

BAB I

PENDAHULUAN

1.1 Latar Belakang

KDD adalah bidang penelitian multidisipliner yang berfokus pada upaya mengekstraksi pengetahuan tersembunyi yang sebelumnya tidak diketahui, namun memiliki potensi untuk digunakan, dari kumpulan data besar. Proses ini melibatkan pendekatan komputasi dan analitis untuk mengungkap wawasan baru dari data(Tsui et al., 2023). Data mining merupakan proses eksplorasi dan analisis data yang bertujuan untuk menemukan pola-pola yang bermanfaat, hubungan tersembunyi, dan informasi yang dapat diambil dari sekumpulan data besar. Proses ini melibatkan tahapan seperti preprocessing data, pemodelan, evaluasi, dan interpretasi hasil untuk membantu dalam pengambilan keputusan yang lebih baik (Hendrastuty, 2024). KDD adalah proses sistematis untuk menemukan pola bermakna dari data besar, dengan data mining sebagai langkah utamanya. Dalam hal ini, K-Means Clustering digunakan untuk mengelompokkan data berdasarkan kesamaan atribut, sehingga menghasilkan wawasan yang bermanfaat.

Penelitian terkait dengan penyakit ayam sebelumnya antara lain adalah study penyakit ayam pedaging di Kota Bogor (Wiedosari et al.,2021) berhasil mengidentifikasi sebanyak 6 penyakit pada ayam, dimana rata-rata di temui pada ayam berumur 11-21 hari (57,5%) dan sering terjadi pada musim penghujan (60%). Kemudian penelitian selanjutnya yaitu diagnosis penyakit pada ayam dengan menggunakan metode certainty factor antara lain dilakukan oleh Bere pada tahun 2021, menunjukan hasil pengujian yang di lakukan terhadap 11 orang responden, yaitu 10 user dan 1 orang pakar, didapatkan hasil 34% user sangat setuju, 44% user setuju, 18% user kurang setuju, dan 4% user tidak setuju. Pengujian yang ditujukan terhadap pakar didapatkan hasil presentase dari 2 penyakit dengan 10 gejala adalah 100% sedangkan di sitem pakar diagnosa penyakit ayam menunjukan presentase 67,744%

pada penyakit gumbro dengan jumlah 5 gejala dan presentase 69,706% pada penyakit mareks dengan jumlah 5 gejala (Bere et al.,2021) . Selanjutnya penelitian yang dilakukan oleh Pratama pada tahun 2024 menunjukan penerapan metode Certainty Factor dalam diagnosis penyakit pada ayam terbukti efektif dengan tingkat akurasi yang tinggi, yaitu sekitar 85%. Sistem pakar berbasis metode ini mampu memberikan diagnosis yang cepat dan akurat, serta memberikan tingkat kepastian yang jelas untuk setiap diagnosis yang dihasilkan (Yanto et al., 2021).

Penelitian lainnya yang menjadi rujukan tentang penyakit pada ayam dengan menggunakan metode forward chaining antara lain seperti yang dilakukan oleh Amriyansah pada tahun 2024 menghasilkan sistem pakar ini diuji dengan menggunakan 33 gejala dan 10 penyakit ayam yang umum terjadi. Hasil pengujian menunjukkan bahwa sistem pakar ini dapat mendiagnosis penyakit ayam dengan akurasi yang tinggi dan memberikan hasil diagnosis secara langsung melalui browser web. Sistem pakar ini juga dapat meningkatkan pemahaman pengguna tentang pembuatan situs web dan penyakit ayam (Amriyansah et al., 2024). Penelitian lainnya juga pernah dilakukan oleh Puadah Hasanah pada tahun 2024 menghasilkan kesimpulan bahwa sistem pakar memiliki 14 rules yang sesuai dengan rules atau aturan metode forward chaining, dilengkapi pengetahuan umum tentang suatu penyakit, saran untuk tindakan yang harus dilakukan setelah melakukan diagnosa pada sistem (Puadah Hasanah & Nanang Durahman, 2024). Selanjutnya penelitian yang menggabungkan antara metode forward chaining dan certainty factor pernah juga dilakukan oleh Irawan pada tahun 2024 menunjukan hasil pengujian blackbox menunjukkan bahwa semua fitur berfungsi dengan baik dan sesuai dengan perancangan. Dan pengujian menggunakan 3 browser 100% berjalan dengan lancar pada setiap browser. Sistem ini juga menunjukkan akurasi 92% dalam mendiagnosis penyakit pada ayam Bangkok. Dan pada pengujian user 100% menyatakan puas (Irawan et al.,2024) . Selain itu juga ada metode lain yang digunakan dalam penelitian mengenai penyakit ayam antara lain metode Convolutional Neural Network (CNN) Densenet 121 (Mintarsih et al., 2024), metode Dempster Shafer (Virga Aulia et al., 2024), dan gabungan antara metode VCIRS dan certainty factor (Andreswari et al., 2024)

Sedangkan penelitian yang menggunakan pendekatan K-Means Clustering antara lain penelitian yang dilakukan Aryanto pada 2024 bahwa hasil penelitian menunjukkan adanya tiga klaster yang signifikan. Cluster 0 mencakup 21 kota, Cluster 1 mencakup 4 kota, dan Cluster 2 mencakup 13 kota. Temuan ini memiliki implikasi

penting untuk perumusan kebijakan pembangunan yang lebih tepat sasaran di tingkat kota di Provinsi Jawa Timur. Hasil penelitian ini juga menyediakan landasan yang kuat bagi perumusan kebijakan pembangunan yang lebih efektif di Provinsi Jawa Timur, memberikan informasi yang diperlukan untuk pengelolaan populasi yang berkelanjutan (Aryanto et al., 2024). Penelitian selanjutnya yang dijadikan referensi dalam penulisan ini adalah penelitian yang dilakukan oleh Nurjannah pada tahun 2024 menyatakan bahwa 38 sampel masuk ke dalam Cluster 1 (C1), 36 sampel masuk ke dalam Cluster 2 (C2), dan 28 sampel masuk ke dalam Cluster 3 (C3). Evaluasi hasil clustering dilakukan dengan menggunakan Box Plot dan Scatter Plot. Box Plot menunjukkan distribusi data yang jelas untuk setiap cluster, memastikan bahwa pengelompokan data sesuai dengan evaluasi statistik. Cluster C1 adalah balita dengan pertumbuhan dan perkembangan yang normal. Klaster C2 menunjukkan tanda-tanda awal masalah pertumbuhan dan perkembangan. Klaster C3 menunjukkan masalah pertumbuhan dan perkembangan yang serius (Nurjannah et al., 2024).

Penelitian selanjutnya adalah penelitian yang dilakukan oleh Sintawati pada tahun 2024 menyatakan hasil pengelompokan K-Means menunjukkan bahwa Jakarta Pusat memiliki status prevalensi tertinggi untuk kasus severely stunting (10,50%) dan stunting (13,01%), sedangkan Jakarta Barat memiliki status prevalensi terendah untuk severely stunting (4,62%) dan stunting (10,22%). Hasil analisis menunjukkan bahwa DKI Jakarta dapat dikelompokkan ke dalam beberapa klaster berdasarkan prevalensi malnutrisi. Cluster dengan prevalensi gizi buruk tertinggi meliputi Jakarta Pusat, sedangkan cluster dengan prevalensi gizi buruk terendah meliputi Jakarta Barat dan Kepulauan Seribu. Penerapan K-Means dalam penelitian ini memberikan pendekatan yang efisien dalam mengidentifikasi kelompok wilayah yang membutuhkan perhatian lebih dalam memerangi malnutrisi pada anak (Sintawati & Mariskhana, 2024). Selanjutnya adalah penelitian oleh Suharyanti pada 2024 mengatakan pada penelitian ini, algoritma K-Means digunakan dengan jumlah cluster optimal sebanyak 3, dan menghasilkan tingkat akurasi sebesar 85%. Hasil penelitian juga menunjukkan bahwa wilayah seperti Jakarta Pusat dan Jakarta Selatan menerima lebih banyak bantuan sosial dibandingkan dengan wilayah lain seperti Jakarta Utara dan Jakarta Timur, yang mencerminkan perbedaan kebutuhan di berbagai wilayah (Suharyanti & Supendar, 2024). Selanjutnya adalah penelitian yang dilakukan oleh Sujjada pada 2024, mengatakan bahwa hasil analisis data menunjukkan terdapat tiga Cluster utama. Cluster pertama memiliki tingkat disabilitas yang tinggi dan mencakup 62 kota dan

kabupaten, Cluster kedua memiliki tingkat disabilitas yang sedang dengan 37 kota dan kabupaten, dan Cluster ketiga memiliki tingkat disabilitas yang rendah dengan 27 kota dan kabupaten. Evaluasi terbaik dengan menggunakan metode Indeks Davies Bouldin menghasilkan dua Klaster, yang menunjukkan kualitas pembagian Klaster yang lebih baik. Hasil penelitian ini memberikan pemahaman yang lebih baik mengenai sebaran disabilitas di Indonesia, yang dapat digunakan sebagai dasar untuk meningkatkan inklusi dan aksesibilitas bagi penyandang disabilitas (Sujjada et al., 2024).

Penelitian selanjutnya adalah penelitian yang dilakukan oleh Tuslaela pada 2024 menyatakan hasil penelitian dengan objek Ketahanan Pangan disimpulkan bahwa, hasil analisis penerapan perhitungan data mining secara manual pada Software Excel dengan menggunakan metode K-Means Clustering, menghasilkan dua jenis clustering berupa C0 yaitu kelompok Luas Lahan dan Produksi Tertinggi dengan jumlah 4 kabupaten: Kabupaten Jember, Kabupaten Ngawi, Kabupaten Bojonegoro dan Kabupaten Lamongan, untuk pengelompokan C1 yaitu kelompok Luas Lahan dan Produksi Terendah dengan 25 kabupaten di Provinsi Jawa Timur (Tuslaela et al., 2024). Penelitian lainnya yang dilakukan oleh Wardani pada 2024, yaitu penelitian ini hanya menggunakan tiga indikator kemiskinan yaitu produk domestik regional bruto, indeks pembangunan manusia, dan tingkat pengangguran. Jumlah cluster yang optimal ditentukan berdasarkan hasil koefisien silhouette. Metode penelitian dimulai dengan pengumpulan dataset, analisis data eksploratori, preprocessing data, dan pengelompokan k-means. Nilai $k = 6$ menghasilkan koefisien silhouette sebesar 0,4135. Penelitian ini menghasilkan enam cluster kabupaten/kota. Cluster 1 terdiri dari 11 kabupaten dan 1 kota, cluster 2 terdiri dari 1 kabupaten dan 2 kota, cluster 3 terdiri dari 4 kabupaten, cluster 4 terdiri dari 7 kabupaten, cluster 5 terdiri dari 4 kota, dan cluster 6 terdiri dari 2 kabupaten dan 1 kota. Variabel produk domestik regional bruto, indeks pembangunan manusia, dan tingkat pengangguran memiliki pengaruh yang besar terhadap hasil cluster (Wardani et al., 2024). Penelitian lainnya yang menjadi referensi dalam penulisan ini antara lain Korelasi Antara Profil Dan Nilai Akademis Siswa (Surohman et al., 2021), Clustering Tingkat Risiko Klasifikasi Lapangan Usaha (KLU) (Nugroho et al., 2022), Klasifikasi Kebutuhan Jumlah Produk Makanan Customer (Istianto & 'Uyun, 2021).

Ayam kampung unggul Balitbangtan sering menghadapi berbagai jenis penyakit yang memengaruhi produktivitas dan kelangsungan hidup mereka. Masalah utamanya adalah sulitnya mengidentifikasi jenis penyakit pada ayam, pola penyakit secara cepat

dan akurat karena banyaknya faktor yang memengaruhi, seperti gejala klinis, lingkungan, dan pola penyebaran penyakit (Firmansyah & Susanto, 2024). Data yang besar dan kompleks ini membuat analisis manual menjadi tidak efisien dan rentan terhadap kesalahan. K-Means Clustering (Ferdy Pangestu et al., 2023) dapat membantu mengelompokkan data penyakit berdasarkan kesamaan gejala, pola penyebaran, atau faktor lingkungan. Dengan algoritma ini, data penyakit ayam dapat dibagi ke dalam beberapa cluster (Pamungkas et al., 2024), di mana setiap cluster mewakili kategori tertentu, seperti jenis penyakit atau tingkat keparahan. Hal ini memungkinkan identifikasi pola penyakit yang lebih cepat dan akurat, sehingga membantu peternak dan peneliti dalam mengambil tindakan pencegahan atau pengobatan yang tepat.

Berdasarkan uraian diatas, untuk mengatasi masalah yang ada maka penulis akan menyusun sebuah penelitian dengan judul “PENERAPAN MACHINE LEARNING METODE K-MEANS CLUSTERING DALAM PENGELOMPOKAN PENYAKIT PADA AYAM KAMPUNG UNGGUL BALITBANGTAN (KUB)”.

1.2 Perumusan Masalah

Dari permasalahan yang dihadapi dalam konteks ini maka dapat disusun rumusan masalah yang dapat disimpulkan untuk menjadi focus pada penelitian ini adalah :

1. Bagaimana penerapan Machine Learning dengan pendekatan metode K-Means Clustering dapat mengelompokkan penyakit pada ayam KUB?
2. Berapa jumlah cluster optimal yang dapat dihasilkan untuk pengelompokan penyakit pada ayam KUB dengan metode k-means?
3. Bagaimana efektivitas metode k-means clustering dalam memberikan wawasan mengenai penyebaran dan pola penyakit pada ayam KUB?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, dapat disimpulkan bahwa yang menjadi fokus tujuan utama dalam penelitian ini adalah :

1. Menerapkan Machine Learning metode K-Means Clustering yang dapat membantu para peternak dalam mengelompokan penyakit pada ayam KUB.

2. Menentukan jumlah cluster optimal untuk pengelompokan penyakit pada ayam KUB.
3. Mengevaluasi efektivitas metode k-means clustering dalam pengelompokan penyakit pada ayam KUB.

1.4 Batasan Masalah

Batasan masalah dilakukan agar peneliti dilakukan lebih terarah dan mencapai sasaran, oleh karena itu Batasan masalah sebagai berikut :

1. Penelitian ini berfokus pada pengelompokan penyakit pada ayam kampung jenis Ayam Kampung Unggul Balitbangtan (KUB).
2. Metode yang akan digunakan pada penelitian ini metode K-Means Clustering.
3. Penelitian ini menggunakan metode K-Means Clustering dengan variabel yang digunakan dalam menentukan pengelompokan penyakit pada Ayam KUB adalah gejala-gejala yang ditemukan selama penelitian dan berdasarkan studi literatur dan wawancara kepada peternak.

1.5 Manfaat Penelitian

Manfaat yang dapat diberikan pada penelitian ini dengan melakukan pengelompokan penyakit pada ayam kampung berbasis artificial intelligence.

1. Peningkatan pemahaman tentang pola penyebaran penyakit pada ayam KUB.
2. Pengelompokan penyakit pada ayam KUB yang lebih efisien.
3. Memberikan manfaat pada bidang ilmu komputer berupa tambahan referensi dalam penelitian – penelitian selanjutnya sehingga bermanfaat terhadap perkembangan Machine Learning dalam bidang peternakan.

1.6 Kontribusi Penelitian

Kontribusi yang diharapkan dari adanya penelitian ini adalah dapat memberikan masukan dalam bidang peternakan dan sistem pakar berupa pengembangan pengetahuan mengenai pengelompokan penyakit pada ayam dan algoritma K-Means Clustering. Sehingga dengan adanya sistem yang dihasilkan dapat membantu para peternak dan dokter hewan dalam melakukan pengelompokan penyakit pada ayam kampung KUB yang mengalami gejala-gejala penyakit yang tampak.

1.7 Sistematika Penulisan

Pembahasan yang lebih rinci tentang penulisan tesis ini penulis membagi menjadi beberapa bab. Tiap bab saling berkaitan dengan sesuai dengan ruang lingkup masalah yang dibahas. Sistematika disesuaikan dengan yang diatur dalam tata penulisan program studi.

BAB I. PENDAHULUAN

Pada BAB I ini berisi tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematis penulisan.

BAB II. LANDASAN TEORI

Pada BAB II menjelaskan tentang landasan teori dan penerapan metode yang berhubungan dengan masalah yang dibahas sesuai dengan judul yang diteliti

BAB III. METODOLOGI PENELITIAN

Pada BAB III menguraikan tentang jenis penelitian, pendekatan yang digunakan, sumber data, lokasi penelitian, metode dan media pengumpulan data.

BAB IV. ANALISA DAN PERANCANGAN

Pada BAB IV ini menjelaskan tentang kondisi objek penelitian. Kondisi yang dijelaskan meliputi gambaran dan langkah – langkah proses yang terjadi pada objek penelitian.

BAB V.IMPLEMENTASI DAN HASIL

Pada BAB V Bab ini berisi tentang implememtasi dan perancangan yang dilakukan serta detail.

BAB VI.PENUTUP

Pada BAB VI ini berisi tentang kesimpulan dari penelitian dan saran untuk penelitian selanjutnya agar dapat dikembangkan atau melanjutkan penelitian yang berkaitan dengan masalah yang akan diteliti sehingga di peroleh hasil yang baik dan bermanfaat bagi masyarakat..