



BAB II

TINJAUAN LITERATUR

Bab ini membahas landasan teori dan tinjauan penelitian terdahulu yang relevan dengan topik penelitian ini. Sumber yang digunakan mencakup jurnal, artikel ilmiah, buku, dan skripsi, yang memberikan wawasan mengenai performa algoritma klasifikasi dalam berbagai aplikasi. Penelitian terdahulu dikaji untuk memahami perkembangan, keunggulan, dan keterbatasan metode klasifikasi yang digunakan, serta untuk mengidentifikasi celah penelitian. Tinjauan ini menjadi dasar dalam menyusun metodologi dan analisis yang spesifik untuk perbandingan metode klasifikasi *K-Nearest Neighbor* dan *Naïve Bayes*[7].

2.1. Penelitian Terdahulu

Penelitian terdahulu merupakan kajian terhadap karya ilmiah yang telah dilakukan sebelumnya dan memiliki relevansi dengan topik penelitian yang sedang dilakukan. Tinjauan terhadap penelitian terdahulu bertujuan untuk mengetahui kontribusi ilmiah yang telah ada, memahami metode dan hasil yang digunakan, serta menemukan celah penelitian yang dapat dikembangkan lebih lanjut. Dalam penelitian ini, beberapa studi terkait penggunaan algoritma *K-Nearest Neighbor* dan *Naïve Bayes* telah dikaji untuk memberikan landasan yang kuat bagi analisis dan pembahasan. Berikut ini adalah beberapa penelitian terdahulu yang relevan dengan topik klasifikasi[8].

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.
 2. Penggunaan tanpa izin untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
 3. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



Tabel 2.1 Penelitian Terdahulu

No	Nama Penulis, Tahun, Judul	Tujuan	Hasil Penelitian
1	Ilyas, 2023, Sistem Pakar Diagnosa Penyakit Sindrom Metabolik pada Rumah Sakit Umum Daerah Tembilahan	Untuk mendeteksi dini penyakit tersebut maka diperlukan aplikasi Sistem pakar dengan metode <i>Forward Chaining</i> sehingga dengan adanya aplikasi ini mempermudah dalam deteksi dini penyakit.	Pengujian pengguna aplikasi, menggunakan uji Statistik dengan Uji Beta dengan tingkat kepercayaan 95% sistem dapat diterima dan dipercaya.
2	Dwi Yuli Prasetyo, 2020, Mesin Pembelajaran Ensemble Untuk Identifikasi Varietas Padi	Menerapkan dan mengevaluasi beberapa algoritma klasifikasi untuk mengidentifikasi varietas padi menggunakan fitur warna dan tekstur.	Berdasarkan tingkat ketelitian masing-masing algoritma, k-Nearest Neighbors memberikan hasil lebih baik dibanding algoritma lainnya, baik dengan maupun tanpa normalisasi data. Terdapat enam sampel Inpara-9 yang diprediksi benar (<i>true positive</i>) dan lima sampel diprediksi salah (<i>false positive</i>). Pada varietas Batang Piaman terdapat delapan sampel yang diprediksi benar (<i>true positive</i>). Pada IR42 terdapat lima sampel yang diprediksi benar.
3	Fatmanisa, 2022, Perbandingan Metode Klasifikasi Sentimen Analisis Penggunaan E-Wallet Menggunakan Algoritma Naïve Bayes dan K-Nearest Neighbor	Untuk mencari algoritma terbaik dalam menentukan klasifikasi sentimen analisis penggunaan <i>e-wallet</i> dengan membandingkan dua algoritma yaitu <i>Naïve Bayes</i> (NB) dan <i>K-Nearest Neighbor</i> (<i>k-NN</i>).	Hasil <i>accuracy</i> dari masing-masing model klasifikasi yaitu NB sebesar 73.03% dan <i>k-NN</i> sebesar 89.44%, <i>precision</i> NB sebesar 21.40% dan <i>k-NN</i> sebesar 65.45%, dan <i>recall</i> NB sebesar 48.32% dan <i>k-NN</i> sebesar 22.25%.

Hak Cipta Diindungi Undang-Undang

Universitas Islam Indragiri

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.
 2. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
 3. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.
- Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

Tabel 2.2 Penelitian Terdahulu (Lanjutan)

No	Nama Penulis, Tahun, Judul	Tujuan	Hasil Penelitian
4	Calvin, 2022, Perbandingan Metode K-Nearest Neighbor dan Naïve Bayes Untuk Klasifikasi Gender Berdasarkan Mata	Untuk klasifikasi Gender berdasarkan mata menggunakan metode K-NN dan Naïve Bayes.	Pada metode k-NN dengan fitur HSV dan HOG dengan Cropping metode K-NN menghasilkan akurasi yang lebih tinggi di bandingkan Naïve Bayes yaitu sebesar 68,61%.
5	Sahar, 2020, Analisis Perbandingan Metode K-Nearest Neighbor dan Naïve Bayes Classifier pada Dataset Penyakit Jantung	Untuk mengetahui apakah termaksud penyakit jantung atau non penyakit jantung dengan menggunakan rumus dari metode K-Nearest Neighbor dan Naïve Bayes Classifier yang menggunakan library scikit learn.	Berdasarkan hasil pengujian, didapatkan tingkat akurasi pada metode K-Nearest Neighbor sebesar 67%, presisi 65%, recall 73%, dan f-measure 96% pada nilai K=250 dan metode jarak Manhattan, tingkat akurasi pada metode jarak Euclidean sebesar 65%, presisi 65%, recall 69%, dan f-measure 67% pada nilai K=250 sedangkan pada metode Naïve Bayes Classifier tingkat akurasi yang didapatkan sebesar 58%, presisi 90%, recall 55%, dan f-measure 68%
6	Veronica, 2024, Perbandingan Metode Klasifikasi Naïve Bayes dan K-Nearest Neighbor pada Data Status Pembayaran Pajak Pertambahan Nilai di Kantor Pelayanan Pajak Pratama Surakarta	Untuk mengoptimasi pengumpulan pajak, menyusun kebijakan yang lebih efektif, meningkatkan ketaatan pajak di masyarakat, dan meningkatkan efisiensi administrasi secara keseluruhan.	Hasil penelitian ini menunjukkan bahwa metode naïve Bayes memiliki performa klasifikasi yang lebih baik dengan nilai APER sebesar 8,31% sedangkan metode K-NN sebesar 13,07%.



Tabel 2.3 Penelitian Terdahulu (Lanjutan)

No	Nama Penulis, Tahun, Judul	Tujuan	Hasil Penelitian
7	Safrida, 2023, Klasifikasi Kematangan Buah Sawit Berdasarkan Fitur Warna, Bentuk dan Tekstur Menggunakan Algoritma K-NN	Untuk memilih fitur terbaik yang mampu melakukan klasifikasi kematangan buah sawit menggunakan Algoritma K-Nearest Neighbors.	Hasil klasifikasi terbaik dalam penelitian ini adalah penerapan pada fitur Warna dengan K = 1, 3, 4, 5 dan 7 sebesar 96.6%, fitur Tektur dengan K = 6 sebesar 66% dan fitur Bentuk dengan K = 1 sebesar 73.3%.
8	Wirawan, 2022, Rancang Bangun Aplikasi dengan Perbandingan Metode k-Nearest Neighbor (KNN) dan Naïve Bayes dalam Klasifikasi Penderita Penyakit Diabetes	Untuk membandingkan kedua algoritma yang memiliki tingkat akurasi yang terbaik.	Hasil nilai akurasi Algoritma KNN dengan K=3 memiliki nilai sebesar 93%, sedangkan algoritma Naïve Bayes memiliki akurasi sebesar 95%.
9	Zalana, 2021, Sistem Informasi Agronomi Tanaman Kelapa pada Industri Hulu dan Hilir di Kabupaten Indragiri Hilir	Menyediakan sarana informasi basis pengetahuan dengan memanfaatkan media teknologi informasi yang dibangun dalam bentuk web view dengan desain yang dibuat sederhana mungkin, sehingga dapat dipahami dan diakses dengan mudah oleh petani kelapa menggunakan smartphone.	Hasil analisis terhadap ICT Literacy petani kelapa di Kabupaten Indragiri Hilir dengan melakukan analisis klasifikasi menggunakan algoritma C4.5, analisis klusterisasi menggunakan algoritma K-Means dan persamaan regresi, diperoleh sumber informasi terkait budidaya kelapa yang diperoleh petani kelapa yaitu melalui penyuluh senilai 7,5%, petani lain senilai 54%, media cetak senilai 22%, media elektronik senilai 28% dan internet senilai 13%.





Tabel 2.4 Penelitian Terdahulu (Lanjutan)

No	Nama Penulis, Tahun, Judul	Tujuan	Hasil Penelitian
10	Ovelina, 2024, Analisis Perbandingan Algoritma Naïve Bayes dengan K-Nearest Neighbor (KNN) Pada Dataset Mobile Price Classification	Untuk memprediksi kisaran harga yang menunjukkan seberapa tinggi harga tersebut berdasarkan fitur-fitur yang tersedia.	Hasil bahwa KNN memiliki akurasi yang lebih tinggi dibandingkan dengan <i>Naïve Bayes</i> dalam memprediksi harga smartphone.
11	Shinta, 2024, Perbandingan Algoritma Dan <i>Knn</i> Dalam Analisis Sentimen Ulasan Pengguna Aplikasi <i>Capcut</i>	Untuk mengevaluasi bagaimana pengaruh variasi rasio pembagian data latih dan uji terhadap kinerja kedua metode dalam analisis sentimen, serta membandingkan keduanya berdasarkan akurasi, presisi, <i>recall</i> dan <i>f1 score</i> .	Menggunakan sembilan rasio pembagian data, ditemukan bahwa rasio 80:20 memberikan kinerja terbaik untuk kedua metode. <i>Naïve Bayes</i> mengungguli <i>KNN</i> dengan akurasi 79.41% dibanding 75.63%. Rasio 50:50 memberikan presisi terbaik untuk kedua metode. Secara keseluruhan, <i>Naïve Bayes</i> menunjukkan performa lebih baik, terutama pada rasio 80:20, menjadikannya pilihan yang lebih tepat untuk analisis sentimen aplikasi <i>CapCut</i> .

Pada Tabel 2.1 diatas dijelaskan bahwa Ilyas (2023) dalam penelitian “Sistem Pakar Diagnosa Penyakit Sindrom Metabolik pada Rumah Sakit Umum Daerah Tembilahan” bertujuan untuk mendeteksi dini penyakit tersebut maka diperlukan aplikasi Sistem pakar dengan metode Forward Chaining sehingga dengan adanya aplikasi ini mempermudah dalam deteksi dini penyakit. Hasil penelitian tersebut adalah pengujian pengguna aplikasi, menggunakan uji Statistik



dengan Uji Beta dengan tingkat kepercayaan 95% sistem dapat diterima dan dipercaya[9].

Dwi Yuli Prasetyo (2020) dalam penelitian "Mesin Pembelajaran Ensemble Untuk Identifikasi Varietas Padi" mengevaluasi algoritma klasifikasi untuk mengidentifikasi varietas padi menggunakan fitur warna dan tekstur. Hasil penelitian menunjukkan bahwa KNN memberikan hasil terbaik dibanding algoritma lainnya, baik dengan maupun tanpa normalisasi data[10].

Fatmanisa (2022) dalam penelitiannya yang berjudul "Perbandingan Metode Klasifikasi Sentimen Analisis Penggunaan E-Wallet Menggunakan Algoritma Naïve Bayes dan K-Nearest Neighbor" bertujuan untuk mencari algoritma terbaik dalam menentukan klasifikasi sentimen analisis penggunaan e-wallet. Dalam penelitian ini, hasil menunjukkan bahwa Naïve Bayes menghasilkan akurasi sebesar 73.03%, presisi 21.40%, dan recall 48.32%. Di sisi lain, KNN memiliki performa yang lebih baik dengan akurasi 89.44%, presisi 65.45%, dan recall 22.25%[11].

Calvin (2022) melakukan penelitian berjudul "Perbandingan Metode K-Nearest Neighbor dan Naïve Bayes Untuk Klasifikasi Gender Berdasarkan Mata". Penelitian ini bertujuan untuk mengklasifikasikan gender berdasarkan fitur visual dari mata menggunakan kedua metode tersebut. Hasil penelitian menunjukkan bahwa metode KNN dengan fitur HSV dan HOG memberikan akurasi sebesar 68.61%, yang lebih tinggi dibandingkan Naïve Bayes[12].

Sahar (2020) dalam studinya yang berjudul "Analisis Perbandingan Metode K-Nearest Neighbor dan Naïve Bayes Classifier pada Dataset Penyakit Jantung" membandingkan kedua algoritma untuk klasifikasi dataset penyakit jantung.

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan tanpa izin untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

3. Plagiarisme juga dilarang dan dapat dikenakan sanksi.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



Penelitian ini menunjukkan bahwa metode KNN dengan jarak Manhattan menghasilkan akurasi sebesar 67%, presisi 65%, dan recall 73%, sedangkan metode Naïve Bayes memiliki akurasi sebesar 58%, presisi 90%, dan recall 55%[6].

Veronica (2024) melakukan penelitian berjudul "Perbandingan Metode Klasifikasi Naïve Bayes dan K-Nearest Neighbor pada Data Status Pembayaran Pajak Pertambahan Nilai di Kantor Pelayanan Pajak Pratama Surakarta". Penelitian ini bertujuan untuk mengoptimalkan pengumpulan pajak dan menyusun kebijakan yang lebih efektif. Hasilnya menunjukkan bahwa Naïve Bayes memiliki performa yang lebih baik dengan nilai APER sebesar 8.31% dibandingkan KNN yang memiliki nilai 13.07%[13].

Safrida (2023) melalui penelitian berjudul "Klasifikasi Kematangan Buah Sawit Berdasarkan Fitur Warna, Bentuk dan Tekstur Menggunakan Algoritma K-NN" berfokus pada memilih fitur terbaik yang mampu melakukan klasifikasi kematangan buah sawit. Hasil penelitian menunjukkan bahwa algoritma KNN menghasilkan akurasi terbaik pada fitur warna sebesar 96.6%, fitur tekstur sebesar 66%, dan fitur bentuk sebesar 73.3%[14].

Wirawan (2022) dalam penelitiannya berjudul "Rancang Bangun Aplikasi dengan Perbandingan Metode K-Nearest Neighbor (KNN) dan Naïve Bayes dalam Klasifikasi Penderita Penyakit Diabetes" membandingkan akurasi kedua algoritma dalam klasifikasi penderita diabetes. Penelitian ini menemukan bahwa KNN dengan $k=3$ menghasilkan akurasi sebesar 93%, sedangkan Naïve Bayes menghasilkan akurasi sebesar 95%[15].

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan tanpa izin untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

3. Plagiarisme juga dilarang dan dapat dikenakan sanksi.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



Zalana (2021) melalui penelitian berjudul "Sistem Informasi Agronomi Tanaman Kelapa pada Industri Hulu dan Hilir di Kabupaten Indragiri Hilir" mengembangkan sistem informasi agronomi kelapa berbasis web. Penelitian ini menggunakan analisis klasifikasi C4.5, klusterisasi K-Means, dan persamaan regresi untuk memberikan informasi agronomi kepada petani kelapa[16].

Ovelina (2024) dalam penelitian berjudul "Analisis Perbandingan Algoritma Naïve Bayes dengan K-Nearest Neighbor (KNN) Pada Dataset Mobile Price Classification" membandingkan kedua algoritma untuk memprediksi kisaran harga smartphone berdasarkan fitur-fitur tertentu. Penelitian ini menemukan bahwa KNN memiliki akurasi lebih tinggi dibandingkan Naïve Bayes dalam memprediksi harga[5].

Shinta (2024) dalam studinya yang berjudul "Perbandingan Algoritma Naïve Bayes dan KNN dalam Analisis Sentimen Ulasan Pengguna Aplikasi CapCut" mengevaluasi pengaruh variasi rasio pembagian data latih dan uji terhadap kinerja kedua metode dalam analisis sentimen. Hasil penelitian menunjukkan bahwa pada rasio data 80:20, Naïve Bayes menghasilkan akurasi 79.41%, yang lebih tinggi dibandingkan KNN dengan akurasi 75.63%[17].

2.2. Rangkuman

Penelitian terdahulu yang telah dikaji menunjukkan bahwa *K-Nearest Neighbor* dan *Naïve Bayes* banyak digunakan dalam berbagai aplikasi klasifikasi data. Fatmanisa (2022) membandingkan kedua metode untuk analisis sentimen e-wallet, menunjukkan bahwa *K-Nearest Neighbors* memiliki keunggulan dalam akurasi (89.44%) dibandingkan Naïve Bayes (73.03%), tetapi Naïve Bayes unggul dalam presisi (21.40%). Calvin (2022) menggunakan metode serupa untuk



klasifikasi gender berdasarkan mata, di mana KNN memberikan hasil yang lebih tinggi dibandingkan Naïve Bayes dengan akurasi 68.61%. Penelitian Sahar (2020) juga menunjukkan keunggulan KNN dalam klasifikasi penyakit jantung dibandingkan Naïve Bayes, terutama saat menggunakan jarak Manhattan[18].

Sebaliknya, Veronica (2024) menemukan bahwa Naïve Bayes memiliki performa lebih baik daripada KNN dalam klasifikasi data status pembayaran pajak, dengan nilai APER lebih rendah (8.31% dibandingkan 13.07%). Safrida (2023) berfokus pada klasifikasi kematangan buah sawit menggunakan KNN dan menunjukkan bahwa algoritma ini menghasilkan akurasi tertinggi untuk fitur warna (96.6%). Penelitian-penelitian lainnya, seperti yang dilakukan oleh Shinta (2024) dan Ovelina (2024), memperkuat pemahaman bahwa performa kedua algoritma sangat bergantung pada jenis dataset, fitur yang digunakan, dan tujuan analisis.

Perbedaan utama antara penelitian terdahulu dan penelitian yang sedang dilakukan terletak pada objek dan tujuan analisisnya. Penelitian ini berfokus pada klasifikasi tingkat kematangan kelapa di daerah Tembilahan kabupaten Indragiri Hilir, yang memiliki karakteristik dataset berbeda, yaitu berbasis fitur visual seperti warna, tekstur, dan bentuk. Selain itu, penelitian ini tidak hanya membandingkan performa KNN dan Naïve Bayes, tetapi juga mengevaluasi parameter dan kondisi optimal untuk kedua algoritma dalam konteks data agrikultur. Pendekatan ini memberikan kontribusi baru dengan menggabungkan hasil-hasil penelitian terdahulu ke dalam konteks spesifik klasifikasi tingkat kematangan kelapa, yang belum banyak dieksplorasi dalam penelitian sebelumnya[19].

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan tanpa izin untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

3. Plagiarisme juga dilarang dan dapat dikenakan sanksi.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



2.3. Kesimpulan

Berdasarkan penelitian terdahulu, metode KNN menunjukkan hasil terbaik dalam klasifikasi berbasis fitur visual, seperti yang ditunjukkan pada penelitian Safrida (2023) dengan judul penelitian “Klasifikasi Kematangan Buah Sawit Berdasarkan Fitur Warna, Bentuk dan Tekstur Menggunakan Algoritma K-NN” dengan akurasi mencapai 96.6% pada fitur warna. Oleh karena itu, penelitian ini akan menjadikan penelitian Safrida sebagai acuan utama untuk membangun pembahasan mengenai klasifikasi tingkat kematangan kelapa di Tembilahan kabupaten Indragiri Hilir[14].



1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.