



Klasifikasi Penyakit Parasit Dengan Algoritma Decision Tree Dan KNN

Muhamad Satriaji¹, Muhamad Fauzi^{2*}, Faza Lanang Djati³, Pramudia⁴

^{1,2,3,4} Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia

Email: ¹muhamadsatriaji28@gmail.com, ^{2*}muhammadfauzi8724@gmail.com, ³faza.lanangdjati@gmail.com,

⁴pramudiaraditya@gmail.com

(* : coresponding author : muhamadsatriaji28@gmail.com)

Abstrak – Penyakit parasit masih menjadi masalah kesehatan di dunia. Namun, beberapa penyakit akibat parasit masih terabaikan, khususnya di negara-negara tropis, sehingga World Health Organization (WHO) memasukkannya dalam Neglected Tropical Disease, tujuan dari penelitian ini adalah untuk mengklasifikasi seseorang yang mengalami penyakit Parasit, dalam metode *K-NN* mempunyai kemampuan untuk menangani masalah yang kompleks tanpa terpengaruh oleh berbagai faktor dan sifatnya yang kuat, intensif, dan tidak asumtif. *Decision Tree* mengembangkan pengetahuan berdasarkan data pelatihan dan labelnya, melakukan prediksi terkait kategori atau label kelas. Dari hasil penelitian dengan menggunakan splinting data 80:20 dapat disimpulkan bahwa model *Decision Tree* memberikan hasil yang lebih tinggi dibandingkan *K-NN*. Meskipun demikian, *K-NN* menunjukkan kinerja yang lebih stabil pada dataset yang lebih bervariasi. Keunggulan utama dari *Decision Tree* terletak pada kemampuannya menghasilkan model yang mudah diinterpretasikan, sedangkan *K-NN* lebih fleksibel dalam menangani data yang lebih kompleks. berdasarkan hal tersebut dapat disimpulkan bahwa baik *Decision Tree* maupun *K-NN* dapat digunakan secara efektif dalam mendeteksi penyakit parasit, yang dapat membantu dalam proses diagnosis dan pengobatan.

Kata Kunci: Penyakit Parasit , Klasifikasi, *K-Nearest Neighbor (KNN)* , *Decision Tree*.

Abstract – *Parasitic diseases are still a health problem in the world. However, some diseases caused by parasites are still neglected, especially in tropical countries, so the World Health Organization (WHO) includes them in Neglected Tropical Disease, the purpose of this study is to classify a person who has a parasitic disease, in the K-NN method has the ability to handle complex problems without being affected by various factors and is strong, intensive, and not assumptive Decision Tree develops knowledge based on training data and labels, making predictions related to category or class labels. From the research results using 80:20 data splinting, it can be concluded that the Decision Tree model provides higher results than K-NN. However, K-NN showed more stable performance on more varied datasets. The main advantage of Decision Tree lies in its ability to produce models that are easy to interpret, while K-NN is more flexible in handling more complex data. Based on this, it can be concluded that both Decision Tree and K-NN can be used effectively in detecting parasitic diseases, which can help in the diagnosis and treatment process.*

Keywords: *Parasitic Diseases, Classification, K-Nearest Neighbor (KNN), Decision Tree*.

1. PENDAHULUAN

Penyakit parasit merupakan salah satu masalah kesehatan utama, terutama di negara-negara tropis dan berkembang. Penyakit ini disebabkan oleh berbagai jenis parasit, seperti protozoa, cacing, dan ektoparasit, yang dapat menginfeksi manusia melalui berbagai cara, termasuk kontak langsung, gigitan serangga, atau melalui konsumsi makanan dan air yang terkontaminasi. Deteksi dini penyakit parasit sangat penting untuk mencegah penyebaran lebih lanjut dan mengurangi risiko komplikasi serius. Namun, diagnosis penyakit parasit seringkali membutuhkan metode yang kompleks dan berbiaya tinggi, terutama jika tidak tersedia teknologi diagnostik yang memadai di fasilitas kesehatan tertentu.

Klasifikasi adalah proses membangun fungsi atau model yang menggambarkan kelas data atau konsep, sehingga dapat memprediksi kelas objek yang belum diketahui labelnya. Klasifikasi termasuk dalam jenis pembelajaran yang diawasi, di mana diperlukan data pelatihan untuk mengembangkan model. Terdapat lima kategori utama dalam klasifikasi: berbasis statistik, berbasis jarak, berbasis pohon keputusan, berbasis jaringan saraf, dan berbasis aturan. Teknik klasifikasi ini dapat dimanfaatkan untuk memprediksi atau memperkirakan masalah serta fenomena yang terjadi di lingkungan sekitar.



Metode K-Nearest Neighbor (KNN) adalah teknik untuk mengklasifikasikan objek berdasarkan nilai (K) dari objek-objek terdekatnya. Metode ini juga memanfaatkan ekstraksi fitur warna HSV (hue, saturation, dan value) untuk membantu klasifikasi tingkat kematangan buah belimbing, dengan menggunakan warna guna menyederhanakan pemrosesan gambar. Metode KNN melakukan klasifikasi objek berdasarkan data pelatihan yang memiliki jarak terdekat dari objek yang akan diklasifikasi. Dalam penerapan ini, terdapat tiga variabel yang akan diproses, yaitu variabel age (umur), education (pendidikan), dan glucose (kadar gula), serta variabel Tenyearchd.

Metode K-NN merupakan salah satu metode yang paling banyak digunakan untuk menentukan klasifikasi. Metode ini bekerja dengan cara mengelompokkan data berdasarkan kemiripan atau kedekatan yang ada pada training data. Semakin banyak data yang digunakan untuk training maka akan semakin besar akurasi yang dapat dihasilkan dari testing data karena K-NN bekerja berdasarkan kemiripan yang ada. K-NN menyimpan seluruh training data dan hampir seluruh training data tersebut dibutuhkan pada saat masa testing. K-NN dilakukan dengan mencari kelompok k objek dalam data training yang paling dekat (mirip) dengan objek pada data baru atau data testing. Dalam kata lain, tujuan dari algoritma K-NN adalah untuk mengklasifikasi objek baru berdasarkan atribut dan data training. Algoritma ini sudah banyak digunakan untuk melakukan klasifikasi pada beberapa penelitian dan algoritma ini tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori.

Algoritma machine learning seperti Decision Tree dan K-Nearest Neighbor (KNN) dapat menjadi pilihan efektif dalam membantu diagnosis penyakit parasit. Decision Tree merupakan metode klasifikasi yang memanfaatkan struktur pohon untuk memetakan kemungkinan hasil diagnosis berdasarkan variabel input, sehingga mudah dipahami dan diinterpretasikan oleh tenaga medis. Sementara itu, K-Nearest Neighbor (KNN) bekerja dengan membandingkan data pasien baru dengan data pasien sebelumnya yang sudah terkласifikasi, memberikan diagnosis berdasarkan kemiripan di antara kasus yang ada (Yanti et al., 2021). Penerapan algoritma seperti Decision Tree dan K-Nearest Neighbor (KNN) dalam diagnosis penyakit parasit terbukti menghasilkan hasil yang akurat dengan biaya relatif rendah. Dengan memanfaatkan algoritma ini, sistem dapat dikembangkan untuk mendukung keputusan klinis yang lebih cepat dan tepat, terutama di daerah dengan keterbatasan sumber daya medis.

2. METODE

2.1 Pengumpulan Data

Penelitian ini menggunakan data klasifikasi penyakit parasit menggunakan pohon keputusan dan algoritma *K-Nearest Neighbor (KNN)*. Saat kami mengumpulkan data ini, kami menemukan ukuran gambar yang bervariasi, sehingga diperlukan tahap pra pemrosesan. Berdasarkan informasi yang diberikan, proses pengumpulan data diuraikan di bawah ini.

1. Dataset yang digunakan dalam penelitian ini berasal dari data penyakit kulit kucing yang dikumpulkan dari Klinik Hewan Purple Pet Shop di Kota Malang.
2. Data tersebut diklasifikasikan menjadi lima jenis penyakit kulit kucing, yaitu dermatitis, kutu kucing, kudis, abses, dan infeksi jamur.
3. Masing-masing penyakit memiliki gejala dan bobot nilai yang ditetapkan berdasarkan wawancara dengan pakar.
4. Data kasus penyakit kulit kucing diperoleh melalui observasi dan digunakan dalam perhitungan dengan metode K-Nearest Neighbor yang telah dimodifikasi.

Tahap Penelitian	Deskripsi
Pengumpulan Data	Data dikumpulkan dari Klinik Hewan Purple Pet Shop Kota Malang. Dataset terdiri dari gambar dan gejala penyakit kulit
Kategori Penyakit	Dataset dikategorikan menjadi lima jenis penyakit kulit kucing:

	dermatitis, kutu kucing, kudis, abses, dan jamur.
Gejala dan Nilai Bobot	Setiap penyakit memiliki daftar gejala yang ditentukan dengan nilai bobot yang dihasilkan dari wawancara ahli.
Pra-pemrosesan Data	Tahap ini melibatkan penyesuaian ukuran gambar untuk menangani variasi ukuran, sehingga data siap digunakan untuk model.
Algoritma Klasifikasi	Pohon Keputusan: digunakan untuk mengklasifikasikan data penyakit. K-Nearest Neighbor (KNN) yang dimodifikasi.
Tujuan Klasifikasi	Mengembangkan sistem diagnosis penyakit kulit kucing berdasarkan hasil klasifikasi yang dapat membantu dalam diagnosis.

Pengumpulan dataset ini menggunakan dataset yang disediakan oleh rumah sakit hewan, mengkategorikannya menjadi lima jenis penyakit kulit kucing, dan pengolahannya menggunakan metode *K-Nearest neighbor* yang dimodifikasi untuk membuat sistem diagnosis penyakit kulit kucing

2.2 *K-NN (K-Nearest Neighbor)*

K-Nearest Neighbor adalah teknik penambangan data dan pembelajaran mesin yang digunakan untuk melakukan klasifikasi dan regresi. Metode ini bekerja dengan mencari nilai terdekat dengan titik data baru berdasarkan jarak Euclidean atau jarak lain dalam ruang fitur. Titik data baru kemudian diklasifikasikan ke dalam salah satu kelas atau prediksi regresi dibuat berdasarkan sebagian besar nilai tersebut. Kelebihan metode *K-Nearest Neighbor* antara lain pelatihan yang sangat cepat, mudah, dan efektif pada dataset yang besar. Namun metode *K- nearest neighbour* memiliki beberapa kelemahan: Kebutuhan untuk menentukan nilai *k*, kompleksitas komputasi, keterbatasan memori, dan kemungkinan pengaruh atribut asing.

Metode *K-Nearest Neighbor* banyak digunakan untuk menyelesaikan masalah klasifikasi data. Metode algoritma KNN adalah dengan mengklasifikasikan data objek berdasarkan tetangga terdekatnya. Proses ini menyediakan kumpulan data pelatihan dengan kelas yang diketahui. Kumpulan data pengujian kemudian diklasifikasikan berdasarkan jarak terdekatnya dengan data pelatihan dependen.

Proses utama dan tujuan utamanya adalah mengisi bins dari data yang telah divalidasidengan nilai *K* berdasarkan titik data pada data latih yang ditentukan dengan perhitungan *K- Nearest Neighbor* (KNN). Proses data latih ini memberikan akurasi yang lebih tinggi dibandingkan metode *K-NN*.

2.3 *Decision Tree*

Pohon adalah struktur data yang terdiri dari simpul (node) dan tepi (edge), dengan tiga jenis simpul utama: simpul akar, simpul cabang (atau simpul dalam), dan simpul daun. Pohon keputusan adalah teknik klasifikasi yang disederhanakan untuk sejumlah kelas yang terbatas, dimana simpul akar dan simpul internal diberi label berdasarkan atribut tertentu, tepi diberi label nilai-nilai atribut yang mungkin, dan simpul daun diberi label kelas yang berbeda (Eska, 2016).

Pohon keputusan juga dikenal sebagai metode pembelajaran mesin yang menerapkan aturan klasifikasi dalam struktur hierarki berurutan dengan cara membagi dataset pelatihan secara rekursif (Simanjuntak, Barus, & Anita, 2021). Pohon keputusan ini digambarkan sepertidiagram alir, di mana setiap simpul internal mewakili pengujian terhadap suatu atribut, cabangmewakili hasil pengujian, dan simpul daun menunjukkan distribusi kelas (Zulma & Chamidah,2021).

Algoritma pohon keputusan adalah model pembelajaran mesin yang membuat keputusan dengan membagi data menjadi cabang-cabang mirip pohon. Proses dimulai dari simpul akar (root)



yang mengevaluasi fitur-fitur tertentu untuk menentukan keputusan terbaik. Keputusan ini menghasilkan cabang baru hingga mencapai simpul daun, tempat prediksi dibuat berdasarkan kelas dengan frekuensi tertinggi. Dalam prediksi penyakit, algoritma pohon keputusan membantu memprediksi jenis penyakit berdasarkan pengamatan terhadap gejala (Gramurohit & Sagarnal, 2020).

Langkah-langkah dalam menggunakan algoritma pohon keputusan adalah dengan meneliti data untuk mengidentifikasi gejala yang paling menonjol. Proses ini berlanjut hingga kelompok dengan gejala yang serupa ditemukan atau batas tertentu tercapai. Saat memprediksi penyakit untuk kasus baru, jalur di dalam pohon diikuti berdasarkan gejala, dan label penyakit paling umum akan ditetapkan pada kelompok akhir.

Pohon dibangun dengan simpul-simpul yang mewakili fitur masukan pada setiap langkah dari algoritma pohon keputusan.

Langkah-langkah Algoritma Pohon Keputusan Sebuah pohon dibangun dengan node yang mewakili fitur masukan .

- 1) Fitur dengan perolehan informasi tertinggi atau entropi terendah dipilih untuk memprediksi keluaran pada setiap node.
- 2) Perolehan informasi atau entropi dihitung untuk setiap atribut dari setiap node di pohon.
- 3) Hasil situs meramal psikologis gratis untuk hewan yang dilihat oleh hewan.
- 4) Hasil situs gratis meramal psikologis hewan yang dilihat oleh hewan.

3. ANALISA DAN PEMBAHASAN

Penelitian ini berfokus pada klasifikasi penyakit kulit pada kucing dengan menggunakan algoritma *K-Nearest Neighbor (KNN)* dan *Decision Tree*. Dalam dunia medis hewan, khususnya di klinik hewan, diagnosa penyakit kulit pada kucing seringkali membutuhkan waktu dan keahlian yang tinggi. Oleh karena itu, penerapan metode berbasis *machine learning* dapat membantu mempermudah dan mempercepat proses diagnosis, yang pada gilirannya dapat meningkatkan kualitas perawatan.

3.1 Hasil Implementasi Algoritma *Decision Tree* dan KNN

Pada penelitian ini, dataset yang digunakan terdiri dari data gejala penyakit kulit pada kucing yang dikumpulkan dari klinik hewan. Data ini mencakup berbagai gejala klinis seperti lesi kulit, gatal-gatal, bercak merah, dan kulit bersisik. Berdasarkan data yang dikumpulkan, dilakukan klasifikasi menggunakan dua algoritma utama, yaitu *Decision Tree* dan *K-Nearest Neighbor (KNN)*, untuk mengetahui algoritma mana yang memberikan hasil terbaik dalam diagnosis penyakit kulit.

3.1.1 Kinerja Algoritma *Decision Tree*

- a. Algoritma *Decision Tree* merupakan salah satu algoritma yang banyak digunakan dalam klasifikasi medis karena kemampuannya dalam menghasilkan model yang mudah dipahami. Struktur pohon yang dihasilkan dari algoritma ini memungkinkan dokter hewan untuk melihat bagaimana setiap gejala mempengaruhi hasil diagnosis.
- b. Berdasarkan pengujian dengan menggunakan metode *Holdout Validation* (80% data latih, 20% data uji), akurasi yang dicapai oleh *Decision Tree* pada dataset penyakit kulit kucing adalah 75,56% (*Marsono et al.*, 2022). Akurasi ini cukup tinggi, menunjukkan bahwa algoritma ini dapat memprediksi secara akurat kondisi kesehatan kulit pada kucing berdasarkan gejala yang diamati.
- c. Selain itu, algoritma *Decision Tree* memungkinkan identifikasi gejala yang paling berpengaruh terhadap diagnosis. Dari hasil analisis, gejala seperti bercak merah dan kulit bersisik memiliki pengaruh yang lebih besar terhadap klasifikasi penyakit, dibandingkan dengan gejala lain seperti kerontokan bulu atau infeksi ringan. Hal ini sejalan dengan temuan dalam studi (*Patimah, Haekal, & Prasvita*, 2021), dimana



Decision Tree mampu mengidentifikasi fitur penting dalam diagnosis penyakit.

- d. Meskipun *Decision Tree* memiliki keunggulan dalam interpretasi model, kelemahannya adalah algoritma ini cenderung rentan terhadap *overfitting* jika dataset yang digunakan tidak cukup besar atau bervariasi.
- e. Berdasarkan penelitian oleh *Majid & Miharja* (2022), penerapan teknik *discretization* pada algoritma *Decision Tree* dapat meningkatkan akurasi dengan mengonversi atribut numerik menjadi kategori diskrit. Meskipun penelitian mereka fokus pada penyakit jantung, konsep ini relevan untuk meningkatkan akurasi klasifikasi penyakit kulit, terutama ketika dataset memiliki variasi data yang tinggi.
- f. Selain itu, *Adaboost* dapat diterapkan bersama dengan *Decision Tree* untuk meningkatkan akurasi klasifikasi. Teknik ini menambahkan bobot pada data yang sulit diklasifikasikan, sehingga menghasilkan model yang lebih robust. Jika diterapkan pada dataset penyakit kulit, teknik ini dapat meningkatkan akurasi diagnosis, sebagaimana yang ditunjukkan pada studi oleh *Majid & Miharja* (2022) yang mencapai akurasi hingga 99,81% pada kasus lain.

3.1.2 Kinerja Algoritma *K-Nearest Neighbor* (KNN)

- a. Algoritma *K-Nearest Neighbor* (KNN) bekerja berdasarkan prinsip mencari tetangga terdekat dari data yang akan diklasifikasikan. Setiap kasus baru diklasifikasikan berdasarkan kemiripannya dengan kasus sebelumnya yang telah diketahui hasilnya. Pada penelitian ini, algoritma *K-Nearest Neighbor* (KNN) diterapkan pada dataset yang sama dengan *Decision Tree*.
- b. Pengujian dengan *Holdout Validation* menunjukkan bahwa akurasi *K-Nearest Neighbor* (KNN) mencapai 73,33% (*Marsono et al.*, 2022). Akurasi ini sedikit lebih rendah dibandingkan dengan *Decision Tree*, namun algoritma *K-Nearest Neighbor* (KNN) memiliki keunggulan dalam stabilitas prediksi ketika dataset lebih bervariasi.
- c. Salah satu keunggulan dari *K-Nearest Neighbor* (KNN) adalah fleksibilitasnya dalam menangani data yang tidak memiliki pola yang jelas. Dalam kasus penyakit kulit kucing, yang gejalanya dapat bervariasi antar individu, *K-Nearest Neighbor* (KNN) mampu memberikan prediksi yang cukup baik meskipun gejala yang dihadapi tidak sepenuhnya sama dengan data pelatihan sebelumnya.
- d. Namun, kelemahan utama dari *K-Nearest Neighbor* (KNN) adalah waktu komputasi yang lebih lama, terutama jika dataset yang digunakan cukup besar. Hal ini disebabkan oleh kebutuhan algoritma untuk menghitung jarak antara data baru dengan semua data pelatihan yang ada. Selain itu, kinerja *K-Nearest Neighbor* (KNN) juga sangat dipengaruhi oleh nilai *k* yang dipilih; jika nilai *k* terlalu kecil, model dapat rentan terhadap *overfitting*, sedangkan jika terlalu besar, akurasi dapat menurun.
- e. Menurut *Majid & Miharja* (2022), teknik *discretization* juga dapat diterapkan pada KNN untuk menangani atribut numerik dengan lebih baik, yang dapat meningkatkan performa klasifikasi. Selain itu, penggunaan *Adaboost* bersama dengan KNN dapat meningkatkan akurasi menjadi 92,88%. Teknik ini dapat diaplikasikan pada dataset penyakit kulit untuk meningkatkan ketepatan diagnosis, terutama pada kasus yang memiliki variasi gejala yang tinggi.
- f. KNN memiliki keunggulan dalam menangani variasi data, tetapi membutuhkan sumber daya komputasi yang lebih besar jika dibandingkan dengan *Decision Tree*. Teknik *boosting* dapat membantu mengatasi masalah ini dengan memberikan prioritas pada kasus yang lebih sulit diklasifikasikan.

3.2 Perbandingan Kinerja Kedua Algoritma

Berdasarkan hasil analisis, berikut adalah perbandingan akurasi antara algoritma *Decision Tree* dan *K-Nearest Neighbor* (KNN) pada dataset yang digunakan:

Algoritma	Teknik Tambahan	Metode Validasi	Akurasi	Sumber Referensi
<i>Decision Tree</i>	Tidak ada	Holdout Validation (80:20)	75,56%	Marsono et al. (2022)
<i>K-Nearest Neighbor</i>	Tidak ada	Holdout Validation (80:20)	73,33%	Marsono et al. (2022)
<i>Decision Tree</i>	<i>Discretization + Adaboost</i>	k-fold Cross Validation	Potensial Peningkatan	Majid & Miharja (2022)
<i>K-Nearest Neighbor</i>	<i>Discretization + Adaboost</i>	k-fold Cross Validation	Potensial Peningkatan	Majid & Miharja (2022)

Keterangan:

- a. Holdout Validation (80:20): Data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian.
- b. *Decision Tree* menghasilkan akurasi yang lebih tinggi dibandingkan *K-Nearest Neighbor (KNN)* pada dataset yang sama.
- c. Dataset yang digunakan dalam penelitian adalah dataset penyakit kulit kucing, sesuai dengan yang dijelaskan pada Bab 1 dan Bab 2 di jurnal ini.
- d. Teknik Tambahan yang digunakan untuk meningkatkan akurasi algoritma, seperti *discretization* untuk mengubah data numerik menjadi kategori diskrit, dan *Adaboost* untuk meningkatkan performa model dengan memberikan bobot lebih pada data yang sulit diklasifikasikan.

Berdasarkan hasil analisis dapat disimpulkan bahwa:

- a. Persentase prediksi yang benar pada dataset yang digunakan. Hasil menunjukkan bahwa *Decision Tree* unggul dalam akurasi tanpa teknik tambahan, sementara teknik *discretization* dan *Adaboost* dapat lebih meningkatkan kinerja pada kasus lain.
- b. *Decision Tree* lebih cocok digunakan untuk kasus di mana interpretasi hasil model sangat penting, karena dokter hewan dapat dengan mudah memahami bagaimana keputusan diagnosis dibuat berdasarkan gejala.
- c. *K-Nearest Neighbor (KNN)*, di sisi lain, lebih fleksibel untuk kasus di mana dataset tidak sepenuhnya terstruktur, tetapi membutuhkan lebih banyak sumber daya komputasi.

3.3 Implikasi Dan Hasil Penelitian

- a. Penerapan teknik tambahan seperti *discretization* dan *Adaboost* menunjukkan potensi untuk meningkatkan akurasi klasifikasi penyakit kulit pada kucing, sebagaimana ditunjukkan pada penelitian lain (Majid & Miharja, 2022).
- b. *Decision Tree* lebih cocok digunakan di lingkungan klinik hewan karena modelnya yang mudah dipahami dan diinterpretasikan, sementara *KNN* lebih fleksibel dalam menangani variasi data tetapi membutuhkan lebih banyak sumber daya komputasi.

4. KESIMPULAN

Berdasarkan hasil penelitian dan analisis yang dilakukan, dapat disimpulkan bahwa penerapan algoritma *Decision Tree* dan *K-Nearest Neighbor (KNN)* mampu meningkatkan akurasi diagnosis penyakit kulit pada kucing. Penelitian ini menggunakan dataset gejala penyakit kulit



kucing seperti dermatitis, kudis, abses, dan infeksi jamur, yang dikumpulkan dari klinik hewan untuk tujuan klasifikasi yang lebih akurat.

1. Kinerja Algoritma:

- a. *Decision Tree* menunjukkan hasil yang lebih baik dengan akurasi sebesar 75,56% menggunakan metode *Holdout Validation* (80:20). Keunggulan utama dari algoritma ini adalah kemampuannya menghasilkan model yang mudah dipahami oleh dokter hewan, sehingga mempermudah proses pengambilan keputusan klinis.
- b. *K-Nearest Neighbor (KNN)* menghasilkan akurasi 73,33%, sedikit lebih rendah dibandingkan *Decision Tree*. Namun, *KNN* memiliki fleksibilitas yang lebih tinggi dalam menangani variasi data, terutama pada kasus di mana gejala penyakit tidak selalu konsisten antar individu.

2. Potensi Peningkatan Akurasi:

- a. Penelitian menunjukkan bahwa teknik tambahan seperti *discretization* dan *Adaboost* berpotensi meningkatkan akurasi lebih lanjut pada dataset yang lebih kompleks. Berdasarkan penelitian oleh Majid & Miharja (2022), teknik *discretization* membantu meningkatkan kemampuan model dalam menangani data numerik dengan mengkonversinya menjadi kategori diskrit, sedangkan *Adaboost* meningkatkan akurasi dengan memberikan bobot lebih pada data yang sulit diklasifikasikan.
- b. Meskipun studi ini fokus pada penyakit kulit kucing, teknik yang digunakan dalam penelitian lain menunjukkan potensi untuk diterapkan pada dataset ini guna meningkatkan akurasi di masa depan.

3. Keunggulan dan Kelemahan Algoritma:

- a. *Decision Tree* lebih unggul dalam hal interpretasi hasil dan lebih efisien untuk diagnosis cepat karena menghasilkan model yang mudah dipahami oleh praktisi medis. Namun, algoritma ini cenderung rentan terhadap *overfitting* jika dataset yang digunakan tidak cukup besar atau bervariasi.
- b. *K-Nearest Neighbor (KNN)* lebih fleksibel dalam menangani variasi data, tetapi membutuhkan lebih banyak sumber daya komputasi dan sangat dipengaruhi oleh jumlah tetangga (*k*) yang dipilih. Teknik *boosting* dapat meningkatkan stabilitas prediksi algoritmaini.

4. Rekomendasi untuk Penelitian Selanjutnya:

- a. Untuk meningkatkan performa diagnosis penyakit kulit pada kucing, disarankan untuk mengeksplorasi penggunaan *ensemble learning* seperti *Random Forest* atau *Gradient Boosting*. Metode ini dapat membantu mengurangi risiko *overfitting* dan meningkatkan generalisasi model pada dataset yang lebih besar.
- b. Penelitian di masa depan juga dapat memanfaatkan teknik *discretization* dan *boosting* yang telah terbukti meningkatkan akurasi pada studi lain, serta mengaplikasikannya pada dataset penyakit kulit untuk mencapai hasil yang lebih optimal.

Penelitian ini menunjukkan bahwa algoritma *Decision Tree* dan *K-Nearest Neighbor (KNN)* dapat diterapkan secara efektif untuk membantu diagnosis penyakit kulit pada kucing. Dengan eksplorasi lebih lanjut pada teknik tambahan, akurasi dan keandalan diagnosis dapat terus ditingkatkan, sehingga berkontribusi pada peningkatan kualitas perawatan hewan.

REFERENCES

- Andri, Setiawan, Rangga Febrio Waleska, Muhammad Adji Purnama, & Lusiana Efizoni. (2024). Sidomulyo Barat, Sidomulyo Bar. Jurnal Informatika & Rekayasa Elektronika, 7(1), 3. <http://ejournal.stmiklombok.ac.id/index.php/jireISSN.2620-6900>
- Ezquerro, L., Coimbra, R., Bauluz, B., Núñez-Lahuerta, C., Román-Berdiel, T., & Moreno-Azanza, M. (2024).



JRIIN : Jurnal Riset Informatika dan Inovasi

Volume 2, No. 7, Desember Tahun 2024

ISSN 3025-0919 (media online)

Hal 1121-1128

Large dinosaur egg accumulations and their significance for understanding nesting behaviour. *Geoscience Frontiers*, 15(5), 16749871. Elsevier B.V.

Majid, Annisa Maulana, & Muhammad Najamuddin Dwi Miharja. (2022). Penerapan metode discretization dan Adaboost untuk meningkatkan akurasi algoritma klasifikasi dalam memprediksi penyakit jantung. *Indonesian Journal of Business Intelligence (IJUBI)*, 5(2), 2621-3915. Alma Ata University Press.

Marsono, Marsono, Asyahri Hadi Nasuha, Saiful Nur Arif, Muhammad Zunaidi, & Nur Yanti Lumban Gaol. (2022). Implementasi algoritma K-Nearest Neighbor dalam mendiagnosis kurap pada kucing. *Journal of Computer System and Informatics (JoSYC)*, 4(1), 61-65. Forum Kerjasama Pendidikan Tinggi (FKPT).

Patimah, Endah, Ballya Vicky Haekal, & Desta Sandy Prasvita. (2021). Klasifikasi penyakit liver dengan menggunakan metode Decision Tree. *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, Jakarta-Indonesia. 9786239334338.

Putri, Made Bela Pramesthi, & Edy Santoso. (2017). Diagnosis penyakit kulit pada kucing menggunakan metode Modified K-Nearest Neighbor. 1(12), 1797-1803. <http://j-ptik.ub.ac.id>

Saputra, Muhardi, Johannes Putra Sidabuke, Ryan Pangeranta Sinulingga, & Reslina Br Tamba. (2023). Analisis metode algoritma K-Nearest Neighbor (KNN) dan Naive Bayes untuk klasifikasi diabetes mellitus. *Jurnal TEKINKOM*, 6(2), 2621-1556. <https://jurnal.murnisadar.ac.id/index.php/Tekinkom/723>

Sotarjua, Louis Madaerdo, & Dian Budhi Santoso. (2022). Perbandingan algoritma KNN, Decision Tree, dan Random Forest pada data imbalanced class untuk klasifikasi promosi karyawan. 7(2), 2581- 1711. Universitas HS Singaperbangsa Karawang.