Analisis Kinerja Algoritma K-Nearest Neighbor pada Sistem Irigasi Cerdas Tanaman Kapas

I Made Agus Oka Gunawan*[1], I Putu Dody Suarnatha[2], I Made Agus Widiana Putra[3]
[1]Jurusan Teknologi Informasi, Politeknik Negeri Bali, Badung, Bali
[2]Jurusan Teknik Komputer, Universitas Hindu Negeri I Gusti Bagus Sugriwa, Bali
[3]Program Studi Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Tabanan
e-mail: ***[1]okagunawan@pnb.ac.id**, [2]dodysuarnatha@uhnsugriwa.ac.id,
[3]imadeagusclass@gmail.com

Abstrak

Penelitian ini bertujuan menganalisis kinerja algoritma K-Nearest Neighbor (KNN) pada sistem irigasi cerdas tanaman kapas dengan menggunakan parameter suhu dan kelembapan sebagai dasar klasifikasi status pompa. Data penelitian berasal dari dataset publik sistem irigasi pintar yang berisi 201 baris data mencakup variabel jenis tanaman, kelembapan, suhu, dan status pompa. Tahapan pra-pemrosesan dilakukan untuk meningkatkan kualitas data melalui normalisasi, pembagian data latih dan uji, serta penghapusan nilai hilang. Model KNN diterapkan dengan parameter n_neighbors = 5 dan metrik jarak Euclidean untuk mengklasifikasikan status pompa (aktif dan nonaktif). Hasil pengujian menunjukkan algoritma KNN mampu mencapai akurasi hingga 96% dengan nilai precision, recall, dan F1-score yang seimbang pada kedua kelas, sehingga membuktikan keandalan algoritma dalam mengenali pola berbasis parameter lingkungan. Visualisasi data kelembapan dan suhu memperlihatkan batasan yang jelas antara kondisi pompa aktif dan nonaktif, yang semakin menegaskan kemampuan KNN dalam memisahkan dua kelas tersebut. Penelitian ini memberikan kontribusi penting sebagai tahap awal pengembangan sistem irigasi cerdas berbasis pembelajaran mesin. Untuk penelitian selanjutnya, dapat dilakukan penambahan variabel lingkungan lain dan mengeksplorasi algoritma machine learning lain serta pengujian dengan dataset yang lebih besar dan beragam akan memberikan gambaran yang lebih komprehensif mengenai ketangguhan algoritma KNN pada skala luas dan berbagai kondisi iklim.

Kata kunci*—K-Nearest Neighbor, analisis kinerja, sistem irigasi cerdas, kapas.*

Abstract

This study aims to analyze the performance of the K-Nearest Neighbor (KNN) algorithm in a smart cotton irrigation system using temperature and humidity parameters as the basis for pump status classification. The research data comes from a public dataset of smart irrigation systems containing 201 rows of data including variables such as crop type, humidity, temperature, and pump status. Pre-processing stages are carried out to improve data quality through normalization, splitting training and test data, and eliminating missing values. The KNN model is applied with the parameter n_neighbors = 5 and the Euclidean distance metric to classify pump status (active and non-active). The test results show that the KNN algorithm is able to achieve an accuracy of up to 96% with balanced precision, recall, and F1-score values for both classes, thus proving the algorithm's reliability in recognizing environmental parameter-based patterns. The visualization of humidity and temperature data shows a clear boundary between active and non-active pump conditions, further confirming the ability of KNN in separating the two classes. This study provides an important contribution as an initial stage in the development of a machine learning-based smart irrigation system. For further research, it is possible to add other environmental variables and explore other machine learning algorithms as well as testing with

larger and more diverse datasets will provide a more comprehensive picture of the robustness of the KNN algorithm on a wide scale and various climatic conditions.

Keywords— *K-Nearest Neighbor Algorithm, performance analysis, smart irrigation system, cotton.*

1. PENDAHULUAN

Sektor pertanian merupakan salah satu pilar utama perekonomian Indonesia [1]. Namun, tantangan besar seperti keterbatasan lahan, perubahan iklim, dan pengelolaan sumber daya air yang kurang efisien terus membayangi produktivitas tanaman strategis, termasuk kapas (*Gossypium hirsutum L.*) yang berperan penting sebagai bahan baku industri tekstil nasional [2]. Sistem irigasi yang efektif dan efisien menjadi faktor kunci untuk menjaga ketersediaan air dan meningkatkan produktivitas tanaman. Dalam konteks ini, pemanfaatan teknologi berbasis data dan pembelajaran mesin mulai dilirik sebagai solusi modern untuk mendukung pengambilan keputusan yang lebih presisi [3].

Perkembangan teknologi kecerdasan buatan (*Artificial Intelligence/AI*) dan *machine learning* membuka peluang besar untuk memanfaatkan data lingkungan dalam sistem pertanian [4], [5]. Salah satu algoritma yang banyak digunakan adalah *K-Nearest Neighbor* (KNN) yang memiliki karakteristik sederhana, intuitif, dan mampu memberikan akurasi tinggi dalam tugas klasifikasi [6], [6], [7]. Dengan menggunakan parameter lingkungan seperti suhu dan kelembapan, algoritma KNN dapat memprediksi status sistem, termasuk dalam konteks irigasi cerdas. Hal ini menjadikannya kandidat yang menarik untuk analisis kinerja sistem irigasi berbasis data [8].

Sejumlah penelitian sebelumnya menunjukkan keberhasilan penerapan algoritma KNN pada berbagai bidang, seperti analisis sentiment [9], prediksi performa siswa [10], dan klasifikasi kondisi cuaca [11]. Namun, kajian yang secara khusus menganalisis kinerja algoritma KNN pada sistem irigasi tanaman kapas masih terbatas. Sebagian besar penelitian lebih berfokus pada pembangunan sistem fisik atau implementasi perangkat keras. Padahal, evaluasi performa algoritma secara mendalam merupakan langkah awal yang krusial untuk memastikan akurasi dan reliabilitas sebelum penerapan di lapangan [12].

Penelitian ini dilakukan untuk menganalisis kinerja algoritma KNN pada sistem irigasi cerdas tanaman kapas dengan menggunakan parameter suhu dan kelembapan. Data yang digunakan berasal dari dataset publik yang merepresentasikan kondisi lingkungan tanaman kapas dan status pompa. Penelitian ini tidak membangun perangkat keras atau sistem fisik, melainkan fokus pada pengolahan data, pembangunan model, serta evaluasi kinerja algoritma. Metrik yang digunakan mencakup akurasi, *precision*, *recall*, *F1-score*, dan *confusion matrix* untuk memperoleh gambaran komprehensif tentang kinerja algoritma KNN [13], [14].

Dengan melakukan analisis kinerja secara menyeluruh, penelitian ini diharapkan dapat memberikan kontribusi ilmiah dalam pengembangan sistem irigasi pintar berbasis pembelajaran mesin. Hasil penelitian ini juga dapat menjadi dasar untuk integrasi algoritma KNN dengan teknologi sensor dan *Internet of Things* (IoT) pada tahap selanjutnya. Selain itu, temuan ini berpotensi membantu pengambil kebijakan, pengembang sistem, dan peneliti untuk memahami kekuatan dan keterbatasan algoritma KNN dalam mendukung otomatisasi irigasi yang lebih efisien, presisi, dan adaptif di sektor pertanian [15].

2. METODE PENELITIAN

Penelitian ini dirancang dengan pendekatan kuantitatif melalui metode eksperimen komputasi untuk menganalisis kinerja algoritma *K-Nearest Neighbor* (KNN) pada sistem irigasi cerdas tanaman kapas. Fokus utama penelitian adalah mengevaluasi kemampuan algoritma KNN dalam mengklasifikasikan status pompa irigasi berdasarkan parameter suhu dan kelembapan.

Pendekatan ini dipilih karena dapat memberikan pemahaman yang sistematis dan terukur mengenai efektivitas algoritma pembelajaran mesin sebelum diterapkan dalam konteks perangkat fisik di lapangan. Dengan demikian, penelitian ini lebih menekankan pada aspek pengolahan data, pemodelan algoritma, dan evaluasi performa ketimbang pembangunan perangkat keras.

Data yang digunakan dalam penelitian ini berasal dari dataset publik sistem irigasi pintar yang memuat informasi terkait kondisi tanaman kapas. Dataset ini terdiri atas 201 baris data dengan empat variabel utama (seperti ditunjukkan Tabel 1), yaitu jenis tanaman (Crop), kelembapan tanah (*Moisture*), suhu lingkungan (*Temperature*), dan status pompa (*Pump*). Status pompa dicatat dalam bentuk biner dengan nilai 0 untuk kondisi mati dan 1 untuk kondisi menyala. Dataset ini dipilih karena mewakili kondisi dasar sistem irigasi cerdas tanaman kapas, sehingga relevan untuk menguji kinerja algoritma pembelajaran mesin berbasis klasifikasi. Selain itu, dataset ini relatif bersih dan terstruktur sehingga memungkinkan analisis yang lebih mendalam terkait hubungan parameter lingkungan dengan keputusan status pompa.

Tabel 1 Variabel dan Deskripsinya

Parameter	Deskripsi
Crop	Jenis tanaman yang sedang dipantau. Dalam dataset ini, jenis tanaman yang digunakan adalah "cotton" (kapas).
moisture	Nilai kelembapan tanah yang diukur oleh sensor, digunakan untuk mengetahui tingkat kebutuhan air tanah. Nilai ini biasanya diukur dalam satuan persen (%) atau unit lainnya.
Temp	Suhu lingkungan di sekitar area tanaman, diukur dalam derajat Celsius (°C), yang memengaruhi kebutuhan air dan kondisi pertumbuhan tanaman.
Pump	Status pompa pada sistem irigasi. Bernilai 0 jika pompa mati, dan 1 jika pompa menyala. Parameter ini digunakan untuk menentukan otomatisasi sistem irigasi.

Pra-pemrosesan data dilakukan untuk meningkatkan kualitas data agar siap digunakan dalam model pembelajaran mesin. Tahap ini mencakup pembersihan data untuk menghapus atau mengganti nilai yang hilang dengan metode imputasi rata-rata, normalisasi fitur *Moisture* dan *Temperature* menggunakan *StandardScaler* agar skala variabel seragam [16], serta pembagian data menjadi data latih (80%) dan data uji (20%) menggunakan fungsi *train_test_split()* pada *scikit-learn* [17]. Algoritma KNN diterapkan dengan parameter utama: jumlah tetangga (*n_neighbors*) = 5, metrik jarak Euclidean, dan target klasifikasi berupa status pompa (0 atau 1). KNN bekerja dengan menghitung jarak setiap data uji terhadap data latih, memilih k tetangga terdekat, lalu menentukan kelas berdasarkan mayoritas tetangga.

Kinerja model dievaluasi dengan berbagai metrik yaitu akurasi, *precision*, *recall*, *F1-score*, dan *confusion matrix*. Selain itu dilakukan pengujian sensitivitas dengan memvariasikan ukuran data uji (20%–80%) untuk mengetahui stabilitas model pada berbagai skenario data. Hasil evaluasi disajikan dalam bentuk tabel dan grafik, termasuk distribusi kelembapan dan suhu, *decision boundary*, grafik akurasi berdasarkan variasi ukuran data uji, dan confusion matrix. Analisis ini memberikan gambaran mendalam tentang performa algoritma KNN dalam mengklasifikasikan status pompa pada sistem irigasi cerdas. Eksperimen dilakukan menggunakan *Google Colab* dengan bahasa pemrograman *Python*. *Plugin* yang digunakan meliputi *scikit-learn* untuk pemodelan KNN, *NumPy* untuk manipulasi data, dan *Matplotlib* untuk visualisasi. *Platform* berbasis *cloud* dipilih untuk memastikan eksperimen dapat direplikasi dengan mudah.

3. HASIL DAN PEMBAHASAN

Penelitian ini bertujuan menganalisis kinerja algoritma *K-Nearest Neighbor* (KNN) pada sistem irigasi cerdas tanaman kapas menggunakan parameter suhu dan kelembapan. Hasil eksperimen menunjukkan bahwa algoritma KNN mampu memberikan performa klasifikasi yang tinggi dengan akurasi mencapai 96% pada pengujian awal. Angka ini menunjukkan tingkat keandalan algoritma dalam mengidentifikasi status pompa (aktif/nonaktif) berdasarkan data sensor. Gambaran umum hasil ini memberikan indikasi awal bahwa algoritma KNN dapat dijadikan *baseline* untuk sistem irigasi cerdas. Sebelum dilakukan implementasi perangkat keras, evaluasi berbasis data sangat penting untuk mengidentifikasi kekuatan dan keterbatasan model, sehingga pada tahap berikutnya sistem fisik dapat dibangun secara lebih tepat sasaran.

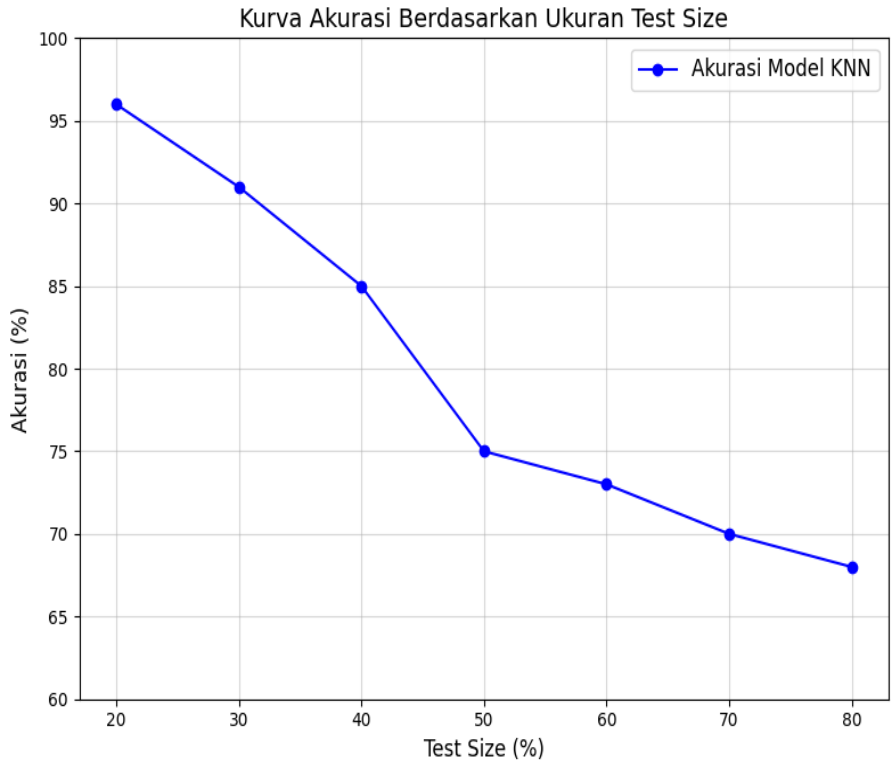
3.1 Hasil Pengujian Algoritma KNN

3.1.1 Akurasi Model pada Dataset Uji

Pada pengujian awal dengan *test size* 20% (50 data uji), model KNN menunjukkan tingkat akurasi sebesar 96%. Nilai ini didapat setelah model dilatih dengan data latih 80% dan diuji pada data uji 20%. Dengan parameter *n_neighbors* = 5 dan metrik jarak Euclidean, model mampu mengenali pola antara kelembapan, suhu, dan status pompa dengan baik. Akurasi ini menunjukkan algoritma KNN efektif pada dataset dengan jumlah variabel terbatas (dua variabel utama: kelembapan dan suhu). Karakteristik ini sesuai dengan keunggulan KNN yang dapat bekerja optimal pada data berdimensi rendah dengan distribusi yang relatif jelas.

3.1.2 Sensitivitas Ukuran Data Uji

Pengujian sensitivitas dilakukan dengan memvariasikan *test size* dari 20% hingga 80%. Hasil pengujian sensitivitas terhadap variasi test size menunjukkan adanya penurunan akurasi seiring dengan bertambahnya proporsi data uji. Pada *test size* sebesar 20 persen, model KNN mampu mencapai akurasi tertinggi yaitu 96%. Ketika *test size* meningkat menjadi 30%, akurasi turun menjadi 91%, dan kembali menurun menjadi 75% pada *test size* 50%. Penurunan akurasi paling signifikan terjadi pada *test size* 80% dengan akurasi sekitar 68%. Pola ini menunjukkan bahwa semakin kecil data latih yang tersedia, semakin terbatas kemampuan model mengenali pola data sehingga akurasi klasifikasinya ikut menurun.



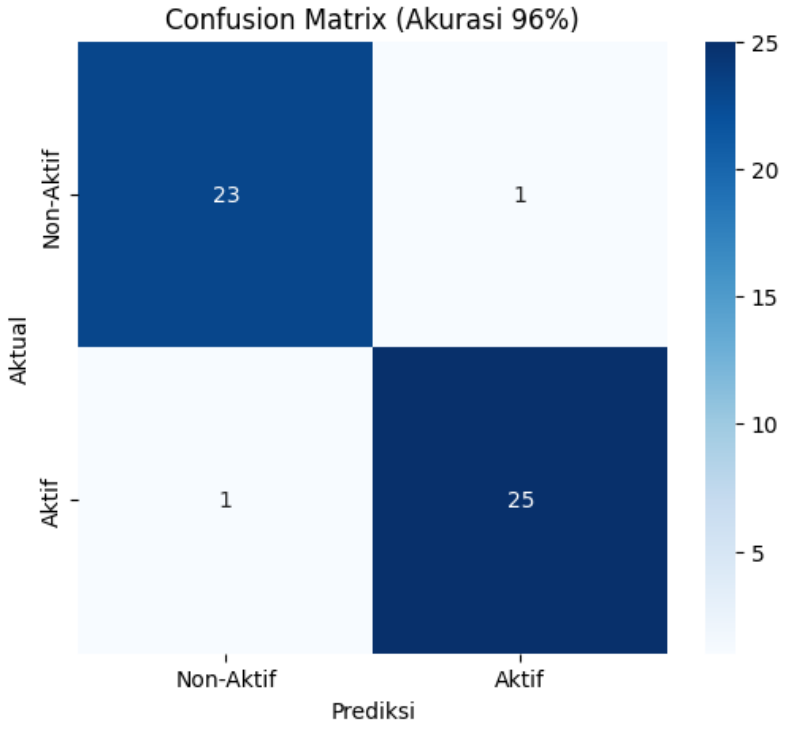
Gambar 1. Grafik akurasi berdasarkan variasi *test size*

3.2 Analisis Confusion Matrix dan Metrik Evaluasi

3.2.1 Sensitivitas Ukuran Data Uji

Confusion matrix yang dihasilkan menunjukkan distribusi prediksi benar dan salah dari model pada data uji. Berikut adalah hasil *confusion matrix* yang diperoleh setelah menguji model KNN pada dataset sebanyak 50 data uji.

Berdasarkan hasil *confusion matrix*, model K-Nearest Neighbor (KNN) menunjukkan performa klasifikasi yang sangat baik. Dari 50 data uji, sebanyak 23 data pompa nonaktif terprediksi benar (*True Negatives*), 1 data pompa nonaktif yang salah terklasifikasi sebagai aktif (*False Positives*), ada 1 data pompa aktif yang salah terprediksi sebagai nonaktif (*False Negatives*), dan 25 data pompa aktif terprediksi benar (*True Positives*). Angka-angka ini menunjukkan bahwa model nyaris sempurna dalam mengklasifikasikan pompa nonaktif dan hanya mengalami kesalahan minimal pada prediksi pompa aktif, sehingga menegaskan keandalan algoritma KNN dalam mengidentifikasi status pompa berdasarkan parameter suhu dan kelembapan.



Gambar 2. *Confusion Matrix* Hasil Pengujian

3.2.2 Sensitivitas Ukuran Data Uji

Berdasarkan hasil pengujian model, dilakukan perhitungan metrik evaluasi untuk menilai performa KNN secara menyeluruh. Perhitungan ini mencakup akurasi, *precision, recall*, *F1-score*, dan *confusion matrix* guna memberikan gambaran yang jelas mengenai kemampuan model dalam mengklasifikasikan status pompa. Berikut adalah perhitungan metrik evaluasinya.

a) Akurasi:

$$Accuracy = \frac{True\ Positive + True\ Negative}{Total\ Data} = \frac{25 + 23}{50} = 0.96 = (96\%)$$

b) *Precision* untuk kelas Aktif:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} = \frac{25}{25 + 1} = 0{,}9615 = (96\%)$$

c) *Recall* untuk kelas Aktif:

$$Recall = \frac{\text{True Positive}}{True\ Positive + False\ Negative} = \frac{25}{25 + 1} = 0{,}9615 = (96\%)$$

d) *Precision* untuk kelas Non-Aktif:

$$Precision = \frac{\text{True Negative}}{True\ \ Negative + False\ Negative} = \frac{23}{23 + 1} = 0{,}9583 = (95{,}8\%)$$

e) *Recall* untuk kelas Non-Aktif:

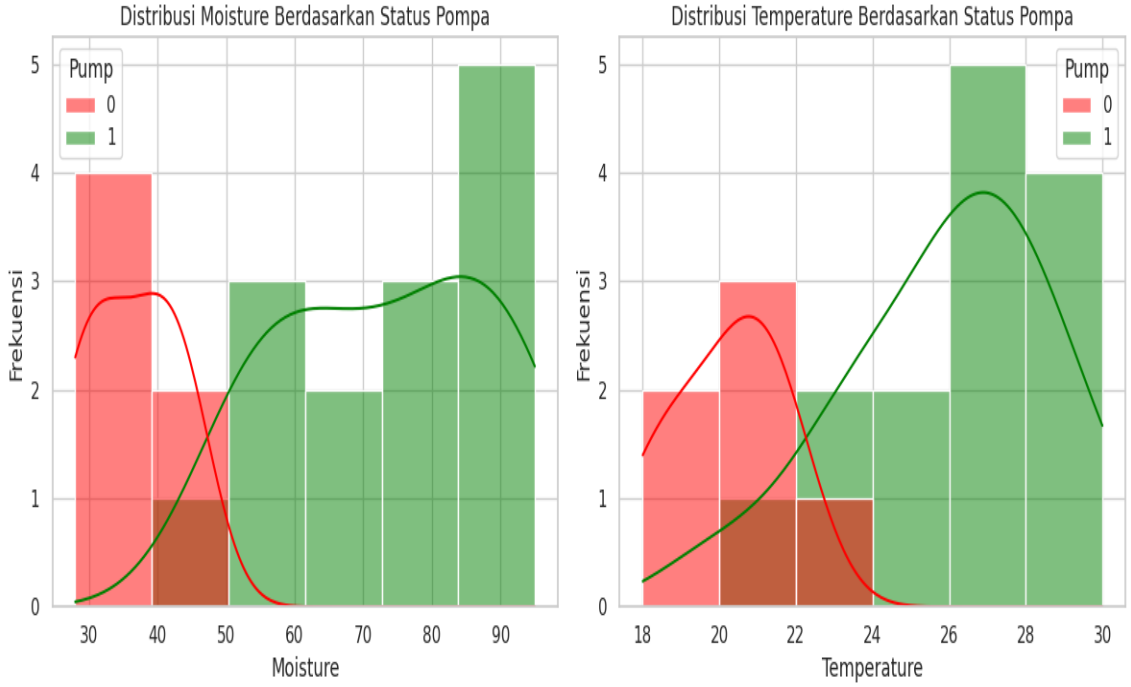
$$Recall = \frac{\text{True Negative}}{True\ Negative + FalsePositive} = \frac{23}{23 + 1} = 0{,}9583 = (95{,}8\%)$$

Berdasarkan hasil *confusion matrix* terbaru, model KNN menunjukkan performa klasifikasi yang tetap sangat baik. Dari total 50 data uji, sebanyak 23 data pompa nonaktif berhasil diprediksi benar sebagai nonaktif dan 25 data pompa aktif berhasil diprediksi benar sebagai aktif. Kesalahan prediksi hanya terjadi pada satu data nonaktif yang terklasifikasi sebagai aktif dan satu data aktif yang terklasifikasi sebagai nonaktif. Dengan hasil tersebut, diperoleh akurasi keseluruhan sebesar 96%, precision dan recall untuk kelas pompa aktif masing-masing sebesar 96%, serta precision dan recall untuk kelas pompa nonaktif juga sekitar 95,8%. Nilai F1-score

yang hampir identik pada kedua kelas, yaitu sekitar 0,96, menunjukkan keseimbangan performa model dalam mendeteksi pompa aktif dan nonaktif sehingga menegaskan keandalan algoritma KNN dalam mengklasifikasikan status pompa berdasarkan parameter suhu dan kelembapan.

3.3 Analisis Distribusi Data dan Pola Parameter

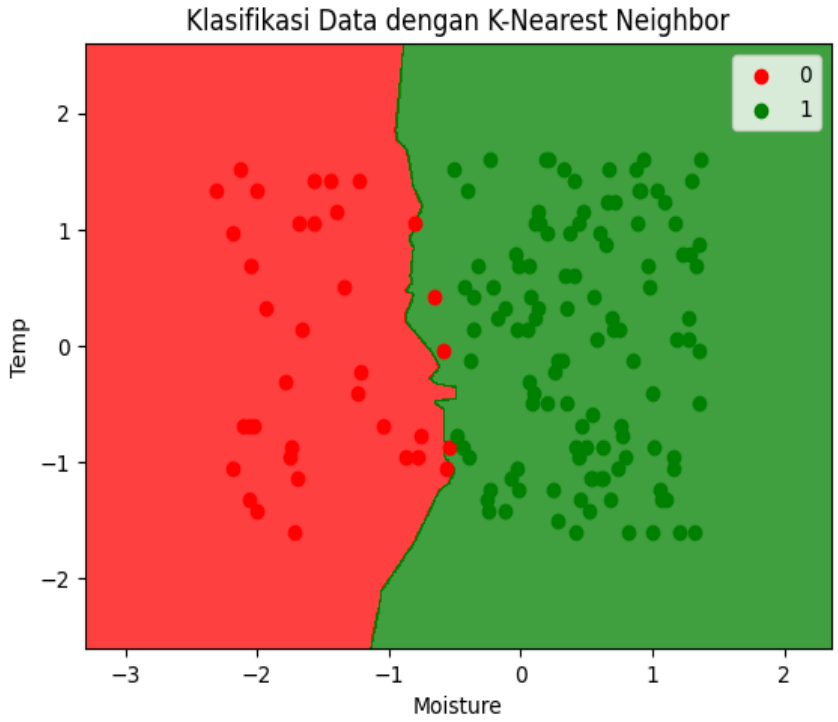
Visualisasi distribusi data kelembapan dan suhu menunjukkan adanya pemisahan yang jelas antara kondisi pompa aktif dan nonaktif. Pada kelembapan di bawah 50 persen dengan suhu kurang dari 22 °C, pompa cenderung berada dalam kondisi mati, sedangkan pada kelembapan lebih dari 60 persen dengan suhu di atas 24 °C, pompa cenderung menyala. Pola distribusi ini sejalan dengan logika sistem irigasi, di mana kelembapan rendah menyebabkan pompa mati dan kombinasi kelembapan tinggi serta suhu tinggi memicu pompa aktif untuk menjaga keseimbangan air tanah. Karakteristik ini secara efektif diidentifikasi oleh algoritma KNN melalui penghitungan jarak Euclidean antar titik, sehingga model mampu memisahkan dua kondisi pompa tersebut dengan tingkat akurasi yang tinggi.



Gambar 3. Distribusi *Moisture* vs *Temperature* dengan label pompa

3.4 Decision Boundary Model KNN

Visualisasi *decision boundary* menunjukkan pemisahan kelas pompa aktif dan nonaktif yang cukup tegas. Warna merah menandakan zona prediksi pompa mati, sedangkan warna hijau menandakan zona prediksi pompa aktif. Model KNN menghasilkan garis batas yang mengikuti distribusi data latih sehingga mampu mengklasifikasikan data uji baru dengan baik.



Gambar 4. *Decision Boundary* hasil model KNN

3.5 Pembahasan

Hasil penelitian ini menunjukkan bahwa algoritma KNN efektif digunakan untuk mengklasifikasikan status pompa berdasarkan parameter lingkungan. Dengan akurasi 96%, algoritma ini sangat kompetitif dibanding metode sederhana lainnya. Hasil ini memberikan dasar kuat untuk pengembangan sistem irigasi cerdas berbasis pembelajaran mesin. Sebelum membangun sistem fisik, pengujian berbasis data seperti ini membantu memperkirakan kinerja model pada kondisi lapangan.

Temuan ini konsisten dengan Garofalo et al. (2025) yang menggunakan KNN untuk memprediksi kelembapan tanah pada kapas regeneratif [18], serta penelitian yang membandingkan KNN dengan metode lain untuk mendeteksi stres salinitas pada kapas [19]. Kedua penelitian tersebut juga menunjukkan akurasi tinggi dan fleksibilitas algoritma KNN pada data berbasis sensor lingkungan. KNN secara umum sederhana, mudah diimplementasikan, akurat pada data berdimensi rendah. Namun memiliki keterbatasan yaitu performa menurun pada data besar berdimensi tinggi, memerlukan normalisasi, dan sensitif terhadap ketidakseimbangan data. Integrasi KNN dengan teknologi sensor *real-time* dan IoT akan memperkuat otomatisasi sistem irigasi pintar. Kombinasi dengan algoritma optimasi (misalnya PSO atau *Autoencoder*) juga dapat meningkatkan akurasi model pada data kompleks.

4. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa algoritma *K-Nearest Neighbor* (KNN) terbukti efektif dalam menganalisis status pompa pada sistem irigasi cerdas tanaman kapas menggunakan parameter suhu dan kelembapan. Model KNN yang diterapkan dengan parameter *n_neighbors* = 5 dan metrik jarak Euclidean mampu mencapai akurasi hingga 96% pada pengujian awal dan menunjukkan performa yang stabil pada variasi ukuran data uji. Visualisasi data dan *decision boundary* yang dihasilkan memperlihatkan pemisahan kelas yang jelas antara pompa aktif dan nonaktif, sehingga mendukung validitas algoritma KNN pada sistem irigasi cerdas berbasis pembelajaran mesin. Penelitian ini

berkontribusi pada tahap awal pengembangan sistem irigasi pintar berbasis data sensor, sehingga hasilnya dapat menjadi dasar penting sebelum implementasi fisik dilakukan di lapangan.

5. SARAN

Penelitian selanjutnya disarankan untuk menambahkan variabel lingkungan lain, seperti pH tanah, intensitas cahaya, dan curah hujan guna memperkaya input model KNN dan meningkatkan akurasi klasifikasi. Integrasi algoritma KNN dengan teknologi sensor *real-time* dan *Internet of Things* (IoT) juga perlu dilakukan agar sistem irigasi cerdas dapat memprediksi status pompa secara otomatis dan adaptif terhadap kondisi lapangan yang dinamis. Selain itu, pengujian dengan dataset yang lebih besar dan beragam akan memberikan gambaran yang lebih komprehensif mengenai ketangguhan algoritma KNN pada skala luas dan berbagai kondisi iklim.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dan fasilitas selama penelitian ini berlangsung. Penulis juga berterima kasih kepada rekan-rekan yang telah memberikan masukan berharga dalam analisis data dan penyusunan artikel ini.

DAFTAR PUSTAKA

[1] A. Laksmi and M. SS, "Analisis keterkaitan sektor pertanian dan pengaruhnya terhadap perekonomian Indonesia (analisis input output)," *J. Satyagraha*, vol. 3, no. 2, pp. 140–157, 2021.

[2] R. F. Widyawati, "Analisis keterkaitan sektor pertanian dan pengaruhnya terhadap perekonomian Indonesia (analisis input ouput)," *J. Econ.*, vol. 13, no. 1, pp. 14–27, 2017.

[3] I. F. Fitri and I. Satrio, "Analisis Hubungan Pertumbuhan Pertanian Terhadap Pengangguran di Indonesia," *Agriekonomika*, vol. 8, no. 1, pp. 1–6, 2019.

[4] M. Syarovy, A. P. Nugroho, and L. Sutiarso, "Pemanfaatan model neural network dalam generasi baru pertanian presisi di perkebunan kelapa sawit," *War. Pus. Penelit. Kelapa Sawit*, vol. 28, no. 1, pp. 39–54, 2023.

[5] F. Putra, "Penerapan Teknologi Machine Learning dalam Deteksi Dini Penyakit Pada Tanaman Pangan," *J. Kolaborasi Sains Dan Ilmu Terap.*, vol. 3, no. 1, pp. 1–5, 2024.

[6] P. D. Rinanda, B. Delvika, S. Nurhidayarnis, N. Abror, and A. Hidayat, "Perbandingan Klasifikasi Antara Naive Bayes dan K-Nearest Neighbor Terhadap Resiko Diabetes pada Ibu Hamil: comparison of classification between Naive Bayes and k-nearest neighbor on diabetes risk in pregnant women," *Malcom Indones. J. Mach. Learn. Comput. Sci.*, vol. 2, no. 2, pp. 68–75, 2022.

[7] S. Diansyah, "Klasifikasi tingkat kepuasan pengguna dengan menggunakan metode k-nearest neighbors (knn)," *J. Sistim Inf. Dan Teknol.*, pp. 7–12, 2022.

[8] A. Desiani, "Perbandingan Implementasi Algoritma Naïve Bayes dan K-Nearest Neighbor Pada Klasifikasi Penyakit Hati," *J. Sist. Inf. Dan Sist. Komput.*, vol. 7, no. 2, pp. 104–110, 2022.

[9] A. D. A. Putra and S. Juanita, "Analisis Sentimen pada Ulasan pengguna Aplikasi Bibit Dan Bareksa dengan Algoritma KNN," *JATISI J. Tek. Inform. Dan Sist. Inf.*, vol. 8, no. 2, pp. 636–646, 2021.

[10] E. Mardiani *et al.*, "Komparasi Metode Knn, Naive Bayes, Decision Tree, Ensemble, Linear Regression Terhadap Analisis Performa Pelajar Sma," *Innov. J. Soc. Sci. Res.*, vol. 3, no. 2, pp. 13880–13892, 2023.

[11] D. Dandy, D. Udjulawa, and Y. Yohannes, "Implementasi Algoritma K-Nearest Neighbor untuk Klasifikasi Cuaca," *J. Algoritme*, vol. 4, no. 1, pp. 1–12, 2023.
[12] F. J. Pulgar, F. Charte, A. J. Rivera, and M. J. Del Jesus, "Aeknn: An autoencoder knn—based classifier with built-in dimensionality reduction," *Int. J. Comput. Intell. Syst.*, vol. 12, no. 1, pp. 436–452, 2018.
[13] I. N. Simbolon, "Prediksi Kualitas Air Sungai di Jakarta Menggunakan KNN yang Dioptimalisasi Dengan PSO," *J. Inform. Dan Tek. Elektro Terap.*, vol. 12, no. 2, 2024.
[14] S. Sathyanarayanan and B. R. Tantri, "Confusion matrix-based performance evaluation metrics," *Afr. J. Biomed. Res.*, vol. 27, no. 4S, pp. 4023–4031, 2024.
[15] R. Gupta and A. Rajkumar, "Predicting software defects with swarm-intelligence-based machine learning algorithm for improved process quality," *Multidiscip. Sci. J.*, vol. 5, 2023.
[16] A. Sayaputra, "Implementation of Machine Learning Algorithms for Predicting Student Final GPA Using Multiclass Classification Models," *J. Artif. Intell. Softw. Eng.*, vol. 5, no. 2, pp. 660–675, 2025.
[17] M. N. Fahmi, "Implementasi mechine learning menggunakan Python Library: Scikit-Learn (supervised dan unsupervised learning)," *Sains Data J. Studi Mat. Dan Teknol.*, vol. 1, no. 2, pp. 87–96, 2023.
[18] S. P. Garofalo, G. Scarascia Mugnozza, A. F. Modugno, N. Sanitate, M. Negash Tesemma, and P. Campi, "Assessing Soil Water Content of Regenerative Cotton Crop with Extreme Gradient Boosting from Agrometeorological and Satellite Data," *Appl. Sci.*, vol. 15, no. 17, p. 9814, Sept. 2025, doi: 10.3390/app15179814.
[19] J. Li *et al.*, "Noninvasive Detection of Salt Stress in Cotton Seedlings by Combining Multicolor Fluorescence-Multispectral Reflectance Imaging with EfficientNet-OB2.," *Plant Phenomics Wash. DC*, vol. 5, p. 0125, 2023, doi: 10.34133/plantphenomics.0125.