

# Analisis Perbandingan Kinerja KNN Regression dan Support Vector Regression dalam Prediksi Keandalan Sistem Tenaga Listrik Berdasarkan Indeks SAIDI–SAIFI

Giovanni Dimas Prenata\*, Ahmad Ridho'i

Program Studi Teknik Elektro, Fakultas Teknologi Elektro dan Informatika Cerdas, Universitas 17 Agustus 1945  
Surabaya

Jl. Semolowaru No.45, Semolowaru, Kec. Sukolilo, Kota Surabaya, Jawa Timur 60119

E-mail: [gprenata@untagsby.ac.id](mailto:gprenata@untagsby.ac.id)\*

Naskah Masuk: 21 Februari 2026; Diterima: 06 Maret 2026; Terbit: 31 Maret 2026

## ABSTRAK

**Abstrak** - Keandalan sistem tenaga listrik merupakan indikator penting dalam menjamin kontinuitas dan kualitas layanan distribusi energi. Indeks SAIDI (*System Average Interruption Duration Index*) dan SAIFI (*System Average Interruption Frequency Index*) secara luas digunakan untuk mengukur durasi dan frekuensi gangguan, namun pendekatan evaluasi berbasis ambang batas konvensional seringkali belum mampu menangkap batas keputusan adaptif pada dataset berukuran kecil. Penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja metode *K-Nearest Neighbor* (KNN) Regression dan *Support Vector Regression* (SVR) dalam memprediksi keandalan sistem distribusi tenaga listrik berbasis indeks SAIDI–SAIFI. Sebanyak 10 data historis digunakan, dengan pembagian 8 data sebagai data latih dan 2 data sebagai data uji. Pada metode KNN Regression, variasi parameter K menunjukkan sensitivitas tinggi terhadap distribusi spasial data. Performa optimal diperoleh pada rentang  $K = 2-5$ , sedangkan pada  $K \geq 6$  terjadi penurunan akurasi akibat hilangnya sifat lokalitas (*loss of locality*) serta dominasi tetangga mayoritas. Sebaliknya, model SVR dengan parameter regularisasi  $\lambda = 0.01$ , zona insensitif  $\varepsilon = 0.03$ , dan pelatihan selama 5000 epoch menunjukkan performa yang lebih stabil. SVR menghasilkan MAE latih sebesar 0.126161, akurasi klasifikasi data latih sebesar 87.5% (7 dari 8 data benar), serta akurasi data uji sebesar 100% (2 dari 2 data benar). Koefisien model yang diperoleh, yaitu  $w_0 = -1.038554$  dan  $w_1 = 0.020590$ , menunjukkan bahwa SAIDI memiliki pengaruh dominan dan negatif terhadap skor keandalan, yang konsisten secara fisis dengan konsep durasi gangguan. Temuan ini menunjukkan bahwa pada dataset berukuran kecil, pendekatan berbasis margin seperti SVR lebih robust dan stabil dibandingkan pendekatan berbasis jarak seperti KNN Regression, sehingga berpotensi menjadi metode yang lebih andal dalam prediksi keandalan sistem distribusi tenaga listrik.

**Kata kunci:** Keandalan Sistem Tenaga Listrik, SAIDI, SAIFI, KNN Regression, Support Vector Regression

## ABSTRACT

**Abstract** - Power system reliability is a critical indicator in ensuring the continuity and quality of electricity distribution services. The System Average Interruption Duration Index (SAIDI) and System Average Interruption Frequency Index (SAIFI) are widely employed to quantify outage duration and frequency; however, conventional threshold-based evaluation approaches are often unable to capture adaptive decision boundaries from limited operational data. This study aims to analyze and compare the performance of *K-Nearest Neighbor* (KNN) Regression and *Support Vector Regression* (SVR) in predicting distribution system reliability based on SAIDI–SAIFI indices. A total of 10 historical records were utilized, consisting of 8 training samples and 2 testing samples. In the KNN Regression model, variations of parameter  $K$  reveal high sensitivity to spatial data distribution. The optimal performance was observed within the range  $K = 2-5$ , whereas for  $K \geq 6$ , prediction accuracy decreased due to loss of locality and majority neighbor dominance. In contrast, the SVR model, trained using a linear epsilon-insensitive formulation with regularization parameter  $\lambda = 0.01$ , insensitive zone  $\varepsilon = 0.03$ , and 5000 training epochs, demonstrated more stable predictive behavior. The SVR achieved a training MAE of 0.126161, training classification accuracy of 87.5% (7 out of 8 samples correctly classified), and testing accuracy of 100% (2 out of 2 samples correctly classified). The resulting model coefficients were  $w_0 = -1.038554$  and  $w_1 = 0.020590$ , indicating that SAIDI exerts a dominant and negative influence on reliability scores, which is physically consistent with outage duration theory. These findings indicate that for small-sized datasets, margin-based approaches such as SVR exhibit greater robustness and stability compared to distance-based approaches such as KNN Regression, suggesting that SVR provides a more reliable predictive framework for distribution system reliability assessment.

**Keywords:** Power System Reliability, SAIDI, SAIFI, KNN Regression, Support Vector Regression

Copyright © 2026 Jurnal Teknik Elektro dan Komputasi (ELKOM)

## 1. PENDAHULUAN

Keandalan sistem distribusi tenaga listrik merupakan salah satu indikator utama kualitas pelayanan energi listrik kepada pelanggan. Parameter yang umum digunakan untuk mengukur tingkat keandalan adalah *System Average Interruption Duration Index* (SAIDI) dan *System Average Interruption Frequency Index* (SAIFI). Kedua indeks ini merepresentasikan durasi dan frekuensi gangguan rata-rata yang dialami pelanggan dalam periode tertentu. Analisis terhadap SAIDI dan SAIFI telah banyak dilakukan untuk mengevaluasi kinerja sistem distribusi dan menentukan strategi peningkatan keandalan [1]. Kajian kuantitatif terhadap indeks tersebut menunjukkan bahwa distribusi spasial gangguan dan karakteristik beban berpengaruh langsung terhadap variasi tingkat keandalan antar wilayah [2].

Berbagai pendekatan komputasional telah diterapkan untuk melakukan prediksi maupun klasifikasi tingkat keandalan berbasis data historis SAIDI–SAIFI. Pendekatan berbasis pembelajaran mesin semakin banyak digunakan karena mampu menangkap pola nonlinier dalam data sistem tenaga [3]. Beberapa penelitian memanfaatkan teknik optimasi dan analisis reliabilitas berbasis indeks untuk mengevaluasi performa jaringan distribusi secara lebih komprehensif [4], termasuk integrasi analisis statistik dan metode komputasi numerik dalam evaluasi indeks keandalan [5].

Metode *K-Nearest Neighbor* (KNN) merupakan salah satu algoritma berbasis jarak yang relatif sederhana namun efektif dalam berbagai kasus prediksi sistem tenaga. Studi sebelumnya menunjukkan bahwa KNN mampu memberikan hasil klasifikasi dan regresi yang baik pada data berdimensi rendah [6]. Namun demikian, performa KNN sangat dipengaruhi oleh pemilihan parameter K dan distribusi spasial data latih. Pada beberapa kasus, peningkatan nilai K menyebabkan hilangnya sensitivitas terhadap karakteristik lokal data, sehingga menurunkan akurasi prediksi [7].

Di sisi lain, *Support Vector Machine* (SVM) dan pengembangannya dalam bentuk *Support Vector Regression* (SVR) menawarkan pendekatan berbasis margin dan regularisasi yang secara teoritis lebih stabil terhadap variasi distribusi data. Metode ini telah diterapkan pada berbagai studi sistem tenaga dan menunjukkan kemampuan generalisasi yang baik, terutama pada dataset berukuran kecil hingga menengah [8]. Prinsip regularisasi pada SVR memungkinkan model membangun fungsi regresi optimal dengan meminimalkan risiko overfitting, sehingga lebih robust dibandingkan pendekatan berbasis jarak murni.

Selain pendekatan berbasis SVM, metode *Artificial Neural Network* (ANN) juga telah banyak digunakan dalam prediksi konsumsi energi dan analisis sistem tenaga. Optimasi bobot ANN menggunakan algoritma metaheuristik seperti *Genetic Algorithm* (GA) dan *Particle Swarm Optimization* (PSO) terbukti mampu meningkatkan akurasi prediksi dibandingkan metode pelatihan konvensional [9]. Penelitian sebelumnya oleh penulis menunjukkan bahwa ANN yang dioptimasi menggunakan metode metaheuristik menghasilkan peningkatan kinerja yang signifikan dalam prediksi konsumsi energi listrik dibandingkan backpropagation standar [10].

Beberapa studi lain juga membandingkan berbagai algoritma pembelajaran mesin dalam konteks sistem tenaga dan menemukan bahwa tidak ada satu metode yang selalu unggul pada seluruh kondisi data [10]. Faktor distribusi spasial, jumlah sampel, serta karakteristik indeks keandalan sangat memengaruhi performa model [11] [2]. Hal ini menunjukkan pentingnya analisis komparatif yang sistematis untuk menentukan metode yang paling sesuai pada kondisi dataset tertentu.

Berdasarkan tinjauan literatur tersebut, dapat diidentifikasi adanya *research gap* pada perbandingan langsung antara pendekatan berbasis jarak (KNN Regression) dan pendekatan berbasis margin-regulasi (SVR) untuk prediksi keandalan sistem distribusi berbasis indeks SAIDI–SAIFI, khususnya pada dataset berukuran kecil. Sebagian besar penelitian berfokus pada satu metode tertentu tanpa melakukan analisis sensitivitas parameter secara mendalam atau evaluasi komparatif terhadap stabilitas model.

Oleh karena itu, penelitian ini bertujuan untuk melakukan analisis komparatif kinerja *KNN Regression* dan *Support Vector Regression* (SVR) dalam memprediksi tingkat keandalan sistem distribusi tenaga listrik berbasis parameter SAIDI dan SAIFI. Kontribusi utama penelitian ini terletak pada evaluasi sensitivitas parameter K pada KNN serta analisis stabilitas dan regularisasi pada SVR, sehingga diperoleh pemahaman yang lebih komprehensif mengenai karakteristik kedua metode dalam konteks dataset keandalan berukuran terbatas.

## 2. KAJIAN PUSTAKA

### 2.1. K-Nearest Neighbor (KNN) Regression

K-Nearest Neighbor (KNN) Regression merupakan metode pembelajaran berbasis jarak (*distance-based learning*) yang termasuk dalam kategori *instance-based learning* atau *lazy learning*. Berbeda dengan metode parametrik yang membangun model eksplisit melalui proses optimasi global, KNN Regression tidak membentuk fungsi model secara eksplisit pada tahap pelatihan, melainkan menyimpan seluruh data latih dan melakukan prediksi berdasarkan kedekatan jarak antara data uji dan data latih [12] [13].

Dalam konteks regresi, nilai prediksi suatu sampel uji ditentukan berdasarkan rata-rata atau rata-rata berbobot dari K tetangga terdekat dalam ruang fitur [12]. Jarak antara data uji  $x$  dan data latih  $x_i$  umumnya dihitung menggunakan jarak Euclidean [12]:

$$d_i = \sqrt{\sum_{j=1}^m (x_j - x_{ij})^2} \quad (1)$$

dimana:

$d_i$  = jarak antara data uji dan data latih ke- $i$

$m$  = jumlah fitur

$x_j$  = nilai fitur ke- $j$  pada data uji

$x_{ij}$  = nilai fitur ke- $j$  pada data latih ke- $i$

Pada KNN Regression sederhana, nilai prediksi dihitung sebagai rata-rata dari K tetangga terdekat [12]:

$$y_{pred} = \frac{1}{K} \sum_{i=1}^K y_i \quad (2)$$

dimana:

$y_{pred}$  = nilai prediksi untuk data uji

$K$  = jumlah tetangga terdekat yang digunakan

$y_i$  = nilai target (label) dari tetangga ke- $i$

$i$  = indeks tetangga terdekat ( $i = 1, 2, \dots, K$ )

Namun untuk meningkatkan sensitivitas terhadap kedekatan lokal, sering digunakan pendekatan *inverse distance weighting* (IDW), di mana kontribusi masing-masing tetangga diberikan bobot yang berbanding terbalik dengan jaraknya [14]:

$$w_i = \frac{1}{d_i} \quad (3)$$

dimana:

$w_i$  = bobot untuk tetangga ke- $i$

$d_i$  = jarak Euclidean antara data uji dan data latih ke- $i$

$i$  = indeks tetangga

Semakin kecil  $d_i$  maka semakin besar nilai  $w_i$ . Sehingga nilai prediksi menjadi [14]:

$$y_{pred} = \frac{\sum_{i=1}^K w_i y_i}{\sum_{i=1}^K w_i} \quad (4)$$

dimana:

$y_i$  = nilai target data latih ke- $i$

$w_i$  = bobot berbasis jarak

$y_{pred}$  = nilai prediksi kontinu

Pendekatan berbobot ini memungkinkan sampel yang lebih dekat memiliki kontribusi lebih besar dibandingkan sampel yang lebih jauh, sehingga model lebih merepresentasikan struktur lokal data.

Parameter utama dalam KNN Regression adalah nilai  $K$ , yang mengontrol keseimbangan antara bias dan variansi model [13]. Nilai  $K$  kecil membuat model sangat sensitif terhadap struktur lokal dan potensi noise, sedangkan nilai  $K$  besar menghasilkan efek perataan (*oversmoothing*) yang dapat menghilangkan karakteristik lokal distribusi data. Oleh karena itu, pemilihan nilai  $K$  menjadi faktor krusial dalam menentukan performa model, khususnya pada dataset berukuran kecil atau distribusi spasial yang tidak seragam.

Dalam aplikasi prediksi kehandalan berbasis indeks SAIDI-SAIFI, KNN Regression menghasilkan nilai kontinu dalam rentang 0–1 yang selanjutnya dapat dikonversi menjadi keputusan klasifikasi menggunakan ambang batas tertentu (misalnya 0,5). Hal ini menjadikan KNN Regression tidak hanya sebagai metode prediksi numerik, tetapi juga sebagai pendekatan klasifikasi berbasis regresi.

## 2.2. Support Vector Regression

*Support Vector Regression* (SVR) merupakan ekstensi dari metode *Support Vector Machine* (SVM) yang dirancang untuk menangani permasalahan regresi dengan prinsip structural risk

minimization [15] [16]. Prinsip ini menyeimbangkan antara minimisasi kesalahan prediksi dan pengendalian kompleksitas model, sehingga menghasilkan kemampuan generalisasi yang baik, terutama pada dataset berukuran kecil.

Secara umum, tujuan utama SVR adalah menentukan suatu fungsi regresi linier yang dapat dinyatakan sebagai [16]:

$$f(x) = w \cdot x + b \quad (5)$$

dimana:

$f(x)$  = fungsi prediksi regresi

$w$  = vektor bobot ( $w_0, w_1, \dots, w_m$ )

$x$  = vektor fitur input

$b$  = parameter bias (*intercept*)

$m$  = jumlah fitur

Fungsi tersebut dirancang untuk memodelkan hubungan fungsional antara variabel input dan output dengan batas kesalahan yang masih berada dalam tingkat toleransi tertentu.

Karakteristik utama dari SVR terletak pada penerapan fungsi kerugian  $\epsilon$ -insensitive loss, yaitu suatu fungsi yang mengabaikan kesalahan prediksi selama nilainya berada dalam rentang toleransi  $\epsilon$ , dan hanya memberikan penalti terhadap kesalahan yang melebihi batas tersebut. Fungsi kerugian ini didefinisikan sebagai [16] [17]:

$$L_\epsilon(y, f(x)) = 0, \text{ jika } |y - f(x)| \leq \epsilon \quad (6)$$

$$L_\epsilon(y, f(x)) = |y - f(x)| - \epsilon, \text{ jika } |y - f(x)| > \epsilon \quad (7)$$

dimana:

$L_\epsilon(y, f(x))$  = fungsi kerugian epsilon-insensitive

$y$  = nilai target aktual

$f(x)$  = nilai prediksi model

$\epsilon$  = parameter toleransi kesalahan

$|y - f(x)|$  = selisih absolut antara nilai aktual dan prediksi

Dengan  $y$  menyatakan nilai target aktual dan  $\epsilon$  sebagai parameter batas toleransi kesalahan. Melalui formulasi tersebut, kesalahan prediksi yang masih berada dalam interval  $\epsilon$  dipandang tidak signifikan sehingga tidak diberikan penalti. Mekanisme ini menjadikan model SVR lebih tahan (*robust*) terhadap gangguan (*noise*) maupun variasi kecil pada data.

Permasalahan optimasi dalam SVR kemudian dirumuskan dengan tujuan untuk meminimalkan fungsi objektif sebagai berikut [16]:

$$\min ( 1/2 ||w||^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) ) \quad (8)$$

dimana:

$||w||^2$  = norma Euclidean kuadrat dari vektor bobot

$C$  = parameter regularisasi

$\xi_i, \xi_i^*$  = variabel slack untuk deviasi positif dan negatif

$n$  = jumlah data latih

dengan permasalahan [16]:

$$y_i - (w \cdot x_i + b) \leq \epsilon + \xi_i \quad (9)$$

$$(w \cdot x_i + b) - y_i \leq \epsilon + \xi_i^* \quad (10)$$

$$\xi_i, \xi_i^* \geq 0 \quad (11)$$

dimana:

$y_i$  = nilai target aktual data ke- $i$

$x_i$  = vektor fitur data ke- $i$

$\epsilon$  = batas toleransi kesalahan

$\xi_i$  = variabel slack untuk kesalahan di atas margin

$\xi_i^*$  = variabel slack untuk kesalahan di bawah margin

$(w \cdot x_i + b)$  = nilai prediksi model untuk data ke- $i$

$\xi_i \geq 0$  dan  $\xi_i^* \geq 0$  adalah variabel slack tidak boleh bernilai negatif karena merepresentasikan besarnya deviasi terhadap batas toleransi.

Pada formulasi tersebut,  $\xi_i$  dan  $\xi_i^*$  merupakan variabel *slack* yang merepresentasikan deviasi atau kesalahan prediksi yang melampaui batas toleransi  $\epsilon$ . Sementara itu, parameter  $C$  berperan sebagai faktor regularisasi yang mengendalikan keseimbangan antara kompleksitas model dan besarnya kesalahan yang masih diperbolehkan [16]. Nilai  $C$  yang tinggi cenderung menghasilkan model dengan fleksibilitas lebih besar sehingga mampu menyesuaikan data pelatihan secara ketat, namun berpotensi mengalami *overfitting*. Sebaliknya, nilai  $C$  yang lebih kecil mendorong terbentuknya model yang lebih sederhana dan memiliki stabilitas generalisasi yang lebih baik.

Dalam penerapan praktis, parameter  $w$  dan  $b$  diperoleh melalui prosedur optimasi iteratif, misalnya dengan pendekatan *stochastic gradient descent* pada fungsi kerugian  $\epsilon$ -insensitive. Proses optimasi ini dijalankan hingga fungsi objektif mencapai kondisi konvergensi, yang umumnya ditandai dengan perubahan nilai kesalahan prediksi yang sangat kecil atau telah memenuhi ambang batas tertentu.

Berdasarkan karakteristik tersebut, SVR dapat dikategorikan sebagai metode regresi yang efektif dalam memodelkan hubungan antara variabel input dan output secara stabil serta presisi. Integrasi antara mekanisme regularisasi yang eksplisit dan fungsi kerugian  $\epsilon$ -insensitive menjadikan SVR sangat relevan untuk aplikasi prediksi konsumsi energi listrik, di mana data umumnya bersifat fluktuatif dan jumlah sampel relatif terbatas, tetapi tetap memerlukan tingkat akurasi yang tinggi.

**3. METODE PENELITIAN**

**3.1. K-Nearest Neighbor (KNN) Regression**

Pada model KNN Regression yang diterapkan dalam penelitian ini, data kehandalan sistem distribusi tenaga listrik direpresentasikan oleh dua fitur utama, yaitu SAIDI dan SAIFI, dengan target berupa label biner kehandalan (1 = handal dan 0 = tidak handal). Sebanyak 10 data digunakan, di mana data ke-1 hingga ke-8 diperlakukan sebagai data latih, sedangkan data ke-9 dan ke-10 digunakan sebagai data uji. Pemisahan ini dilakukan secara deterministik untuk menguji kemampuan generalisasi model terhadap data yang tidak terlibat dalam pembentukan struktur neighborhood.

Untuk menjaga konsistensi skala antar fitur, SAIDI dan SAIFI pada data latih dapat dinormalisasi menggunakan pendekatan *Min-Max*. Nilai minimum dan maksimum yang diperoleh dari data latih kemudian digunakan kembali untuk mentransformasikan data uji, sehingga proses evaluasi tetap bebas dari kebocoran informasi (*data leakage*). Dengan demikian, struktur jarak yang terbentuk benar-benar merepresentasikan distribusi spasial data latih.

Setelah tahap persiapan data, ditentukan parameter  $K$  yang merepresentasikan jumlah tetangga terdekat yang digunakan dalam proses prediksi. Pada penelitian ini, nilai  $K$  divariasikan dari 1 hingga 8 untuk mengamati sensitivitas model terhadap ukuran neighborhood. Untuk setiap data uji, jarak antara data uji dan seluruh data latih dihitung, kemudian diurutkan dari yang terkecil hingga terbesar. Proses pengurutan ini menghasilkan struktur kedekatan lokal yang menjadi dasar pembentukan prediksi.

Berbeda dengan KNN klasifikasi sederhana yang menggunakan mekanisme *voting* mayoritas, model yang digunakan dalam penelitian ini menerapkan pendekatan regresi berbobot. Setiap tetangga memberikan kontribusi terhadap nilai prediksi sesuai dengan tingkat kedekatannya terhadap data uji, sehingga tetangga yang lebih dekat memiliki pengaruh yang lebih besar dibandingkan tetangga yang lebih jauh. Pendekatan ini menghasilkan nilai prediksi kontinu dalam rentang 0 hingga 1 yang mencerminkan derajat kecenderungan kehandalan berdasarkan struktur lokal data.



Gambar 1. Diagram alir K-Nearest Neighbor (KNN) Regression

Nilai kontinu tersebut kemudian dipetakan menjadi keputusan klasifikasi menggunakan ambang batas 0.5. Skor di atas atau sama dengan 0.5 diklasifikasikan sebagai handal, sedangkan skor di bawah 0.5 diklasifikasikan sebagai tidak handal. Dengan mekanisme ini, KNN Regression berfungsi sebagai pembentuk skor keputusan berbasis kedekatan lokal, yang kemudian dikonversi menjadi keputusan klasifikasi biner.

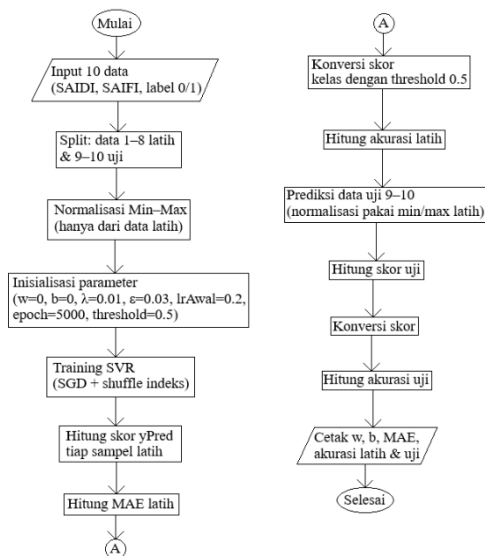
Kinerja model dievaluasi menggunakan *Mean Squared Error (MSE)* dan *Mean Absolute Error (MAE)* untuk mengukur deviasi numerik antara skor prediksi dan label aktual. Selain itu, dilakukan analisis sensitivitas terhadap variasi nilai K untuk mengidentifikasi konfigurasi parameter yang memberikan keseimbangan terbaik antara representasi lokal dan stabilitas prediksi. Seluruh nilai prediksi, keputusan akhir, serta metrik evaluasi dianalisis secara komprehensif guna menentukan rentang K optimal dalam konteks prediksi kehandalan berbasis parameter SAIDI dan SAIFI.

### 3.2. Support Vector Regression

Pada penelitian ini digunakan pendekatan *Support Vector Regression (SVR)* linear berbasis *e-insensitive loss* yang diadaptasi untuk menghasilkan skor kontinu kehandalan, kemudian dikonversi menjadi keputusan klasifikasi melalui mekanisme ambang (*threshold*). Sebanyak 10 data digunakan sebagai dataset, yang terdiri dari dua fitur utama yaitu SAIDI dan SAIFI, serta label kehandalan biner (1 = handal, 0 = tidak handal). Data dibagi menjadi dua bagian, yaitu data 1–8 sebagai data latih dan data 9–10 sebagai data uji. Sebelum proses pelatihan dilakukan, fitur SAIDI dan SAIFI pada data latih dinormalisasi menggunakan metode *Min-Max scaling* agar berada dalam rentang [0,1], sehingga perbedaan skala antar fitur tidak mempengaruhi proses optimasi parameter. Nilai minimum dan maksimum yang diperoleh dari data latih kemudian digunakan kembali untuk menormalisasi data uji guna menjaga konsistensi dan mencegah kebocoran informasi.

Model SVR diformulasikan dalam bentuk fungsi linear  $y_{pred} = w \cdot x_{norm} + b$ , dengan  $w$  sebagai vektor bobot dan  $b$  sebagai bias. Parameter awal diinisialisasi dengan  $w = 0$  dan  $b = 0$ . Proses pelatihan dilakukan menggunakan *Stochastic Gradient Descent (SGD)* dengan regularisasi L2 sebesar  $\lambda = 0.01$  untuk menjaga kestabilan model dan menghindari overfitting. Zona insensitif ditetapkan sebesar  $\epsilon = 0.03$ , yang berarti pembaruan parameter hanya dilakukan apabila nilai absolut error  $y - y_{pred}$  melebihi nilai epsilon tersebut. Laju pembelajaran awal ditetapkan sebesar 0.2 dan secara bertahap mengalami penurunan mengikuti jumlah iterasi global agar konvergensi lebih stabil. Proses training dijalankan hingga 5000 epoch dengan pengacakan indeks data latih pada setiap epoch untuk menghindari bias urutan data.

Selama proses pelatihan, setiap sampel latih dihitung nilai prediksinya, kemudian dievaluasi menggunakan *Mean Absolute Error (MAE)* untuk memantau kinerja regresi. Setelah pelatihan selesai, model menghasilkan parameter akhir berupa nilai bobot dan bias yang merepresentasikan hyperplane regresi di ruang fitur ternormalisasi. Karena output SVR berupa nilai kontinu, skor prediksi selanjutnya dikonversi menjadi kelas kehandalan menggunakan ambang batas 0.5. Skor prediksi yang lebih besar atau sama dengan 0.5 dikategorikan sebagai handal, sedangkan skor di bawah 0.5 dikategorikan sebagai tidak handal. Akurasi latih kemudian dihitung dengan membandingkan kelas hasil konversi terhadap label asli pada data 1–8.



Gambar 2. Diagram alir Support Vector Regression

Pada tahap pengujian, data 9–10 dinormalisasi menggunakan parameter Min–Max dari data latih, kemudian dihitung skor SVR menggunakan parameter  $w$  dan  $b$  yang telah diperoleh. Skor tersebut kembali dikonversi menjadi kelas menggunakan *threshold* yang sama, dan akurasi uji dihitung berdasarkan kesesuaian antara kelas prediksi dan label aktual. Hasil akhir yang ditampilkan mencakup nilai bobot akhir, bias, MAE latih, akurasi latih, serta akurasi uji, yang secara keseluruhan menggambarkan performa model SVR dalam memetakan hubungan SAIDI–SAIFI terhadap status kehandalan jaringan listrik.

#### 4. HASIL DAN PEMBAHASAN

##### 4.1. Perfoma K-Nearest Neighbor (KNN) Regression

Hasil dan pembahasan memaparkan hasil penelitian ataupun analisis yang diperoleh. Berbagai fakta serta fenomena yang dianggap penting dapat dijabarkan lebih lanjut pada bagian ini. Setelah itu, dilanjutkan dengan pembahasan secara mendalam dengan menyebutkan temuan atau kepioniran gagasan serta signifikansinya.

Metode *K-Nearest Neighbor* (KNN) Regression berbobot diterapkan untuk memprediksi kategori kehandalan sistem distribusi tenaga listrik berdasarkan dua parameter utama, yaitu SAIDI dan SAIFI. Dataset terdiri dari 8 data latih dan 2 data uji (data ke-9 dan data ke-10). Nilai  $K$  divariasikan dari 1 hingga 8 untuk mengamati sensitivitas model terhadap jumlah tetangga yang dipertimbangkan dalam proses prediksi.

Nilai prediksi dihitung menggunakan pendekatan *inverse distance weighting*:

$$y_{pred} = \frac{\sum_{i=1}^K \frac{1}{d_i} y_i}{\sum_{i=1}^K \frac{1}{d_i}} \quad (12)$$

dengan  $d_i$  adalah jarak Euclidean antara data latih dan data uji, sedangkan  $y_{i_i}$  merupakan label target (0 = tidak handal, 1 = handal). Pendekatan ini menghasilkan nilai kontinu dalam rentang 0–1, yang kemudian dikonversi menjadi keputusan klasifikasi menggunakan ambang batas 0.5. Dengan demikian, analisis ini tidak hanya menilai hasil klasifikasi, tetapi juga mengevaluasi dinamika numerik dari nilai prediksi regresi.

Tabel 1. Hasil Prediksi KNN Regression pada Variasi  $K$  (Data Uji ke-9 dan ke-10)

$K$	$y_{pred}$ (Uji-9)	Keputusan	$y_{pred}$ (Uji-10)	Keputusan	MSE	MAE
1	0.000000	Tidak Handal	0.000000	Tidak Handal	0.000000	0.000000
2	0.000000	Tidak Handal	0.000000	Tidak Handal	0.000000	0.000000
3	0.220525	Tidak Handal	0.018050	Tidak Handal	0.024478	0.119287
4	0.360091	Tidak Handal	0.035141	Tidak Handal	0.065450	0.197616
5	0.455897	Tidak Handal	0.051195	Tidak Handal	0.105231	0.253546
6	0.525920	Handal ✘	0.066382	Tidak Handal	0.140499	0.296151
7	0.579973	Handal ✘	0.081090	Tidak Handal	0.171472	0.330531
8	0.622325	Handal ✘	0.094961	Tidak Handal	0.198153	0.358643

Data uji ke-9 memiliki label aktual 0 (tidak handal). Hasil prediksi menunjukkan peningkatan progresif nilai  $y_{pred}$  seiring bertambahnya nilai  $K$ . Pada  $K=1$  dan  $K=2$ , nilai prediksi adalah 0.000000 karena dua tetangga terdekat memiliki label 0 dan bobotnya sepenuhnya mendominasi perhitungan. Ketika  $K$  meningkat menjadi 3, nilai  $y_{pred}$  naik menjadi 0.220525, kemudian meningkat menjadi 0.360091 pada  $K=4$  dan 0.455897 pada  $K=5$ . Pada rentang ini, meskipun tetangga dengan label 1 mulai masuk ke dalam perhitungan, kontribusi kumulatif bobot kelas 0 masih cukup besar sehingga nilai prediksi tetap berada di bawah ambang batas 0.5 dan klasifikasi masih sesuai dengan kondisi aktual.

Namun, perubahan signifikan terjadi pada  $K=6$ , di mana nilai  $y_{pred}$  meningkat menjadi 0.525920. Nilai ini melewati ambang batas 0.5 sehingga sistem mengubah keputusan menjadi “Handal”, yang berarti terjadi kesalahan klasifikasi. Peningkatan berlanjut pada  $K=7$  dan  $K=8$ , dengan nilai masing-masing 0.579973 dan 0.622325. Kenaikan ini bukan bersifat acak, melainkan merupakan akumulasi sistematis dari bobot kelas 1 yang semakin mendominasi ketika jumlah tetangga yang dipertimbangkan semakin banyak.

Secara struktural, setelah proses pengurutan jarak, dua tetangga terdekat memiliki label 0, sedangkan enam tetangga berikutnya memiliki label 1. Ketika  $K \leq 5$ , model masih mempertahankan karakteristik lokal karena hanya sebagian tetangga kelas 1 yang masuk dalam perhitungan. Akan tetapi, ketika  $K \geq 6$ , mayoritas tetangga dalam *neighborhood* menjadi kelas 1. Secara matematis, kondisi ini menyebabkan:

$$\sum w_i y_i > 0.5 \sum w_i \quad (13)$$

yang berarti kontribusi kelas 1 melebihi separuh total bobot. Dengan kata lain, model mengalami pergeseran dari keputusan berbasis kedekatan lokal menuju kecenderungan mayoritas dalam ruang fitur. Fenomena ini dapat dipahami sebagai hilangnya sifat lokalitas (*loss of locality*), di mana parameter K yang terlalu besar menyebabkan model tidak lagi merepresentasikan struktur lokal data uji, melainkan dipengaruhi oleh distribusi global data latihan.

Selain itu, pola peningkatan nilai prediksi menunjukkan bahwa tambahan setiap tetangga baru selalu meningkatkan nilai  $y_{pred}$ , meskipun besarnya kenaikan semakin kecil. Hal ini menunjukkan bahwa bobot tambahan dari tetangga yang lebih jauh tetap memberikan kontribusi positif terhadap kelas 1, sehingga secara bertahap mendorong prediksi melewati ambang keputusan.

Berbeda dengan data uji ke-9, data uji ke-10 menunjukkan stabilitas prediksi untuk seluruh nilai K. Nilai  $y_{pred}$  hanya meningkat dari 0.000000 pada K=1 menjadi 0.094961 pada K=8, tetap jauh di bawah ambang 0.5. Stabilitas ini disebabkan oleh keberadaan satu tetangga yang sangat dekat dengan jarak 1.800023 dan bobot sebesar 0.555548. Bobot ini jauh lebih besar dibandingkan bobot tetangga lainnya yang hanya berkisar sekitar 0.01.

Dominasi bobot tersebut menyebabkan kontribusi kelas 0 hampir sepenuhnya menentukan hasil akhir. Secara matematis dapat dinyatakan bahwa:

$$w_1 \gg w_2, w_3, \dots, w_8 \quad (14)$$

Akibatnya, meskipun nilai K diperbesar dan tetangga dengan label 1 mulai masuk ke dalam perhitungan, kontribusi kumulatifnya tidak cukup besar untuk menggeser nilai prediksi melewati ambang batas. Fenomena ini menunjukkan bahwa dalam KNN berbobot, jarak yang sangat kecil dapat menghasilkan efek dominasi bobot (*weight dominance effect*), yang menjadikan keputusan sangat stabil terhadap variasi K.

Hasil eksperimen ini menunjukkan bahwa pemilihan nilai K memiliki implikasi langsung terhadap keseimbangan antara bias dan sensitivitas model. Pada K kecil, model sangat bergantung pada beberapa tetangga terdekat sehingga keputusan benar-benar merepresentasikan struktur lokal data. Namun, risiko yang muncul adalah sensitivitas tinggi terhadap noise. Sebaliknya, pada K besar, model menjadi lebih stabil terhadap variasi kecil jarak, tetapi berpotensi mengalami bias terhadap distribusi mayoritas dalam dataset latihan.

Dalam studi ini, nilai  $K \geq 6$  menyebabkan peningkatan bias terhadap kelas 1 pada data uji ke-9, sehingga terjadi kesalahan prediksi. Hal ini menunjukkan bahwa meskipun jumlah data latihan relatif seimbang (empat kelas 0 dan empat kelas 1), distribusi spasialnya dalam ruang fitur tidak simetris. Posisi data uji ke-9 relatif lebih dekat secara kolektif terhadap kelompok kelas 1, sehingga ketika K diperbesar, efek mayoritas menjadi dominan.

Sebaliknya, data uji ke-10 berada sangat dekat dengan satu titik kelas 0, sehingga struktur lokalnya kuat dan tidak terpengaruh oleh peningkatan K. Hal ini mengindikasikan bahwa sensitivitas KNN tidak hanya bergantung pada distribusi label, tetapi juga pada distribusi geometris dalam ruang fitur.

Tabel 2. Rekomendasi Nilai K

Rentang K	Status	Keterangan
1-2	Sangat Lokal	Stabil & Akurat
3-5	Optimal	Masih benar
6-8	Over-smoothing	Terjadi Bias Mayoritas

Secara keseluruhan, hasil ini menegaskan bahwa *KNN Regression* untuk prediksi kehandalan berbasis SAIDI-SAIFI sangat dipengaruhi oleh pemilihan parameter K dan distribusi spasial data latihan. Nilai K optimal dalam eksperimen ini berada pada rentang  $2 \leq K \leq 5$ , di mana model masih mempertahankan sifat lokal tanpa mengalami dominasi mayoritas global. Kegagalan pada  $K = 6-8$  bukan merupakan kesalahan algoritmik, melainkan konsekuensi matematis dari mekanisme akumulasi bobot dalam lingkungan yang mayoritasnya memiliki label berbeda dari data uji.

Temuan ini memperlihatkan pentingnya analisis sensitivitas parameter dalam penerapan KNN Regression untuk sistem kehandalan distribusi tenaga listrik, terutama ketika jumlah data latihan terbatas dan distribusi spasial tidak seragam.

#### 4.2. Perfoma Support Vector Regression

Pada eksperimen Model SVR, model dilatih menggunakan 8 data latih (Data-1 s.d. Data-8) dengan dua fitur input SAIDI dan SAIFI, serta target biner label kehandalan (1 = handal, 0 = tidak handal). Selama proses pelatihan hingga epoch 5000, nilai *mean absolute error* (MAE) pada ruang prediksi berada pada rentang 0.144785 sampai 0.126161, dengan tren menurun. Secara spesifik, MAE tercatat 0.136018 (epoch 500), 0.137806 (epoch 1000), 0.144785 (epoch 1500), lalu membaik menjadi 0.126774 (epoch 3500) dan mencapai 0.126161 pada epoch 5000. Penurunan MAE ini menunjukkan bahwa model SVR berhasil mengurangi rata-rata deviasi absolut antara nilai target ( $Y_{true}$ ) dan skor prediksi ( $Y_{pred}$ ) walaupun target yang digunakan berupa label 0/1, karena pada Model SVR label biner diperlakukan sebagai target regresi dan kemudian dipetakan menjadi kelas menggunakan ambang keputusan (*threshold*).

Tabel 3. Perkembangan MAE Selama Proses Training model SVR

Epoch	MAE Latih	Perubahan MAE (%)
500	0.136018	-
1000	0.137806	+1.31%
1500	0.144785	+5.07%
3500	0.126774	-12.45%
5000	0.126161	-0.48%

Model akhir menghasilkan parameter SVR:  $w_0 = -1.038554$ ,  $w_1 = 0.020590$ , dan  $b = 1.029143$  dengan  $\epsilon = 0.03$  dan  $\lambda = 0.01$ , serta ambang klasifikasi *threshold* kelas = 0.5. Secara matematis, skor regresi yang dihasilkan dapat dipandang sebagai fungsi linear terhadap fitur yang telah dinormalisasi, yaitu kecenderungan skor kehandalan dihitung dari kombinasi SAIDI dan SAIFI. Tanda koefisien menunjukkan interpretasi yang konsisten dengan makna fisis indeks kehandalan: koefisien SAIDI bernilai negatif besar ( $-1.038554$ ) sehingga kenaikan SAIDI (durasi gangguan rata-rata) cenderung menurunkan skor kehandalan, sedangkan koefisien SAIFI sangat kecil dan positif (0.020590) sehingga pengaruh SAIFI terhadap skor relatif jauh lebih lemah pada model ini. Nilai bias  $b = 1.029143$  berperan sebagai penggeser keputusan, membuat skor awal cenderung berada pada area “handal” ketika kombinasi SAIDI–SAIFI masih kecil, namun dapat turun melewati ambang saat SAIDI meningkat.

Tabel 4. Parameter Akhir Model SVR

Parameter	Nilai
$w_0$ (SAIDI)	-1.038554
$w_1$ (SAIFI)	0.020590
$b$	1.029143
$\epsilon$ (epsilon)	0.03
$\lambda$ (lambda)	0.01
Threshold Klasifikasi	0.5

Berdasarkan evaluasi pada data latih (1–8), model mencapai akurasi klasifikasi 87.5%, yang berarti 7 dari 8 data diklasifikasikan benar ketika skor SVR dikonversi menjadi kelas menggunakan ambang 0.5. Pada output ditampilkan bahwa Data-1 memiliki  $Y_{true} = 0$  namun skor prediksi  $Y_{pred} = 0.857431$  sehingga diputuskan sebagai kelas = 1 (handal); ini merupakan satu kasus yang menyebabkan akurasi latih tidak mencapai 100%. Sementara itu, pada data lain model konsisten menghasilkan skor yang sesuai dengan label: misalnya Data-2 ( $Y_{true} = 1$ ) diprediksi dengan skor 1.008191 (kelas 1), Data-3 ( $Y_{true} = 1$ ) skor 1.008275 (kelas 1), Data-4 ( $Y_{true} = 1$ ) skor 0.991297 (kelas 1), dan Data-6 ( $Y_{true} = 0$ ) skor sangat rendah 0.001788 sehingga diputuskan sebagai kelas 0 (tidak handal). Temuan ini memperlihatkan bahwa Model SVR tidak hanya “menghafal” data, tetapi belajar membentuk pemisah berbasis skor yang cukup stabil; adanya satu kesalahan pada latih (akurasi 87.5%) juga bisa dipahami sebagai indikasi model tidak dipaksa overfit sempurna, terutama karena pendekatan yang dipakai adalah regresi + ambang.

Pada tahap pengujian, Model SVR dievaluasi pada data uji (9–10) dan memperoleh akurasi uji 100% (2 dari 2 benar). Data-9 (SAIDI = 173.9, SAIFI = 0.66) menghasilkan skor SVR  $-0.642222$  dan diputuskan sebagai kelas  $Pred = 0$ , sesuai dengan label asli = 0. Data-10 (SAIDI = 105.8, SAIFI = 0.85) menghasilkan skor 0.018993 dan juga diputuskan kelas  $Pred = 0$ , kembali sesuai dengan label

asli. Dua hasil uji ini konsisten dengan logika domain: SAIDI yang sangat tinggi (173.9 maupun 105.8) mencerminkan durasi gangguan yang besar sehingga wajar dipetakan sebagai “tidak handal”. Secara keseluruhan, kombinasi MAE latih ~0.126, akurasi latih 87.5%, dan akurasi uji 100% pada data 9–10 menunjukkan bahwa Model SVR telah bekerja sebagai pendekatan SVR-klasifikasi berbasis skor yang mampu memodelkan relasi SAIDI–SAIFI terhadap label kehandalan dengan performa yang cukup kuat pada dataset kecil yang digunakan

#### 4.3. Perbandingan Kinerja KNN Regression dan Support Vector Regression

Penelitian ini membandingkan dua pendekatan pembelajaran mesin untuk prediksi kehandalan sistem tenaga listrik berbasis indeks SAIDI dan SAIFI, yaitu *KNN Regression* dan *Support Vector Regression* (SVR). *KNN Regression* bekerja berdasarkan prinsip kedekatan jarak antar sampel, sedangkan SVR membangun fungsi regresi berbasis margin dengan regularisasi.

Pada eksperimen KNN Regression yang telah dilakukan sebelumnya, diperoleh bahwa pemilihan parameter K sangat mempengaruhi kinerja model. Untuk nilai K kecil (misalnya K=1 hingga K=3), model masih mampu mengikuti pola distribusi data latih dengan baik. Namun ketika K ditingkatkan menjadi K=6 hingga K=8, performa model menurun dan menghasilkan prediksi yang tidak sesuai dengan label kehandalan aktual. Hal ini menunjukkan bahwa pada dataset berukuran kecil, peningkatan nilai K menyebabkan efek *oversmoothing*, sehingga model kehilangan sensitivitas terhadap karakteristik lokal distribusi SAIDI–SAIFI. Secara konseptual, semakin besar K, model semakin mendekati rata-rata global, sehingga titik-titik ekstrem (misalnya SAIDI sangat tinggi) tidak lagi terpisah secara tegas.

Tabel 5. Perbandingan Kuantitatif KNN Regression dan SVR

Metode	Parameter Utama	Nilai Parameter	MAE Latih	Akurasi Latih	Akurasi Uji	Stabilitas
KNN	K	1	0.000000	100%	100%	Sangat sensitif terhadap noise
KNN	K	2	0.000000	100%	100%	Stabil
KNN	K	3	0.119287	100%	100%	Baik
KNN	K	4	0.197616	100%	100%	Baik
KNN	K	5	0.253546	100%	100%	Mulai oversmoothing
KNN	K	6	0.296151	Turun	Salah prediksi	Tidak stabil
KNN	K	7	0.330531	Turun	Salah prediksi	Tidak stabil
KNN	K	8	0.358643	Turun	Salah prediksi	Tidak stabil
SVR	$\lambda = 0.01$ $\varepsilon = 0.03$	5000 epoch	0.126161	87.5%	100%	Stabil

Berbeda dengan *KNN Regression* yang menggunakan parameter K sebagai jumlah tetangga terdekat, *Support Vector Regression* (SVR) tidak memiliki parameter K. SVR membangun fungsi regresi berbasis margin dan regularisasi, dengan parameter utama berupa  $\lambda$  (*regularization parameter*) dan  $\varepsilon$  (*epsilon-insensitive zone*). Oleh karena itu, perbandingan parameter K pada KNN tidak memiliki padanan langsung pada SVR.

Hasil eksperimen menunjukkan bahwa KNN sangat sensitif terhadap variasi K, di mana nilai optimal berada pada rentang K = 2–5. Ketika  $K \geq 6$ , terjadi peningkatan MAE hingga 0.358643 serta kesalahan klasifikasi pada data uji ke-9. Sebaliknya, SVR dengan  $\lambda = 0.01$  dan  $\varepsilon = 0.03$  menunjukkan performa yang lebih stabil dengan MAE latih 0.126161 dan akurasi uji 100%. Hal ini mengindikasikan bahwa pendekatan berbasis margin lebih robust terhadap variasi distribusi spasial data dibandingkan pendekatan berbasis jarak.

Sebaliknya, metode SVR menunjukkan karakteristik yang berbeda. Dengan parameter  $\lambda = 0.01$  dan  $\varepsilon = 0.03$ , serta pelatihan hingga 5000 epoch, model menghasilkan MAE latih stabil pada kisaran 0.126–0.144. Akurasi klasifikasi berbasis threshold 0.5 pada data latih mencapai 87.5% (7 dari 8 data benar), sementara pada data uji (data 9–10) diperoleh akurasi 100%. Secara parameter, model menghasilkan bobot  $w_0 = -1.038554$  dan  $w_1 = 0.020590$ , yang menunjukkan bahwa SAIDI memiliki pengaruh dominan terhadap skor kehandalan dibandingkan SAIFI. Koefisien negatif pada SAIDI konsisten secara fisis, karena peningkatan durasi gangguan memang menurunkan tingkat kehandalan.

Perbedaan utama antara kedua metode terlihat pada stabilitas model terhadap perubahan parameter. *KNN Regression* sangat sensitif terhadap pemilihan K (pada K=6–8 model gagal merepresentasikan pola data dengan benar). Sebaliknya, SVR lebih stabil karena menggunakan

prinsip regularisasi dan optimasi berbasis margin, sehingga pembentukan fungsi keputusan tidak bergantung langsung pada rata-rata tetangga lokal. Pada dataset kecil seperti penelitian ini (10 data), pendekatan *margin-based* SVR cenderung lebih *robust* dibanding pendekatan *distance-based* KNN ketika distribusi data tidak merata.

Namun demikian, perlu dicatat bahwa akurasi uji 100% pada SVR diperoleh hanya dari dua data uji, sehingga belum cukup untuk menyimpulkan generalisasi model secara luas. Dalam konteks dataset kecil, baik *KNN Regression* maupun SVR menunjukkan performa yang baik, tetapi SVR memberikan interpretasi parameter yang lebih jelas dan stabil secara matematis.

## 5. KESIMPULAN

Penelitian ini membandingkan metode *KNN Regression* dan *Support Vector Regression* (SVR) untuk prediksi kehandalan sistem tenaga listrik berbasis indeks SAIDI dan SAIFI. Hasil eksperimen menunjukkan bahwa *KNN Regression* sangat dipengaruhi oleh pemilihan parameter K, di mana peningkatan K hingga 6–8 menyebabkan penurunan ketepatan prediksi akibat efek perataan berlebihan pada dataset berukuran kecil. Sementara itu, metode SVR dengan parameter  $\lambda = 0.01$  dan  $\varepsilon = 0.03$  menghasilkan MAE latih sekitar 0.126 serta akurasi klasifikasi latih sebesar 87.5%, dan mampu mencapai akurasi 100% pada dua data uji yang tersedia.

Secara metodologis, SVR menunjukkan kestabilan yang lebih baik dibanding *KNN Regression* karena memanfaatkan prinsip margin dan regularisasi dalam pembentukan fungsi keputusan. Koefisien model SVR juga memberikan interpretasi kuantitatif yang konsisten dengan makna fisik indeks kehandalan, di mana SAIDI memiliki pengaruh dominan terhadap skor keputusan. Oleh karena itu, pada dataset kecil berbasis SAIDI–SAIFI seperti pada penelitian ini, pendekatan SVR dapat dianggap lebih *robust* dibanding *KNN Regression*, meskipun pengujian dengan dataset yang lebih besar tetap diperlukan untuk validasi generalisasi model.

## REFERENSI

- [1] I. Hajar and M. H. Pratama, "Tenaga Listrik Pada Penyulang Cahaya PT . PLN ( Persero )," *J. Ilm.*, vol. 10, no. 1, pp. 70–77, 2018.
- [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] I. A. Lubis, D. Erivianto, and Z. Tharo, "Analysis Of Power Distribution System Reliability Using System Average Interruption Duration Index ( SAIDI) and System Average Interruption Frequency Index( SAIFI) On Feeder KR04," *Int. J. Econ. Technol. Soc. Sci.*, vol. 5, no. 2, pp. 211–223, 2024.
- [4] D. Wahyudi, "Evaluasi Keandalan Sistem Distribusi Tenaga Listrik Berdasarkan SAIDI Dan SAIFI Pada PT. PLN (Persero) Rayon Kakap," *J. Tek. Elektro*, vol. 1, no. 8, pp. 1–7, 2016.
- [5] M. Asrul Sani, A. Pengalaman Tarigan, and S. Anisah, "Optimization and Reliability Analysis Based on Saidi, Saifi Values on 20KV Distribution System Channels at PT. PLN ULP Labuhan Haji," vol. 2, no. 1, pp. 298–308, 2025, [Online]. Available: <https://proceeding.pancabudi.ac.id/index.php/ICDSET/>
- [6] N. Lembang, Y. N. Nantan, R. Rusliadi, N. Husnah, and Y. La Elo, "Analysis of Saidi and Saifi Stability in the Distribution System At Pt. Pln (Persero) Ulp Fakfak," *J. Media Elektr.*, vol. 21, no. 3, pp. 191–202, 2024, doi: 10.59562/metrik.v21i3.3501.
- [7] R. Adolph, "Analisis Kehandalan Sistem Penyaluran Energi Listrik Pada Penyulang Bj-02 Pada Pt. Pln (Persero) Ulp Kuala," *Syntax Lit. Indones. Sci. J.*, vol. 7, no. 9, pp. 1–23, 2022.
- [8] D. Sorokin, "Machine learning application for power systems reliability assessment," *E3S Web Conf.*, vol. 461, pp. 1–5, 2023, doi: 10.1051/e3sconf/202346101032.
- [9] G. D. Prenata and N. A. Basyarach, "Prediksi Impedansi Rele menggunakan metode PSO sebagai validasi pengukuran rele," pp. 181–186, 1945.
- [10] A. R. Prenata G D, "EMITOR : Jurnal Teknik Elektro," vol. Vol 25, No, pp. 39–46, 2025, doi: <https://doi.org/10.23917/emitor.v25i3.12719>.
- [11] G. D. Prenata, "Klasifikasi Keandalan Sistem Distrbusi 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.
- [12] David W. Aha, Dennis Kibler, and M. K. Albert, "Instance-Based Learning Algorithms DAVID," *Mach. Learn.*, vol. 6, no. Springer, pp. 37–66, 1991.
- [13] T. M. Cover and P. E. Hart, "Nearest Neighbor Pattern Classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, 1967, doi: 10.1109/TIT.1967.1053964.
- [14] D. Shepard, "A two-dimensional interpolation function for irregularly-spaced data," *Proc. 1968 23rd ACM Natl. Conf. ACM 1968*, no. January 1968, pp. 517–524, 1968, doi: 10.1145/800186.810616.

- [15] V. Vapnik, "The Nature of Statistical Learning Theory, Second Edition," 1995.
- [16] H. Drucker, C. J. C. Surges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," *Adv. Neural Inf. Process. Syst.*, no. x, pp. 155–161, 1997.
- [17] S. Sande and M. L. Privalsky, "Identification of TRACs (T3 receptor-associating cofactors), a family of cofactors that associate with, and modulate the activity of, nuclear hormone receptors.," *Mol. Endocrinol.*, vol. 10, no. 7, pp. 813–825, 1996, doi: 10.1210/mend.10.7.8813722.