

TEMPERATURE CLASSIFICATION IN HIGH VOLTAGE ELECTRICAL EQUIPMENT USING K-NEAREST NEIGHBOR

by Giovanni Dimas Prenata

Submission date: 11-Jul-2025 09:29AM (UTC+0200)

Submission ID: 2713243022

File name: JurnalTESLA_SuhuKNN_v5.docx (1.16M)

Word count: 3943

Character count: 24120

TEMPERATURE CLASSIFICATION IN HIGH VOLTAGE ELECTRICAL EQUIPMENT USING K-NEAREST NEIGHBOR

KLASIFIKASI SUHU PADA PERALATAN LISTRIK TEGANGAN TINGGI MENGGUNAKAN K-NEAREST NEIGHBOR

Giovanni Dimas Prenata¹, Aswanda Harja², Izzah Aula Wardah³

¹Program Studi Teknik Elektro, Fakultas Teknik, Universitas 17 Agustus 1945 Surabaya, Indonesia
Email: gprenata@untag-sby.ac.id

²Program Studi Teknik Elektro, Fakultas Teknik, Universitas 17 Agustus 1945 Surabaya, Indonesia
Email: aswanda.harja23@gmail.com

³Program Studi Teknik Elektro, Fakultas Teknik, Universitas 17 Agustus 1945 Surabaya, Indonesia
Email: iwardah@untag-sby.ac.id

Received: Revised: Accepted:
DOI: <https://doi.org/10.24912/tesla.v25i2.27065>

Abstract

The application of Artificial Intelligence (AI) for classifying safe or hazardous conditions in high-voltage electrical equipment using thermal images is a relatively new and emerging approach. Thermal imaging provides a visual representation of temperature distribution across an object's surface by utilizing color differences. High-temperature areas are typically represented in white, indicating potential hazards such as overheating or insulation failure. In contrast, non-white regions such as blue, green, or purple correspond to lower temperatures and are generally considered safe. This study employs feature extraction techniques to process thermal image data and obtain quantitative parameters, specifically the percentage of white and non-white areas within each image. These parameters serve as the input features for classification purposes. The dataset used in this research consists of 10 thermal images for training and 2 thermal images for testing. The classification process utilizes the K-Nearest Neighbor (K-NN) algorithm, with Euclidean Distance employed to measure the similarity between the test data and training samples. The images are then ranked based on proximity values, and classification is performed using different values of K. Testing was conducted by varying the value of K from 1 to 7 to assess the impact on classification performance. The results demonstrate that this method achieved a maximum classification accuracy of 71.42%. These findings suggest that AI-based image classification has significant potential for enhancing predictive maintenance and early detection of abnormal conditions in high-voltage electrical equipment through automated thermal image analysis.

Keywords: Thermal photos, danger, safety, classification, K-Nearest Neighbor.

Abstrak

Pemanfaatan kecerdasan buatan (Artificial Intelligence/AI) dalam klasifikasi kondisi aman atau berbahaya pada peralatan bertegangan tinggi berdasarkan citra termal merupakan pendekatan yang relatif baru dan masih jarang diterapkan secara luas. Citra termal memungkinkan visualisasi distribusi suhu pada permukaan objek melalui perbedaan warna. Area yang menunjukkan suhu tinggi biasanya divisualisasikan dalam warna putih dan diidentifikasi sebagai area dengan potensi bahaya, seperti adanya komponen yang mengalami pemanasan abnormal. Sebaliknya, warna selain putih seperti biru, hijau, atau ungu mengindikasikan suhu yang lebih rendah, sehingga dikategorikan sebagai kondisi yang aman. Dalam penelitian ini, dilakukan proses ekstraksi fitur dari citra termal untuk memperoleh data kuantitatif berupa persentase luas area berwarna putih dan non-putih. Parameter ini selanjutnya digunakan sebagai dasar dalam proses klasifikasi. Dataset terdiri dari 10 gambar yang digunakan sebagai data pelatihan dan 2 gambar sebagai data pengujian. Metode klasifikasi yang digunakan adalah K-Nearest Neighbor (K-NN), dengan pengukuran jarak antar data dilakukan menggunakan rumus Euclidean Distance. Setiap data uji dibandingkan terhadap data latih berdasarkan nilai kedekatan jarak tersebut, lalu dilakukan klasifikasi berdasarkan nilai K tertentu. Pengujian dilakukan dengan variasi nilai K mulai dari 1 hingga 7. Hasil dari

proses klasifikasi menunjukkan bahwa pendekatan ini mampu mencapai tingkat akurasi sebesar 71,42%. Temuan ini menunjukkan potensi AI dalam membantu pengambilan keputusan kondisi peralatan listrik berbasis analisis citra thermal.

Kata Kunci: Foto thermal, bahaya, aman, klasifikasi, K-Nearest Neighbor.

PENDAHULUAN

Penerapan kecerdasan buatan dalam sektor ketenagalistrikan merupakan sebuah inovasi yang masih tergolong baru, tetapi memiliki potensi besar dalam meningkatkan efisiensi dan keakuratan berbagai proses dalam sistem tenaga listrik. Kecerdasan buatan memungkinkan berbagai konsep dan metode dalam ketenagalistrikan diterapkan dengan lebih cepat, efisien, dan akurat dibandingkan metode konvensional. Dengan adanya teknologi ini, berbagai tantangan dalam pengelolaan dan distribusi tenaga listrik dapat diatasi dengan pendekatan berbasis data serta algoritma pembelajaran mesin yang canggih.

Salah satu penelitian yang membuktikan manfaat kecerdasan buatan dalam ketenagalistrikan dilakukan oleh Prenata, yang berfokus pada klasifikasi keandalan sistem distribusi tenaga listrik di PT PLN (Persero) UP3 Surabaya Selatan. Dalam penelitian tersebut, Prenata mengembangkan tiga metode utama dalam kecerdasan buatan, yaitu Single Perceptron [1], Support Vector Machine (SVM) [2], dan K-Nearest Neighbor (KNN) [3]. Single Perceptron merupakan model jaringan saraf tiruan paling sederhana yang mampu mengklasifikasikan data berdasarkan pemetaan linear. SVM bekerja dengan mencari hyperplane optimal yang memisahkan data dalam ruang berdimensi tinggi, sedangkan KNN mengklasifikasikan data berdasarkan kedekatannya dengan sejumlah titik data latih tertentu.

Hasil penelitian menunjukkan bahwa metode K-Nearest Neighbor (KNN) mampu mencapai tingkat akurasi 100% dalam proses klasifikasi keandalan sistem distribusi tenaga listrik. Keakuratan ini dicapai untuk berbagai nilai K, mulai dari $K = 1$ hingga $K = 7$, yang berarti bahwa metode ini secara konsisten dapat mengidentifikasi tingkat keandalan sistem dengan sangat baik. Keunggulan KNN dalam klasifikasi ini disebabkan oleh kemampuannya dalam mengelompokkan data berdasarkan kedekatan dengan data latih, sehingga memberikan hasil yang presisi dalam menentukan apakah suatu sistem tergolong andal atau tidak.

Hasil keseluruhan penelitian ini mengonfirmasi bahwa penerapan metode kecerdasan buatan dalam klasifikasi keandalan sistem tenaga listrik dapat memberikan hasil yang sangat presisi. Dengan akurasi yang mencapai 100%, baik pada KNN, Single Perceptron, maupun SVM, metode berbasis AI ini dapat diandalkan untuk meningkatkan efisiensi dan keakuratan dalam menganalisis keandalan sistem distribusi tenaga listrik, sehingga membantu dalam pengambilan keputusan yang lebih tepat dalam pengelolaan jaringan listrik.

Penelitian lain yang membahas penerapan kecerdasan buatan (Artificial Intelligence/AI) dalam bidang ketenagalistrikan, yang dilakukan oleh Prenata, berfokus pada upaya validasi hasil pengukuran rele dalam sistem tenaga listrik [4]. Rele merupakan salah satu komponen penting dalam sistem proteksi tenaga listrik yang berfungsi untuk mendeteksi gangguan dan mengisolasi bagian jaringan yang bermasalah guna mencegah dampak yang lebih luas.

Penelitian terkait penerapan metode K-Nearest Neighbor (KNN) telah dilakukan dalam berbagai bidang, menunjukkan efektivitasnya dalam proses klasifikasi dan prediksi. Beberapa penelitian yang telah dilakukan di antaranya adalah studi yang dilakukan oleh Khairi, Ghazali, dan Hidayah [5], yang menerapkan KNN untuk klasifikasi masyarakat pra-sejahtera di Desa Sapikerep, Kecamatan Sukapura. Dalam penelitian ini, mereka menggunakan beberapa nilai K untuk melihat tingkat akurasi metode KNN dalam mengklasifikasikan kelompok masyarakat pra-sejahtera berdasarkan kriteria tertentu. Hasil yang diperoleh menunjukkan bahwa untuk $K =$

3, akurasi klasifikasi mencapai 97,36%, sedangkan untuk $K = 5$ dan $K = 7$, tingkat akurasinya meningkat menjadi 98,68%.

Selain itu, penelitian yang dilakukan oleh Nanda, Pudjiantoro, dan Sabrina juga menerapkan metode KNN, namun dalam konteks yang berbeda, yaitu prediksi curah hujan di Kota Bandung [6]. Dengan mempertimbangkan berbagai faktor meteorologi yang berpengaruh terhadap curah hujan, mereka menggunakan KNN dengan $K = 5$ dan memperoleh tingkat akurasi sebesar 86,199%. Meskipun akurasi yang diperoleh tidak setinggi penelitian sebelumnya dalam klasifikasi masyarakat pra-sejahtera, metode KNN tetap menunjukkan keandalannya dalam menangani permasalahan prediksi berbasis data cuaca.

Penelitian lain yang juga menggunakan metode KNN dilakukan oleh Choli, Handayani, dan Ardianita dalam proses klasifikasi seleksi penerima beasiswa [7]. Dalam penelitian ini, mereka menerapkan metode KNN untuk membantu menentukan kandidat yang berhak menerima beasiswa berdasarkan sejumlah kriteria tertentu, seperti prestasi akademik, kondisi ekonomi, dan faktor pendukung lainnya. Hasil yang diperoleh menunjukkan bahwa metode KNN mampu mengklasifikasikan calon penerima beasiswa dengan tingkat akurasi sebesar 90,5%, yang membuktikan bahwa metode ini dapat diandalkan dalam proses seleksi berbasis data.

Dari berbagai penelitian di atas, dapat disimpulkan bahwa metode K-Nearest Neighbor (KNN) memiliki fleksibilitas dan efektivitas yang tinggi dalam berbagai bidang aplikasi, baik dalam klasifikasi sosial, prediksi cuaca, maupun seleksi penerima beasiswa. Dengan tingkat akurasi yang relatif tinggi pada berbagai skenario penggunaan, KNN dapat menjadi salah satu metode pilihan dalam berbagai penerapan kecerdasan buatan dan analisis data untuk menghasilkan keputusan yang lebih tepat dan akurat.

KAJIAN PUSTAKA

Metode K-Nearest Neighbor (KNN) merupakan salah satu teknik klasifikasi yang bekerja dengan membandingkan data uji terhadap sejumlah data latih berdasarkan perhitungan jarak di antara keduanya. Prinsip utama dari metode ini adalah bahwa kondisi data uji sangat dipengaruhi oleh data latih yang paling dekat dengannya, di mana kedekatan ini diukur menggunakan suatu metrik jarak tertentu. Semakin dekat suatu data uji dengan sekelompok data latih yang memiliki label tertentu, semakin besar kemungkinan bahwa data uji tersebut akan diklasifikasikan ke dalam kategori yang sama dengan data latih tersebut.

Dalam metode KNN, jumlah data latih yang dipertimbangkan dalam menentukan klasifikasi data uji disebut sebagai nilai K . Nilai K ini menunjukkan berapa banyak tetangga terdekat yang akan diperhitungkan dalam menentukan kelas atau kategori dari data uji yang sedang dianalisis. Pemilihan nilai K menjadi faktor krusial dalam metode ini, karena jika nilai K terlalu kecil, maka klasifikasi bisa menjadi terlalu sensitif terhadap outlier atau data yang tidak representatif. Sebaliknya, jika K terlalu besar, klasifikasi bisa menjadi terlalu umum dan kehilangan detail spesifik dari pola yang ada.

Nilai K dapat berupa bilangan genap maupun ganjil, namun dalam banyak kasus, penggunaan nilai ganjil lebih disarankan. Hal ini karena nilai K yang ganjil dapat membantu menghindari kemungkinan terjadinya jumlah suara (vote) yang sama dalam proses klasifikasi. Sebagai contoh, jika $K = 3$, maka klasifikasi dilakukan berdasarkan mayoritas dari tiga tetangga terdekat, yang memastikan bahwa selalu ada keputusan akhir yang jelas. Namun, jika $K = 4$, terdapat kemungkinan bahwa dua tetangga memiliki satu label dan dua lainnya memiliki label yang berbeda, sehingga keputusan

klasifikasi menjadi kurang pasti. Oleh karena itu, pemilihan nilai K yang optimal perlu dilakukan dengan mempertimbangkan karakteristik data serta kebutuhan spesifik dari sistem yang sedang dikembangkan.

Untuk menentukan jarak antara data uji dan data latih, penelitian ini menerapkan metode perhitungan menggunakan rumus Euclidean Distance, yang merupakan salah satu metrik jarak yang paling umum digunakan dalam algoritma KNN. Euclidean Distance mengukur jarak lurus antara dua titik dalam ruang dimensi tertentu dan dapat dihitung menggunakan rumus sebagai berikut:

$$d = \sqrt{\sum_{i=1}^k (X_i - Y_i)^2} \quad (1)$$

Dalam rumus di atas, p dan q merepresentasikan dua titik dalam ruang multidimensi, sedangkan koordinat (x, y, z, ...) adalah nilai-nilai fitur dari data yang dibandingkan. Dengan menggunakan metode ini, sistem dapat mengidentifikasi seberapa dekat suatu data uji dengan sejumlah data latih yang sudah diketahui labelnya, sehingga dapat menentukan klasifikasi yang paling sesuai berdasarkan pola yang ada.


Dengan demikian, metode KNN menawarkan pendekatan yang intuitif dan efektif dalam klasifikasi berbasis jarak, serta memiliki fleksibilitas dalam menangani berbagai jenis data dengan beragam karakteristik. Meskipun sederhana dalam konsepnya, pemilihan parameter seperti nilai K dan metode perhitungan jarak tetap menjadi aspek penting yang perlu diperhatikan untuk mendapatkan hasil klasifikasi yang optimal dan akurat.


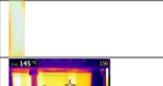
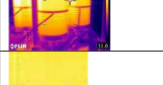


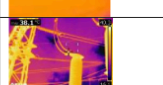
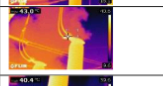
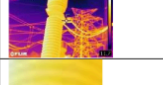

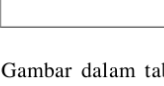
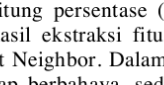
METODE PENELITIAN

Langkah pertama dalam proses ini adalah mengumpulkan data yang akan digunakan sebagai data latih dan data uji. Data yang dikumpulkan berupa gambar thermal yang berisi informasi suhu dari objek yang diamati. Berbeda dengan gambar konvensional yang menampilkan warna asli suatu objek, gambar thermal menggunakan warna untuk merepresentasikan suhu di setiap bagian objek. Warna gelap menunjukkan bahwa suhu pada bagian tersebut rendah atau dingin, yaitu di bawah 50°C, sedangkan warna terang menunjukkan suhu yang lebih tinggi atau panas, mendekati 160°C.

Data latih dan data uji terdiri dari gambar thermal berbagai objek, termasuk contact panel, housing lightning arrester, serta clamp circuit breaker pada bagian atas dan bawah. Setiap gambar thermal kemudian dipecah menjadi beberapa bagian kecil untuk memperbanyak jumlah sampel yang digunakan dalam analisis.

Tabel 1 Foto thermal

| Data | Foto thermal | Nama objek | % Putih | % Non Putih | Kategori |
|------|---|--|---------|-------------|----------|
| 1 |  | Bagian dari contact panel (data latih 0) | 18.5556 | 81.4444 | Aman (0) |

| | | | | | |
|----|---|--|----------|---------|------------|
| 2 |  | Bagian dari contact panel (data latih 1) | 36.5556 | 70.1111 | Bahaya (1) |
| 3 |  | Bagian dari contact panel (data latih 2) | 56.5476 | 43.4524 | Bahaya (1) |
| 4 |  | Housing Lightning arrester (data uji 2) | 2,8 | 97,2 | ?? |
| 5 |  | Bagian dari housing Lightning arrester (data latih 3) | 0 | 100 | Aman (0) |
| 6 |  | Bagian dari housing Lightning arrester (data uji 1) | 80,3 | 19,8 | ?? |
| 7 |  | Clamp Circuit Breaker bagian atas (data latih 4) | 0 | 100 | Aman (0) |
| 8 |  | Clamp Circuit Breaker bagian bawah (data latih 5) | 1.12989 | 98.9585 | Aman (0) |
| 9 |  | Clamp Circuit Breaker bagian bawah (data latih 6) | 1.40157 | 98.5984 | Aman (0) |
| 10 |  | Bagian Clamp Circuit Breaker bagian bawah (data latih 7) | 1.63177 | 98.6321 | Aman (0) |
| 11 |  | Bagian Clamp Circuit Breaker bagian bawah (data latih 8) | 0.093985 | 99.906 | Aman (0) |
| 12 |  | Bagian Clamp Circuit Breaker bagian bawah (data latih 9) | 0.127714 | 99,8768 | Aman (0) |

Gambar dalam tabel sebelumnya akan melalui proses ekstraksi fitur untuk menghitung persentase (%) warna putih dan persentase (%) warna selain putih. Data hasil ekstraksi fitur tersebut kemudian digunakan untuk melatih model K-Nearest Neighbor. Dalam pelatihan ini, jika persentase warna putih $\geq 30\%$, kondisi dianggap berbahaya, sedangkan jika persentasenya $< 30\%$, kondisi dikategorikan aman. Tahap berikutnya adalah menguji model K-Nearest Neighbor yang telah menghitung jarak antara data uji dengan masing-masing data latih menggunakan rumus 1. Hasil perhitungan akan disimpan lalu diurutkan.

Data uji ini sebelumnya juga telah melalui proses ekstraksi fitur untuk memperoleh persentase warna putih dan warna selain putih. Dengan menentukan nilai K (tingkat ketentangan) maka hasil klasifikasi K-Nearest Neighbor bisa mengklasifikasikan data uji termaksud kategori aman/berbahaya dengan menghitung jumlah data yang sudah diurutkan sebelumnya.

HASIL DAN PEMBAHASAN

Sebelum data latih dan data uji dipergunakan, proses ekstraksi fitur dilakukan untuk mendapat persentase warna putih dan persentase warna selain putih. Tahapan pengujian menggunakan 2 data uji. Untuk data uji pertama jarak antara data uji dan data latih sebagai berikut :

Tabel 2 Jarak antara data uji pertama dengan masing-masing data latih

| Terhadap | Nilai X1 | Nilai Y1 | Nilai X2 | Nilai Y2 | Jarak | Kategori |
|------------------|----------|----------|----------|----------|---------|------------|
| Data latih ke- 0 | 18.5556 | 81.4444 | 83.2647 | 20.439 | 88.9323 | Aman (0) |
| Data latih ke- 1 | 36.5556 | 70.1111 | 83.2647 | 20.439 | 68.1841 | Bahaya (1) |
| Data latih ke- 2 | 56.5476 | 43.4524 | 83.2647 | 20.439 | 35.2622 | Bahaya (1) |
| Data latih ke- 3 | 0 | 100 | 83.2647 | 20.439 | 116.342 | Aman (0) |
| Data latih ke- 4 | 0 | 100 | 83.2647 | 20.439 | 115.962 | Aman (0) |
| Data latih ke- 5 | 1.12989 | 98.9585 | 83.2647 | 20.439 | 113.629 | Aman (0) |
| Data latih ke- 6 | 1.40157 | 98.5984 | 83.2647 | 20.439 | 113.183 | Aman (0) |
| Data latih ke- 7 | 1.63177 | 98.6321 | 83.2647 | 20.439 | 113.04 | Aman (0) |
| Data latih ke- 8 | 0.093985 | 99.906 | 83.2647 | 20.439 | 115.032 | Aman (0) |
| Data latih ke- 9 | 0.127714 | 99.8768 | 83.2647 | 20.439 | 117.574 | Aman (0) |

Pada tabel diatas, nilai X1 dan Y1 adalah nilai % warna putih dan % warna selain putih pada data latih. Sedangkan nilai X2 dan Y2 adalah nilai % warna putih dan % warna selain putih pada data uji. Sedangkan kategori merupakan kondisi klasifikasi gambar thermal. Lalu data tabel 2 diurutkan dari yang bernilai kecil menuju bernilai besar.

Tabel 3 Pengurutan Jarak antara data uji pertama dengan masing-masing data latih

| Urutan | Jarak | Kategori |
|--------|---------|------------|
| 1 | 35.2622 | Bahaya (1) |
| 2 | 68.1841 | Bahaya (1) |
| 3 | 88.9323 | Aman (0) |
| 4 | 113.04 | Aman (0) |
| 5 | 113.183 | Aman (0) |
| 6 | 113.629 | Aman (0) |
| 7 | 115.032 | Aman (0) |
| 8 | 115.962 | Aman (0) |
| 9 | 116.342 | Aman (0) |
| 10 | 117.574 | Aman (0) |

Berdasarkan tabel diatas yang telah diurutkan berdasarkan kedekatan jarak, tahap pengujian dapat dilakukan dengan menggunakan berbagai nilai K untuk menentukan kategori data uji. Nilai K ini menunjukkan jumlah data tetangga terdekat yang digunakan dalam proses klasifikasi.

- Jika $K = 1$, maka hanya data dengan urutan pertama yang digunakan. Jika data ini termasuk dalam kategori bahaya (1), maka hasil klasifikasi juga bahaya.
- Jika $K = 2$, maka dua data pertama digunakan. Jika kedua data ini termasuk dalam kategori bahaya (1,1), maka hasil klasifikasi tetap bahaya.
- Jika $K = 3$, tiga data pertama diperhitungkan. Jika terdapat dua data kategori bahaya (1,1) dan satu data kategori aman (0), maka hasil akhir mengikuti kategori terbanyak, yaitu bahaya.

- Jika $K = 4$, empat data pertama digunakan. Jika terdapat dua data kategori bahaya (1,1) dan dua data kategori aman (0,0), maka terjadi imbang, sehingga sistem gagal mengidentifikasi kategori dengan pasti.
- Jika $K = 5$, lima data pertama diperhitungkan. Jika terdapat dua data kategori bahaya (1,1) dan tiga data kategori aman (0,0,0), maka sistem mengklasifikasikan data sebagai aman, karena kategori ini lebih dominan.
- Jika $K = 6$, enam data pertama dipertimbangkan. Dengan dua data kategori bahaya (1,1) dan empat data kategori aman (0,0,0,0), hasil klasifikasi tetap aman, karena jumlah data kategori aman lebih banyak.
- Jika $K = 7$, tujuh data pertama digunakan. Jika terdapat dua data kategori bahaya (1,1) dan lima data kategori aman (0,0,0,0,0), maka sistem menetapkan bahwa data termasuk dalam kategori aman.

Dari hasil klasifikasi ini, dapat disimpulkan bahwa semakin besar nilai K , semakin besar kemungkinan sistem memilih kategori dengan jumlah data terbanyak. Hal ini menunjukkan bahwa pemilihan nilai K yang optimal sangat penting agar sistem dapat memberikan klasifikasi yang lebih akurat, terutama ketika data dalam kategori tertentu lebih sedikit dibandingkan kategori lainnya.

Data pada tabel diatas akan diurutkan menjadi :

Tabel 4 Jarak antara data uji kedua dengan masing-masing data latih

| Terhadap | Nilai X1 | Nilai Y1 | Nilai X2 | Nilai Y2 | Jarak | Kategori |
|-------------------|----------|----------|-----------|----------|----------|------------|
| Data latih ke- 1 | 18.5556 | 81.4444 | 0.0886525 | 102.039 | 27.6616 | Aman (0) |
| Data latih ke- 2 | 36.5556 | 70.1111 | 0.0886525 | 102.039 | 48.4688 | Bahaya (1) |
| Data latih ke- 3 | 56.5476 | 43.4524 | 0.0886525 | 102.039 | 81.3634 | Bahaya (1) |
| Data latih ke- 4 | 0 | 100 | 0.0886525 | 102.039 | 0.35533 | Aman (0) |
| Data latih ke- 5 | 0 | 100 | 0.0886525 | 102.039 | 0.893994 | Aman (0) |
| Data latih ke- 6 | 1.12989 | 98.9585 | 0.0886525 | 102.039 | 3.2517 | Aman (0) |
| Data latih ke- 7 | 1.40157 | 98.5984 | 0.0886525 | 102.039 | 3.68258 | Aman (0) |
| Data latih ke- 8 | 1.63177 | 98.6321 | 0.0886525 | 102.039 | 3.74011 | Aman (0) |
| Data latih ke- 9 | 0.093985 | 99.906 | 0.0886525 | 102.039 | 2.133 | Aman (0) |
| Data latih ke- 10 | 0.127714 | 99.8768 | 0.0886525 | 102.039 | 1.53748 | Aman (0) |

Data pada tabel diatas akan diurutkan menjadi :

Tabel 5 Pengurutan Jarak antara data uji kedua dengan masing-masing data latih

| Urutan | Jarak | Kategori |
|--------|----------|------------|
| 1 | 0.35533 | Aman (0) |
| 2 | 0.893994 | Aman (0) |
| 3 | 1.53748 | Aman (0) |
| 4 | 2.133 | Aman (0) |
| 5 | 3.2517 | Aman (0) |
| 6 | 3.68258 | Aman (0) |
| 7 | 3.74011 | Aman (0) |
| 8 | 27.6616 | Aman (0) |
| 9 | 48.4688 | Bahaya (1) |
| 10 | 81.3634 | Bahaya (1) |

Berdasarkan tabel diatas yang telah diurutkan berdasarkan kedekatan jarak, pengujian

dapat dilanjutkan dengan menggunakan data uji kedua. Proses ini bertujuan untuk menentukan kategori data berdasarkan jumlah tetangga terdekat yang dianalisis, dengan berbagai nilai K yang menunjukkan jumlah data latih yang digunakan dalam klasifikasi.

- Jika $K = 1$, hanya data dengan urutan pertama yang digunakan. Data ini termasuk dalam kategori aman (0), sehingga hasil klasifikasi untuk data uji juga aman.
- Jika $K = 2$, dua data pertama dipertimbangkan. Jika kedua data ini termasuk dalam kategori aman (0,0), maka hasil klasifikasi tetap aman.
- Jika $K = 3$, tiga data pertama diperhitungkan. Jika ketiga data ini termasuk dalam kategori aman (0,0,0), maka sistem mengklasifikasikan data uji sebagai aman, karena seluruh data tetangga termasuk dalam kategori tersebut.
- Jika $K = 4$, empat data pertama digunakan. Dengan seluruh data tetangga termasuk dalam kategori aman (0,0,0,0), sistem tetap mengidentifikasi data sebagai aman.
- Jika $K = 5$, lima data pertama digunakan dalam perhitungan. Karena seluruhnya masih dalam kategori aman (0,0,0,0,0), hasil klasifikasi tidak berubah, yaitu aman.
- Jika $K = 6$, enam data pertama diperhitungkan. Dengan seluruh data tetangga masih tergolong dalam kategori aman (0,0,0,0,0,0), hasil klasifikasi tetap aman.
- Jika $K = 7$, tujuh data pertama digunakan. Karena semua data ini termasuk dalam kategori aman (0,0,0,0,0,0,0), maka sistem tetap mengidentifikasi data sebagai aman.

Dari hasil pengujian ini, terlihat bahwa semakin banyak data latih yang dipertimbangkan (nilai K lebih besar), semakin stabil hasil klasifikasi yang diperoleh. Jika seluruh data latih yang menjadi tetangga terdekat memiliki kategori yang sama, maka sistem akan lebih mudah dan konsisten dalam mengklasifikasikan data uji ke dalam kategori yang benar. Hal ini menunjukkan pentingnya distribusi data latih yang seimbang agar proses klasifikasi dapat berjalan dengan lebih optimal.

Aplikasi K-Nearest Neighbor (KNN) dalam penelitian ini dikembangkan menggunakan bahasa pemrograman C++, yang dikenal sebagai bahasa pemrograman berorientasi objek dengan performa tinggi. Untuk proses pengembangan dan implementasi, digunakan Code::Blocks versi 12.11 sebagai lingkungan pengembangannya. Code::Blocks 12.11 merupakan Integrated Development Environment (IDE) berbasis open source yang menyediakan berbagai fitur untuk mempermudah pengkodean, debugging, serta pengelolaan proyek pemrograman C++. IDE ini mendukung berbagai compiler, salah satunya adalah GNU GCC (GNU Compiler Collection), yang digunakan dalam pembuatan aplikasi KNN ini. Keunggulan utama dari penggunaan Code::Blocks 12.11 adalah fleksibilitasnya dalam menangani proyek-proyek berbasis C++ serta kemampuannya untuk berjalan di berbagai sistem operasi, termasuk Windows, Linux, dan macOS. Selain itu, karena bersifat gratis dan open source, pengguna dapat dengan mudah mengunduh dan menggunakan aplikasi ini tanpa perlu biaya lisensi. Dengan memanfaatkan Code::Blocks 12.11 serta compiler GNU GCC, aplikasi KNN yang dikembangkan dapat berjalan dengan efisien dan optimal, baik dalam melakukan perhitungan jarak antar data maupun dalam proses klasifikasi menggunakan algoritma KNN.

KESIMPULAN

Berikut hasil pengujian menggunakan 2 data uji dan menggunakan nilai $K = 1$, $K = 2$, $K = 3$, $K = 4$, $K = 5$, $K = 6$ dan $K = 7$.

Tabel 6 Hasil pengujian

| Data uji ke- | Nilai K | Kategori | Kesesuaian |
|--------------|---------|----------|------------|
| 1 | 1 | Bahaya | Sesuai |
| | 2 | Bahaya | Sesuai |

| | | | |
|--|---|------------------------|--------------|
|  | 3 | Bahaya | Sesuai |
| | 4 | Gagal mengidentifikasi | Tidak sesuai |
| | 5 | Aman | Tidak sesuai |
| | 6 | Aman | Tidak sesuai |
| | 7 | Aman | Tidak sesuai |
| 2  | 1 | Aman | Sesuai |
| | 2 | Aman | Sesuai |
| | 3 | Aman | Sesuai |
| | 4 | Aman | Sesuai |
| | 5 | Aman | Sesuai |
| | 6 | Aman | Sesuai |
| | 7 | Aman | Sesuai |

Berdasarkan hasil perhitungan dan analisis pada tabel di atas, metode K-Nearest Neighbor (KNN) menghasilkan tingkat akurasi klasifikasi sebesar 71,42% untuk data uji 1 dan akurasi 100% untuk data uji 2. Persentase akurasi ini menunjukkan bahwa model KNN yang digunakan dalam penelitian ini memiliki keterbatasan dalam mengenali dan mengklasifikasikan data dengan tepat. Salah satu faktor utama yang menyebabkan tingkat akurasi tersebut adalah ketidakseimbangan jumlah data latih, khususnya pada kategori bahaya. Jumlah data latih yang lebih sedikit dalam kategori ini mengakibatkan model kesulitan dalam mengenali pola dan karakteristik dari data berisiko tinggi/bahaya. Akibatnya, ketika parameter K (jumlah tetangga terdekat) diatur ke nilai yang lebih besar dari 3, hasil klasifikasi cenderung menjadi kurang akurat dan tidak sesuai dengan kondisi sebenarnya. Kekurangan data latih dalam kategori bahaya membuat model cenderung mengalami bias terhadap kategori aman, terutama ketika jumlah tetangga yang dipertimbangkan dalam klasifikasi meningkat. Oleh karena itu, untuk meningkatkan akurasi, diperlukan penambahan jumlah data latih pada kategori bahaya agar distribusi data lebih seimbang dan model dapat melakukan klasifikasi dengan lebih optimal.

Klasifikasi kondisi bahaya/aman menggunakan metode K-Nearest Neighbor (KNN) ini jika dibandingkan dengan metode lain seperti Artificial Neural Network (ANN) Single Perceptron dan Support Vector Machine (SVM) menghasilkan akurasi yang berbeda. Pada metode Single Perceptron menghasilkan akurasi 100% dengan bobot neuron 1 sebesar 6,47129, neuron 2 sebesar -5,93147 dan neuron 3 sebesar 0,0579922. Hal ini terjadi karena konsep pada KNN berbeda dengan ANN yaitu pembaharuan bobot pada semua neuron berdasarkan error karena selisih antara output terhadap threshold. Sedangkan pada metode SVM, garis hyperplane yang didapat yaitu $y = -1,12672x + 54,6005$ dengan akurasi 90% untuk data latih. Garis tersebut merupakan batas linear (linear boundry) yang memisahkan kondisi bahaya/aman pada data latih. Ketika diimplementasikan untuk data uji mampu mengklasifikasikan dengan akurasi 100%. Konsep KNN berbeda dengan SVM yaitu mencari titik tengah antar data latih lalu mencari nilai gradien dan konstanta c berdasarkan jarak terdekat yang berbeda kategori menggunakan teori jarak Euclidean Distance.

DAFTAR PUSTAKA

- [1] G. D. Prenata, "Klasifikasi Keandalan Sistim Distribusi Tenaga Listrik Di Pt. Pln (Persero) Up3 Surabaya Selatan Menggunakan Metode Single Perceptron," *VoteTEKNIKA*, vol. 12, no. 1, 2024.

- [2] G. D. Prenata, "Klasifikasi Keandalan Sistem Distribusi Tenaga Listrik Di Pt. Pln (Persero) Up3 Surabaya Selatan Menggunakan Support Vector Machine (Svm)," *J. Tek. Elektro*, vol. 16, no. 2, pp. 62–70, 2023, doi: 10.9744/jte.16.2.62-70.
- [3] G. D. Prenata, "Klasifikasi Keandalan Sistem Distribusi Tenaga Listrik Di Pt. Pln (Persero) Up3 Surabaya Selatan Menggunakan Metode K-Nearest Neighbor (Knn)," *J. Inform. dan Tek. Elektro Terap.*, vol. 11, no. 3s1, 2023, doi: 10.23960/jitet.v11i3s1.3397.
- [4] G. D. Prenata and N. A. Basyarach, "Prediksi Impedansi Rele menggunakan metode PSO sebagai validasi pengukuran rele," pp. 181–186, 1945.
- [5] A. Khairi, A. F. Ghozali, and A. D. N. Hidayah, "Implementasi K-Nearest Neighbor (KNN) untuk Mengklasifikasi Masyarakat Pra-Sejahtera Desa Sapikerep Kecamatan Sukapura," *TRILOGI J. Ilmu Teknol. Kesehatan, dan Hum.*, vol. 2, no. 3, pp. 319–323, 2021, doi: 10.33650/trilogi.v2i3.2878.
- [6] D. M. Nanda, T. H. Pudjiantoro, and P. N. Sabrina, "Metode K-Nearest Neighbor (KNN) dalam Memprediksi Curah Hujan di Kota Bandung," *Semin. Nas. Tek. Elektro, Sist. Informasi, dan Tek. Inform.*, pp. 387–393, 2022, [Online]. Available: <https://doi.org/10.31284/p.snestik.2022.2750>
- [7] S. R. Cholil, T. Handayani, R. Prathivi, and T. Ardianita, "Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa," *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 6, no. 2, pp. 118–127, 2021, doi: 10.31294/ijcit.v6i2.10438.

TEMPERATURE CLASSIFICATION IN HIGH VOLTAGE ELECTRICAL EQUIPMENT USING K-NEAREST NEIGHBOR

ORIGINALITY REPORT

| | | | |
|------------------|------------------|--------------|----------------|
| 14% | 11% | 6% | 3% |
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
|----|---|-----|
| 1 | journal.untar.ac.id Internet Source | 3% |
| 2 | journal.eng.unila.ac.id Internet Source | 1% |
| 3 | Submitted to Tarumanagara University Student Paper | 1% |
| 4 | digilib.uns.ac.id Internet Source | 1% |
| 5 | ojs.unsiq.ac.id Internet Source | 1% |
| 6 | pt.scribd.com Internet Source | <1% |
| 7 | Iustisia Natalia Simbolon. "PREDIKSI KUALITAS AIR SUNGAI DI JAKARTA MENGGUNAKAN KNN YANG DIOPTIMALISASI DENGAN PSO", Jurnal Informatika dan Teknik Elektro Terapan, 2024 Publication | <1% |
| 8 | jurnal.untag-sby.ac.id Internet Source | <1% |
| 9 | text-id.123dok.com Internet Source | <1% |
| 10 | Dadang Iskandar Mulyana, Siti Nurhaliza. "Klasifikasi Jenis Kucing Menggunakan Algoritma K-Nearest Neighbor", Jurnal JTİK | <1% |

(Jurnal Teknologi Informasi dan Komunikasi),
2025

Publication

11 Umar D. Akor, Muhammad Rezky Fiesta Payu, La Ode Nashar. "Penerapan Metode I-CHAID Menggunakan SMOTE pada Data Tidak Seimbang untuk Klasifikasi Durasi Studi Mahasiswa", Jambura Journal of Mathematics, 2025

Publication

12 fdocumenti.com <1 %
Internet Source

13 jtiik.ub.ac.id <1 %
Internet Source

14 repository.uin-suska.ac.id <1 %
Internet Source

15 Taufik Rahman, Abdul Ahmad Sahroni. "Penerapan K-Means untuk Pengelompokan Hasil Belajar Informatika", The Indonesian Journal of Computer Science, 2025

Publication

16 eprints.uty.ac.id <1 %
Internet Source

17 www.mdpi.com <1 %
Internet Source

18 blog.teknokrat.ac.id <1 %
Internet Source

19 etd.repository.ugm.ac.id <1 %
Internet Source

20 lib.ui.ac.id <1 %
Internet Source

| | | |
|----|--|------|
| 21 | www.nature.com Internet Source | <1 % |
| 22 | Muhammad Andi Hermawan, Ahmad Faqih, Gifthera Dwilestari. "IMPLEMENTASI AKURASI MODEL NAIVE BAYES MENGGUNAKAN SMOTE DALAM ANALISIS SENTIMEN PENGGUNA APLIKASI BRIMO", Jurnal Informatika dan Teknik Elektro Terapan, 2025 Publication | <1 % |
| 23 | es.scribd.com Internet Source | <1 % |
| 24 | de.scribd.com Internet Source | <1 % |
| 25 | ejournal.unuja.ac.id Internet Source | <1 % |
| 26 | eprints.itn.ac.id Internet Source | <1 % |
| 27 | id.scribd.com Internet Source | <1 % |
| 28 | moam.info Internet Source | <1 % |
| 29 | repositorio.fjp.mg.gov.br Internet Source | <1 % |
| 30 | rumahjurnal.or.id Internet Source | <1 % |
| 31 | www.ioinformatic.org Internet Source | <1 % |
| 32 | Fahmy Syahputra, Elsa Sabrina, Anjelita, Muhammad Ilham, Rahmansyah Angga Saputra, Zherina Br Sitepu. "Penerapan Ai Di | <1 % |

Platform LinkedIn Learning Untuk Meningkatkan Keterampilan Profesional Mahasiswa", Jurnal Pengabdian Masyarakat dan Riset Pendidikan, 2025

Publication

33

Giovanni Dimas Prenata. "KLASIFIKASI KEANDALAN SISTIM DISTRBUSI TENAGA LISTRIK DI PT. PLN (PERSERO) UP3 SURABAYA SELATAN MENGGUNAKAN METODE K-NEAREST NEIGHBOR (KNN)", Jurnal Informatika dan Teknik Elektro Terapan, 2023

Publication

<1%

34

Siti Raysyah Raysyah, Veri Arinal, Dadang Iskandar Mulyana. "KLASIFIKASI TINGKAT KEMATANGAN BUAH KOPI BERDASARKAN DETEKSI WARNA MENGGUNAKAN METODE KNN DAN PCA", JSil (Jurnal Sistem Informasi), 2021

Publication

<1%

35

myteks.files.wordpress.com

Internet Source

<1%

Exclude quotes Off

Exclude matches Off

Exclude bibliography On