AnoaTIK: Jurnal Teknologi Informasi dan Komputer e-ISSN: **2987-7652**
Vol. 4, No. 1, Juni 2026 Hal. 38-44

KLASIFIKASI INDEKS PEMBANGUNAN MANUSIA DENGAN METODE K-NEAREST NEIGHBOR DAN SUPPORT VECTOR MACHINE DI SULAWESI TENGGARA

Gusti Arviana Rahman[1], Sulfikram[*2]
[1,2]Universitas Halu Oleo
Email: [1]arviana.rahman@uho.ac.id , [2]sulfikram360@gmail.com
***** *Penulis Korenspondensi*

Abstrak

Indeks Pembangunan Manusia (IPM) merupakan indikator yang digunakan untuk menggambarkan kualitas pembangunan manusia melalui dimensi kesehatan, pendidikan, dan standar hidup yang layak. Perbedaan kondisi sosial, ekonomi, dan pembangunan antarwilayah menyebabkan variasi nilai IPM sehingga diperlukan metode yang mampu mengelompokkan wilayah berdasarkan kategori IPM secara tepat. Penelitian ini bertujuan untuk membandingkan kinerja algoritma *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM) dalam mengklasifikasikan kategori IPM di Provinsi Sulawesi Tenggara. Data yang digunakan terdiri atas 17 kabupaten/kota dengan atribut Angka Harapan Hidup (AHH), Harapan Lama Sekolah (HLS), Rata-rata Lama Sekolah (RLS), dan Pengeluaran Per Kapita (PPP). Tahapan penelitian meliputi preprocessing data, pembentukan kategori IPM, normalisasi data, pembagian data menjadi data latih dan data uji, serta proses klasifikasi menggunakan algoritma KNN dan SVM. Pengujian dilakukan pada KNN dengan nilai K=3, K=5, dan K=7, sedangkan SVM menggunakan kernel linear dan kernel Radial Basis Function (RBF). Kinerja model dievaluasi menggunakan confusion matrix dan nilai akurasi. Hasil penelitian menunjukkan bahwa algoritma KNN dengan nilai K=3 memperoleh akurasi tertinggi sebesar 100%, sedangkan SVM dengan kernel linear dan kernel RBF masing-masing memperoleh akurasi sebesar 83,33%. Berdasarkan hasil tersebut, algoritma KNN menunjukkan performa yang lebih baik dibandingkan SVM dalam mengklasifikasikan kategori Indeks Pembangunan Manusia di Provinsi Sulawesi Tenggara.

Kata kunci: *Indeks Pembangunan Manusia, K-Nearest Neighbor, Support Vector Machine, klasifikasi, machine learning.*

Abstract

The Human Development Index (HDI) is an indicator used to measure the quality of human development based on the dimensions of health, education, and a decent standard of living. Differences in social, economic, and development conditions across regions result in variations in HDI values, making it necessary to apply an appropriate method for classifying HDI categories. This study aims to compare the performance of the K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) algorithms in classifying HDI categories in Southeast Sulawesi Province. The dataset used consists of 17 regencies/cities in Southeast Sulawesi, with Life Expectancy (AHH), Expected Years of Schooling (HLS), Mean Years of Schooling (RLS), and Per Capita Expenditure (PPP) as predictor variables. The research stages include data preprocessing, HDI category formation, data normalization, dataset splitting into training and testing sets, and classification using KNN and SVM algorithms. The KNN algorithm was tested using K values of 3, 5, and 7, while SVM was implemented using linear and Radial Basis Function (RBF) kernels. Model performance was evaluated using a confusion matrix and accuracy metrics. The results showed that KNN with K=3 achieved the highest accuracy of 100%, while SVM with linear and RBF kernels each achieved an accuracy of 83.33%. These findings indicate that the KNN algorithm outperformed SVM in classifying Human Development Index categories in Southeast Sulawesi Province.

Keywords: *Human Development Index, K-Nearest Neighbor, Support Vector Machine, classification, machine learning.*

1. PENDAHULUAN

Indeks Pembangunan Manusia (IPM) merupakan indikator yang digunakan untuk mengukur tingkat keberhasilan pembangunan manusia berdasarkan tiga dimensi utama, yaitu kesehatan, pendidikan, dan standar hidup layak. IPM diperkenalkan oleh United Nations Development Programme (UNDP) sebagai ukuran untuk menggambarkan kualitas hidup masyarakat melalui kemampuan memperoleh akses terhadap layanan kesehatan, pendidikan, dan kebutuhan ekonomi. Dalam konteks pembangunan nasional, IPM menjadi salah satu indikator strategis yang digunakan untuk mengevaluasi keberhasilan pembangunan daerah serta mendukung proses perencanaan dan pengambilan kebijakan pemerintah. Menurut Badan

Pusat Statistik (BPS), IPM diklasifikasikan ke dalam empat kategori yaitu rendah (IPM < 60), sedang (60 ≤ IPM < 70), tinggi (70 ≤ IPM < 80), dan sangat tinggi (IPM ≥ 80)[1]. Selain itu, UNDP menegaskan bahwa HDI merupakan ukuran komposit pembangunan manusia yang merepresentasikan dimensi umur panjang dan sehat, pengetahuan, serta standar hidup layak [2].

Nilai Indeks Pembangunan Manusia (IPM) di setiap wilayah Indonesia menunjukkan perbedaan yang dipengaruhi oleh berbagai faktor, seperti kondisi sosial, ekonomi, pendidikan, dan kesehatan. Ketimpangan pembangunan antar daerah menyebabkan kualitas sumber daya manusia juga bervariasi pada setiap provinsi maupun kabupaten/kota. Oleh karena itu, diperlukan suatu metode yang dapat mengelompokkan wilayah berdasarkan kategori IPM guna memberikan informasi yang mendukung proses evaluasi dan penyusunan strategi pembangunan yang lebih terarah[3].

Kemajuan teknologi informasi yang didukung oleh perkembangan machine learning telah meningkatkan pemanfaatan berbagai metode analisis data, salah satunya adalah klasifikasi. Klasifikasi merupakan teknik dalam data mining yang bertujuan untuk mengelompokkan data ke dalam kategori tertentu berdasarkan karakteristik yang dimilikinya. Pada pendekatan *supervised learning*, proses klasifikasi dilakukan menggunakan data yang telah memiliki label sehingga model dapat mempelajari pola dari data tersebut dan menggunakannya untuk memprediksi kelas pada data baru[4].

Salah satu algoritma klasifikasi yang banyak diterapkan adalah K-Nearest Neighbor (KNN). Metode ini mengelompokkan suatu objek berdasarkan kedekatannya dengan sejumlah data tetangga terdekat pada data latih. Tingkat kedekatan antar data umumnya dihitung menggunakan metode Euclidean Distance, sehingga kelas suatu objek ditentukan berdasarkan mayoritas kelas dari tetangga terdekatnya. KNN banyak digunakan karena memiliki konsep yang sederhana, mudah diterapkan, serta mampu memberikan hasil klasifikasi yang baik pada berbagai jenis data[5].

Selain KNN, metode Support Vector Machine (SVM) juga banyak digunakan dalam berbagai penelitian klasifikasi. SVM bekerja dengan membentuk hyperplane optimal yang mampu memisahkan data dari kelas yang berbeda dengan margin maksimum. Metode ini memiliki kemampuan yang baik dalam menangani data berdimensi tinggi dan sering menghasilkan tingkat akurasi yang tinggi pada berbagai kasus klasifikasi sehingga banyak diterapkan dalam penelitian machine learning[6].

Beberapa penelitian terdahulu telah menerapkan metode machine learning dalam proses klasifikasi Indeks Pembangunan Manusia. Penelitian Pamungkas dan Widiyanto menggunakan metode Support Vector Machine untuk mengklasifikasikan IPM di Indonesia tahun 2022 dan menunjukkan bahwa metode tersebut mampu menghasilkan performa klasifikasi yang baik. Penelitian Pratiwi dan Wijayanto membandingkan metode K-Nearest *Neighbor* dan *Support Vector Machine* dalam klasifikasi IPM kabupaten/kota di Pulau Jawa serta memperoleh hasil bahwa metode SVM menghasilkan akurasi sebesar 88,89% dan memiliki performa yang lebih baik dibandingkan KNN. Selain itu, penelitian Safitri menggunakan metode KNN untuk mengklasifikasikan IPM Provinsi Sumatera Selatan dengan variabel Umur Harapan Hidup (UHH), Rata-rata Lama Sekolah (RLS), Harapan Lama Sekolah (HLS), dan Pengeluaran Per Kapita. Penelitian lainnya menunjukkan bahwa metode machine learning seperti KNN, SVM, Random Forest, dan XGBoost mampu menghasilkan tingkat akurasi yang tinggi dalam klasifikasi IPM di Indonesia[7].

Meskipun penelitian mengenai klasifikasi Indeks Pembangunan Manusia (IPM) telah banyak dilakukan dengan berbagai metode machine learning, kajian yang secara khusus membahas wilayah Provinsi Sulawesi Tenggara masih belum banyak ditemukan. Padahal, setiap kabupaten dan kota di Sulawesi Tenggara memiliki karakteristik pembangunan yang berbeda, baik dari aspek pendidikan, kesehatan, maupun ekonomi, sehingga menghasilkan tingkat IPM yang beragam. Kondisi tersebut menjadikan klasifikasi kategori IPM penting untuk dilakukan guna memperoleh gambaran mengenai pola pembangunan manusia di setiap wilayah. Oleh karena itu, penelitian ini bertujuan untuk mengklasifikasikan kategori IPM di Provinsi Sulawesi Tenggara menggunakan metode K-Nearest Neighbor (KNN) dan Support Vector Machine (SVM). Variabel yang digunakan dalam penelitian meliputi Angka Harapan Hidup (AHH), Harapan Lama Sekolah (HLS), Rata-rata Lama Sekolah (RLS), dan Pengeluaran Per Kapita (PPP). Selanjutnya, kinerja kedua metode dievaluasi dan dibandingkan berdasarkan nilai akurasi serta confusion matrix untuk menentukan metode yang memiliki performa terbaik dalam mengklasifikasikan kategori IPM.

2. METODE PENELITIAN

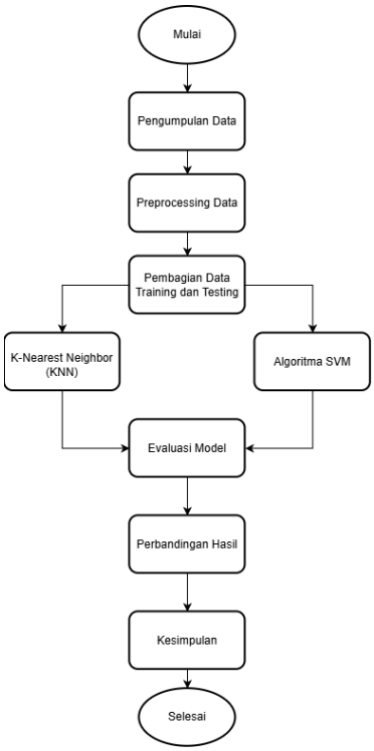
2.1. Data Penelitian

Penelitian ini memanfaatkan data sekunder yang bersumber dari publikasi Badan Pusat Statistik (BPS) Provinsi Sulawesi Tenggara. Data yang digunakan mencakup indikator-indikator penyusun Indeks Pembangunan Manusia (IPM) pada 17 kabupaten/kota di Provinsi Sulawesi Tenggara. Pemilihan indikator tersebut didasarkan pada perannya sebagai komponen utama dalam perhitungan IPM yang dapat merepresentasikan tingkat pembangunan manusia di suatu wilayah. Rujukan data juga disesuaikan dengan publikasi IPM Provinsi Sulawesi Tenggara 2024 yang memuat perkembangan indikator IPM kabupaten/kota [8].

Variabel yang digunakan dalam penelitian ini terdiri atas Angka Harapan Hidup (AHH), Harapan Lama Sekolah (HLS), Rata-rata Lama Sekolah (RLS), dan Pengeluaran Per Kapita (PPP) yang berfungsi sebagai variabel independen atau atribut prediktor. Adapun variabel dependen yang digunakan adalah kategori Indeks Pembangunan Manusia (IPM).

Dalam penelitian ini, kategori IPM dibagi menjadi dua kelompok, yaitu IPM rendah dan IPM tinggi. Pengelompokan tersebut didasarkan pada nilai batas (threshold) IPM sebesar 70. Kabupaten/kota dengan nilai IPM di bawah 70 dimasukkan ke dalam kategori rendah, sedangkan wilayah yang memiliki nilai IPM sebesar 70 atau lebih diklasifikasikan ke dalam kategori tinggi. Pembentukan dua kategori ini bertujuan untuk menyederhanakan proses klasifikasi serta menciptakan distribusi data yang lebih seimbang pada dataset penelitian.

2.2. Tahapan Penelitian



***Gambar 1.1** Flowchart Penelitian*

Tahapan penelitian yang dilakukan pada penelitian ini ditunjukkan pada Gambar 1.1. Penelitian diawali dengan proses pengumpulan data yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Sulawesi Tenggara. Data yang digunakan terdiri atas indikator-indikator yang berkaitan dengan Indeks Pembangunan Manusia (IPM), yaitu Angka Harapan Hidup (AHH), Harapan Lama Sekolah (HLS), Rata-rata Lama Sekolah (RLS), dan Pengeluaran Per Kapita (PPP).

Setelah proses pengumpulan data, dilakukan tahap preprocessing data untuk mempersiapkan data sebelum digunakan dalam proses klasifikasi. Tahap ini meliputi pembentukan kategori IPM, transformasi data kategorik menjadi data numerik, normalisasi data, serta pembagian data menjadi data latih dan data uji.

Selanjutnya, proses klasifikasi dilakukan menggunakan dua metode machine learning, yaitu K-Nearest Neighbor (KNN) dan Support Vector Machine (SVM). Kedua metode digunakan untuk mengklasifikasikan kategori IPM berdasarkan atribut yang telah ditentukan. Hasil klasifikasi dari masing-masing metode kemudian dievaluasi untuk mengetahui tingkat kinerjanya.

Tahap evaluasi model dilakukan menggunakan confusion matrix dan nilai akurasi. Hasil evaluasi dari metode KNN dan SVM selanjutnya dibandingkan untuk menentukan metode yang memiliki performa terbaik dalam mengklasifikasikan kategori Indeks Pembangunan Manusia di Provinsi Sulawesi Tenggara. Tahap terakhir adalah penyusunan kesimpulan berdasarkan hasil pengujian dan analisis yang telah dilakukan.

2.3. Preprocessing Data

Tahap preprocessing merupakan salah satu tahapan penting yang dilakukan sebelum proses klasifikasi untuk memastikan data yang digunakan telah siap dan sesuai untuk dianalisis. Pada tahap ini dilakukan beberapa proses, meliputi pembentukan label kategori IPM, konversi data kategorik ke dalam bentuk numerik, normalisasi data, serta pembagian dataset menjadi data latih dan data uji**.** Penggunaan tahapan preprocessing dan pemodelan pada penelitian ini selaras dengan praktik implementasi algoritma machine learning yang tersedia dalam pustaka scikit-learn [9].

Label kategori IPM yang awalnya berupa data kategorik dikonversi menjadi data numerik menggunakan metode label encoding. Proses konversi ini diperlukan karena algoritma *machine learning* hanya dapat mengolah data dalam bentuk numerik dan tidak dapat memproses data teks secara langsung. Setelah itu, dilakukan normalisasi data menggunakan metode StandardScaler. Tujuan normalisasi adalah untuk menyeragamkan skala setiap atribut sehingga perbedaan rentang nilai antarvariabel tidak menyebabkan suatu atribut memiliki pengaruh yang lebih dominan dibandingkan atribut lainnya.

Setelah seluruh data dinormalisasi, dataset kemudian dibagi menjadi dua bagian, yaitu data latih (*training data*) dan data uji (*testing data*) menggunakan metode train-test split. Proporsi pembagian yang digunakan adalah 70% untuk data latih dan 30% untuk data uji. Data latih dimanfaatkan untuk membangun dan melatih model klasifikasi, sedangkan data uji digunakan untuk mengukur kinerja model dalam mengklasifikasikan data baru yang tidak terlibat pada proses pelatihan sebelumnya.

2.4. Metode K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) merupakan salah satu algoritma klasifikasi yang termasuk ke dalam kelompok supervised learning. Prinsip kerja KNN didasarkan pada asumsi bahwa data yang memiliki karakteristik serupa cenderung berada dalam kelas yang sama. Oleh karena itu, proses klasifikasi dilakukan dengan mencari sejumlah data tetangga terdekat dari suatu data uji berdasarkan ukuran jarak tertentu. Secara konseptual, dasar metode nearest neighbor diperkenalkan oleh Cover dan Hart melalui pendekatan klasifikasi berdasarkan kedekatan pola[10].

Pada penelitian ini, ukuran jarak yang digunakan adalah Euclidean Distance karena merupakan metode yang paling umum digunakan dalam algoritma KNN. Jarak antara dua data dihitung menggunakan Persamaan (1).

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}, \quad (1)$$

di mana:
$(d(x,y))$ = jarak antara dua data
(x_i) = nilai atribut ke-i pada data pertama
(y_i) = nilai atribut ke-i pada data kedua
(n) = jumlah atribut

Setelah seluruh jarak dihitung, algoritma akan memilih sejumlah tetangga terdekat sesuai nilai K yang telah ditentukan. Kelas yang paling banyak muncul pada tetangga tersebut akan digunakan sebagai hasil klasifikasi data uji. Pada penelitian ini dilakukan pengujian menggunakan nilai K=3, K=5, dan K=7 untuk mengetahui nilai K yang menghasilkan performa klasifikasi terbaik.

2.5. Metode Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan algoritma klasifikasi supervised learning yang bekerja dengan membentuk hyperplane optimal sebagai batas pemisah antar kelas. Tujuan utama SVM adalah menemukan hyperplane yang memiliki margin maksimum sehingga model mampu membedakan kelas data secara lebih optimal dan memiliki kemampuan generalisasi yang baik terhadap data baru. Formulasi awal SVM dikembangkan oleh Cortes dan Vapnik sebagai metode klasifikasi berbasis margin maksimum [11].

SVM memanfaatkan support vector, yaitu titik-titik data yang berada paling dekat dengan hyperplane. Titik-titik tersebut berperan penting dalam menentukan posisi hyperplane yang optimal. Semakin besar margin yang terbentuk antara dua kelas, maka semakin baik kemampuan model dalam melakukan klasifikasi.

Pada penelitian ini digunakan dua jenis kernel, yaitu kernel linear dan kernel Radial Basis Function (RBF). Kernel linear digunakan untuk menangani data yang dapat dipisahkan secara linear, sedangkan kernel RBF digunakan untuk menangani data yang memiliki pola nonlinier. Penggunaan kedua kernel tersebut bertujuan untuk membandingkan performa SVM pada karakteristik data yang berbeda sehingga dapat diketahui kernel yang memberikan hasil klasifikasi terbaik.

2.6. Evaluasi Model

Evaluasi model dilakukan untuk mengetahui tingkat kinerja metode K-Nearest Neighbor dan Support Vector Machine dalam mengklasifikasikan kategori Indeks Pembangunan Manusia. Pada penelitian ini evaluasi dilakukan menggunakan confusion matrix dan nilai akurasi.

Confusion matrix digunakan untuk menunjukkan jumlah prediksi yang benar maupun salah pada setiap kelas. Melalui confusion matrix dapat diketahui jumlah True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Informasi tersebut digunakan untuk menilai kemampuan model dalam membedakan data pada masing-masing kategori.

Selain confusion matrix, penelitian ini juga menggunakan nilai akurasi sebagai ukuran utama performa model. Akurasi menunjukkan persentase jumlah prediksi yang benar terhadap seluruh data yang diuji. Nilai akurasi dihitung menggunakan Persamaan (2).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (2)$$

Model dengan nilai akurasi tertinggi dianggap memiliki performa terbaik dalam mengklasifikasikan kategori Indeks Pembangunan Manusia di Provinsi Sulawesi Tenggara. Hasil evaluasi dari metode KNN dan SVM kemudian dibandingkan untuk menentukan metode yang paling sesuai digunakan pada dataset penelitian.

3. HASIL DAN PEMBAHASAN

3.1. Deskripsi Dataset

Dataset yang digunakan dalam penelitian ini terdiri atas 17 kabupaten/kota di Provinsi Sulawesi Tenggara. Data diperoleh dari publikasi Badan Pusat Statistik (BPS) dan mencakup empat variabel yang digunakan sebagai atribut prediktor, yaitu Angka Harapan Hidup (AHH), Harapan Lama Sekolah (HLS), Rata-rata Lama Sekolah (RLS), dan Pengeluaran Per Kapita (PPP). Variabel target yang digunakan adalah kategori Indeks Pembangunan Manusia (IPM) yang dibagi menjadi dua kelas, yaitu rendah dan tinggi.

Berdasarkan hasil pengelompokan kategori IPM, diperoleh 8 data yang termasuk kategori rendah dan 9 data yang termasuk kategori tinggi. Distribusi kelas yang relatif seimbang tersebut dapat membantu model klasifikasi dalam mempelajari pola data secara lebih baik sehingga mengurangi potensi bias terhadap salah satu kelas.

3.2. Hasil Processing Data

Tabel 3. 1 Distribusi Kategori IPM

Kategori IPM	Jumlah Data
Rendah	8
Tinggi	9
Total	17

Tahap preprocessing dilakukan sebelum proses klasifikasi untuk mempersiapkan data agar sesuai dengan kebutuhan algoritma machine learning. Proses yang dilakukan meliputi pembentukan kategori IPM, transformasi label menggunakan label encoding, normalisasi data menggunakan StandardScaler, serta pembagian data menjadi data latih dan data uji.

Hasil pembentukan kategori IPM menunjukkan bahwa terdapat 8 wilayah yang termasuk kategori rendah dan 9 wilayah yang termasuk kategori tinggi. Selanjutnya, kategori tersebut diubah ke dalam bentuk numerik dengan nilai 0 untuk kategori rendah dan nilai 1 untuk kategori tinggi. Data kemudian dinormalisasi untuk menyamakan skala antar atribut sehingga proses klasifikasi dapat dilakukan secara lebih optimal.

3.2. Hasil Klasifikasi Dengan KNN

Tabel 3. 2 Hasil Pengujian KNN

Nilai K	Akurasi
3	100%
5	83,33%
7	83,33%

Metode K-Nearest Neighbor (KNN) diuji dengan menggunakan tiga variasi nilai parameter K, yaitu K = 3, K = 5, dan K = 7. Pengujian ini dilakukan untuk menentukan nilai K yang mampu menghasilkan kinerja klasifikasi terbaik pada dataset penelitian.

Berdasarkan hasil pengujian yang ditunjukkan pada Tabel 2, nilai K = 3 memberikan tingkat akurasi tertinggi, yaitu sebesar 100%. Sementara itu, nilai K = 5 dan K = 7 masing-masing menghasilkan akurasi sebesar 83,33%. Berdasarkan hasil tersebut, dapat disimpulkan bahwa K = 3 merupakan nilai parameter yang paling optimal dalam proses klasifikasi pada penelitian ini.

Hasil akurasi yang tinggi menunjukkan bahwa pola data pada kategori IPM dapat dibedakan dengan baik oleh algoritma KNN. Selain itu, ukuran dataset yang relatif terbatas membuat penggunaan nilai K yang lebih kecil lebih efektif dalam mempertahankan karakteristik data di lingkungan terdekatnya. Sebaliknya, penggunaan nilai K yang lebih besar cenderung mengurangi sensitivitas model terhadap pola lokal sehingga dapat menurunkan tingkat akurasi klasifikasi.

3.2. Hasil Klasifikasi Menggunakan Support Vector Machine (SVM)

Pengujian metode Support Vector Machine dilakukan menggunakan dua jenis kernel, yaitu kernel linear dan kernel Radial Basis Function (RBF). Penggunaan kedua kernel tersebut bertujuan untuk mengetahui kemampuan SVM dalam mengklasifikasikan kategori Indeks Pembangunan Manusia (IPM) berdasarkan karakteristik data yang digunakan.

Tabel 3. 3 Hasil Pengujian SVM

Kernel	Akurasi
Linear	83,33%
RBF	83,33%

Berdasarkan hasil pengujian yang ditampilkan pada Tabel 3, kernel linear dan kernel RBF memperoleh nilai akurasi yang identik, yaitu sebesar 83,33%. Temuan ini mengindikasikan bahwa kedua jenis kernel memiliki performa yang hampir sama dalam melakukan klasifikasi kategori IPM pada dataset yang digunakan dalam penelitian.

Hasil confusion matrix menunjukkan bahwa dari total enam data uji, lima data berhasil diprediksi sesuai dengan kelas sebenarnya, sedangkan satu data mengalami kesalahan klasifikasi. Dengan demikian, model SVM dapat dikatakan mampu mengidentifikasi pola pada data dengan cukup baik, walaupun masih terdapat beberapa kasus yang belum dapat diklasifikasikan secara tepat.

3.2. Perbandingan Kinerja KNN dan SVM

Perbandingan kinerja metode K-Nearest Neighbor dan Support Vector Machine dilakukan berdasarkan nilai akurasi yang diperoleh dari masing-masing model. Hasil perbandingan ditunjukkan pada Tabel 4.

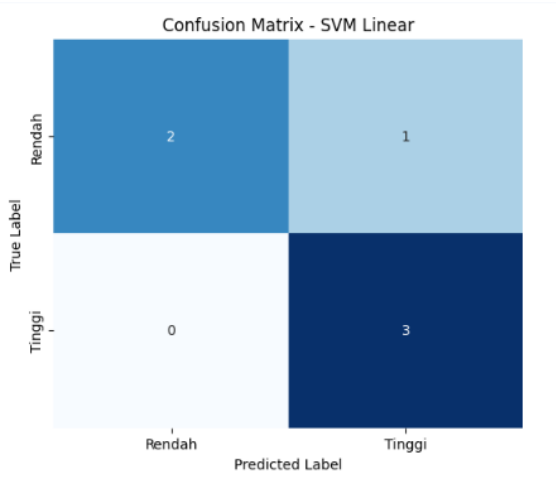
Tabel 3. 4 Perbandingan Akurasi KNN dan SVM

Metode	Parameter	Akurasi
KNN	K=3	100%
KNN	K=5	83,33%
KNN	K=7	83,33%
SVM	Linear	83,33%
SVM	RBF	83,33%

Berdasarkan hasil pengujian yang telah dilakukan, algoritma K-Nearest Neighbor (KNN) dengan nilai K=3 memperoleh tingkat akurasi tertinggi, yaitu sebesar 100%. Di sisi lain, algoritma Support Vector Machine (SVM) dengan kernel linear maupun kernel RBF menghasilkan akurasi yang sama, yaitu sebesar 83,33%.

Temuan tersebut menunjukkan bahwa KNN memiliki kemampuan yang lebih baik dalam mengidentifikasi pola pada dataset yang digunakan dibandingkan SVM. Kinerja KNN yang tinggi kemungkinan dipengaruhi oleh ukuran dataset yang relatif kecil, sehingga metode yang mengandalkan kedekatan antar data dapat bekerja secara lebih efektif. Dalam kondisi tersebut, data dengan karakteristik yang mirip cenderung berada dalam kelas yang sama, sehingga proses klasifikasi menjadi lebih akurat.

Sementara itu, SVM menghasilkan akurasi yang lebih rendah dibandingkan KNN. Walaupun demikian, tingkat akurasi sebesar 83,33% menunjukkan bahwa SVM masih mampu melakukan klasifikasi dengan cukup baik. Perbedaan performa antara kedua metode tersebut mengindikasikan bahwa karakteristik data IPM Kabupaten/Kota di Provinsi Sulawesi Tenggara lebih cocok diklasifikasikan menggunakan pendekatan K-Nearest Neighbor dibandingkan Support Vector Machine.



Gambar 2. 1 Confusion Matrix SVM

Berdasarkan confusion matrix pada gambar 2.1, model SVM berhasil mengklasifikasikan dua data kategori rendah dan tiga data kategori tinggi dengan benar. Terdapat satu data kategori rendah yang diklasifikasikan sebagai kategori tinggi sehingga menghasilkan tingkat akurasi sebesar 83,33%.

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma K-Nearest Neighbor (KNN) dan Support Vector Machine (SVM) mampu digunakan untuk mengklasifikasikan kategori Indeks Pembangunan Manusia (IPM) di Provinsi Sulawesi Tenggara dengan memanfaatkan variabel Angka Harapan Hidup (AHH), Harapan Lama Sekolah (HLS), Rata-rata Lama Sekolah (RLS), dan Pengeluaran Per Kapita (PPP).

Hasil pengujian menunjukkan bahwa algoritma KNN dengan nilai K=3 memberikan kinerja terbaik dengan tingkat akurasi mencapai 100%. Di sisi lain, algoritma SVM dengan kernel linear maupun kernel RBF memperoleh akurasi yang sama, yaitu sebesar 83,33%. Temuan ini menunjukkan bahwa KNN memiliki kemampuan klasifikasi yang lebih baik dibandingkan SVM pada dataset yang digunakan dalam penelitian ini.

Berdasarkan hasil tersebut, metode K-Nearest Neighbor dapat dipertimbangkan sebagai salah satu pendekatan yang efektif untuk mengklasifikasikan kategori IPM di Provinsi Sulawesi Tenggara. Untuk penelitian selanjutnya, disarankan menggunakan dataset yang lebih besar, menambahkan variabel yang relevan dengan pembangunan manusia, serta melakukan perbandingan dengan algoritma klasifikasi lainnya agar diperoleh hasil yang lebih komprehensif dan optimal.

5. DAFTAR PUSTAKA

[1] I. A. A. S. Pratiwi and A. W. Wijayanto, "Klasifikasi Indeks Pembangunan Manusia dengan Metode K-Nearest Neighbor dan Support Vector Machine di Pulau Jawa," *J. Ilmu Komput.*, vol. 15, no. 1, pp. 8–21, 2022.

[2] U. N. D. Programme, "Human Development Report 2023/2024: Breaking the Gridlock: Reimagining Cooperation in a Polarized World," United Nations Development Programme, New York, 2024. [Online]. Available: https://hdr.undp.org/system/files/documents/global-report-document/hdr2023-24reporten.pdf

[3] C. A. Pamungkas and W. W. Widiyanto, "Klasifikasi Indeks Pembangunan Manusia Di Indonesia Tahun 2022 Dengan Support Vector Machine," *J. Ilm. Sist. Inf. dan Ilmu Komput.*, vol. 2, no. 3, pp. 139–145, 2023, doi: 10.55606/juisik.v3i1.407.

[4] N. K. A. P. S. Dewi, A. W. Wijayanto, and J. A. Nursiyono, "Comparison of Machine Learning Algorithms in Classifying Districts/Cities in Indonesia According to the Human Development Index (HDI) in 2021," *J. Sains, Nalar, dan Apl. Teknol. Inf.*, vol. 4, no. 1 SE-Articles, pp. 26–33, Jan. 2025, doi: 10.20885/snati.v4.i1.4.

[5] Fauziah, M. A. Tiro, and Ruliana, "Comparison of k-Nearest Neighbor (k-NN) and Support Vector Machine (SVM) Methods for Classification of Poverty Data in Papua," *ARRUS J. Math. Appl. Sci.*, vol. 2, no. 2, pp. 83–91, 2022, doi: 10.35877/mathscience741.

[6] M. Y. Darsyah, "Klasifikasi Indeks Pembangunan Manusia (Ipm) Dengan," no. October 2017, 2020.

[7] I. Safitri, A. Satria, and R. M. Badri, "Implementasi Algoritma K-Nearest Neighbor Dalam Klasifikasi Indeks Pembangunan Manusia Provinsi Sumatera

Selatan Tahun 2023," vol. 4, no. 2, pp. 768–775, 2024.

[8] B. P. S. P. S. Tenggara, "Indeks Pembangunan Manusia Provinsi Sulawesi Tenggara 2024," BPS Provinsi Sulawesi Tenggara, Kendari, 2025. [Online]. Available: https://sultra.bps.go.id/id/publication/2025/05/28/7472c8d999354ec5742f1298/indeks-pembangunan-manusia-provinsi-sulawesi-tenggara-2024.html

[9] F. Pedregosa, R. Weiss, and M. Brucher, "Scikit-learn : Machine Learning in Python," vol. 12, pp. 2825–2830, 2011.

[10] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, 1967, doi: 10.1109/TIT.1967.1053964.

[11] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995, doi: 10.1007/BF00994018.