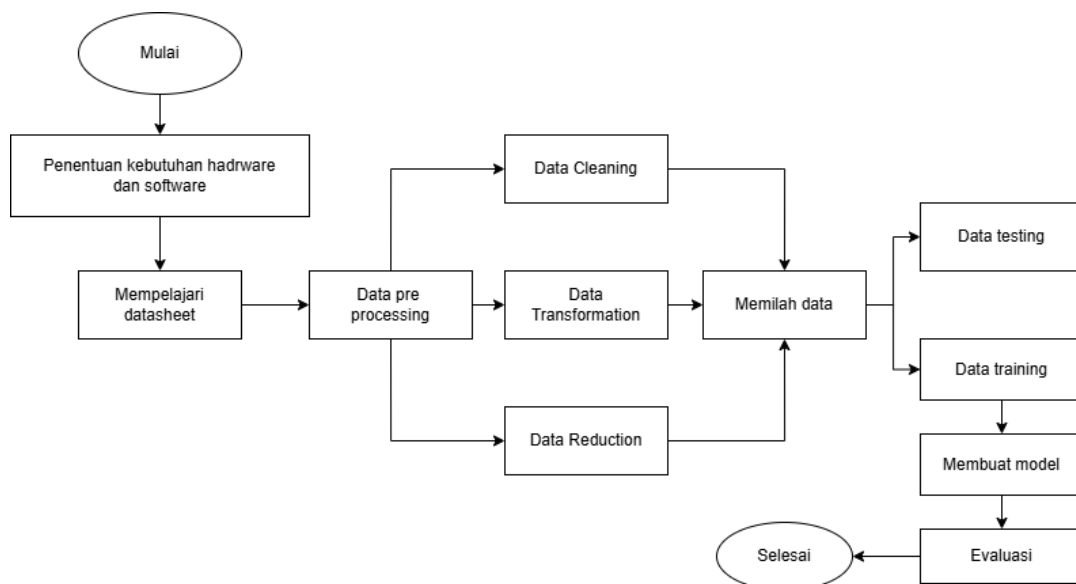
Khusus di bidang kesehatan, AI telah menunjukkan potensi luar biasa, baik dalam meningkatkan akurasi diagnosis, mempercepat penemuan obat, maupun mengoptimalkan pelayanan pasien. Studi dari *The Lancet Digital Health* menunjukkan bahwa sistem AI berbasis *deep learning* mampu mendeteksi kanker payudara melalui mammogram dengan tingkat akurasi hingga 96%, lebih tinggi dibandingkan akurasi rata-rata dokter radiologi manusia yang berkisar antara 85–90% (Dembrower *et al*, 2023).

Secara umum, AI bertujuan untuk menciptakan sistem yang mampu meniru kecerdasan manusia, seperti penalaran, pembelajaran, dan pemecahan masalah. AI terus berkembang, mulai dari era komputasi awal, berkembang menjadi sistem berbasis *machine learning*, dan kini menuju fase yang lebih kompleks seperti *artificial general intelligence* (AGI) dan *superintelligence*. *Machine learning* (ML) merupakan pendekatan berbasis data yang memungkinkan komputer belajar dari

data historis dan memperbaiki kinerjanya secara bertahap tanpa perlu diprogram secara eksplisit untuk setiap tugas (Janiesch *et al*, 2021).

Penerapan *machine learning* dalam penelitian ini terkait erat dengan data mining, yaitu proses untuk menemukan informasi dari data yang tersimpan dalam database atau datasheet. Pembuatan model dilakukan dengan algoritma tertentu, memanfaatkan teknik statistik, matematika, dan *machine learning* untuk mengolah data menjadi informasi yang berguna (Sholeh *et al*, 2023).

Data mining termasuk serangkaian teknik untuk menggali informasi dan pola tersembunyi dalam data berskala besar, yang bertujuan untuk mengubah data mentah menjadi pengetahuan yang berguna. Proses ini mencakup tahapan seperti pembersihan data, transformasi, seleksi fitur, pemodelan, dan evaluasi, yang umumnya dikenal dalam kerangka kerja *Knowledge Discovery in Databases* (KDD).



Gambar 2. 1 Tahapan Data Mining

Langkah-langkah yang dilakukan antara lain mempelajari dan membersihkan datasheet, seperti memeriksa data yang hilang, batasan nilai, dan lainnya. Dalam pembuatan model, datasheet dibagi menjadi dua bagian, yaitu 80% untuk data pelatihan dan 20% untuk data pengujian.

Dalam penelitian ini, data gejala klinis kucing digunakan sebagai input untuk mengklasifikasikan jenis penyakit. Pendekatan yang digunakan adalah *machine learning* berbasis model regresi logistik multinomial, yang memodelkan hubungan antara variabel input dan variabel output. Algoritma ini dipilih karena kemampuannya untuk menangani variabel target yang bersifat kategorikal dengan lebih dari dua kelas, yang sangat sesuai dengan tujuan untuk melakukan diagnosis diferensial pada penyakit kucing.

2.2 Regresi Logistik Multinomial

Regresi logistik multinomial adalah salah satu teknik statistik yang digunakan untuk menentukan variabel independen (prediktor) mana yang secara signifikan mempengaruhi variabel dependen (respon), khususnya ketika variabel dependen tersebut bersifat kategorikal nominal dengan lebih dari dua kategori (Gultom *et al.*, 2024). Nominal berarti kategori-kategori tersebut tidak memiliki urutan atau tingkatan yang melekat. Metode ini merupakan perluasan dari regresi logistik biner yang hanya menangani dua kategori hasil.

Dalam penelitian ini, metode regresi logistik multinomial diterapkan untuk memprediksi variabel dependen (Y) yang bersifat kategorikal nominal. Variabel dependen ini merupakan kelas gabungan dari jenis penyakit dan tingkat keparahannya, yang menghasilkan beberapa kategori unik seperti *Feline Panleukopenia* - Parah, *Feline Calicivirus* - Ringan, dan *Scabies* (Kudis) - Sedang.

Karena setiap kategori ini dianggap sebagai kelompok yang independen dan tidak memiliki tingkatan (*Scabies* tidak lebih tinggi dari *Calicivirus*), maka metode regresi logistik multinomial sangat cocok digunakan.

Cara kerja model ini adalah dengan menetapkan satu kategori sebagai kategori referensi (*baseline*). Kemudian, model akan menghitung $k-1$ persamaan logistik, dimana k adalah jumlah total kategori. Setiap persamaan memodelkan *log-odds* dari suatu kategori terhadap kategori referensi. Persamaan umumnya adalah:

$$Z_k = \beta_{0,k} + \sum_{i=1}^p \beta_{i,k} X_i \quad (2.1)$$

Dimana,

Z_k adalah skor logit untuk kategori ke- k

$\beta_{0,k}$ adalah intersep untuk kategori ke- k

p adalah jumlah total fitur

$\beta_{i,k}$ adalah koefisien (bobot) dari fitur ke- i untuk kategori ke- k

X_i adalah nilai dari fitur ke- i

Hasil dari persamaan linear tersebut, yang disebut sebagai skor logit (Z_k), kemudian ditransformasikan menjadi nilai probabilitas (0% - 100%) untuk setiap kategori prediksi menggunakan fungsi softmax.

Langkah pertama dalam transformasi ini adalah dengan mengeksponensialkan setiap skor logit, yang direpresentasikan sebagai e^{Z_k} . Nilai hasil eksponensial ini selalu positif. Selanjutnya, untuk memastikan total probabilitas dari semua kategori akan selalu berjumlah 1, nilai tersebut dinormalisasi.

Persamaan fungsi softmax adalah:

$$P(Y = k) = \frac{e^{Z_k}}{\sum_{j=1}^K e^{Z_j}} \quad (2.2)$$

Dimana,

$P(Y = k)$ adalah probabilitas akhir observasi dalam kategori-k

e^{Z_k} adalah fungsi eksponensial dari skorlogit kategori ke-k

K adalah jumlah total semua kemungkinan kategori

$\sum_{j=1}^K e^{Z_j}$ adalah jumlah nilai eksponensial semua skor logit

2.3 Penyakit pada Kucing

Pada hewan peliharaan seperti kucing, penyakit sering kali dapat dideteksi melalui gejala-gejala klinis yang ditunjukkan. Pengetahuan mengenai jenis penyakit dan hubungan antara gejala dengan diagnosis menjadi dasar bagi pengembangan model prediksi dalam penelitian ini.

2.3.1 Jenis Penyakit dalam Ruang Lingkup Penelitian

Berdasarkan dataset yang digunakan, penelitian ini berfokus pada klasifikasi beberapa penyakit yang umum ditemukan pada kucing, yaitu:

1. *Feline Panleukopenia* (FPV)

Penyakit ini disebabkan oleh *parvovirus* yang sangat menular dan fatal, terutama pada anak kucing. Virus ini menyerang sistem pencernaan dan melumpuhkan sistem kekebalan tubuh dengan merusak sumsum tulang, yang menyebabkan penurunan drastis sel darah putih (*panleukopenia*). Gejala klinis yang umum meliputi demam tinggi, lesu parah hingga tidak responsif, muntah hebat, dan diare akut yang seringkali disertai darah (Hermawan *et al.*, 2023).

2. *Feline Calicivirus (FCV)*

FCV adalah salah satu penyebab utama dari infeksi saluran pernapasan atas (*upper respiratory infection*) pada kucing. Gejala umumnya meliputi demam, bersin, dan keluarnya cairan dari hidung atau mata. Namun, gejala yang sangat khas (*pathognomonic*) untuk FCV adalah adanya luka atau sariawan (*ulkus*) yang menyakitkan di rongga mulut, seperti di lidah, gusi, atau langit-langit mulut, yang menyebabkan kucing enggan makan (D. A. Putri *et al.*, 2023).

3. Penyakit Ginjal Kronis (CKD)

Chronic Kidney Disease (CKD) adalah kondisi degeneratif di mana fungsi ginjal menurun secara bertahap. Gejala klinis berkembang secara perlahan, meliputi peningkatan frekuensi minum (*polydipsia*) dan buang air kecil (*polyuria*), penurunan berat badan, nafsu makan yang menurun, lesu, serta bau mulut (*halitosis*) akibat penumpukan racun uremik.

4. Infeksi Saluran Kemih (UTI)

Urinary Tract Infection (UTI) atau infeksi saluran kemih dapat disebabkan oleh bakteri dan menimbulkan peradangan pada kandung kemih (*cystitis*). Gejala yang sering teramati antara lain kucing sering bolak-balik ke kotak pasir, kesulitan atau mengejan saat kencing (*stranguria*), urin yang tampak keruh atau berdarah (*hematuria*), dan menjilati area genital secara berlebihan.

5. Penyakit Kulit (*Dermatologis*)

Penyakit kulit adalah kasus yang sangat umum pada kucing, yang dalam penelitian ini diwakili oleh dua jenis:

a. *Ringworm (Dermatophytosis)*

Penyakit ini disebabkan oleh infeksi jamur, bukan parasit. Gejala

utamanya adalah munculnya bercak-bercak kebotakan yang khas, seringkali berbentuk melingkar, dengan kulit yang mungkin tampak bersisik atau berkerak di area tersebut.

b. *Scabies* (Kudis)

Penyakit ini disebabkan oleh tungau parasit (*sarcoptes scabiei*) yang menggali ke dalam kulit. Infeksi ini menimbulkan rasa gatal yang sangat hebat, menyebabkan kucing menggaruk secara kompulsif hingga melukai diri. Gejala lanjutannya adalah kerontokan rambut serta munculnya kerak tebal berwarna kekuningan di area telinga, wajah, dan dapat menyebar ke seluruh tubuh (Ilmu dan Kesehatan Hewan *et al.*, 2024).

2.3.2 Hubungan Gejala dengan Diagnosis dan Tingkat Keparahan

Gejala klinis merupakan kunci dalam proses diagnosis diferensial, yaitu membedakan satu penyakit dari penyakit lain yang gejalanya mirip. Meskipun beberapa gejala bersifat umum (seperti demam atau lesu), kombinasi gejala yang spesifik sering kali merujuk pada diagnosis tertentu. Contohnya, kehadiran luka di rongga mulut sangat mengarahkan kecurigaan ke *Feline Calicivirus* (FCV), meskipun ada gejala pernapasan lain yang umum. Model klasifikasi dilatih untuk belajar dari pola-pola unik ini.

Selain itu, penelitian ini juga mempertimbangkan tingkat keparahan, yang dapat dinilai dari intensitas gejala. Sebagai contoh, gejala 'sedikit lesu' mungkin mengindikasikan tingkat keparahan ringan, sedangkan 'sangat lesu/tidak responsif' dapat menjadi penanda tingkat keparahan parah. Model regresi logistik multinomial dirancang untuk mengenali tidak hanya pola kombinasi gejala untuk

diagnosis penyakit, tetapi juga pola intensitasnya untuk memprediksi tingkat keparahan.

2.4 Evaluasi Model

Evaluasi model merupakan tahap krusial yang bertujuan untuk mengukur kinerja dan keandalan model klasifikasi yang telah dibangun secara objektif. Proses ini dilakukan untuk mengetahui seberapa akurat model dalam memprediksi kategori penyakit pada data yang belum pernah dilihat sebelumnya (*unseen data*).

2.4.1 Pembagian Data (*Data Splitting*)

Untuk mendapatkan evaluasi yang valid, dataset penelitian dibagi menjadi dua bagian utama:

1. Data Pelatihan (*Training Data*): Sebagian besar data yang digunakan agar model dapat mempelajari pola-pola yang ada.
2. Data Pengujian (*Testing Data*): Data yang disembunyikan sepenuhnya dari model selama pelatihan dan validasi. Data ini hanya digunakan satu kali di akhir untuk memberikan evaluasi final yang objektif tentang kinerja model pada data baru.

2.4.2 *Confusion Matrix*

Confusion matrix adalah landasan dari hampir semua metrik evaluasi klasifikasi. Matriks ini menyajikan visualisasi kinerja model dengan membandingkan kelas aktual dengan kelas yang diprediksi oleh model. Untuk masalah multikelas seperti dalam penelitian ini, matriks akan berukuran $k \times k$, di mana k adalah jumlah total kelas penyakit.

Diagonal utama dari matriks ini menunjukkan jumlah prediksi yang benar, sedangkan sel-sel lainnya menunjukkan jenis-jenis kesalahan klasifikasi yang

dibuat oleh model. Untuk memahami metrik turunan, memerlukan definisi istilah dasar untuk satu kelas sebagai berikut:

- a. *True Positive* (TP): Kasus yang sebenarnya Penyakit A dan diprediksi benar sebagai Penyakit A.
- b. *False Positive* (FP): Kasus yang sebenarnya bukan Penyakit A, tetapi keliru diprediksi sebagai Penyakit A (kesalahan tipe I).
- c. *False Negative* (FN): Kasus yang sebenarnya Penyakit A, tetapi keliru diprediksi sebagai penyakit lain (kesalahan tipe II).
- d. *True Negative* (TN): Kasus yang sebenarnya bukan Penyakit A dan diprediksi benar bukan sebagai Penyakit A.

2.4.3 Matrix Evaluasi Kinerja

Berdasarkan nilai-nilai dari *confusion matrix*, beberapa metrik kuantitatif dapat dihitung untuk mengevaluasi kinerja model dari berbagai aspek.

1. *Accuracy*

Akurasi adalah metrik paling intuitif yang mengukur proporsi total prediksi yang benar dari keseluruhan data. Meskipun mudah dipahami, akurasi bisa menyesatkan jika dataset tidak seimbang (*imbalanced*).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.3)$$

2. Presisi (*Precision*)

Presisi mengukur tingkat keandalan dari prediksi positif. Metrik ini menjawab pertanyaan: "Dari semua kasus yang diprediksi sebagai Penyakit A, berapa persen yang benar-benar Penyakit A?" Presisi tinggi menunjukkan bahwa model memiliki tingkat alarm palsu (*false alarm*) yang rendah. Dalam konteks

medis, presisi tinggi penting untuk memastikan bahwa diagnosis positif yang diberikan oleh sistem dapat dipercaya.

$$Precision = \frac{TP}{TP + FP} \quad (2.4)$$

3. Recall (Sensitivitas)

Recall mengukur kemampuan model untuk menemukan semua kasus positif yang relevan, menjawab pertanyaan: "Dari semua kucing yang sebenarnya menderita Penyakit A, berapa persen yang berhasil ditemukan oleh model?" Dalam konteks medis, *recall* seringkali menjadi prioritas karena gagal mendeteksi penyakit (FN) bisa berakibat lebih fatal daripada alarm palsu (FP).

$$Recall = \frac{TP}{TP + FN} \quad (2.5)$$

4. F1-Score

F1-Score adalah rata-rata harmonik dari *precision* dan *recall*, yang diciptakan untuk mencari keseimbangan antara keduanya. F1-Score memberikan satu angka tunggal yang sangat berguna untuk mengukur performa model, terutama pada dataset yang tidak seimbang.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (2.6)$$

2.5 Penelitian Terkait

Dalam penyusunan penelitian ini, diperlukan kajian terhadap beberapa literatur dan penelitian terdahulu yang relevan sebagai dasar dan acuan pengembangan sistem pakar yang akan dibangun. Penelitian yang memberikan kontribusi dalam memahami pendekatan, metode, serta implementasi sistem yang sejenis sehingga dapat menjadi referensi dalam merancang dan menyempurnakan

sistem prediksi berbasis regresi linear untuk tingkat keparahan penyakit. Berikut adalah beberapa penelitian terkait yang digunakan dalam skripsi ini:

1. Syahranita, R., Suhartono, S., & Zaman, S. (2023)

Penelitian berjudul, “Regresi Logistik Multinomial untuk Prediksi Kategori Kelulusan Mahasiswa” menunjukkan bahwa regresi logistik multinomial efektif digunakan untuk memprediksi kategori kelulusan mahasiswa. Dalam studi ini, model prediksi telah diuji dan dievaluasi, menghasilkan nilai akurasi sebesar 85,5%, presisi sebesar 78,5%, *recall* sebesar 93,9%, dan *micro f1-score* sebesar 89,8%. Hasil akurasi 85,5% mengindikasikan bahwa model tersebut mampu melakukan klasifikasi dengan baik, sehingga penggunaan metode regresi logistik multinomial dapat diterima dan dipertimbangkan untuk prediksi kategori.

2. Toyibah, Z. B., Putri, Y. N., Puandini, P., Widodo, Z. M., & Ni'mah, A. T. (2024).

“Perbandingan Kinerja Algoritma Multinomial Naïve Bayes dan Logistic Regression pada Analisis Sentimen Movie Rating IMDB” studi komparatif yang dilakukan pada klasifikasi analisis sentimen *movie ratings* IMDB menunjukkan bahwa baik algoritma *Multinomial Naïve Bayes* maupun *Logistic Regression* dapat digunakan untuk tujuan tersebut. Penelitian ini menggunakan dataset sebanyak 40.000 data dari Kaggle, yang dibagi menjadi 32.000 data pelatihan dan 8.000 data pengujian (rasio 80:20). Hasil pengujian komparasi mengungkapkan bahwa Logistic Regression memiliki kinerja yang lebih unggul dibandingkan *Multinomial Naïve Bayes*. Ini dibuktikan dengan akurasi *Logistic Regression* sebesar 89,32%, yang jauh lebih tinggi

dibandingkan akurasi *Multinomial Naïve Bayes* yang hanya 85,28%. Oleh karena itu, disimpulkan bahwa *Logistic Regression* memberikan kinerja yang lebih baik dalam konteks analisis sentimen ini.

3. Gunawan, N. O., & Nasib, S. K. (2025)

Penelitian yang berjudul, “Analisis Regresi Logistik Multinomial untuk Menentukan Faktor-Faktor yang Mempengaruhi Jenis Penyakit pada Mahasiswa (Studi Kasus: Mahasiswa Program Studi Statistika, Jurusan Matematika, Universitas Negeri Gorontalo)” mengidentifikasi faktor-faktor yang memengaruhi jenis penyakit pada mahasiswa Program Studi Statistika, Jurusan Matematika, Universitas Negeri Gorontalo.

4. Gusti Agung Widian Wira Brata & Iwan Wahyudin (2025)

Penelitian ini berjudul "Pengembangan Model Klasifikasi Penyakit Kucing Menggunakan K-Nearest Neighbor untuk Identifikasi Penyakit Menular/Tidak Menular". Penelitian ini mengembangkan model klasifikasi menggunakan algoritma K-Nearest Neighbor (KNN) untuk membedakan penyakit kucing yang menular dan tidak menular berdasarkan gejala klinis yang diamati. Model ini bertujuan untuk mendukung diagnosis dini dan meningkatkan akurasi identifikasi penyakit pada hewan peliharaan.

5. Putri, F. C., Andriyati, A., & Rohaeti, E. (2022)

Penelitian ini berjudul, “Analisis Regresi Logistik Multinomial pada Faktor-Faktor yang Mempengaruhi Status Pasien Covid-19 di Kota Depok” yang menganalisis faktor-faktor yang memengaruhi status pasien COVID-19 di Kota Depok. Peluang yang didapatkan untuk masing-masing status pasien

COVID-19 di Kota Depok adalah 0,000000036 untuk kategori terkonfirmasi, 0,999 untuk kategori kontak erat, dan 0,00000048 untuk kategori suspek.

6. Ermawati, E., Iknas, R., & Kurniawan, B. A. (2023)

Penelitian berjudul, “Klasifikasi Penderita Anemia Menggunakan Metode Regresi Logistik” mengklasifikasi faktor-faktor penyebab anemia pada remaja wanita. Atribut yang digunakan dalam penelitian ini meliputi *ferritin serum*, STfR, dan riwayat penyakit kronis. Hasil penelitian menunjukkan bahwa *ferritin serum* dan STfR memiliki pengaruh signifikan dalam mengklasifikasi anemia. Model yang dibangun memiliki tingkat akurasi sebesar 79,23%, presisi 80,18%, dan *recall* 96,39%. Hal ini mengindikasikan bahwa model tersebut cukup baik dalam mengklasifikasikan remaja putri yang terdiagnosis anemia dan tidak anemia.

BAB III

METODOLOGI PENELITIAN

3.1 Metode dan Tahapan Penelitian

Penelitian ini termasuk dalam penelitian kuantitatif, dengan pendekatan prediktif menggunakan regresi logistik multinomial. Penelitian ini bertujuan untuk memprediksi tingkat keparahan penyakit kucing (variabel dependen kategorikal nominal) berdasarkan gejala klinis yang diamati. Untuk mencapai tujuan tersebut, penelitian ini mengimplementasikan algoritma regresi logistik multinomial sebagai teknik analisis data utama. Data yang digunakan bersifat campuran (numerik dan kategorikal), terdiri dari variabel-variabel prediktor dan variabel target (tingkat keparahan penyakit dalam kategori: ringan, sedang, parah) yang bersumber dari data sekunder rekam medis.

Untuk mencapai tujuan penelitian secara efektif, dirancang sebuah alur kerja yang terdiri dari beberapa tahapan sistematis. Tahapan ini mencakup seluruh proses mulai dari studi awal, perancangan, implementasi, hingga evaluasi akhir. Secara garis besar, meliputi: (1) Pengumpulan Data, (2) Perancangan Sistem dan Model, (3) Implementasi Sistem, dan (4) Pengujian dan Evaluasi.

3.2 Metode Pengumpulan Data

Tahap pengumpulan data merupakan implementasi dari perencanaan untuk memperoleh bahan baku utama penelitian. Data yang digunakan dalam penelitian ini merupakan gabungan dari data primer yang diperoleh secara langsung dari klinik hewan dan data sekunder yang berasal dari repositori daring (*online*) publik.

3.2.1 Sumber Data Penelitian

Pengumpulan data dilakukan dari tiga jenis sumber utama untuk mendapatkan dataset yang beragam dan komprehensif:

1. Data dari Klinik Hewan

Data primer berupa rekam medis pasien diperoleh dari dua klinik hewan di kota yang berbeda, yaitu: *Fafa Pet Clinic* dan *Petshop* dan Klinik Hewan *Mitra Pet - Vet*.

2. Data dari Repositori Daring (*online*)

Untuk memperkaya variasi dan jumlah data, penelitian ini juga memanfaatkan beberapa dataset publik dari platform Kaggle dan GitHub. Secara rinci, data sekunder yang digunakan adalah:

- a. Dataset *Animal Disease Prediction* dari Kaggle
- b. Dataset *Animal Disease* dari Kaggle
- c. Dataset dari repositori GitHub (*dwinuragustina/CatDiagnostic*)

Perlu dicatat bahwa dataset sekunder ini berisi data penyakit hewan secara umum. Oleh karena itu, dilakukan proses pra-pemrosesan awal yang meliputi pembersihan dan penyaringan untuk mengambil hanya data yang spesifik untuk spesies kucing (*Felis catus*) dan yang relevan dengan enam penyakit dalam ruang lingkup penelitian ini.

3.2.2 Waktu dan Lokasi Penelitian

Penelitian ini berlangsung selama periode Maret hingga Juli 2025. Proses penelitian mencakup studi literatur, pengumpulan data, pengembangan model, hingga analisis hasil. Lokasi penelitian ini tidak terikat pada satu wilayah geografis, melainkan didefinisikan berdasarkan lokasi sumber data. Data klinis yang

digunakan dalam penelitian ini diperoleh dari dua klinik hewan di kota yang berbeda, yaitu:

1. Fafa *Pet Clinic* dan *Petshop*, yang beralamat di Jl. Sutrisno No.732, Sukaramai I, Kec. Medan Area, Kota Medan.
2. Klinik Hewan Mitra *Pet - Vet*, yang beralamat di Jl. Kyai Abdul Karim No.50, Rungkut Menanggal, Kec. Gn. Anyar, Surabaya.

Pengumpulan data dari Fafa *Pet Clinic* dilakukan secara langsung, sementara pengumpulan data dari Klinik Hewan Mitra *Pet* di Surabaya dilakukan secara daring (*online*) melalui komunikasi via aplikasi pesan dan serah terima data melalui email.

3.2.3 Populasi dan Sampel

Populasi adalah keseluruhan kelompok individu, objek, atau data yang memiliki karakteristik sejenis yang menjadi fokus penelitian. Populasi dalam penelitian ini adalah seluruh data rekam medis pasien kucing yang pernah didiagnosis menderita salah satu dari enam jenis penyakit yang menjadi ruang lingkup penelitian (FPV, FCV, CKD, UTI, *Ringworm*, dan *Scabies*).

Sampel adalah bagian dari populasi yang diambil untuk diamati secara langsung dan dianggap dapat mewakili keseluruhan populasi. Teknik pengambilan sampel yang digunakan dalam penelitian ini adalah *purposive sampling*. *Purposive sampling* adalah teknik penentuan sampel dengan pertimbangan atau kriteria tertentu yang relevan dengan tujuan penelitian. Kriteria inklusi (kriteria sampel yang diterima) yang ditetapkan adalah sebagai berikut:

1. Rekam medis harus mencatat diagnosis akhir yang termasuk dalam salah satu dari enam penyakit yang diteliti.

2. Data mengenai gejala klinis yang teramati harus tercatat dengan jelas dan lengkap.
3. Informasi mengenai usia kucing dan tingkat keparahan penyakit harus tersedia.

Berdasarkan teknik pengambilan sampel dan kriteria yang telah ditetapkan, dari keseluruhan populasi rekam medis yang ada, berhasil diperoleh sampel akhir sebanyak 501 data rekam medis yang valid dan siap untuk digunakan dalam tahap pra-pemrosesan dan pemodelan.

3.2.4 Variabel Penelitian

Data yang dikumpulkan dalam penelitian ini terdiri dari variabel dependen (variabel target yang diprediksi) dan variabel independen (fitur atau prediktor).

1. Variabel Dependen (Y)

Variabel dependen dalam penelitian ini adalah Tren Tingkat Keparahan Penyakit, yang merupakan variabel kategorikal nominal dengan tiga level kelas, yaitu: Ringan – Sedang – Parah

2. Variabel Independen (X)

Variabel independen adalah fitur-fitur yang digunakan untuk memprediksi variabel dependen. Fitur ini didasarkan pada 6 jenis penyakit dan 76 gejala klinis yang umum terjadi pada kucing. Daftar penyakit dan gejala yang menjadi ruang lingkup penelitian ini disajikan pada Tabel 3.1 dan Tabel 3.2.

Tabel 3. 1 Daftar Penyakit yang Menjadi Ruang Lingkup Penelitian

No.	Nama Penyakit
1	Penyakit Ginjal Kronis (CKD)
2	Panleukopenia
3	Infeksi Saluran Kemih (UTI)

4	Feline Calicivirus
5	Scabies
6	Ringworm

Tabel 3. 2 Daftar Gejala Klinis yang Digunakan sebagai Fitur

No.	Nama Gejala
1	Air liur berlebih
2	Anemia (gusi pucat)
3	Anoreksia (tidak mau makan sama sekali)
4	Banyak kencing (poliuria)
5	Banyak minum (polidipsia)
6	Bau mulut (halitosis)
7	Beberapa bercak kebotakan melingkar
8	Bersin
9	Bersin-bersin
10	Bulu kusam
11	Bulu rapuh di area tertentu
12	Dehidrasi
13	Dehidrasi berat
14	Dehidrasi parah
15	Dehidrasi ringan
16	Demam
17	Demam ringan
18	Demam tinggi
19	Diare (tanpa darah)
20	Diare berdarah
21	Gatal (tidak selalu)
22	Gatal hebat
23	Gatal ringan di tepi telinga
24	Gatal sangat parah hingga melukai diri sendiri
25	Hidung berair (ingus bening/kuning)
26	Hipotermia
27	Hipotermia (tahap akhir)
28	Infeksi bakteri sekunder
29	Infeksi kulit sekunder (bakteri)
30	Kebotakan luas
31	Kencing di luar kotak pasir
32	Kerak kekuningan
33	Kerontokan rambut di area yang gatal
34	Keropeng di tepi telinga, wajah, dan leher
35	Keropeng tebal dan menyebar di banyak area (kepala, telinga, kaki)
36	Keropeng tebal dan menyebar ke seluruh tubuh

37	Kolaps
38	Kulit kemerahan
39	Kulit menebal
40	Kulit sangat menebal dan berkerut
41	Lesu
42	Lesu parah
43	Leukopenia berat (dari data lab)
44	Luka di mulut (stomatitis uremik)
45	Luka di mulut/gusi (stomatitis)
46	Luka kecil di lidah
47	Luka parah di mulut dan hidung
48	Mata sedikit berair
49	Menangis saat kencing
50	Menjilati area genital
51	Muntah
52	Muntah parah
53	Muntah sering
54	Muntah sesekali
55	Nafsu makan berkurang
56	Nafsu makan menurun signifikan
57	Nafsu makan sedikit menurun
58	Nafsu makan turun karena mulut sakit
59	Penurunan berat badan karena stres dan tidak nyaman
60	Penurunan berat badan ringan
61	Perut tegang dan membesar
62	Radang paru-paru (pneumonia)
63	Radang sendi (limping syndrome)
64	Sakit perut hebat saat dipalpasi
65	Sangat kesakitan
66	Sangat lesu/lemah
67	Sangat lesu/tidak responsif
68	Satu atau dua bercak kecil kebotakan
69	Sedikit berkerak
70	Sedikit kemerahan
71	Sedikit lesu
72	Sering ke kotak pasir
73	Sulit bernapas
74	Sulit/mengejan saat kencing (stranguria)
75	Tidak bisa kencing sama sekali (obstruksi total - darurat!)
76	Urin tampak keruh/sedikit berdarah (hematuria)

3.3 Metode Perancangan Sistem dan Model

Ini merupakan tahap perancangan sebelum proses implementasi (*coding*) dilakukan. Perancangan ini mencakup arsitektur sistem, alur pemrosesan data, perancangan model, hingga desain antarmuka pengguna.

3.3.1 Arsitektur Sistem

Sistem dirancang dengan arsitektur klien-server sederhana. Klien adalah peramban web pengguna yang mengakses aplikasi. Server adalah aplikasi web yang dibangun menggunakan *framework* Flask, yang bertugas menerima permintaan HTTP, memproses data input, dan mengirimkan kembali hasil prediksi.

3.3.2 Pra-pemrosesan Data

Tahap ini bertujuan untuk mengubah data rekam medis mentah menjadi sebuah dataset yang terstruktur, dan siap untuk diolah oleh algoritma. Proses ini dirancang melalui beberapa langkah:

a. Pembersihan Data (*Data Cleaning*)

Merancang proses untuk menangani data yang tidak lengkap (*missing values*) atau tidak konsisten untuk memastikan kualitas dan integritas dataset.

b. Generalisasi Gejala

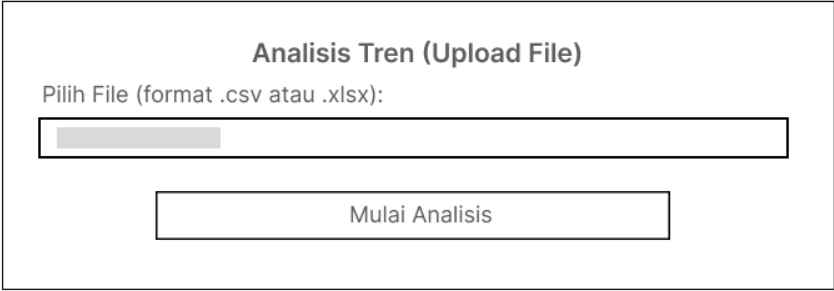
Merancang strategi *feature engineering* untuk mengurangi kompleksitas data. Berbagai deskripsi gejala spesifik yang memiliki makna serupa (contoh: 'Lesu', 'Sangat lesu/lemah') dipetakan ke dalam satu kategori fitur yang lebih umum (contoh: 'Lesu (Umum)').

c. Transformasi Fitur

Proses mengubah setiap fitur gejala yang sudah digeneralisasi menjadi representasi biner (0 dan 1) melalui metode yang secara konseptual setara dengan *One-Hot Encoding*.

3.3.3 Perancangan Antarmuka Pengguna (*User Interface*)

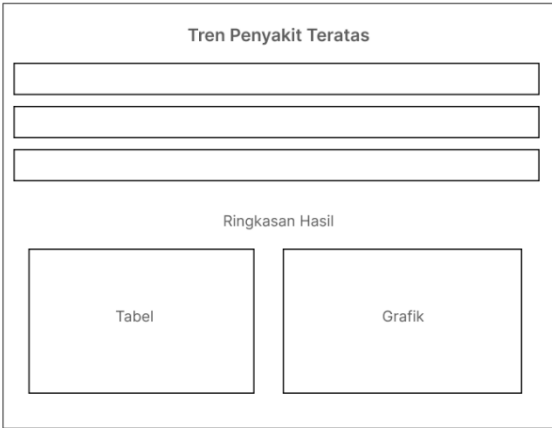
Antarmuka pengguna (UI) dirancang dengan prinsip minimalis dan intuitif untuk memastikan sistem dapat digunakan dengan mudah oleh berbagai kalangan, termasuk pemilik kucing tanpa latar belakang teknis. Fokus utama desain adalah pada alur kerja yang sederhana dan penyajian informasi yang jelas. Visualisasi dari rancangan tersebut diwujudkan dalam bentuk *wireframe* untuk Halaman Analisis, Halaman Hasil Analisis, Halaman Input Manual, dan Halaman Hasil Manual, seperti yang ditunjukkan pada gambar-gambar berikut.



Analisis Tren (Upload File)

Pilih File (format .csv atau .xlsx):

Gambar 3. 1 Halaman Analisis



Tren Penyakit Teratas

Ringkasan Hasil

Tabel

Grafik

Gambar 3. 2 Halaman Hasil Analisis

This wireframe represents the input page of a cat disease prediction application. It features a large empty rectangular box at the top for a user profile or image. Below it is a text input field with the placeholder text "Masukkan usia kucing". Underneath is a dropdown menu labeled "Pilih Gejala". The main section contains two columns of five checkboxes, each followed by a text input field, for selecting symptoms. At the bottom is a button labeled "Prediksi".

Masukkan usia kucing

Pilih Gejala

☐ ☐

☐ ☐

☐ ☐

☐ ☐

☐ ☐

Prediksi

Gambar 3. 3 Wireframe Halaman Input

This wireframe represents the result page of the application. It starts with a box labeled "Hasil Penyakit". Below it is a text area labeled "Text". A button labeled "Analisis Prediksi" is positioned above a container that holds a small text input field and a table with six rows. At the bottom is a button labeled "Prediksi Lagi".

Hasil Penyakit

Text

Analisis Prediksi

Prediksi Lagi

Gambar 3. 4 Wireframe Halaman Hasil

Halaman utama aplikasi ini berfungsi sebagai pusat interaksi pengguna dengan sistem prediksi. Terdapat dua fitur utama yang disajikan: Analisis Tren melalui unggah file dan Prediksi Manual untuk kasus tunggal. Bagian Analisis Tren memungkinkan pengguna mengunggah dataset dalam format CSV atau Excel untuk mendapatkan gambaran umum tren penyakit. Sementara itu, bagian Prediksi Manual menyediakan antarmuka interaktif berupa kolom isian untuk usia kucing dan sebuah daftar *checklist* yang komprehensif untuk berbagai gejala klinis. Di bagian atas, disajikan pula informasi performa model, termasuk skor akurasi dan tautan untuk melihat *Confusion Matrix* sebagai bukti validitas. Setiap bagian dilengkapi dengan tombol *call-to-action* yang jelas untuk memulai proses analisis.

Halaman Hasil Analisis dirancang secara dinamis untuk menyajikan wawasan prediktif baik secara agregat maupun individual. Untuk analisis tren melalui unggah file, halaman ini menampilkan sorotan utama berupa tren penyakit teratas, lengkap dengan deskripsi klinis dan saran penanganan yang disesuaikan dengan tingkat keparahannya. Di bawahnya, ringkasan tren yang lebih luas disajikan dalam bentuk tabel dan grafik distribusi untuk memberikan gambaran umum yang komprehensif, diikuti oleh tabel rincian interaktif untuk setiap kasus.

Untuk prediksi manual satu kasus, halaman hasil berfokus pada transparansi dan interpretasi. Secara hierarkis, di bawahnya disajikan detail analisis koefisien yang merinci bobot pengaruh dari setiap gejala yang dipilih pengguna. Desain ini bertujuan agar pengguna tidak hanya menerima hasil prediksi, tetapi juga mendapatkan wawasan dasar mengenai faktor-faktor klinis yang paling berkontribusi terhadap diagnosis tersebut.

3.4 Evaluasi dan Validasi Model

Model prediksi dibangun menggunakan algoritma regresi logistik multinomial. Dataset yang telah diproses dibagi menjadi 80% data latih dan 20% data uji (`random_state=42`) untuk memastikan hasil yang dapat direplikasi. Model dilatih secara eksklusif menggunakan data latih untuk mempelajari pola hubungan antara gejala dan diagnosis penyakit.

Kinerja model yang telah dilatih akan diukur menggunakan data uji. Evaluasi bertujuan untuk menilai seberapa akurat model dapat memprediksi tingkat keparahan pada data yang belum pernah dilihat sebelumnya. Metrik evaluasi yang akan digunakan, yang sesuai untuk masalah klasifikasi multinomial, adalah akurasi (*accuracy*), presisi (*precision*), *recall*, F1-Score, dan matriks konfusi (*confusion matrix*).

BAB IV

HASIL DAN PEMBAHASAN

4.1 Implementasi Sistem

Sistem prediksi ini dirancang dengan dua fungsionalitas utama untuk memenuhi kebutuhan pengguna yang berbeda: fitur Analisis Tren untuk analisis data dalam jumlah besar dan fitur Prediksi Manual untuk diagnosis kasus individual. Alur kerja sistem dimulai ketika pengguna mengakses halaman utama yang menyediakan dua mode interaksi.

Pertama, Analisis Tren yang memungkinkan pengguna untuk mengunggah file dataset (dalam format CSV atau Excel) berisi data beberapa kucing sekaligus. Fitur ini dirancang untuk memproses data secara *batch* dan menyajikan gambaran umum tren penyakit. Kedua, sistem juga dilengkapi fitur Prediksi Manual, di mana pengguna dapat memasukkan data berupa usia kucing (dalam tahun) dan memilih satu atau lebih gejala klinis yang teramati dari daftar *checklist* yang telah disediakan. Tampilan antarmuka untuk kedua fitur input data ini dapat dilihat pada Gambar 4.1 dan Gambar 4.2.

Analisis Tren (Upload File)

Pilih File (format .csv atau .xlsx):

Choose File

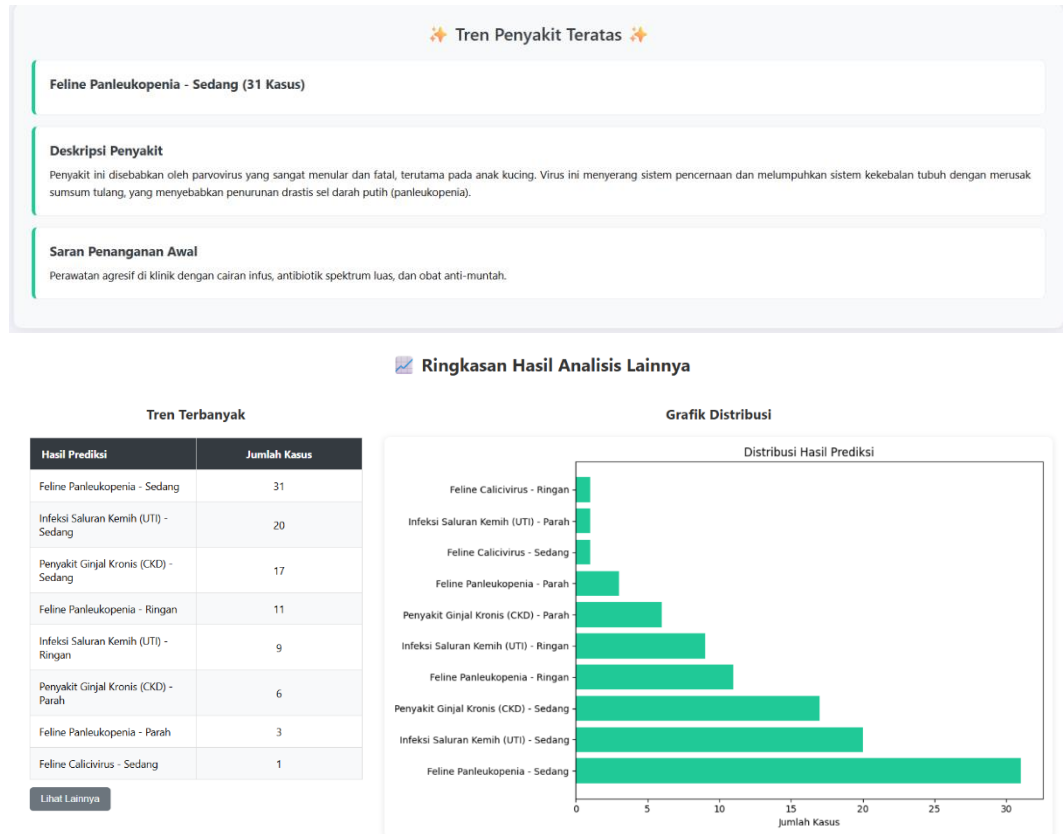
No file chosen

Mulai Analisis

Gambar 4. 1 Halaman Analisis

Gambar 4.1 adalah tampilan antarmuka untuk fitur Analisis Tren. Pada halaman ini, pengguna dapat mengunggah file data dengan format .csv atau .xlsx

yang berisi rekam data gejala klinis. Proses analisis data secara massal dimulai ketika pengguna menekan tombol "Mulai Analisis".



Gambar 4. 2 Halaman Hasil Analisis

Gambar 4.2 menampilkan ringkasan hasil analisis setelah sistem memproses file data yang diunggah. Halaman ini menyajikan dua bentuk output utama. Di sisi kiri, terdapat tabel "Tren Terbanyak" yang mengurutkan hasil prediksi (kombinasi penyakit dan tingkat keparahan) dari yang paling banyak hingga paling sedikit ditemukan dalam dataset. Di sisi kanan, "Grafik Distribusi" menyajikan data yang sama dalam format diagram batang horizontal, memungkinkan pengguna untuk secara visual membandingkan jumlah kasus pada setiap kategori prediksi dengan cepat.

Prediksi Penyakit Kucing

Usia Kucing (Tahun):

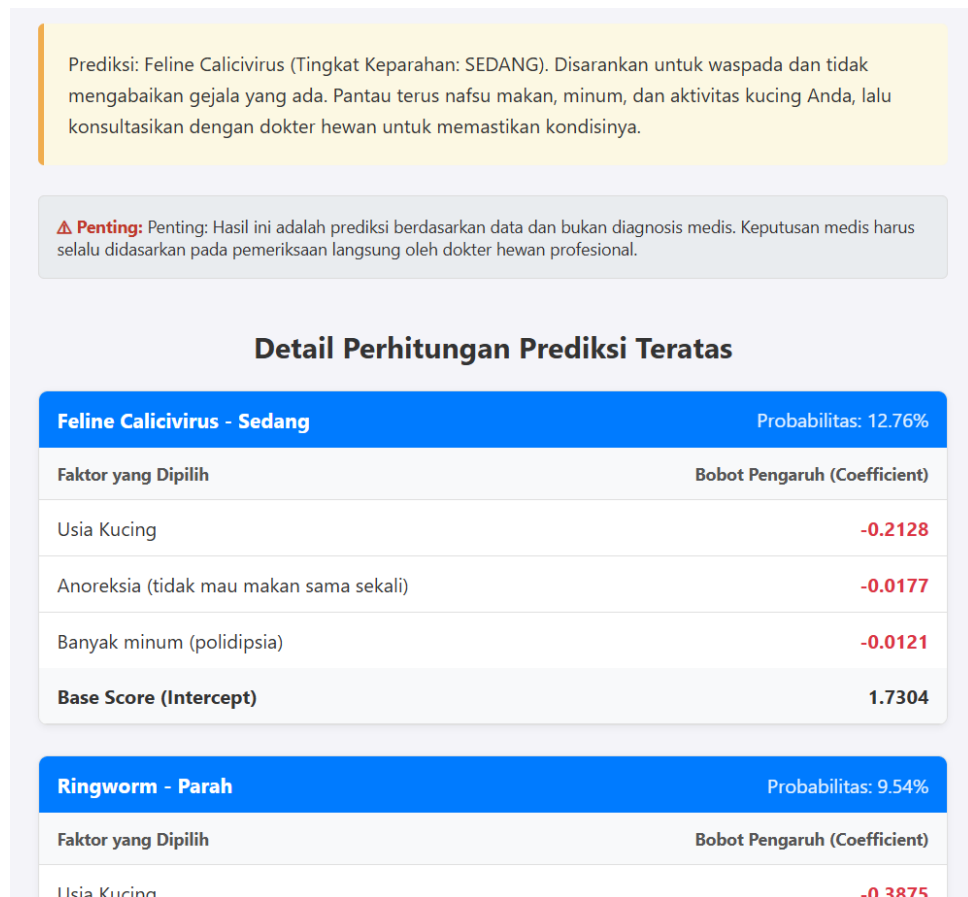
Pilih Gejala yang Dialami:

☐ Anoreksia (tidak mau makan sama sekali)	☐ Banyak kencing (poliuria)
☐ Banyak minum (polidipsia)	☐ Bau mulut (halitosis)
☐ Beberapa bercak kebotakan melingkar	☐ Bersin
☐ Bersin-bersin	☐ Bulu kusam
☐ Bulu rapuh di area tertentu	☐ Dehidrasi
☐ Dehidrasi berat	☐ Dehidrasi parah
☐ Dehidrasi ringan	☐ Demam

Prediksi Penyakit

Gambar 4. 3 Halaman Input

Pengguna menginput usia beserta gejala yang dialami kucing, jika semua gejala telah sesuai, pengguna dapat memprediksi penyakit dengan menekan tombol prediksi penyakit. Setelah pengguna menekan tombol tersebut, data input tersebut akan dikirim dan diproses oleh sistem di sisi server. Model regresi logistik multinomial kemudian menghitung probabilitas untuk setiap kelas penyakit. Hasilnya, sistem akan menampilkan halaman hasil prediksi yang berisi tiga kemungkinan penyakit teratas dengan persentase probabilitas tertinggi. Selain itu, untuk prediksi teratas, sistem juga menampilkan detail faktor-faktor yang berpengaruh beserta bobot koefisiennya untuk memberikan transparansi terhadap hasil model.



Gambar 4. 4 Halaman Hasil

Setelah pengguna memasukkan data usia dan gejala lalu menekan tombol prediksi, sistem akan menampilkan halaman hasil. Halaman ini menyajikan output dari model regresi logistik multinomial yang telah memproses data input secara *real-time*. Penting untuk ditekankan bahwa hasil yang ditampilkan merupakan prediksi berdasarkan pola data dan tidak dapat menggantikan diagnosis medis profesional dari dokter hewan.

Pengembangan sistem ini melibatkan beberapa teknologi utama. Pada sisi server (*backend*), sistem dibangun menggunakan bahasa pemrograman Python. Logika aplikasi web dan penanganan rute ditangani oleh *micro-framework* Flask. Model klasifikasi regresi logistik multinomial itu sendiri diimplementasikan,

dilatih, dan digunakan untuk prediksi dengan bantuan library *machine learning* Scikit-learn (sklearn). Sementara itu, antarmuka pengguna (*frontend*) yang responsif dan mudah digunakan dibangun menggunakan teknologi web standar, yaitu HTML, CSS, dan sedikit JavaScript untuk interaktivitas.

4.2 Pengujian dan Evaluasi Model

Pengujian dan evaluasi model bertujuan untuk mengukur performa dan keandalan model regresi logistik multinomial secara kuantitatif dalam memprediksi penyakit kucing berdasarkan gejala klinis.

4.2.1 Skenario Pengujian

Pengujian model dilakukan dengan sebuah skenario yang terstruktur untuk memastikan hasil yang didapatkan objektif dan terukur. Skenario pengujian yang dilakukan adalah sebagai berikut:

1. Dataset Pengujian

Model diuji menggunakan dataset independen yang tidak digunakan selama proses pelatihan, yaitu *testing set*. Berdasarkan laporan klasifikasi yang dihasilkan, diketahui bahwa jumlah data yang digunakan untuk pengujian adalah sebanyak 101 sampel kasus penyakit. Dataset ini berisi fitur (gejala klinis dan usia) serta label (diagnosis penyakit asli) yang akan digunakan sebagai pembandingan untuk hasil prediksi model.

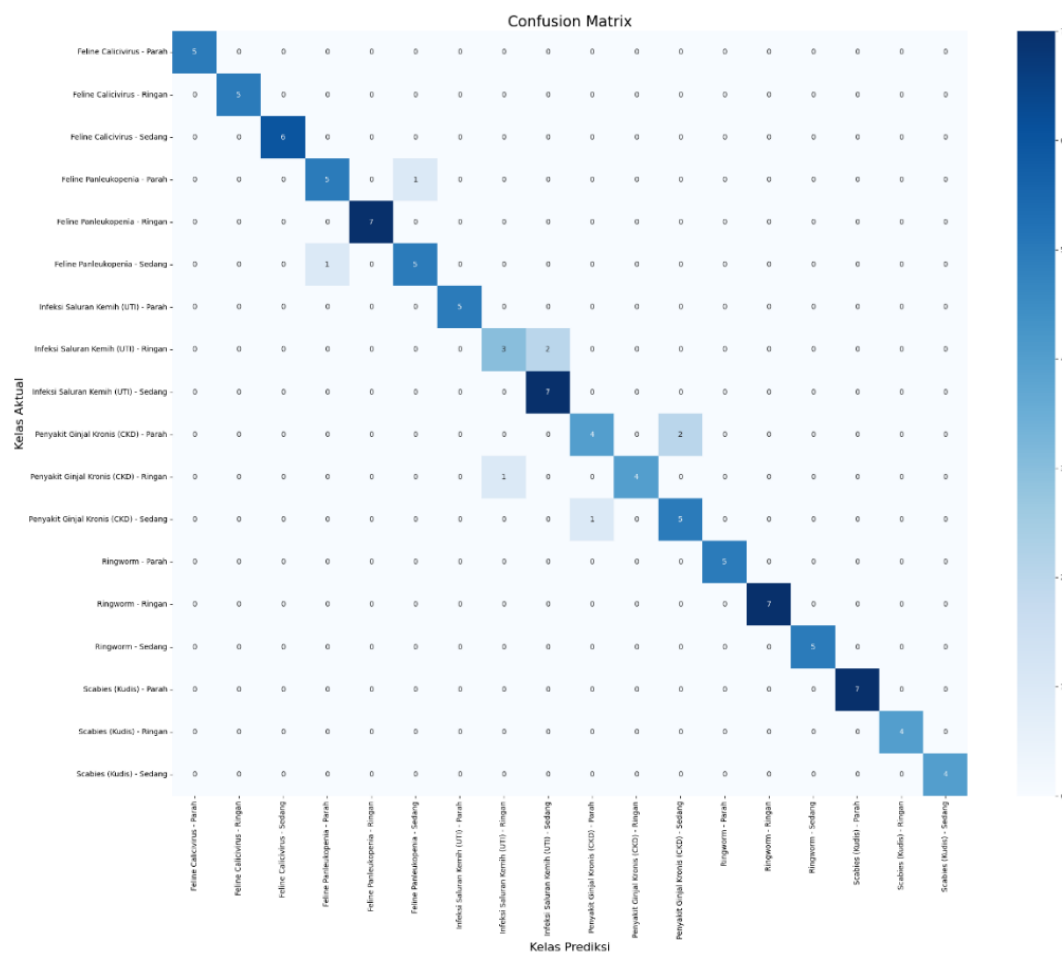
2. Pembagian Data

Dataset utama yang digunakan dalam penelitian ini berisi total 501 sampel. Sebelum tahap pelatihan, dataset tersebut dibagi menjadi dua bagian secara acak dengan rasio 80% data latih dan 20% data uji. Dengan demikian, sebanyak 400 sampel digunakan untuk melatih model, dan 101 sampel sisanya

digunakan khusus untuk tahap pengujian guna mengukur performa model pada data yang belum pernah dilihat sebelumnya.

4.2.2 Hasil Pengujian Model

Berdasarkan skenario pengujian yang telah dijelaskan pada sub-bab sebelumnya, pengujian dilakukan terhadap model regresi logistik multinomial menggunakan 101 sampel data uji. Proses pengujian ini menghasilkan grafik *confusion matrix* dan evaluasi kinerja model yang merinci performa model untuk setiap kelas penyakit.



Gambar 4. 5 Grafik Confusion Matrix

Sumbu vertikal (kelas aktual) merepresentasikan diagnosis penyakit yang sebenarnya, sedangkan sumbu horizontal (kelas prediksi) merepresentasikan hasil tebakan model. Nilai yang tinggi di sepanjang garis diagonal menunjukkan jumlah prediksi yang benar, yang menandakan performa model secara umum sangat baik. Kesalahan klasifikasi yang terjadi (nilai di luar diagonal) menunjukkan pola yang jelas, di mana model cenderung keliru antara tingkat keparahan yang berdekatan, seperti memprediksi ringan sebagai sedang untuk jenis penyakit yang sama, yang mengindikasikan kemiripan gejala pada level tersebut.

Tabel 4. 1 Evaluasi Kinerja Model

Kelas	Precision	Recall	F1-Score	Support
Feline Calicivirus - Parah	1	1	1	5
Feline Calicivirus - Ringan	1	1	1	5
Feline Calicivirus - Sedang	1	1	1	6
Feline Panleukopenia - Parah	0.83	0.83	0.83	6
Feline Panleukopenia - Ringan	1	1	1	7
Feline Panleukopenia - Sedang	0.83	0.83	0.83	6
Infeksi Saluran Kemih (UTI) - Parah	1	1	1	5
Infeksi Saluran Kemih (UTI) - Ringan	0.75	0.6	0.67	5
Infeksi Saluran Kemih (UTI) - Sedang	0.78	1	0.88	7
Penyakit Ginjal Kronis (CKD) - Parah	0.8	0.67	0.73	6
Penyakit Ginjal Kronis (CKD) - Ringan	1	0.8	0.89	5
Penyakit Ginjal Kronis (CKD) - Sedang	0.71	0.83	0.77	6
Ringworm - Parah	1	1	1	5
Ringworm - Ringan	1	1	1	7
Ringworm - Sedang	1	1	1	5
Scabies (Kudis) - Parah	1	1	1	7
Scabies (Kudis) - Ringan	1	1	1	4
Scabies (Kudis) - Sedang	1	1	1	4
Accuracy			0.92	101
Macro Avg	0.93	0.92	0.92	101
Weighted Avg	0.92	0.92	0.92	101

Pengujian model pada 101 data uji menghasilkan performa yang sangat baik dengan tingkat akurasi keseluruhan mencapai 92%. Hal ini didukung oleh nilai *macro average* dan *weighted average* untuk F1-Score yang juga berada di angka 0.92, menunjukkan bahwa model memiliki kemampuan generalisasi yang baik di berbagai kelas penyakit.

Model menunjukkan performa sempurna (F1-Score 1.00) pada beberapa kelas penyakit seperti 'Feline Calicivirus' (Ringan & Sedang), 'Ringworm' (semua tingkat keparahan), dan 'Scabies' (semua tingkat keparahan). Hal ini menandakan bahwa gejala untuk penyakit-penyakit tersebut sangat khas sehingga model dapat mempelajarinya dengan sangat baik tanpa terjadi kesalahan klasifikasi pada data uji.

Analisis lebih dalam menunjukkan beberapa karakteristik menarik pada kelas tertentu. Sebagai contoh, pada kelas 'Infeksi Saluran Kemih (UTI) - Sedang', model mencapai *recall* 1.00 namun presisi 0.78. Ini berarti model mampu mengidentifikasi semua kasus nyata UTI - Sedang, namun terkadang keliru mengklasifikasikan penyakit lain sebagai UTI - Sedang. Sebaliknya, pada kelas 'Penyakit Ginjal Kronis (CKD) - Parah', model memiliki *recall* yang lebih rendah yaitu 0.67. Hal ini mengindikasikan bahwa model cenderung melewati beberapa kasus nyata CKD - Parah, yang kemungkinan disebabkan oleh kemiripan gejalanya dengan penyakit lain dalam dataset.

4.3 Analisis Prediksi Model

Untuk membuktikan alur kerja prediksi dari awal hingga akhir, sebuah kasus akan diuraikan secara mendetail. Proses ini akan mengikuti satu set input gejala spesifik, memverifikasi perhitungannya secara manual menggunakan rumus regresi

logistik multinomial, dan membandingkannya dengan hasil yang ditampilkan oleh sistem untuk menunjukkan konsistensi implementasi model. Sebelum membedah satu kasus prediksi secara rinci, kinerja model akan dievaluasi terlebih dahulu secara keseluruhan. Alat evaluasi utama yang digunakan adalah *confusion matrix*, yang merangkum performa prediksi model di semua kelas.

4.3.1 Confusion Matrix

Tabel 4. 2 Rincian Perhitungan Metrik Kinerja

Kelas	TP	FP	FN	Perhitungan Presisi (TP / (TP+FP))	Perhitungan Recall (TP / (TP+FN))	Perhitungan F1-Score ($2 * (P*R) / (P+R)$)	Support
FCV - Parah	5	0	0	$5 / (5+0) = 1.0$	$5 / (5+0) = 1.0$	$2*(1*1) / (1+1) = 1.0$	5
FCV - Ringan	5	0	0	$5 / (5+0) = 1.0$	$5 / (5+0) = 1.0$	$2*(1*1) / (1+1) = 1.0$	5
FCV - Sedang	6	0	0	$6 / (6+0) = 1.0$	$6 / (6+0) = 1.0$	$2*(1*1) / (1+1) = 1.0$	6
FPV - Parah	5	1	1	$5 / (5+1) = 0.83$	$5 / (5+1) = 0.83$	$2*(0.83*0.83) / (0.83+0.83) = 0.83$	6
FPV - Ringan	7	0	0	$7 / (7+0) = 1.0$	$7 / (7+0) = 1.0$	$2*(1*1) / (1+1) = 1.0$	7
FPV - Sedang	5	1	1	$5 / (5+1) = 0.83$	$5 / (5+1) = 0.83$	$2*(0.83*0.83) / (0.83+0.83) = 0.83$	6
UTI - Parah	5	0	0	$5 / (5+0) = 1.0$	$5 / (5+0) = 1.0$	$2*(1*1) / (1+1) = 1.0$	5
UTI - Ringan	3	1	2	$3 / (3+1) = 0.75$	$3 / (3+2) = 0.60$	$2*(0.75*0.6) / (0.75+0.6) = 0.67$	5
UTI - Sedang	7	2	0	$7 / (7+2) = 0.78$	$7 / (7+0) = 1.0$	$2*(0.78*1) / (0.78+1) = 0.88$	7
CKD - Parah	4	1	2	$4 / (4+1) = 0.80$	$4 / (4+2) = 0.67$	$2*(0.8*0.67) / (0.8+0.67) = 0.73$	6
CKD - Ringan	4	0	1	$4 / (4+0) = 1.0$	$4 / (4+1) = 0.80$	$2*(1*0.8) / (1+0.8) = 0.89$	5
CKD - Sedang	5	2	1	$5 / (5+2) = 0.71$	$5 / (5+1) = 0.83$	$2*(0.71*0.83) / (0.71+0.83) = 0.77$	6
Ringworm - Parah	5	0	0	$5 / (5+0) = 1.0$	$5 / (5+0) = 1.0$	$2*(1*1) / (1+1) = 1.0$	5
Ringworm - Ringan	7	0	0	$7 / (7+0) = 1.0$	$7 / (7+0) = 1.0$	$2*(1*1) / (1+1) = 1.0$	7

Ringworm - Sedang	5	0	0	$5 / (5+0) = 1.0$	$5 / (5+0) = 1.0$	$2*(1*1) / (1+1) = 1.0$	5
Kudis - Parah	7	0	0	$7 / (7+0) = 1.0$	$7 / (7+0) = 1.0$	$2*(1*1) / (1+1) = 1.0$	7
Kudis - Ringan	4	0	0	$4 / (4+0) = 1.0$	$4 / (4+0) = 1.0$	$2*(1*1) / (1+1) = 1.0$	4
Kudis - Sedang	4	0	0	$4 / (4+0) = 1.0$	$4 / (4+0) = 1.0$	$2*(1*1) / (1+1) = 1.0$	4
Total	93	8	8	16.7	16.56	16.6	101

Accuracy

$$Accuracy = (\sum TP) / (Total\ seluruh\ data) = 93/101 = 0.9207 \approx 0.92$$

Macro Avarage

$$Precision = (\sum Precision) / (Total\ Kelas) = 16.7/18 = 0.9277 \approx 0.93$$

$$Recall = (\sum Recall) / (Total\ Kelas) = 16.56/18 = 0.92$$

$$F1 - Score = (\sum F1 - Score) / (Total\ Kelas) = 16.6/18 = 0.9222 \approx 0.92$$

Weighted Average

$$Precision = (Precision * Support) / Total\ Support$$

$$\begin{aligned}
&= ((1 * 5) + (1 * 5) + (1 * 6) + (0.83 * 6) + (1 * 7) + \\
&\quad (0.83 * 6) + (1 * 5) + (0.75 * 5) + (0.78 * 7) + (0.80 * \\
&\quad 6) + (1 * 5) + (0.71 * 6) + (1 * 5) + (1 * 7) + (1 * 5) + \\
&\quad (1 * 7) + (1 * 4) + (1 * 4)) / 101 \\
&= 92.23 / 101 \\
&= 0.9131 \approx 0.91
\end{aligned}$$

$$Recall = (Total\ TP) / (Total\ seluruh\ data) = 93/101 = 0.9207 \approx 0.92$$

$$F1 - Score = (F1 - Score * Support) / (Total\ Support)$$

$$\begin{aligned}
&= ((1 * 5) + (1 * 5) + (1 * 6) + (0.83 * 6) + (1 * 7) + \\
&\quad (0.83 * 6) + (1 * 5) + (0.67 * 5) + (0.88 * 7) + \\
&\quad (0.73 * 6) + (0.89 * 5) + (0.77 * 6) + (1 * 5) + (1 * 7) + \\
&\quad (1 * 5) + (1 * 7) + (1 * 4) + (1 * 4)) \\
&= 91.92/101 \\
&= 0.9100 \approx 0.91
\end{aligned}$$

Berdasarkan perhitungan di atas, performa model klasifikasi menunjukkan hasil yang sangat baik, dengan nilai akurasi keseluruhan mencapai 0.92. Nilai *Macro Average* yang juga berada di angka 0.92 - 0.93 mengindikasikan bahwa model memiliki kemampuan yang merata dalam mengenali setiap kelas tanpa memandang jumlah sampelnya.

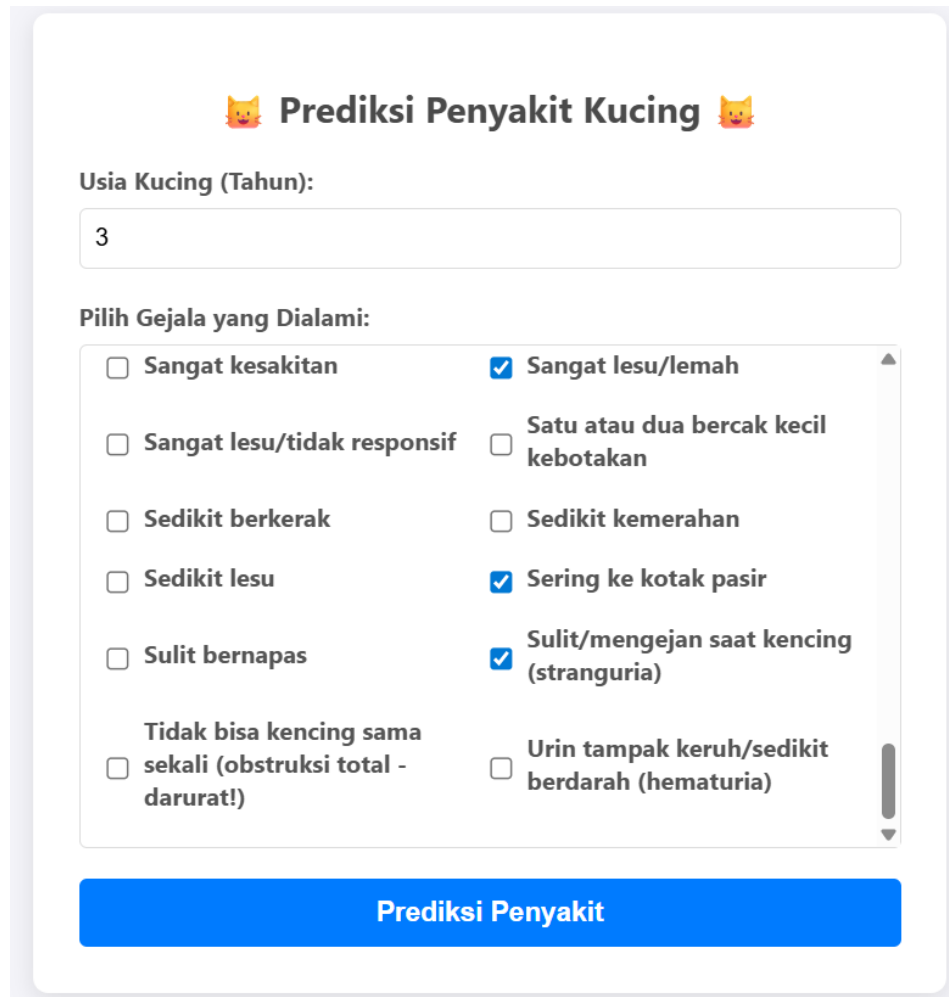
Perlu dicatat bahwa terdapat perbedaan minor antara hasil perhitungan manual *Weighted Average* untuk *Precision* dan *F1-Score* (0.91) dengan nilai yang tertera pada laporan klasifikasi sistem (0.92). Perbedaan ini disebabkan oleh akumulasi galat pembulatan (*rounding error*) yang timbul saat menghitung rata-rata dari nilai-nilai (seperti 0.83 atau 0.67) yang sudah dibulatkan terlebih dahulu. Perhitungan internal oleh sistem (*library Scikit-learn*) menggunakan presisi *floating-point* penuh sehingga hasilnya lebih akurat.

Perhitungan rinci di atas memvalidasi secara definitif bahwa nilai *precision* dan *recall* pada laporan klasifikasi (Tabel 4.1) adalah akurat, karena seluruhnya terbukti sesuai dengan data aktual dan prediksi dari *confusion matrix*.

4.3.2 Studi Kasus Prediksi

Untuk menguji dan mendemonstrasikan alur kerja sistem secara praktis, sebuah studi kasus spesifik dipilih. Kasus ini dirancang untuk mewakili skenario

penggunaan nyata, di mana serangkaian input gejala akan diproses oleh model untuk kemudian diverifikasi dan dianalisis secara mendalam.



🐱 Prediksi Penyakit Kucing 🐱

Usia Kucing (Tahun):

3

Pilih Gejala yang Dialami:

☐ Sangat kesakitan	☒ Sangat lesu/lemah
☐ Sangat lesu/tidak responsif	☐ Satu atau dua bercak kecil kebotakan
☐ Sedikit berkerak	☐ Sedikit kemerahan
☐ Sedikit lesu	☒ Sering ke kotak pasir
☐ Sulit bernapas	☒ Sulit/mengejan saat kencing (stranguria)
☐ Tidak bisa kencing sama sekali (obstruksi total - darurat!)	☐ Urin tampak keruh/sedikit berdarah (hematuria)

Prediksi Penyakit

Gambar 4. 6 Input Studi Kasus Kucing

Data input yang digunakan dalam studi kasus ini adalah sebagai berikut:

1. Usia Kucing: 3 Tahun
2. Gejala yang Dipilih:
 - a. Sangat lesu/lemah
 - b. Sering ke kotak pasir
 - c. Sulit/mengejan saat kencing (stranguria)

Setelah pengguna memasukkan data input, sistem akan memproses informasi tersebut saat tombol "Prediksi" ditekan. Model regresi logistik multinomial kemudian menghitung probabilitas untuk setiap kelas penyakit dan menampilkan hasilnya pada halaman output, seperti yang terlihat pada Gambar 4.7.

Infeksi Saluran Kemih (UTI) - Sedang		Probabilitas: 25.15%
Faktor yang Dipilih	Bobot Pengaruh (Coefficient)	
Usia Kucing	0.2483	
Sangat lesu/lemah	-0.0946	
Sering ke kotak pasir	-0.1881	
Sulit/mengejan saat kencing (stranguria)	1.8017	
Base Score (Intercept)	-0.2333	

Infeksi Saluran Kemih (UTI) - Ringan		Probabilitas: 15.39%
Faktor yang Dipilih	Bobot Pengaruh (Coefficient)	
Usia Kucing	0.3391	
Sangat lesu/lemah	-0.1116	
Sering ke kotak pasir	1.7797	
Sulit/mengejan saat kencing (stranguria)	-0.1270	
Base Score (Intercept)	-1.0186	

Feline Calicivirus - Sedang		Probabilitas: 7.85%
Faktor yang Dipilih	Bobot Pengaruh (Coefficient)	
Usia Kucing	-0.2128	
Sangat lesu/lemah	-0.0105	
Sering ke kotak pasir	-0.0943	
Sulit/mengejan saat kencing (stranguria)	-0.1208	
Base Score (Intercept)	1.7304	

Gambar 4. 7 Hasil Studi Kasus Kucing

Gambar di atas menampilkan hasil prediksi dari sistem klasifikasi penyakit kucing, yang menyajikan diagnosis teratas beserta probabilitasnya. Prediksi utama yang dihasilkan adalah ‘Infeksi Saluran Kemih (UTI) – Sedang’ dengan probabilitas 25.15%.

Untuk setiap prediksi, disajikan pula tabel rincian yang menunjukkan faktor-faktor gejala yang diinput beserta bobot pengaruh (koefisien) masing-masing. Bobot ini mengindikasikan seberapa besar kontribusi setiap gejala terhadap kemungkinan diagnosis tersebut. Pada prediksi ‘UTI – Sedang’, gejala ‘Sulit/mengejan saat kencing’ memiliki bobot positif yang sangat tinggi (+1.8017), menandakan bahwa gejala ini adalah indikator terkuat untuk penyakit tersebut.

4.3.3 Verifikasi Perhitungan Manual

Untuk memvalidasi hasil sistem, dilakukan perhitungan manual untuk skor logit dari prediksi teratas berdasarkan rumus teori. Dari Gambar 4.7, diperoleh bobot (koefisien) dan intercept sebagai berikut:

1. Perhitungan skor logit

‘UTI – Sedang’, untuk perhitungan dimulai dari skor dasar (*intercept*) -0.2333, lalu disesuaikan dengan bobot usia kucing (+0.2483) dan gejala terpilih: sangat lesu/lemah (-0.0946), sering ke kotak pasir (-0.1881), dan sulit/mengejan saat kencing (+1.8017).

$$\begin{aligned} \text{logit} = & \text{Intercept} + (\text{Bobot}_{\text{Usia}} \times \text{Nilai}_{\text{Usia}}) + (\text{Bobot}_{\text{SangatLesu}} \times \\ & \text{Nilai}_{\text{SangatLesu}}) + (\text{Bobot}_{\text{SeringKeKotakPasir}} \times \\ & \text{Nilai}_{\text{SeringKeKotakPasir}}) + (\text{Bobot}_{\text{SulitKencing}} \times \text{Nilai}_{\text{SulitKencing}}) \end{aligned}$$

$$\begin{aligned}
&= -0.2333 + (0.2483 \times 3) + (-0.0946 \times 1) + (-0.1881 \times 1) + (1.8017 \times 1) \\
&= 2.0306
\end{aligned}$$

‘UTI – Ringan’, untuk perhitungan dimulai dari skor dasar (*intercept*) -1.0186, lalu disesuaikan dengan bobot usia kucing (+0.3391) dan gejala terpilih: sangat lesu/lemah (-0.1116), sering ke kotak pasir (+1.7797), dan sulit/mengejan saat kencing (-0.1270).

$$\begin{aligned}
\text{logit} &= \text{Intercept} + (\text{Bobot}_{\text{Usia}} \times \text{Nilai}_{\text{Usia}}) + (\text{Bobot}_{\text{SangatLesu}} \times \text{Nilai}_{\text{SangatLesu}}) + (\text{Bobot}_{\text{SeringKeKotakPasir}} \times \text{Nilai}_{\text{SeringKeKotakPasir}}) + (\text{Bobot}_{\text{SulitKencing}} \times \text{Nilai}_{\text{SulitKencing}}) \\
&= -1.0186 + (0.3391 \times 3) + (-0.1116 \times 1) + (1.7797 \times 1) + (-0.1270 \times 1) \\
&= -1.0186 + 1.0173 - 0.1116 + 1.7797 - 0.1270 \\
&= 1.5398
\end{aligned}$$

‘FCV – Sedang’, untuk skor dasar (*intercept*) +1.7304, lalu disesuaikan dengan bobot usia kucing (-0.2128) dan gejala terpilih: sangat lesu/lemah (-0.0105), sering ke kotak pasir (-0.0943), dan sulit/mengejan saat kencing (-0.1208).

$$\begin{aligned}
\text{logit} &= \text{Intercept} + (\text{Bobot}_{\text{Usia}} \times \text{Nilai}_{\text{Usia}}) + (\text{Bobot}_{\text{SangatLesu}} \times \text{Nilai}_{\text{SangatLesu}}) + (\text{Bobot}_{\text{SeringKeKotakPasir}} \times \text{Nilai}_{\text{SeringKeKotakPasir}}) + (\text{Bobot}_{\text{SulitKencing}} \times \text{Nilai}_{\text{SulitKencing}}) \\
&= 1.7304 + (-0.2128 \times 3) + (-0.0105 \times 1) + (-0.0943 \times 1) + (-0.1208 \times 1)
\end{aligned}$$

$$\begin{aligned}
 &= 1.7304 - 0.6384 - 0.0105 - 0.0943 - 0.1208 \\
 &= 0.8664
 \end{aligned}$$

2. Hitung nilai eksponensial (e^Z) untuk setiap skor logit

$$\text{UTI - Sedang: } e^{\text{logit}} = e^{2.0306} = 7.619$$

$$\text{UTI - Ringan: } e^{\text{logit}} = e^{1.5398} = 4.663$$

$$\text{FCV - Sedang: } e^{\text{logit}} = e^{0.8664} = 2.378$$

3. Jumlahkan semua nilai eksponensial untuk mendapatkan nilai penyebut

Nilai penyebut (*denominator*) D dalam perhitungan *softmax* diperoleh dari penjumlahan total nilai eksponensial dari skor logit untuk seluruh 18 kelas penyakit yang ada dalam model. Sistem secara otomatis menghitung setiap skor logit ini menggunakan set koefisien uniknya masing-masing, yang kemudian dieksponensialkan dan dijumlahkan untuk menghasilkan nilai akhir D. Berikut akan dirincikan perhitungan skor logit (Z_k) dan nilai eksponensial (e^{Z_k}) untuk setiap dari 18 kelas.

Tabel 4. 3 Perhitungan Denominator Softmax

Kelas	Perhitungan Skor Logit (Z_k)	Hasil Z_k	Nilai Eksponensial (e^{Z_k})
FCV - Parah	$1.1391 + (-0.1251 * 3) + (-0.0129) + (-0.0854) + (-0.1610)$	0.5045	1.6562
FCV - Ringan	$0.9386 + (-0.1412 * 3) + (-0.0101) + (-0.0664) + (-0.0820)$	0.3565	1.4283
FCV - Sedang	$1.7304 + (-0.2128 * 3) + (-0.0105) + (-0.0943) + (-0.1208)$	0.8664	2.3783
FPV - Parah	$2.2248 + (-0.6485 * 3) + (-0.0016) + (-0.0288) + (-0.0617)$	0.1872	1.2059
FPV - Ringan	$1.7358 + (-0.5963 * 3) + (-0.0004) + (-0.0217) + (-0.0500)$	-0.1252	0.8823
FPV - Sedang	$1.6937 + (-0.5936 * 3) + (-0.0024) + (-0.0212) + (-0.1248)$	-0.2355	0.7902

UTI - Parah	$-0.1715 + (0.2402 * 3) + (-0.1498) + (-0.1889) + (-0.1462)$	0.0642	1.0663
UTI - Ringan	$-1.0186 + (0.3391 * 3) + (-0.1116) + (1.7797) + (-0.1270)$	1.5398	4.6637
UTI - Sedang	$-0.2333 + (0.2483 * 3) + (-0.0946) + (-0.1881) + (1.8017)$	2.0306	7.6187
CKD - Parah	$-4.5122 + (0.7603 * 3) + (1.0536) + (-0.2096) + (-0.0985)$	-1.4858	0.2263
CKD - Ringan	$-4.3686 + (0.7129 * 3) + (-0.2125) + (-0.1687) + (-0.0842)$	-2.6953	0.0675
CKD - Sedang	$-4.3837 + (0.7001 * 3) + (-0.2702) + (-0.1517) + (-0.1501)$	-2.8554	0.0575
Ringworm - Parah	$1.7658 + (-0.3875 * 3) + (-0.0025) + (-0.0480) + (-0.0821)$	0.4707	1.6011
Ringworm - Ringan	$1.8213 + (-0.4185 * 3) + (-0.0020) + (-0.0448) + (-0.0799)$	0.4391	1.5513
Ringworm - Sedang	$1.5862 + (-0.3359 * 3) + (-0.0033) + (-0.0502) + (-0.0806)$	0.4444	1.5596
Kudis - Parah	$0.0232 + (0.1789 * 3) + (-0.0688) + (-0.1575) + (-0.1312)$	0.2024	1.2243
Kudis - Ringan	$0.1027 + (0.1257 * 3) + (-0.0478) + (-0.1264) + (-0.1116)$	0.1940	1.2141
Kudis - Sedang	$-0.0738 + (0.1538 * 3) + (-0.0524) + (-0.1281) + (-0.1100)$	0.0971	1.1020
Total			30.2936

Dari rincian pada tabel di atas, dapat kita lihat bahwa hasil penjumlahan semua nilai eksponensial adalah $30,2936 \approx 30.29$. Nilai D inilah yang akan digunakan sebagai penyebut untuk menghitung probabilitas akhir setiap penyakit.

4. Hitung Probabilitas Akhir

$$\text{Probabilitas (UTI – Sedang)} = \frac{e^{\text{logit}}}{D} = \frac{7.619}{30.29} = 0.2515 \text{ atau } 25.15\%$$

$$\text{Probabilitas (UTI – Ringan)} = \frac{e^{\text{logit}}}{D} = \frac{4.663}{30.29} = 0.1539 \text{ atau } 15.39\%$$

$$\text{Probabilitas (FCV – Sedang)} = \frac{e^{\text{logit}}}{D} = \frac{2.378}{30.29} = 0.0785 \text{ atau } 7.85\%$$

Hasil perhitungan manual menghasilkan tiga skor logit yang berbeda untuk tiga prediksi teratas, yaitu 2.0306 untuk ‘UTI – Sedang’, 1.5398 untuk ‘UTI – Ringan’, dan 0.8664 untuk ‘FCV – Sedang’. Skor-skor inilah yang kemudian diolah oleh fungsi *softmax* (bersama skor logit dari 15 penyakit lainnya) untuk menghasilkan probabilitas akhir yang ditampilkan oleh sistem. Urutan skor logit yang dihasilkan dari perhitungan manual (UTI Sedang > UTI Ringan > FCV Sedang) terbukti sesuai dengan urutan probabilitas yang ditampilkan oleh sistem (25.15% > 15.39% > 7.85%). Hal ini membuktikan bahwa sistem prediksi telah mengimplementasikan logika model secara akurat.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah diuraikan pada bab-bab sebelumnya mengenai sistem prediksi tingkat keparahan penyakit kucing menggunakan metode regresi logistik multinomial, maka dapat ditarik kesimpulan sebagai berikut:

1. Penelitian ini telah berhasil membangun sebuah model klasifikasi fungsional menggunakan metode Regresi Logistik Multinomial yang mampu memprediksi tingkat keparahan penyakit kucing (misalnya: Ringan, Sedang, Parah) berdasarkan sekumpulan input gejala klinis.
2. Model yang dihasilkan menunjukkan performa yang andal dalam proses evaluasi. Pada pengujian dengan 101 data uji, model mencapai tingkat akurasi keseluruhan sebesar 92%. Analisis metrik F1-Score, presisi, dan *recall* juga menunjukkan kemampuan model yang kuat dalam mengidentifikasi sebagian besar kelas, meskipun ditemukan tantangan dalam membedakan tingkat keparahan yang gejalanya tumpang tindih.
3. Model telah berhasil diimplementasikan ke dalam sebuah purwarupa sistem berbasis web yang fungsional dengan dua fitur utama:
 - a. Fitur Analisis Tren, yang mampu memproses sekumpulan data (dalam format .csv atau .xlsx) untuk mengidentifikasi pola dan menyajikan laporan berupa tabel tren terbanyak dan visualisasi grafik distribusi, yang secara efektif menjawab tujuan penelitian untuk menganalisis tren.

- b. Fitur Prediksi Manual, yang mampu menerima input gejala untuk satu kasus dan menyajikan hasil prediksi tingkat keparahan secara individual.

5.2 Saran

Berdasarkan hasil dan keterbatasan dari penelitian yang telah dilakukan, terdapat beberapa saran yang dapat dijadikan acuan untuk pengembangan sistem dan penelitian selanjutnya:

1. Menyempurnakan *dataset* dengan menambah volume dan variasi data, serta mengeksplorasi fitur-fitur tambahan di luar gejala klinis, seperti data demografis (ras, berat badan), riwayat vaksinasi, atau hasil tes laboratorium untuk meningkatkan akurasi dan keandalan model.
2. Melakukan studi komparatif dengan menggunakan algoritma *machine learning* lain seperti *Support Vector Machine* (SVM), *Random Forest*, atau model *deep learning*. Hal ini bertujuan untuk menemukan metode yang paling optimal dan berpotensi memberikan performa yang lebih tinggi untuk dataset spesifik ini.
3. Menambahkan fitur yang lebih interaktif pada halaman hasil analisis tren, perbandingan tren antar periode waktu, kemampuan untuk mengeksport laporan analisis ke dalam format PDF, dan mengembangkan versi *mobile* (Android/iOS) dari sistem untuk meningkatkan kemudahan akses bagi pemilik hewan dan praktisi di lapangan.

DAFTAR PUSTAKA

- Chishti, S., Jaggi, K. R., Saini, A., Agarwal, G., & Ranjan, A. (2020). Artificial Intelligence-Based Differential Diagnosis: Development and Validation of a Probabilistic Model to Address Lack of Large-Scale Clinical Datasets. *Journal of Medical Internet Research*, 22(4). <https://doi.org/10.2196/17550>
- Dembrower, K., Crippa, A., Colón, E., Eklund, M., & Strand, F. (2023). Artificial intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study. *The Lancet Digital Health*, 5(10), e703–e711. [https://doi.org/10.1016/S2589-7500\(23\)00153-X](https://doi.org/10.1016/S2589-7500(23)00153-X)
- Fatimah Indrianti, N., Kania Ningsih, A., & Ilyas, R. (2024). IMPLEMENTASI DATA MINING UNTUK KLASIFIKASI PENYAKIT GAGAL GINJAL KRONIS MENGGUNAKAN METODE K-NEAREST NEIGHBOR. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 8, Issue 2).
- Gultom, Z. A., Hamidy Hazidar, A., Basri, M., Berkendara, P., Logistik, R., Pelanggaran, ;, & Lintas, L. (2024). *Factors Influencing Safety Consciousness and Violations Among Licensed Drivers Kata kunci* (Vol. 7, Issue 4). <http://Jiip.stkipyapisdompu.ac.id>
- Gunawan, N. O., & Nasib, S. K. (2025). Analisis Regresi Logistik Multinomial untuk Menentukan Faktor-Faktor yang Mempengaruhi Jenis Penyakit pada Mahasiswa (Studi Kasus: Mahasiswa Program Studi Statistika, Jurusan Matematika, Universitas Negeri Gorontalo). In *Research Review Jurnal Ilmiah Multidisiplin* (Vol. 4, Issue 1).

- Gusti Agung Widiani Wira Brata, I., & Iwan Wahyudin, M. (2025). PENGEMBANGAN MODEL KLASIFIKASI PENYAKIT KUCING MENGGUNAKAN K-NEAREST NEIGHBOR UNTUK IDENTIFIKASI PENYAKIT MENULAR/TIDAK MENULAR. *Journal of Computer Science and Information Technology*.
- Hermawan, I. P., Darantika, G., Tage, R. A., Desiandura, K., & Wardhani, H. C. P. (2023). Studi Kasus : Kesembuhan Kasus Feline Panleukopenia pada Kucing Mocca di Surabaya. *JURNAL KAJIAN VETERINER*, 11(1), 10–18. <https://doi.org/10.35508/jkv.v11i1.7627>
- Ibnas, R., & Agung Kurniawan, B. (n.d.). Klasifikasi Penderita Anemia Menggunakan Metode Regresi Logistik. In *Jurnal Matematika dan Statistika serta Aplikasinya* (Vol. 11, Issue 2).
- Ilmu dan Kesehatan Hewan, J., Ari Andhita Badraresta Arnaya, K., Kayati Widyastuti, S., & Gede Soma, I. (2024). *VETERINARY SCIENCE AND MEDICINE JOURNAL TREATMENT OF PARASITE INFECTION IN DOMESTIC CAT Penanganan Infeksi parasit pada kucing lokal*. <https://doi.org/10.24843/vsmj.2024.v6.i02.p09>
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). *Machine learning and deep learning*. <https://doi.org/10.1007/s12525-021-00475-2/Published>
- Putri, D. A., Anthara, M. S., & Batan, I. W. (2023). Laporan Kasus: Infeksi Saluran Pernapasan Atas pada Kucing Kampung. *Indonesia Medicus Veterinus*, 12(3), 451–461. <https://doi.org/10.19087/imv.2023.12.3.451>
- Putri, F. C., Andriyati, A., & Rohaeti, E. (2022). Analisis Regresi Logistik Multinomial pada Faktor-Faktor yang Mempengaruhi Status Pasien Covid-19

- di Kota Depok. *Jurnal Matematika Integratif*, 18(2), 103.
<https://doi.org/10.24198/jmi.v18.n2.40058.103-114>
- Sari, A., & Putri, D. U. (2023). PENERAPAN REGRESI LINEAR BERGANDA UNTUK MEMPREDIKSI DIABETES SECARA DINI. *Jurnal Teknologi Komputer Dan Sistem Informasi* Februari, 2023(1), 32–39.
<http://jurnal.goretanpena.com/index.php/teknisi>
- Sholeh, M., Kumalasari Nurnawati, E., & Lestari, U. (2023). Penerapan Data Mining dengan Metode Regresi Linear untuk Memprediksi Data Nilai Hasil Ujian Menggunakan RapidMiner. In *Jurnal Informatika Sunan Kalijaga* (Vol. 8, Issue 1). <https://archive.ics.uci.edu/ml/datasheets.php>.
- Sukma, I., Petrus, M., Catur Sakti Kendari, S., & Abdullah, J. (2020). *SISTEM PAKAR PENYAKIT KUCING MENGGUNAKAN METODE FORWARD CHAINING BERBASIS WEB*. 5(1), 327275–327276.
- Syahrani, R., & Zaman, S. (2023). Regresi Logistik Multinomial untuk Prediksi Kategori Kelulusan Mahasiswa. In *Jurnal Informatika Sunan Kalijaga* (Vol. 8, Issue 2). MEI.
- Toyibah, Z. B., Putri, Y. N., Puandini, P., Widodo, Z. M., & Ni'mah, A. T. (2024). Perbandingan Kinerja Algoritma Multinomial Naïve Bayes dan Logistic Regression pada Analisis Sentimen Movie Ratings IMDB. *Jurnal Ilmiah Edutic : Pendidikan Dan Informatika*, 10(2), 181–189.
<https://doi.org/10.21107/edutic.v10i2.28150>
- Weng, Y., Wu, J., Kelly, T., & Johnson, W. (2024). *Comprehensive Overview of Artificial Intelligence Applications in Modern Industries*.

Widayani, W., Solikhah, I., & Syafrianto, A. (2023). Sistem Pakar Diagnosis Penyakit Pada Kucing Dengan Metode Certainty Factor. *Jurnal Informatika Komputer, Bisnis Dan Manajemen*, 20(2).
<https://doi.org/10.61805/fahma.v20i2.33>

1. Pelatihan model

```
# train_model.py

import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
import joblib
import seaborn as sns
import matplotlib.pyplot as plt
import re

print("Memulai proses training dan evaluasi model...")

# 1. Membaca Data
try:
    df_main = pd.read_csv('data-klinis-kucing-500fix.csv')
    print("File CSV berhasil dibaca.")
except FileNotFoundError:
    print("Error: File 'data-klinis-kucing-500fix.csv' tidak ditemukan.")
    exit()

# 2. DEFINISI GENERALISASI GEJALA
# Kamus untuk memetakan gejala spesifik ke kategori umum
symptom_generalization_map = {

    'Demam': 'Demam (Umum)', 'Demam ringan': 'Demam (Umum)', 'Demam tinggi': 'Demam (Umum)',
    # Kategori Lesu/Lemah
}

# 4. Pembagian Data, Training, dan Evaluasi
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42, stratify=y)
print(f>Data dibagi: {len(X_train)} training, {len(X_test)} testing.")

model = LogisticRegression(multi_class='multinomial', solver='lbfgs', C=0.5, max_iter=2000)
model.fit(X_train, y_train)
print("Model berhasil dilatih.")

print("\n--- HASIL EVALUASI MODEL ---")
y_pred = model.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
print(f"\n1. Akurasi Model: {accuracy:.2%}")
print("\n2. Laporan Klasifikasi:")
print(classification_report(y_test, y_pred, zero_division=0))

# 5. Membuat Confusion Matrix
try:
    plt.figure(figsize=(22, 18))
    cm = confusion_matrix(y_test, y_pred, labels=model.classes_)
    sns.heatmap(cm, annot=True, fmt='d', cmap='Blues', xticklabels=model.classes_, yticklabels=model.classes_)
    plt.title('Confusion Matrix', fontsize=20)
    plt.ylabel('Kelas Aktual', fontsize=16)
    plt.xlabel('Kelas Prediksi', fontsize=16)
    plt.tight_layout()
    plt.savefig('confusion_matrix.png')
    print("\n3. Gambar 'confusion_matrix.png' berhasil disimpan.")
except Exception as e:
    print(f"\nGagal membuat confusion matrix: {e}")
```

Gambar A. 1 Kode Pelatihan dan Evaluasi Model

2. Hasil Nilai Koefisien dan Intercept

```
# inspect_model.py
import joblib
import sys

print("Memuat model yang sudah dilatih...")
try:
    model = joblib.load('model.joblib')
    model_columns = joblib.load('model_columns.joblib')
except FileNotFoundError:
    print("Model tidak ditemukan! Pastikan Anda sudah menjalankan 'train_model.py' terlebih dahulu.")
    sys.exit()

# Daftar semua kategori penyakit yang diprediksi
all_classes = model.classes_

# Koefisien (beta) dan intercept untuk setiap kelas
all_coefficients = model.coef_
all_intercepts = model.intercept_

# Nama file untuk output
output_filename = 'model_coefficients.txt'

# Menggunakan 'with open' untuk menulis ke file
# 'w' berarti write (menimpa file jika sudah ada)
# encoding='utf-8' untuk memastikan semua karakter tertulis dengan benar
with open(output_filename, 'w', encoding='utf-8') as f:
    print("--- INILAH KOEFISIEN ( $\beta$ ) DAN INTERCEPT ANDA ---", file=f)
    print("\nFile ini berisi semua nilai yang dibutuhkan untuk perhitungan manual.", file=f)
```

Gambar A. 2 Kode Koefisien dan Intercept

```
=====
Kategori Penyakit: Feline Calicivirus - Sedang
=====

Intercept ( $\beta_0$ ): 1.7304

Koefisien ( $\beta$ ) untuk setiap Gejala/Fitur:
- Usia Kucing Tahun: -0.2128
- Anoreksia (tidak mau makan sama sekali): -0.0177
- Banyak kencing (poliuria): -0.0247
- Banyak minum (polidipsia): -0.0121
- Bau mulut (halitosis): -0.0053
- Beberapa bercak kebotakan melingkar: -0.1065
- Bersin: 0.9154
- Bersin-bersin: 1.3614
- Bulu kusam: -0.0067
- Bulu rapuh di area tertentu: -0.2193
- Dehidrasi: -0.2914
- Dehidrasi berat: -0.0998
- Dehidrasi parah: -0.0020
- Dehidrasi ringan: -0.0088
- Demam: 1.0407
- Demam ringan: -0.3294
- Demam tinggi: -0.5046
- Diare (tanpa darah): -0.1864
- Diare berdarah: -0.1117
```

```
=====
Kategori Penyakit: Feline Calicivirus - Ringan
=====
```

```
Intercept ( $\beta_0$ ): 0.9386
```

```
Koefisien ( $\beta$ ) untuk setiap Gejala/Fitur:
```

- Usia Kucing Tahun: -0.1412
- Anoreksia (tidak mau makan sama sekali): -0.0157
- Banyak kencing (poliuria): -0.0210
- Banyak minum (polidipsia): -0.0105
- Bau mulut (halitosis): -0.0056
- Beberapa bercak kebotakan melingkar: -0.0690
- Bersin: 1.6551
- Bersin-bersin: -0.3054
- Bulu kusam: -0.0070
- Bulu rapuh di area tertentu: -0.1311
- Dehidrasi: -0.1566
- Dehidrasi berat: -0.0633
- Dehidrasi parah: -0.0022
- Dehidrasi ringan: -0.0093
- Demam: -0.3199
- Demam ringan: -0.0691
- Demam tinggi: -0.1136
- Diare (tanpa darah): -0.0714

```
=====
Kategori Penyakit: Feline Calicivirus - Parah
=====
```

```
Intercept ( $\beta_0$ ): 1.1391
```

```
Koefisien ( $\beta$ ) untuk setiap Gejala/Fitur:
```

- Usia Kucing Tahun: -0.1251
- Anoreksia (tidak mau makan sama sekali): -0.0218
- Banyak kencing (poliuria): -0.0294
- Banyak minum (polidipsia): -0.0148
- Bau mulut (halitosis): -0.0123
- Beberapa bercak kebotakan melingkar: -0.0836
- Bersin: -0.2358
- Bersin-bersin: -0.0996
- Bulu kusam: -0.0133
- Bulu rapuh di area tertentu: -0.1553
- Dehidrasi: -0.3403
- Dehidrasi berat: -0.0629
- Dehidrasi parah: -0.0025
- Dehidrasi ringan: -0.0165
- Demam: 0.5678
- Demam ringan: -0.1714
- Demam tinggi: 0.9919
- Diare (tanpa darah): -0.2788
- Diare berdarah: -0.0763
- Gatal (tidak selalu): -0.1397
- Gatal hebat: -0.0585

```
=====
Kategori Penyakit: Infeksi Saluran Kemih (UTI) - Ringan
=====
```

```
Intercept ( $\beta_0$ ): -1.0186
```

```
Koefisien ( $\beta$ ) untuk setiap Gejala/Fitur:
```

- Usia Kucing Tahun: 0.3391
- Anoreksia (tidak mau makan sama sekali): -0.1558
- Banyak kencing (poliuria): -0.1790
- Banyak minum (polidipsia): -0.1028
- Bau mulut (halitosis): -0.0806
- Beberapa bercak kebotakan melingkar: -0.0411
- Bersin: -0.1499
- Bersin-bersin: -0.0668
- Bulu kusam: -0.0843
- Bulu rapuh di area tertentu: -0.0509
- Dehidrasi: -0.1716
- Dehidrasi berat: -0.0173
- Dehidrasi parah: -0.0378
- Dehidrasi ringan: -0.0899
- Demam: -0.1294
- Demam ringan: -0.0249
- Demam tinggi: -0.0503
- Diare (tanpa darah): -0.0174
- Diare berdarah: -0.0223
- Gatal (tidak selalu): -0.0536

Gambar A. 3 Hasil Koefisien dan Intercept

```
--- PERHITUNGAN DIMULAI ---
Data Input:
- Usia Kucing: 3 Tahun
- Gejala: ['Sangat lesu/lemah', 'Sering ke kotak pasir', 'Sulit/mengejan saat kencing (stranguria)']
-----

Nilai Denominator (D): 30.293556

Tabel Hasil Perhitungan Lengkap:
Kategori Penyakit      Nilai Z (Logit)  Nilai e^Z  Probabilitas (P)
0  Infeksi Saluran Kemih (UTI) - Sedang      2.0306      7.618656      0.251494
1  Infeksi Saluran Kemih (UTI) - Ringan      1.5398      4.663657      0.153949
2      Feline Calicivirus - Sedang      0.8664      2.378333      0.078510
3      Feline Calicivirus - Parah      0.5045      1.656157      0.054670
4      Ringworm - Parah      0.4707      1.601115      0.052853
5      Ringworm - Sedang      0.4444      1.559554      0.051481
6      Ringworm - Ringan      0.4391      1.551310      0.051209
7      Feline Calicivirus - Ringan      0.3565      1.428322      0.047149
8      Scabies (Kudis) - Parah      0.2024      1.224338      0.040416
9      Scabies (Kudis) - Ringan      0.1940      1.214096      0.040078
10     Feline Panleukopenia - Parah      0.1872      1.205868      0.039806
11     Scabies (Kudis) - Sedang      0.0971      1.101971      0.036376
12     Infeksi Saluran Kemih (UTI) - Parah      0.0642      1.066306      0.035199
13     Feline Panleukopenia - Ringan     -0.1252      0.882320      0.029126
14     Feline Panleukopenia - Sedang     -0.2355      0.790176      0.026084
15     Penyakit Ginjal Kronis (CKD) - Parah -1.4858      0.226321      0.007471
16     Penyakit Ginjal Kronis (CKD) - Ringan -2.6953      0.067522      0.002229
17     Penyakit Ginjal Kronis (CKD) - Sedang -2.8554      0.057533      0.001899
-----

--- KESIMPULAN PREDIKSI ---
Berdasarkan input baru, prediksi utamanya adalah:
-> Infeksi Saluran Kemih (UTI) - Sedang
    Dengan probabilitas sebesar 25.15%
```

B. Lampiran Wawancara

Wawancara 1

Narasumber : Drh. Dian Delima

Tanggal : 21 Juli 2025

Tempat : Fafa Pet Clinic dan Petshop, Medan

Cintami : "Permisi, Dok. Selamat pagi. Maaf ganggu lagi, hehe."

Drh. Dian : "Pagi, masuk aja. Gimana skripsinya, udah sidang?"

Cintami : "Belum, Dok. Judul saya diubah sama dosen, jadi metodenya ganti."

Drh. Dian : "Loh, ganti gimana? Jadi nggak bikin sistem pakar lagi?"

Cintami : "Bukan, Dok. Saya diminta menggunakan data mining. Jadi, bukan meminta dokter yang mendiagnosa lagi, tapi butuh contoh kasusnya langsung."

Drh. Dian : "Oalaah... Jadi sekarang butuh data pasien gitu maksudnya?"

Cintami : "Betul, Dok. Saya butuh data riwayat pasien: umurnya berapa, akhirnya didiagnosis apa, terus gejalanya apa aja. Makin banyak contoh kasusnya, makin bagus."

Drh. Dian : "Boleh-boleh."

Cintami : "Makasih Dok!"

Drh. Dian : "Sama-sama."



Wawancara 2

Narasumber : Louisa Christine Hartanto, S.I.Kom., M.Si.

Metode : Percakapan via Aplikasi Pesan Instan

Tanggal : 14 Juli 2025

Cintami : "Selamat siang, Ibu Christine. Saya Cintami Arifina, mahasiswa tingkat akhir. Saat ini saya sedang menyusun skripsi dengan topik prediksi tingkat keparahan penyakit kucing menggunakan metode data mining. Mohon izin bertanya, Bu, apakah saya boleh menggunakan data dari klinik yang Ibu kelola untuk keperluan penelitian saya?"

Ibu Christine : "Halo, Fina. Tentu saja. Wah, topiknya menarik ya, menggunakan data mining untuk analisis klinis. Data spesifik seperti apa yang kamu butuhkan?"

Cintami : "Terima kasih banyak, Bu. Saya membutuhkan data anonim mengenai jenis penyakit kucing, gejala-gejala yang muncul, serta usia dan jenis kelaminnya jika ada. Tentu semua informasi pribadi seperti nama pemilik dan hewan akan dihilangkan."

Ibu Christine : "Begini, Fina, semua data di klinik kami terintegrasi dalam sistem manajemen bernama Kreloses. Data tersebut tercatat detail per pasien, jadi kami tidak memiliki ringkasan umumnya. Tim dokter di sini harus membuka rekam medisnya satu per satu untuk mengumpulkan data yang kamu perlukan."

Cintami : "Wah, kalau begitu, apakah tidak terlalu merepotkan tim Dokter di sana, Bu?"

Ibu Christine : "Memang perlu waktu, tapi untuk penelitian tidak apa-apa, kita usahakan. Kira-kira kamu butuh data dari periode kapan?"

Cintami : "Jika memungkinkan, Bu, data dari tahun 2023-2024. Namun jika terlalu banyak, data dari Januari hingga Juni 2025 saja sudah sangat membantu."

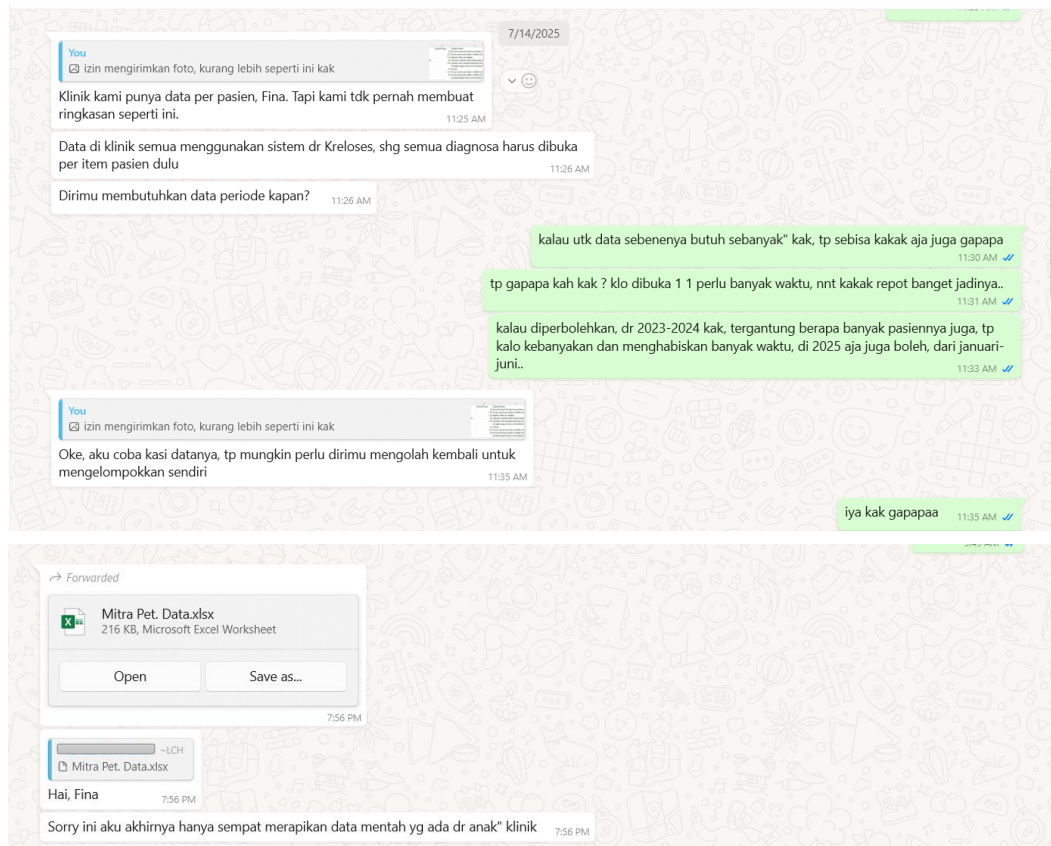
Ibu Christine : "Oke, saya coba koordinasikan dengan tim dokter di klinik, ya. Nanti data mentahnya akan kami siapkan, tapi kamu perlu mengolah dan mengelompokkannya lagi sendiri."

Cintami : "Tentu, Bu, tidak masalah sama sekali. Itu sudah sangat luar biasa. Untuk memudahkan, apa perlu saya bantu buat template spreadsheet-nya, Bu?"

Ibu Christine : "Boleh, itu ide yang bagus. Tolong kirimkan saja templatnya ke email saya. Mungkin datanya baru bisa siap minggu depan ya, bagaimana?"

Cintami : "Siap, Bu, tidak apa-apa. Saya tunggu kabarnya. Sekali lagi, terima kasih banyak atas kesempatan dan bantuannya, Bu."

Ibu Christine : "Sama-sama, Fina. Semoga sukses dengan skripsinya."



C. Lampiran Berkas

1. SK Pembimbing



**PENETAPAN DOSEN PEMBIMBING
PROPOSAL/SKRIPSI MAHASISWA
NOMOR : 576/IL3-AU/UMSU-09/F/2025**

Assalamu'alaikum Warahmatullahi Wabarakatuh

Dekan Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Muhammadiyah Sumatera Utara, berdasarkan Persetujuan permohonan judul penelitian Proposal / Skripsi dari Ketua / Sekretaris.

Program Studi : Teknologi Informasi
Pada tanggal : 28 April 2025

Dengan ini menetapkan Dosen Pembimbing Proposal / Skripsi Mahasiswa.

Nama : Cintami Arifina Ramadhani
NPM : 2109020102
Semester : VIII (Delapan)
Program studi : Teknologi Informasi
Judul Proposal / Skripsi : Implementasi Sistem Pakar Diagnosa Penyakit Kucing Domestik Menggunakan Forward Chaining Dengan Output Tingkat Kepastian

Dosen Pembimbing : Zuli Agustina Gultom, M.Si

Dengan demikian di izinkan menulis Proposal / Skripsi dengan ketentuan

1. Penulisan berpedoman pada buku panduan penulisan Proposal / Skripsi Fakultas Ilmu Komputer dan Teknologi Informasi UMSU
2. Pelaksanaan Sidang Skripsi harus berjarak 3 bulan setelah dikeluarkannya Surat Penetapan Dosen Pembimbing Skripsi.
3. **Proyek Proposal / Skripsi** dinyatakan " **BATAL** " bila tidak selesai sebelum Masa Kadalua tanggal : **28 April 2026**
4. Revisi judul.....

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Ditetapkan di : Medan
Pada Tanggal : 30 Syawal 1446 H
28 April 2025M



DR. A. Khowarizmi, M.Kom.
NIDN: 0127099201

Cc. File



2. Berita Acara Skripsi



MAJELIS PENDIDIKAN TINGGI PENELITIAN & PENGEMBANGAN PIMPINAN PUSAT MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH SUMATERA UTARA
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI

UMSU Terakreditasi A Berdasarkan Keputusan Badan Akreditasi Nasional Perguruan Tinggi No. 89/SK/BAN-PT/Akred/PT/III/2019
 Pusat Administrasi: Jalan Mukhtar Basri No. 3 Medan 20238 Telp. (061) 6622400 - 66224567 Fax. (061) 6625474 - 6631003

<https://fkip.umsu.ac.id> fkip@umsu.ac.id [umsu.medan](https://www.facebook.com/umsu.medan) [umsu.medan](https://www.instagram.com/umsu.medan) [umsu.medan](https://www.youtube.com/channel/UCqj8K8K8K8K8K8K8K8K8K8K) [umsu.medan](https://www.linkedin.com/company/umsu.medan)

Berita Acara Pembimbingan Skripsi

Nama Mahasiswa : Cintami Arifina R. Program Studi : Teknologi Informasi
 NPM : 2109020102 Konsentrasi :
 Nama Dosen Pembimbing : Zuli Agustina Gultom, M.Si Judul Penelitian : Prediksi tren tingkat keparahan penyakit kencing berdasarkan gejala klinis menggunakan regresi logistik multinomial.

Item	Hasil Evaluasi	Tanggal	Paraf Dosen
26/07-25	Penyesuaian Judul berdasarkan dataset		
28/07-25	Prototipe dasar (menampilkan)		
30/07-25	Pembuatan model regresi logistik		
1/8-25	Analisis perhitungan confusion matrix		
2/8-25	Perbaikan tabel di bab 4		
4/8-25	Mengubah beberapa error di aplikasi		
5/8-25	Mencari bukti D dr perhitungan menu		
6/8-25	ACC bab 4 & 5		

Medan, 6 August 2025

Diketahui oleh :

Ketua Program Studi
Pembimbing

Sistem Informasi

Teknologi Informasi

Disetujui oleh :

Dosen

(.....)
Zuli Agustina Gultom, M.Si

3. Perubahan Judul Skripsi



PERUBAHAN TOPIK/JUDUL PENELITIAN

Nomor agenda

Tanggal persetujuan

Topik yang disetujui Program Studi

: Implementasi sistem pakar diagnosis penyakit kucing domestik menggunakan forward chaining dan certainty factor dgn output tingkat kepastian

Nama Dosen pembimbing

: Zuli Agustina Gutom, M.Si

Judul yang disetujui Dosen Pembimbing

: Prediksi tren tingkat keparahan penyakit kucing berdasarkan gejala klinis menggunakan regresi logistik multinomial

Medan Agustus 2025

Disahkan oleh

Ketua Program Studi
Teknologi Informasi

FATMA SARI HUTAGALUNG, M.KOM

Persetujuan

Dosen Pembimbing

ZULI AGUSTINA GUTOM, M.Si

