

**Implementasi Algoritma Decision Tree untuk Klasifikasi
Pemerlu Pelayanan Kesejahteraan Sosial dengan Ekspansi
Fitur Berbasis Waktu**

Tugas Akhir

**diajukan untuk memenuhi salah satu syarat
memperoleh gelar sarjana pada
Program Studi Sains Data
Fakultas Informatika Universitas Telkom**

1305210061

Hafizh Riyandi Fawwaz



**Program Studi Sarjana Sains Data
Fakultas Informatika
Universitas Telkom
Bandung
2025**

LEMBAR PENGESAHAN

Implementasi Algoritma Decision Tree untuk Klasifikasi Pemerlu Pelayanan Kesejahteraan Sosial Dengan Ekspansi Fitur Berbasis Waktu

Implementation Of Decision Tree Algorithm For Classification Of Social Welfare Service Needers With Time-Based Feature Expansion

NIM : 1305210061

Hafizh Riyandi Fawwaz

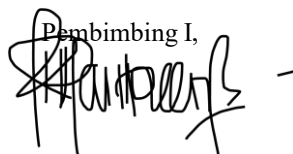
Tugas akhir ini telah diterima dan disahkan untuk memenuhi sebagian syarat memperoleh gelar pada Program Studi Sarjana Sains Data

Fakultas Informatika

Universitas Telkom

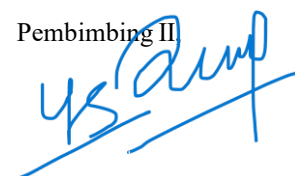
Bandung, 13 Agustus 2025

Menyetujui

Pembimbing I,


Sri Suryani Prasetyowati, S.Si., M.Si

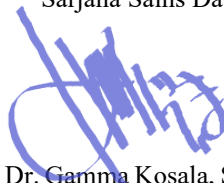
NIP: 99750003

Pembimbing II


Dr. Yuliant Sibaroni

NIP: 00750036

Ketua Program Studi
Sarjana Sains Data,



Dr. Gamma Kosala, S.Si

NIP: 21920013

LEMBAR PERNYATAAN

Dengan ini saya, Hafizh Riyandi Fawwaz, menyatakan sesungguhnya bahwa Tugas Akhir saya dengan judul “Implementasi Algoritma Decision Tree Untuk Klasifikasi Pemerlu Pelayanan Kesejahteraan Sosial Dengan Ekspansi Fitur Berbasis Waktu” beserta dengan seluruh isinya adalah merupakan hasil karya sendiri, dan saya tidak melakukan penjiplakan yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan. Saya siap menanggung resiko/sanksi yang diberikan jika di kemudian hari ditemukan pelanggaran terhadap etika keilmuan dalam Laporan TA atau jika ada klaim dari pihak lain terhadap keaslian karya,

Bandung, 11 Juli 2025

Yang Menyatakan



Hafizh Riyandi Fawwaz

1305210061

Implementasi Algoritma Decision Tree Untuk Klasifikasi Pemerlu Pelayanan Kesejahteraan Sosial Dengan Ekspansi Fitur Berbasis Waktu

1st Hafizh Riyandi Fawwaz, 2nd Sri Suryani Prasetyowati, 3rd Yuliant Sibaroni

Fakultas Informatika, Universitas Telkom, Bandung, Indonesia

hafizhrf@student.telkomuniversity.ac.id, srisuryani@telkomuniversity.ac.id, yuliant@telkomuniversity.ac.id

Abstrak

Jumlah Pemerlu Pelayanan Kesejahteraan Sosial (PPKS) di Provinsi Jawa Tengah yang tinggi, dan berkontribusi besar terhadap angka kemiskinan nasional, memerlukan metode identifikasi yang akurat serta mudah dipahami oleh pengambil kebijakan. Metode identifikasi yang ada saat ini sering terkendala akurasi dan pembaruan data, sehingga berisiko menimbulkan kesalahan identifikasi dan kebijakan yang kurang tepat sasaran. Penelitian ini mengusulkan penggunaan algoritma Decision Tree dengan pendekatan ekspansi fitur berbasis waktu untuk meningkatkan akurasi klasifikasi PPKS. Data penelitian mencakup 35 kabupaten/kota pada periode 2013–2022, dengan berbagai fitur-fitur sosial ekonomi. Ekspansi fitur dilakukan dengan menambahkan data historis dari 1 hingga 9 tahun sebelumnya, kemudian memilih fitur terbaik menggunakan metode SelectKBest. Pengujian dilakukan pada sembilan variasi model (t-1 hingga t-9) untuk melihat pola dari waktu ke waktu, sehingga membuat prediksi yang lebih akurat. Hasil menunjukkan bahwa model t-3 (39 fitur) dan t-5 memiliki performa terbaik dengan akurasi hingga 91%, presisi 94%, *recall* 83%, dan skor F1 86%. Visualisasi spasial prediksi tahun 2023–2031 memperlihatkan pola klasifikasi yang cenderung stabil, dengan wilayah kategori tinggi terkonsentrasi di bagian barat dan tengah provinsi. Penelitian ini berkontribusi dengan mengembangkan model klasifikasi yang akurat, mudah dipahami, dan berbasis data, sehingga dapat membantu pemerintah daerah memprioritaskan intervensi di wilayah berisiko tinggi serta fokus pada fitur penting seperti tingkat pendidikan, indeks pembangunan manusia, dan pengeluaran per kapita.

Kata Kunci : Decision Tree, Pemerlu Pelayanan Kesejahteraan Sosial, Ekspansi Fitur Berbasis Waktu, Klasifikasi, Visualisasi Spasial

Abstract

The high number of Social Welfare Service Needers (PPKS) in Central Java Province, which significantly contributes to the national poverty rate, requires an identification method that is both accurate and easy to interpret for policymakers. Existing identification methods often face challenges related to accuracy and data updates, which can lead to misidentification and less effective policy decisions. This study proposes the use of the Decision Tree algorithm with a time-based feature expansion approach to improve the accuracy of PPKS classification. The dataset covers 35 districts/cities from 2013 to 2022, incorporating various socio-economic features. Feature expansion is carried out by adding historical data from one to nine years prior, followed by selecting the most relevant features using the SelectKBest method. Experiments are conducted on nine model variations (t-1 to t-9) to capture temporal patterns, thereby enhancing prediction accuracy. The results indicate that the t-3 model (39 features) and t-5 model achieve the best performance, with accuracy reaching 91%, precision 94%, recall 83%, and F1-score 86%. Spatial visualization of predictions for 2023–2031 shows a relatively stable classification pattern, with high-category areas concentrated in the western and central regions of the province. This study contributes by developing an accurate, interpretable, and data-driven classification model to assist local governments in prioritizing interventions in high-risk areas, with a focus on key features such as education level, human development index, and per capita expenditure.

Keywords : Decision Tree, Social Welfare Service Needers, Time-Based Feature Expansion, Classification, Spatial Visualization

I. PENDAHULUAN

Kesejahteraan sosial merupakan hak dasar manusia yang fundamental bagi tercapainya kehidupan yang layak dan bermartabat. Kondisi ini diwujudkan dalam terpenuhinya kebutuhan dasar manusia, seperti sandang, pangan, papan, kesehatan, pendidikan, pekerjaan, dan keamanan [1]. Kesejahteraan sosial erat kaitannya dengan stabilitas dan kemajuan bangsa, di mana individu yang sejahtera dapat berkontribusi secara optimal dalam pembangunan nasional [2].

Pemerintah Indonesia telah berkomitmen untuk meningkatkan kesejahteraan sosial masyarakat melalui berbagai kebijakan dan program. Namun, masih terdapat berbagai tantangan yang dihadapi, seperti masih banyak masyarakat yang hidup di bawah garis kemiskinan dan belum memiliki akses yang memadai terhadap kebutuhan dasar, anggaran pemerintah untuk program kesejahteraan sosial masih terbatas, sehingga belum dapat menjangkau seluruh masyarakat yang membutuhkan., dan masih terdapat program kesejahteraan sosial yang kurang efektif dalam mencapai target dan Sasarannya [3].

Pada tahun 2022, Jawa Tengah memiliki sekitar 4,5 juta Pemerlu Pelayanan Kesejahteraan Sosial (PPKS). Pada periode yang sama, tingkat kemiskinan nasional tercatat sebesar 9,57% atau setara dengan 26,23 juta orang, dengan sekitar 3,8 juta di antaranya berada di Jawa Tengah [4]. Pelayanan kesejahteraan sosial merupakan salah satu upaya strategis pemerintah untuk mengatasi berbagai tantangan dalam mencapai kesejahteraan sosial. Pelayanan ini bertujuan untuk memberikan bantuan kepada individu yang memerlukan, seperti orang miskin, orang tua terlantar, anak yatim piatu, dan penyandang disabilitas, meningkatkan akses masyarakat terhadap kebutuhan dasar, membantu individu dalam mengembangkan potensi diri dan kemandirian, dan menciptakan kondisi sosial yang kondusif dan adil [5].

Pemerintah harus mengidentifikasi individu yang memerlukan untuk menjalankan program pelayanan kesejahteraan sosial. Salah satu metode yang dapat digunakan untuk melakukan identifikasi ini adalah klasifikasi [6]. Klasifikasi adalah proses 2 pengelompokan data ke dalam kategori-kategori tertentu. Dalam konteks pelayanan kesejahteraan sosial, klasifikasi dapat digunakan untuk mengidentifikasi individu yang memerlukan berdasarkan profil mereka [7]. Metode identifikasi yang biasa digunakan saat ini, seperti menggunakan data kependudukan dan data program bantuan sosial, memiliki beberapa kelemahan, antara lain data kependudukan dan data 1 program bantuan sosial seringkali tidak diperbaharui dan tidak akurat, sehingga dapat menyebabkan kegagalan identifikasi serta tidak mencakup semua fitur yang relevan dalam menentukan apakah seseorang memerlukan pelayanan kesejahteraan sosial [8].

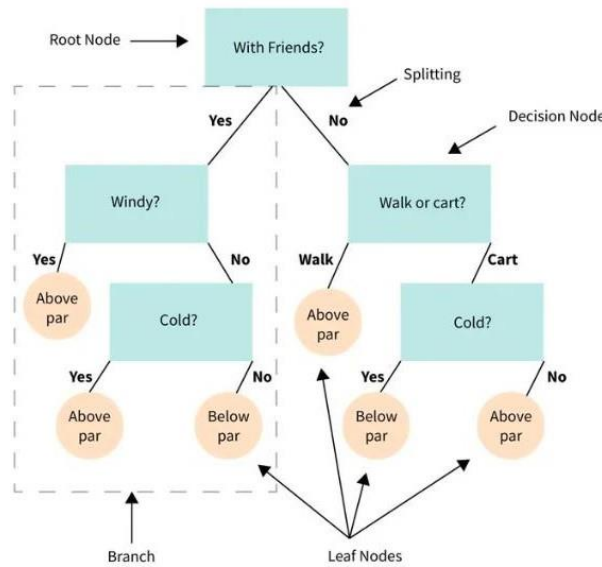
Salah satu metode klasifikasi yang dapat digunakan untuk mengidentifikasi individu yang membutuhkan pelayanan kesejahteraan sosial adalah algoritma decision tree. Beberapa kelebihan algoritma ini termasuk model klasifikasi yang dihasilkannya mudah dipahami dan diinterpretasikan, dan relatif tahan terhadap gangguan data [9]. Algoritma decision tree ini efektif dalam menangani data yang memiliki banyak variabel, seperti data profil individu yang digunakan dalam klasifikasi ini [10].

Ekspansi fitur berbasis waktu, yang melibatkan penambahan fitur-fitur yang terkait dengan waktu ke dalam model klasifikasi, dapat meningkatkan akurasi model. Fitur-fitur berbasis waktu ini dapat berupa lama pengangguran, lama tinggal di daerah terpencil, dan lain sebagainya [11]. Penelitian ini disebabkan oleh tingginya jumlah PPKS di Jawa Tengah yang berkontribusi signifikan terhadap jumlah penduduk miskin nasional, sehingga diperlukan metode analisis yang akurat dan mudah diinterpretasikan oleh pengambil kebijakan. Data sosial ekonomi yang tersedia lengkap dari tahun 2013 hingga 2022 sehingga diperlukan penerapan teknik ekspansi fitur berbasis waktu untuk meningkatkan kinerja klasifikasi. Penelitian ini bertujuan untuk mengembangkan model klasifikasi PPKS menggunakan algoritma Decision Tree, mengidentifikasi fitur-fitur penting yang berpengaruh terhadap klasifikasi PPKS, meningkatkan akurasi prediksi dengan ekspansi fitur berbasis waktu, dan memberikan dasar pengambilan keputusan berbasis data. Penelitian ini diharapkan dapat memberikan kontribusi yaitu pengembangan model klasifikasi PPKS yang dapat menjadi referensi analisis spasial serta menjadi acuan pemerintah dalam pengambilan keputusan PPKS yang efektif dan berfokus kepada fitur-fitur pentingnya saja agar mampu memberikan bantuan sosial secara akurat. Penerapan algoritma decision tree dengan ekspansi fitur berbasis waktu diharapkan dapat meningkatkan akurasi identifikasi individu yang memerlukan pelayanan kesejahteraan sosial. Penambahan fitur-fitur yang relevan, seperti total penduduk, upah minimum, rata-rata jumlah anggota keluarga, presentase kemiskinan, dan lain sebagainya, diharapkan dapat memberikan informasi yang lebih lengkap tentang profil individu dan meningkatkan akurasi klasifikasi.

II. Tinjauan Pustaka

Pemerlu Pelayanan Kesejahteraan Sosial (PPKS) adalah individu, keluarga, atau kelompok yang mengalami kesulitan untuk memenuhi kebutuhan dasar mereka, seperti makanan, pakaian, kesehatan, pendidikan, dan yang memerlukan intervensi dari lembaga pemerintah atau sosial. Identifikasi PPKS penting dalam perumusan kebijakan bantuan yang akurat dan tepat sasaran. Sebagai contoh, studi di Jerman menggunakan algoritma Decision Tree untuk mendeteksi kebutuhan subjektif akan bantuan bagi lansia dalam menjalani aktivitas sehari-hari (ADL dan IADL), dan menunjukkan bahwa gejala depresi merupakan fitur utama kebutuhan yang belum terpenuhi bagi kelompok ini. Di Indonesia, penelitian ini menggunakan Decision Tree untuk mengklasifikasikan tingkat kemiskinan, mengidentifikasi variabel ekonomi dan demografi yang dominan, serta membangun model prediksi yang dapat mendukung pencegahan kemiskinan [12]. Algoritma decision tree adalah metode supervised learning berbentuk pohon keputusan, yang mudah diinterpretasi dan ideal untuk klasifikasi. Model ini terdiri dari node akar (*root*), node internal, dan daun (*leaf*), di mana setiap jalur dari akar ke daun membentuk aturan keputusan (*if-then*) [13]. Algoritma decision tree dipilih karena

hasilnya berupa aturan yang jelas mudah dimengerti, dan mudah diinterpretasikan. Sebaliknya, naive bayes menggunakan perhitungan probabilitas yang lebih sulit dipahami secara langsung [14].



Gambar 2. 1 Arsitektur Decision Tree

Pada **Gambar 2. 1** menggambarkan arsitektur dari algoritma Decision Tree. Proses dimulai dari root node, yaitu pertanyaan utama “*With Friends?*”. Setiap jawaban dari node ini akan diarahkan ke decision node berikutnya, seperti “*Windy?*” atau “*Walk or cart?*”, yang berfungsi untuk membagi data berdasarkan kondisi tertentu. Jalur dari satu node ke node lain disebut branch, sedangkan hasil akhir dari keputusan ditunjukkan pada leaf node, seperti “*Above par*” atau “*Below par*”. Setiap jalur dari atas ke bawah dalam pohon ini mencerminkan sebuah aturan logika sederhana dalam bentuk jika-maka (*if-then*). Misalnya, jika tidak bersama teman → memilih cart → tidak dingin → maka hasilnya adalah “*Above par*”. Struktur seperti ini menunjukkan bagaimana algoritma Decision Tree membentuk keputusan secara bertahap berdasarkan kondisi-kondisi yang mudah dimengerti. Representasi visual ini membantu memahami bagaimana proses klasifikasi berjalan secara sistematis dan logis.

Pemilihan atribut saat pembentukan pohon menggunakan metrik seperti Entropy, Information Gain yang dihitung berdasarkan rumus berikut:

$$Entropy(S) = - \sum p_i \log_2 p_i \quad (1)$$

$$Gain(S, A) = Entropy(S) - \sum \frac{|S_v|}{|S|} Entropy(S_v) \quad (2)$$

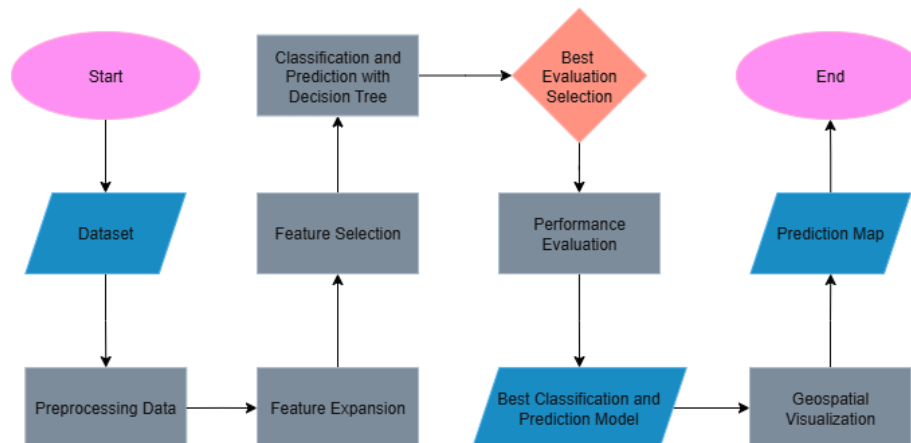
Selanjutnya, algoritma decision tree C4.5 untuk mengkategorikan masyarakat yang berisiko kelaparan di negara berkembang. Data survei yang digunakan mencakup informasi tentang profil setiap orang, seperti usia, jenis kelamin, pendapatan, dan akses ke makanan. Hasil penelitian menunjukkan bahwa algoritma decision tree C4.5 dapat menghasilkan model klasifikasi yang akurat dengan akurasi sebesar 85%. [15]. Penerapan algoritma decision tree untuk mengklasifikasikan individu yang berisiko kelaparan di Desa Turirejo, Kedamean Gresik: penelitian ini mengidentifikasi masyarakat di Desa Turirejo, Kedamean Gresik yang membutuhkan bantuan sosial dengan menggunakan algoritma pohon keputusan CART. Data yang digunakan adalah data dari program bantuan sosial dan kependudukan. Hasil penelitian menunjukkan bahwa algoritma pohon keputusan CART dapat menghasilkan model klasifikasi dengan akurasi sebesar 92% dibandingkan dengan model tradisional yang digunakan oleh pemerintah [5]. Penerapan algoritma C4.5 yang dikombinasikan dengan normalisasi s-core mampu mengklasifikasikan penerima bantuan Program Keluarga Harapan (PKH) dengan akurasi sebesar 94% [16]. Algoritma decision tree C4.5 untuk klasifikasi tingkat kesejahteraan keluarga pada Desa Tiga Dolok. Dengan penggunaan pohon keputusan (decision tree) memberikan gambaran secara jelas prioritas atribut mana yang digunakan sebagai penentu keputusan. Pohon keputusan (decision tree) berupa tingkatan atribut dari kategori kesejahteraan yang ada [17]. Combinatorial optimization and reasoning with graph neural networks : penelitian ini menggunakan fitur-fitur berbasis waktu untuk meningkatkan akurasi model klasifikasi churn rate pelanggan di industri telekomunikasi. Hasil penelitian menunjukkan bahwa penambahan fitur-fitur berbasis waktu dapat meningkatkan akurasi model klasifikasi secara signifikan [18]. Mengklasifikasikan status sosial ekonomi warga berdasarkan atribut seperti

jumlah anggota keluarga, pendidikan, dan pekerjaan. Dengan menggunakan metode decision tree berhasil mengidentifikasi faktor-faktor yang paling berpengaruh dan memberikan masukan untuk membuat program pemberdayaan ekonomi [19]. Untuk data yang tidak seimbang, biasanya tidak perlu menggunakan teknik oversampling pada algoritma decision tree karena decision tree secara alami mampu menangani perbedaan jumlah kecil antar kelas. Penggunaan oversampling justru memungkinkan peningkatan kompleksitas model tanpa meningkatkan kinerja secara signifikan [20],

III. Perancangan Sistem

A. Flowchart

Dalam konteks penelitian ini, alur umum pemodelan sistem digambarkan melalui *flowchart*, yang disajikan pada **Gambar 3. 1** berikut:

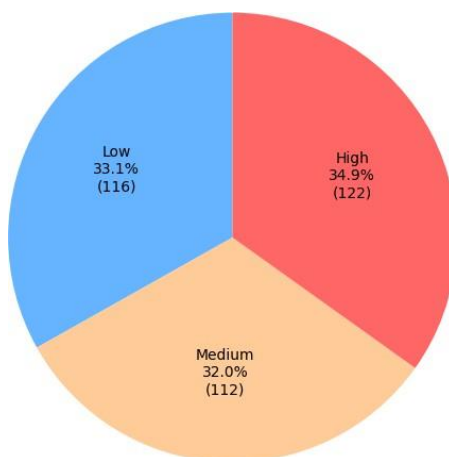


Gambar 3. 1 Desain Sistem Pemodelan Klasifikasi dan Prediksi

Desain sistem dalam penelitian ini melibatkan penggunaan algoritma Decision Tree untuk memprediksi klasifikasi Pemerlu Pelayanan Kesejahteraan Sosial (PPKS) di 35 kabupaten/kota di provinsi Jawa Tengah, berdasarkan data dari 1 hingga 9 tahun sebelumnya. Model terbaik dipilih melalui evaluasi kinerja, dan hasil prediksi kemudian divisualisasikan secara geospasial untuk menunjukkan bagaimana PPKS tersebar di seluruh wilayah provinsi Jawa Tengah.

B. Dataset

Penelitian ini menggunakan data pemerlu pelayanan kesejahteraan sosial yang diperoleh dari Dinas Sosial Provinsi Jawa Tengah dan Badan Pusat Statistik (BPS) Provinsi Jawa Tengah. Data mencakup 35 kabupaten/kota selama periode 2013 hingga 2022 dengan total 14 fitur. Berdasarkan **Gambar 3. 2** total jumlah data sebanyak 350 yang meliputi distribusi data masing-masing kelas di antaranya kelas '0' sebanyak 116 data, kelas '1' sebanyak 112 data, dan kelas '2' sebanyak 122 data.



Gambar 3. 2 Distribusi Kelas PPKS

Kemudian, atribut yang digunakan meliputi jumlah penduduk, persentase kemiskinan, rata-rata lama sekolah, Indeks Pembangunan Manusia (IPM), dan jumlah lapangan kerja, yang seluruhnya berperan penting dalam memengaruhi distribusi Pemerlu Pelayanan Kesejahteraan Sosial (PPKS) [21], [22]. Seluruh atribut memiliki tipe data numerik yaitu integer dan float. Atribut data yang digunakan berdasarkan tabel berikut:

Tabel 3. 1 Data Atribut

Atribut	Deskripsi
X1	Total Penduduk
X2	Upah Minimum
X3	Rata-rata jumlah anggota keluarga
X4	Presentase Kemiskinan
X5	Jumlah Pengangguran
X6	Jumlah orang yang bekerja
X7	Lulusan SD
X8	Lulusan SMP
X9	Lulusan SMA
X10	Lulusan Perguruan Tinggi
X11	Rata-rata Lama Sekolah
X12	Indek Pembangunan Manusia Pria
X13	Indeks Pembangunan Manusia Perempuan
X14	Pengeluaran per kapita per orang (rupiah)
Y	Prediksi angka PPKS untuk setiap kabupaten/kota di Jawa Tengah per tahun

Karena belum terdapat klasifikasi resmi dari pemerintah terkait jumlah PPKS, penelitian ini mengusulkan target kelas dengan menggunakan metode pengelompokan berdasarkan frekuensi yang sama [23]. Metode ini dilakukan dengan mengurutkan nilai PPKS (jumlah penerima bantuan) dan membaginya ke dalam tiga interval yang seimbang, sehingga setiap interval memiliki jumlah data yang setara dalam rentang tahun 2013 hingga 2022.

Tabel 3. 2 Pelabelan Data

Rentang	Kategori	Kelas
$0 \leq \text{PPKS} \leq 78.500$	Rendah	0
$78.500 < \text{PPKS} \leq 146.000$	Sedang	1
$\text{PPKS} > 146.000$	Tinggi	2

C. Preprocessing Data

Pada tahap *preprocessing* data, dilakukan serangkaian langkah untuk memastikan kualitas dan kesiapan data sebelum digunakan dalam proses klasifikasi. Data yang digunakan telah melalui pemeriksaan nilai kosong dan tidak ditemukan *missing value* pada seluruh atribut, sehingga tidak diperlukan imputasi data. Tipe data juga telah sesuai, terdiri dari nilai numerik bertipe integer dan float. Karena, algoritma decision tree tidak sensitif terhadap skala fitur, jadi tidak ada proses standarisasi atau normalisasi. Selain itu, label target klasifikasi adalah kolom "Kelas", yang dibentuk melalui pengelompokan frekuensi yang sama menjadi tiga kelas seimbang berdasarkan distribusi jumlah PPKS dari tahun 2013 hingga 2022. Fitur yang digunakan terdiri dari variabel x1 hingga x14. Tujuan dari tahapan ini adalah untuk memastikan bahwa data bersih, terstruktur, dan optimal untuk proses pemodelan.

D. Feature Expansion

Ekspansi fitur dilakukan melalui proses seleksi fitur dan iterasi untuk memperluas sejumlah variabel [24]. Proses ini digunakan dalam membangun model klasifikasi untuk memprediksi jumlah Penerima Pelayanan Kesejahteraan Sosial (PPKS) di wilayah Jawa Tengah. Ekspansi menggabungkan data historis dari tahun-tahun sebelumnya, yaitu mulai dari 1 hingga 9 tahun sebelum tahun target (t), dengan variabel k merepresentasikan jumlah atribut yang digunakan dalam data latih agar model bisa melihat pola dari waktu ke waktu, sehingga membuat prediksi yang lebih akurat.

Tabel 3. 3 Pelabelan Kelas Data

Model t-k	Kombinasi	Atribut Data Latih	Target
t-1	1	2021	2022
	2	2020	2021
	:	:	:
	9	2013	2014
t-2	1	2020, 2021	2022
	2	2019, 2020	2021
	:	:	:
	8	2013, 2014	2015
:			
t-9	1	2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021	2022

Dalam penelitian ini, proses ekspansi fitur dilakukan menggunakan pustaka *SelectKBest* dari Scikit-learn, yang terbukti meningkatkan akurasi serta kinerja model prediksi [25]. Metode ini berfungsi untuk memilih k fitur terbaik berdasarkan skor tertinggi yang dihitung menggunakan fungsi $f_classif$. Fungsi $f_classif$ sendiri merupakan implementasi dari ANOVA F-test yang digunakan untuk mengevaluasi relevansi setiap fitur terhadap variabel target secara statistik [26]. Prinsipnya yaitu mengukur seberapa signifikan hubungan antara setiap fitur dengan kelas target dengan membandingkan variasi antar kelompok (antara kelas) dan variasi dalam kelompok (nilai dalam setiap kelas). Fitur-fitur dengan nilai F-score yang lebih tinggi sangat berpengaruh terhadap proses klasifikasi [27]. Berikut rumus ANOVA F-test yang digunakan untuk menghitung skor F menggunakan *SelectKBest* dengan fungsi $f_classif$:

$$F = \frac{\frac{1}{C-1} \sum_{i=1}^C N_i (\bar{x}_i - \bar{x})^2}{\frac{1}{C-1} \sum_{i=1}^C \sum_{j=1}^{N_i} (\bar{x}_{ij} - \bar{x}_i)^2} \quad (3)$$

E. Decision Tree

Algoritma decision tree adalah algoritma pembelajaran mesin yang digunakan untuk membangun model klasifikasi atau regresi [28]. Algoritma ini juga alat yang ampuh untuk tugas klasifikasi dan regresi karena kesederhanaan, kemampuan interpretasi, dan fleksibilitasnya dalam menangani berbagai jenis data [29]. Decision Tree dibangun dengan mempartisi data secara rekursif menjadi beberapa subset berdasarkan nilai variabel masukan. Tujuannya adalah untuk membuat model dapat memprediksi nilai variabel target secara akurat dengan mempelajari aturan keputusan sederhana yang disimpulkan dari beberapa fitur data. Prosesnya melibatkan pemilihan atribut terbaik di setiap langkah untuk membagi data menjadi subkumpulan yang semurni mungkin terhadap variabel target [30].

F. Evaluasi

Evaluasi model sebagai penyempurnaan dari model yang telah dilatih agar prediksinya menjadi lebih akurat. Proses evaluasi yang digunakan pada penelitian ini yaitu Confusion Matrix. Confusion Matrix terdiri dari 4 istilah yaitu nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) [31]

Tabel 3. 4 Confusion Matrix Multikelas [32]

		Predicted Class			
		C_1	C_2	...	C_N
True Class	C_1	$C_{1,1}$	FP	...	$C_{1,N}$
	C_2	FN	TP	...	FN

	C_N	$C_{N,1}$	FP	...	$C_{N,N}$

$$Accuracy = \frac{\sum_{i=1}^N TP(C_i)}{\sum_{i=1}^N \sum_{j=1}^N C_{ij}} \quad (4)$$

$$Precision = \frac{TP(C_i)}{TP(C_i) + FP(C_i)} \quad (5)$$

$$Recall = \frac{TP(C_i)}{TP(C_i) + FN(C_i)} \quad (6)$$

$$F1 - Score = 2 \frac{Recall(C_i) Precision(C_i)}{Recall(C_i) + Precision(C_i)} \quad (7)$$

IV. Hasil dan Diskusi

A. Hasil

Pada bagian ini yaitu hasil implementasi dan pengujian model klasifikasi menggunakan algoritma Decision Tree dengan pendekatan ekspansi fitur berbasis waktu. Hasil mencakup performa model terhadap data uji serta analisis akurasi dan efektivitas prediksi jumlah PPKS. Pembahasan juga mencakup interpretasi hasil dan relevansinya dalam pemetaan wilayah prioritas pelayanan sosial. Pada **Tabel 4. 1** di bawah ini merupakan performa terbaik dari model Decision Tree:

Tabel 4. 1 Performa Terbaik Model Decision Tree

t-k	Model	Accuracy	Precision	Recall	F1-Score
t-1	1A	73%	88%	67%	69%
t-2	2B	82%	90%	75%	79%
t-3	3B (29 fitur)	82%	90%	75%	79%
	3B (39 fitur)	91%	94%	83%	86%
t-4	4B	82%	90%	75%	79%
	4E	73%	78%	80%	73%
t-5	5B	91%	94%	83%	86%
t-6	6A	73%	75%	78%	73%
	6C	73%	71%	78%	73%
t-7	7A	82%	82%	81%	80%
t-8	8B	82%	90%	75%	79%
t-9	9A	82%	78%	90%	78%

Berdasarkan **Tabel 4. 1** tersebut hasil perbandingan beberapa model dari t-1 hingga t-9 menunjukkan adanya variasi performa berdasarkan jumlah fitur dan jenis model yang digunakan. Pada ekspansi fitur t-1, model 1A menghasilkan akurasi sebesar 73% dengan presisi tinggi 88%, namun memiliki *recall* rendah 67%. Memasuki t-2, model 2B menunjukkan peningkatan performa dengan akurasi 82% serta keseimbangan metrik evaluasi, yaitu presisi 90% dan *recall* 75%. Performa terbaik secara keseluruhan terjadi pada t-3, khususnya pada model 3B dengan 39 fitur. Hasil ini menunjukkan bahwa penambahan jumlah fitur memberikan kontribusi positif terhadap peningkatan performa model. Pada t-4, model 4B mempertahankan akurasi 82% dengan metrik evaluasi yang seimbang, sementara model 4E mengalami penurunan kinerja dengan akurasi 73%. Selanjutnya, model 5B pada t-5 kembali menunjukkan performa tinggi, dengan akurasi 91% dan skor F1 mencapai 86%, yang menjadi salah satu nilai tertinggi dalam seluruh eksperimen. Namun, performa model mulai menurun pada t-6, di mana model 6A dan 6C hanya mencapai akurasi 73%.

Sementara pada t-7 hingga t-9, performa model cenderung stabil dengan akurasi 82%, meskipun terdapat sedikit perubahan pada nilai presisi dan *recall*. Secara keseluruhan, hasil ini mengindikasikan bahwa pemilihan jumlah fitur dan kombinasi parameter model berpengaruh signifikan terhadap performa klasifikasi. Model dengan jumlah fitur optimal seperti 3B (39 fitur) dan 5B mampu memberikan prediksi yang lebih akurat dan konsisten dibandingkan model lainnya.

Tabel 4. 2 Hasil Klasifikasi dari Model Terpilih Decision Tree

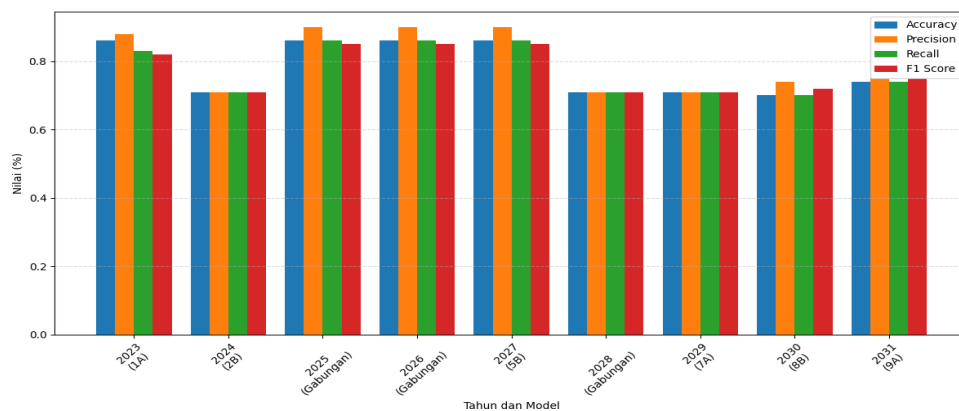
Tahun	Model	Accuracy	Precision	Recall	F1-Score
2023	1A	86%	88%	83%	82%
2024	2B	71%	71%	71%	71%
2025	Gabungan	86%	90%	86%	85%
2026	Gabungan	86%	90%	86%	85%
2027	5B	86%	90%	86%	85%
2028	Gabungan	71%	71%	71%	71%
2029	7A	71%	71%	71%	71%
2030	8B	70%	74%	70%	72%
2031	9A	74%	75%	74%	75%

Pada **Tabel 4. 2** menunjukkan bahwa model-model yang ditampilkan merupakan pilihan terbaik dari setiap ekspansi fitur. Pemilihan ini didasarkan pada evaluasi manual terhadap metrik akurasi, presisi, *recall*, dan skor F1. Proses ini membuktikan bahwa setiap model yang dipilih mewakili performa terbaik pada tahap ekspansi fiturnya masing-masing.

Tabel 4. 3 Fitur Terpilih dari Model Decision Tree

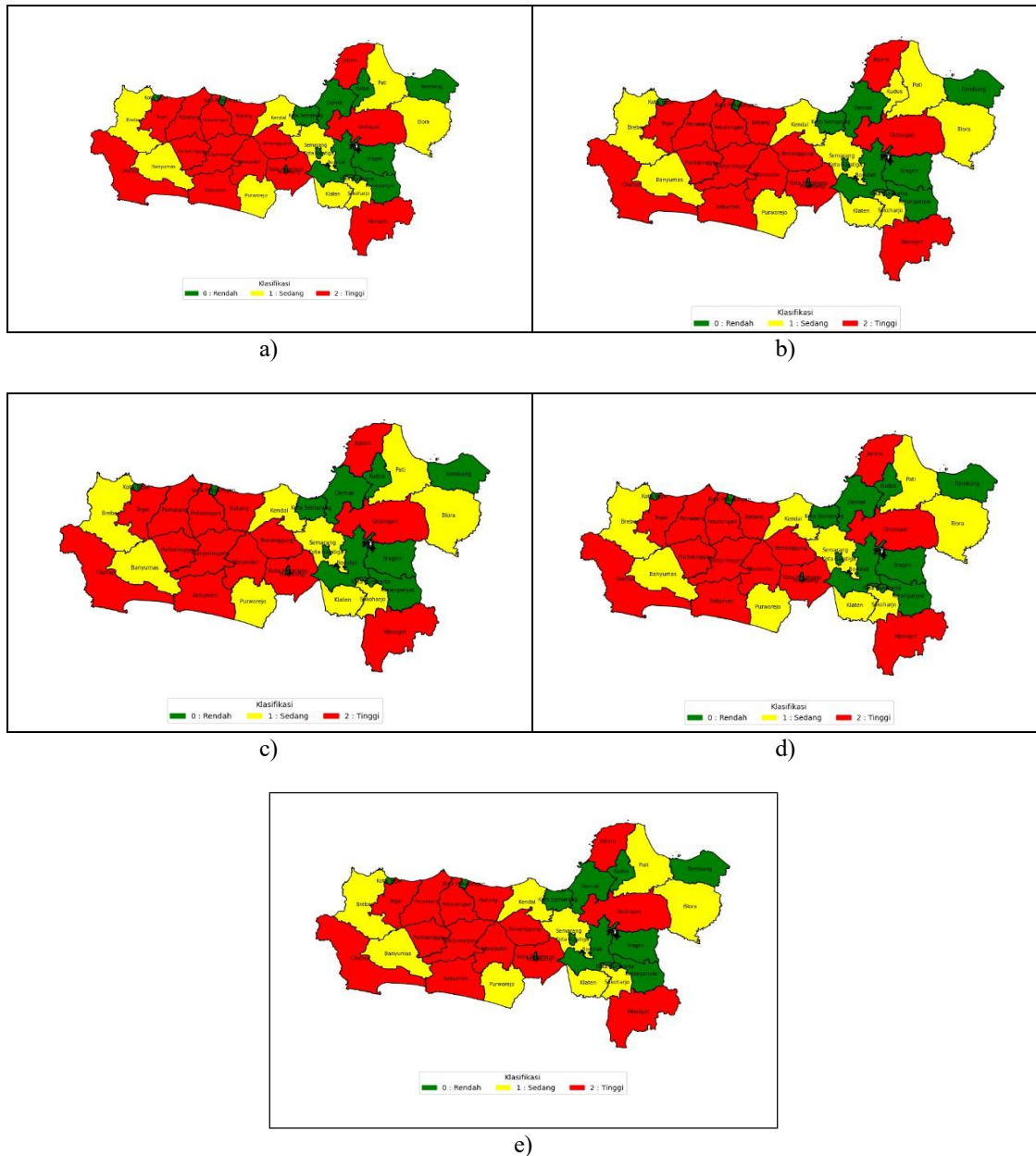
t-k	Model	Fitur
t-1	1A	x7, x9, x11, x12, x13, x14
t-2	2B	x1, x4, x5, x6, x7, x9, x10, x11, x12, x13, x14, x15, x18, x19, x20, x21, x23, x24, x25, x26, x27, x28
t-3	Gabungan	x1, x2, x3, x4, x5, x6, x7, x9, x10, x11, x12, x13, x14, x15, x17, x18, x19, x20, x21, x23, x24, x25, x26, x27, x28, x29, x30, x31, x32, x33, x34, x35, x36, x37, x38, x39, x40, x41, x42
t-4	Gabungan	x1, x3, x4, x5, x6, x7, x9, x10, x11, x12, x13, x14, x15, x17, x18, x19, x20, x21, x23, x24, x25, x26, x27, x28, x29, x32, x33, x34, x35, x38, x39, x40, x41, x42, x43, x46, x47, x48, x49, x51, x52, x53, x54, x55, x56
t-5	5B	x1, x3, x4, x5, x6, x7, x9, x10, x11, x12, x13, x14, x15, x17, x18, x19, x20, x21, x23, x24, x25, x26, x27, x28, x29, x31, x32, x33, x34, x35, x38, x39, x40, x41, x42, x43, x45, x46, x47, x48, x49, x51, x52, x53, x54, x55, x56, x57, x58, x59, x60, x61, x62, x64, x65, x66, x67, x68, x69, x70
t-6	Gabungan	x1, x5, x6, x7, x9, x11, x12, x13, x14, x15, x19, x20, x21, x23, x25, x26, x27, x28, x29, x32, x34, x35, x37, x38, x39, x40, x41, x42, x43, x46, x47, x52, x53, x54, x56, x57, x58, x60, x61, x63, x65, x67, x68, x69, x70, x71, x75, x76, x79, x81, x82, x84
t-7	7A	x7, x9, x11, x12, x14, x21, x23, x25, x26, x28, x35, x37, x39, x40, x42, x53, x54, x56, x63, x65, x67, x68, x70, x81, x84, x91, x93, x95, x96, x98
t-8	8B	x1, x9, x10, x11, x12, x13, x14, x23, x24, x25, x26, x27, x28, x39, x40, x41, x42, x49, x51, x52, x53, x54, x55, x56, x66, x67, x70, x79, x80, x81, x82, x83, x84, x93, x94, x95, x96, x97, x98, x99, x107, x108, x109, x110, x111, x112
t-9	9A	x28, x40, x68, x82, x96, x110, x121

Tabel 4. 3 Beberapa fitur yang muncul dari model yang terpilih Decision Tree. Hal ini dapat digunakan sebagai prioritas pemerintah tentang kondisi yang mempengaruhi PPKS terutama di Provinsi Jawa Tengah.

**Gambar 4. 1** Histogram Metrik Evaluasi

Pada **Gambar 4. 1** visualiasi dengan Histogram tampak lebih jelas yang memperlihatkan perbandingan empat metrik utama secara berdampingan per tahun, sehingga pola penurunan dan kenaikan performa model antar tahun dapat diamati dengan lebih mudah. Secara keseluruhan, model gabungan tahun 2025 dan 2026 serta model tahun 2027 terbukti memberikan hasil terbaik dan paling stabil dibandingkan model lainnya.

Akurasi tinggi menunjukkan banyak prediksi yang benar, tetapi jika data tidak seimbang, dapat memberikan gambaran yang kurang akurat. *Recall* menunjukkan jumlah kasus positif yang berhasil yang dikenali oleh model, sementara presisi menunjukkan jumlah prediksi positif yang benar. Skor F1 digunakan untuk menilai keseimbangan antara presisi dan *recall*. Nilai skor F1 yang tinggi menunjukkan bahwa model memprediksi positif dengan tepat dan mampu menemukan sebagian besar kasus positif. Oleh karena itu, skor F1 sangat cocok digunakan pada data yang tidak seimbang.



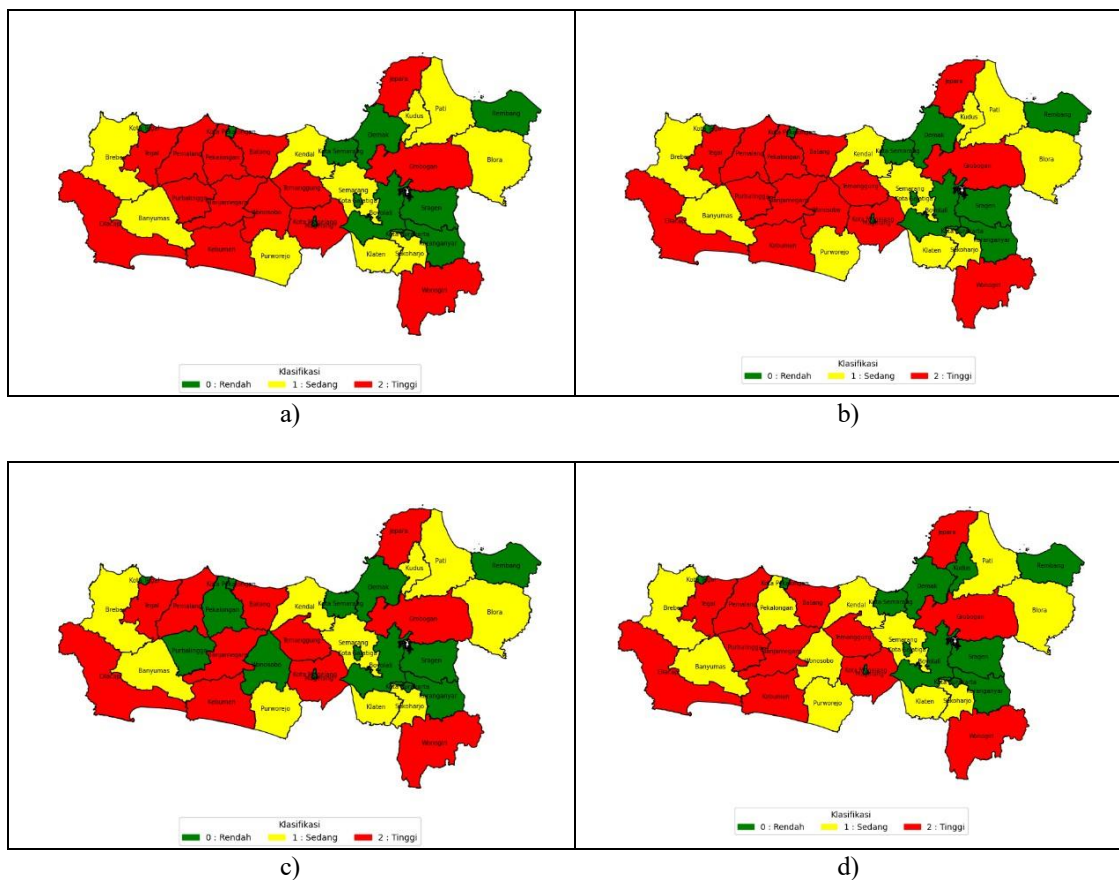
Gambar 4. 2 Sebaran prediksi PPKS Provinsi Jawa Tengah tahun **a** 2023, **b** 2024, **c** 2025, **d** 2026, **e** 2027

Pada **Gambar 4.2** secara umum, persebaran prediksi PPKS (Penerima Pelayanan Kesejahteraan Sosial) di Provinsi Jawa Tengah menunjukkan pola yang cukup stabil dari tahun 2023 hingga 2027. Setiap wilayah diklasifikasikan ke dalam tiga kategori: rendah (hijau), sedang (kuning), dan tinggi (merah).

Pada tahun 2023, sebagian besar wilayah di bagian barat dan selatan, seperti Banyumas, Cilacap, Kebumen, serta beberapa kabupaten di pesisir utara seperti Pekalongan dan Batang, masuk dalam kategori tinggi. Wilayah dengan klasifikasi rendah tampak terkonsentrasi di bagian timur, seperti Rembang, Sragen, dan Karanganyar. Memasuki tahun 2024, pola klasifikasi tidak mengalami perubahan drastis.

Namun, beberapa daerah seperti Temanggung dan Semarang mulai menunjukkan perbaikan, berpindah dari kategori tinggi ke sedang atau rendah. Ini mengindikasikan adanya upaya perbaikan yang mulai menunjukkan hasil di sebagian wilayah.

Pada tahun 2025, wilayah yang termasuk dalam klasifikasi tinggi masih mendominasi, terutama di bagian tengah dan selatan provinsi. Meski demikian, wilayah seperti Kendal, Kota Semarang, dan Karanganyar tetap berada dalam kategori rendah, menandakan stabilitas di daerah-daerah tersebut. Tren berlanjut pada tahun 2026, dengan sebaran kategori tinggi yang relatif serupa dengan tahun-tahun sebelumnya. Namun, wilayah seperti Temanggung dan Demak tampak menunjukkan potensi perbaikan, meskipun masih berada di kategori sedang atau tinggi. Hingga tahun 2027, klasifikasi wilayah secara umum masih stabil. Daerah-daerah seperti Rembang, Sragen, dan Karanganyar tetap konsisten berada di kategori rendah. Sementara itu, sebagian besar wilayah barat dan tengah provinsi tetap berada dalam kategori tinggi.



Gambar 4. 3 Sebaran prediksi PPKS Provinsi Jawa Tengah tahun **a** 2028, **b** 2029, **c** 2030, **d** 2031

Pada tahun 2028, persebaran wilayah dengan klasifikasi tinggi (merah) masih mendominasi di bagian barat dan selatan Jawa Tengah, seperti Cilacap, Banyumas, Kebumen, dan Batang. Wilayah dengan klasifikasi rendah (hijau) tetap konsisten berada di bagian timur, seperti Rembang, Sragen, dan Karanganyar. Sebagian wilayah tengah seperti Temanggung dan Kendal menunjukkan kecenderungan membaik dengan klasifikasi sedang (kuning). Memasuki tahun 2029, pola klasifikasi secara umum masih serupa dengan tahun sebelumnya. Wilayah-wilayah dengan klasifikasi tinggi tidak mengalami banyak perubahan, menunjukkan bahwa tantangan sosial di daerah tersebut masih cukup kuat. Beberapa daerah dengan klasifikasi rendah tetap stabil, namun belum ada ekspansi signifikan ke wilayah tengah atau barat.

Perubahan mulai tampak pada tahun 2030, di mana sejumlah wilayah seperti Pekalongan, Purbalingga, dan Banjarnegara yang sebelumnya termasuk dalam kategori tinggi mulai berpindah ke kategori rendah (hijau). Ini menandakan adanya perbaikan di wilayah tengah, meskipun daerah barat seperti Cilacap dan Kebumen masih bertahan di kategori tinggi.

Pada tahun 2031, persebaran kembali menyerupai tahun-tahun sebelumnya. Wilayah dengan klasifikasi tinggi masih terkonsentrasi di bagian tengah dan barat. Sementara itu, daerah-daerah seperti Rembang, Karanganyar, Sragen, dan Blora terus menunjukkan kestabilan dalam klasifikasi rendah, menjadikannya sebagai wilayah yang relatif konsisten dalam pengelolaan pelayanan sosial.

B. Diskusi

Berdasarkan hasil evaluasi performa model klasifikasi Decision Tree, terlihat bahwa model gabungan secara konsisten memberikan hasil terbaik pada beberapa tahun berturut-turut. Model gabungan pada t-3, t-4, dan model t-5 menghasilkan akurasi tertinggi sebesar 86%, dengan presisi dan *recall* mencapai 90% dan 86%, serta skor F1 85%. Hal ini menunjukkan kestabilan dan keunggulan pendekatan penggabungan fitur dari berbagai model dalam menghasilkan prediksi yang lebih akurat dan seimbang.

Terlihat pada tahun 2013 hingga 2022, klasifikasi PPKS di Provinsi Jawa Tengah menunjukkan pola yang relatif stabil, dengan hanya sedikit perubahan signifikan antar wilayah. Namun, prediksi yang dilakukan untuk tahun 2023 hingga 2031 menggunakan model Decision Tree dengan ekspansi fitur berbasis waktu menunjukkan adanya dinamika klasifikasi di sejumlah daerah. Beberapa wilayah mengalami pergeseran ke kategori yang lebih rendah, sementara wilayah lain tetap berada pada kategori tinggi.

Fenomena ini menjadi fitur penting bagi pemerintah daerah Provinsi Jawa Tengah dalam melaksanakan kebijakan sosial yang lebih terarah. Wilayah-wilayah dengan klasifikasi tinggi perlu menjadi prioritas dalam intervensi pelayanan sosial secara berkelanjutan. Fitur-fitur penting seperti tingkat kemiskinan, jumlah pengangguran, dan pengeluaran per kapita diperkirakan memiliki kontribusi besar terhadap klasifikasi tinggi yang terjadi di wilayah-wilayah tersebut.

Selain itu, beberapa fitur dominan seperti jumlah lulusan pendidikan dasar hingga tinggi, rata-rata lama sekolah, serta indeks pembangunan manusia dan pengeluaran per kapita, secara konsisten muncul pada model-model terbaik. Fitur-fitur ini dapat dijadikan acuan dalam perumusan strategi peningkatan kesejahteraan sosial berbasis wilayah. Dengan memfokuskan program intervensi pada fitur-fitur tersebut, pemerintah dapat mengoptimalkan alokasi sumber daya dan mempercepat penurunan jumlah PPKS secara lebih efektif dan berkelanjutan.

V. Kesimpulan

Penelitian ini berhasil mengembangkan model klasifikasi PPKS di Provinsi Jawa Tengah menggunakan algoritma Decision Tree dengan pendekatan ekspansi fitur berbasis waktu. Model gabungan pada t-3 hingga t-5, hasil dari proses ekspansi fitur dan seleksi menggunakan *SelectKBest*, memberikan performa terbaik dengan akurasi hingga 91% serta nilai presisi, *recall*, dan skor F1 yang konsisten tinggi. Fitur-fitur dominan yang berpengaruh terhadap klasifikasi meliputi tingkat pendidikan, indeks pembangunan manusia, dan pengeluaran per kapita, sesuai dengan tujuan untuk mengidentifikasi variabel yang paling memengaruhi klasifikasi. Berdasarkan visualisasi spasial, sebagian besar wilayah bagian timur cenderung memiliki klasifikasi rendah, sedangkan wilayah barat dan tengah lebih banyak berada di kategori tinggi. Prediksi tahun 2023–2031 menunjukkan pola klasifikasi yang cukup stabil, namun beberapa wilayah masih memerlukan perhatian khusus. Penelitian ini memberikan dasar pengambilan keputusan berbasis data bagi pemerintah daerah untuk memprioritaskan intervensi pada wilayah berklasifikasi tinggi serta memperbaiki fitur-fitur penting yang menjadi faktor utama dalam klasifikasi.

ACKNOWLEDGMENT

Penulis menyampaikan terima kasih kepada Universitas Telkom atas segala dukungan sarana dan prasarana yang telah diberikan selama proses penyusunan tugas akhir ini. Terima kasih juga disampaikan kepada Badan Pusat Statistik (BPS) dan Dinas Sosial Provinsi Jawa Tengah atas bantuan dalam penyediaan data yang sangat penting bagi kelancaran penelitian ini. Akhir dari penelitian ini, penulis berharap dapat memberikan manfaat bagi banyak pihak dan menjadi awal kontribusi kecil menuju sesuatu yang lebih besar di masa mendatang.

Daftar Pustaka

- [1] Nurdin, N. (2023). Kesejahteraan Sosial dalam Persepektif Demokrasi dan Otoritarian: Analisis Perbandingan Inggris dan Singapura. *PARAPOLITIKA: Journal of Politics and Democracy Studies*, 4(2), 140-167.
- [2] Masrukhan, M., Judijanto, L., Timisela, S. I., Yasmita, I. G. A. L., & Suprayitno, A. (2023). Evaluasi Dampak Program Kesejahteraan Sosial terhadap Kemiskinan dan Ketimpangan Sosial: Perspektif Ekonomi. *Jurnal Cahaya Mandalika ISSN 2721-4796 (online)*, 3(2), 1193-1199.
- [3] Wahyudi, W., & Lamama, A. (2023). Kebijakan Ekonomi Kreatif Dalam Meningkatkan Kesejahteraan Sosial. *Jurnal Mirai Management*, 8(1), 604-614.
- [4] I. Machdi, Statistik Indonesia 2023, vol. 1101001. 2023.

- [5] M. C. Letsoin and G. Prayitno, "Sistem Informasi Arsip Data Kepegawaian Menggunakan Metode Agile Development Dinas Kesejahteraan Sosial," vol. 10, no. 3, pp. 120–130, 2023, [Online]. Available: <http://jurnal.mdp.ac.id>
- [6] Khaerunisa, K., Purnamasari, H. ., & P, L. D. M. . (2024). INOVASI PELAYANAN PUBLIK BERBASIS E-GOVERNMENT DALAM PENINGKATAN LAYANAN KESEJAHTERAAN SOSIAL STUDI : WEBSITE SINYAMAN DINSOSKU. *Community Development Journal : Jurnal Pengabdian Masyarakat*, 5(2), 3201–3206. <https://doi.org/10.31004/cdj.v5i2.26791>
- [7] Sarah Maulina Octavia Malau, & Erni Asneli Asbi. (2023). Dampak Pengimplementasian Program Pembelajaran Langsung di Lingkungan Lembaga Kesejahteraan Sosial Anak. *Jurnal Pendidikan Sosial Dan Konseling*, 1(3), 1078–1085. Retrieved from <https://jurnal.ittc.web.id/index.php/jpdsk/article/view/470>
- [8] A. Putri and W. Utami, "Pelayanan Kesejahteraan Sosial terhadap Anak Terlantar di Panti Sosial Sos Children Villages," *Attractive : Innovative Education Journal*, vol. 5, no. 1, 2020, [Online]. Available: <https://www.attractivejournal.com/index.php/aj/>
- [9] Elvatisha, N. .-, Astuti, P. .-, & Manar, D. G. (2023). ANALISIS KINERJA UNIT PELAYANAN KESEJAHTERAAN SOSIAL ANAK INTEGRATIF (UPKSAI) KOTA SEMARANG. *Journal of Politic and Government Studies*, 12(4), 357-374. Retrieved from <https://ejournal3.undip.ac.id/index.php/jpgs/article/view/41398>
- [10] Moni, N., & Susilo, K. D. (2023). *Evaluasi pelayanan kesejahteraan sosial kepada pengemis dan gelandangan di Pondok Sosial Keputih Kecamatan Sukolilo Kota Surabaya*. Retrieved from <https://ejournal.unitomo.ac.id>
- [11] F. Yanna Zega, A. Chanra, Y. Tanjung, and U. Muhammadiyah Sumatera Utara, "PELAYANAN KESEJAHTERAAN SOSIAL BAGI KELUARGA MISKIN DALAM MENINGKATKAN PEMBANGUNAN MASYARAKAT DESA HUMENE KECAMATAN GUNUNGSITOLI IDANOI KOTA GUNUNGSITOLI," 2023. [Online]. Available: <https://doi.org/XX..XXXXX/ARIMA>
- [12] M. Wilda Al -Aluf and Z. Fatah, "Klasifikasi Algoritma Decision Tree Untuk Tingkat Kemiskinan Di Indonesia," *J Comput Sci Technol*, vol. 3, pp. 55–62, doi: 10.59435/jocstec.v3i1.404.
- [13] A. H. Nasrullah, "IMPLEMENTASI ALGORITMA DECISION TREE UNTUK KLASIFIKASI PRODUK LARIS," vol. 7, no. 2, 2021, [Online]. Available: <http://ejournal.fikom-unasman.ac.id>
- [14] Mandasari, R. D., & Hartana, H. (2024). Implementation Of Decision Tree Algorithm For Classification Of Eligibility In Social Assistance Fund Distribution. *TIERS Information Technology Journal*, 5(1), 34-40.
- [15] Rahmayanti, F., Lovita, A., & Muntari, S. (2023). Klasifikasi Kelayakan Penerima Bantuan Sosial Dengan Algoritma Decision Tree. *ESCAF*, 1002-1007.
- [16] Wahyuni, S., Asrianda, A., & Retno, S. (2025). Classification of Family Hope Program Assistance Recipients Using the C4. 5 Algorithm with Z-Score Normalization (Case Study in Atu Lintang District). *ITEJ (Information Technology Engineering Journals)*, 10(1), 160-173.
- [17] S. F. Damanik, A. Wanto, and I. Gunawan, "Penerapan Algoritma Decision Tree C4.5 untuk Klasifikasi Tingkat Kesejahteraan Keluarga pada Desa Tiga Dolok." [Online]. Available: <https://ejournal.catuspata.com/index.php/jkdn/index>
- [18] Cappart, Q., Chételat, D., Khalil, E. B., Lodi, A., Morris, C., & Veličković, P. (2023). Combinatorial optimization and reasoning with graph neural networks. *Journal of Machine Learning Research*, 24(130), 1-61.
- [19] Nengsi, E. P. S., Komalla, D., Wulandari, A., Lorensya, C. N., & Aziz, M. F. (2025). Socio-Economic Status Classification of Neighborhood Residents Using the Decision Tree Algorithm. *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, 4(3), 2565-2569.
- [20] Rao, J., & Ross, K. A. (1998, June). Reusing invariants: A new strategy for correlated queries. In *Proceedings of the 1998 ACM SIGMOD international conference on Management of data* (pp. 37-48).

- [21] Utami, E. S., Setyawan, Y., & Noeryanti, N. (2022). Klasifikasi Kabupaten/Kota Di Indonesia Berdasarkan Tingkatkedalaman Dan Keparahan Kemiskinan Menggunakan Naive Bayes Classifier Dan K-Nearest Neighbor. *Pros. Semin. Nas. Apl. Sains Teknol.*
- [22] Finaliamartha, D., Supriyadi, D., & Fitriana, G. F. (2022). Penerapan metode jaringan syaraf tiruan backpropagation untuk prediksi tingkat kemiskinan di provinsi jawa tengah. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 9(4), 751-760.
- [23] H. Haifa Zahrah, S. Sa'adah, and R. Rismala, "The Foreign Exchange Rate Prediction Using Long-Short Term Memory," *International Journal on Information and Communication Technology (IJoICT)*, vol. 6, no. 2, pp. 94–105, Jan. 2021, doi: 10.21108/ijoict.2020.62.538.
- [24] R. I. Essa, S. Prasetyowati, and Y. Sibaroni, "Performance of ANN and RNN in Predicting the Classification of Covid-19 Diseases based on Time Series Data," *Jurnal Riset Komputer*, vol. 10, no. 1, pp. 2407–389, 2023, doi: 10.30865/jurikom.v10i1.5557.
- [25] T. Desyani, A. Saifudin, and Y. Yulianti, "Feature Selection Based on Naive Bayes for Caesarean Section Prediction," in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing Ltd, Aug. 2020. doi: 10.1088/1757-899X/879/1/012091.
- [26] E. Ashok, S. S. Prasetyowati, and Y. Sibaroni, "DHF Incidence Rate Prediction Based on Spatial-Time with Random Forest Extended Features," *Jurnal RESTI*, vol. 6, no. 4, pp. 612–623, Aug. 2022, doi: 10.29207/resti.v6i4.4268.
- [27] Abdumalikov, S., Kim, J., & Yoon, Y. (2024). Performance analysis and improvement of machine learning with various feature selection methods for EEG-based emotion classification. *Applied Sciences*, 14(22), 10511.
- [28] A. T. Nurani, A. Setiawan, and B. Susanto, "Perbandingan Kinerja Regresi Decision Tree dan Regresi Linear Berganda untuk Prediksi BMI pada Dataset Asthma," *Jurnal Sains dan Edukasi Sains*, vol. 6, no. 1, pp. 34–43, May 2023, doi: 10.24246/juses.v6i1p34-43.
- [29] D. Assyakurrohim, D. Ikham, R. A. Sirodj, and M. W. Afgani, "Metode Studi Kasus dalam Penelitian Kualitatif," *Jurnal Pendidikan Sains dan Komputer*, vol. 3, no. 01, pp. 1–9, Dec. 2022, doi: 10.47709/jpsk.v3i01.1951.
- [30] M. Sari, H. Rachman, N. Juli Astuti, M. Win Afgani, and R. Abdullah Siroj, "Explanatory Survey dalam Metode Penelitian Deskriptif Kuantitatif," *Jurnal Pendidikan Sains dan Komputer*, vol. 3, no. 01, pp. 10–16, Dec. 2022, doi: 10.47709/jpsk.v3i01.1953.
- [31] Alam, S., & Sulisty, M. I. (2023). Analisis Sentimen Berdasarkan Ulasan Pengguna Aplikasi MyPertamina Pada Google Playstore Menggunakan Metode Naïve Bayes. *Storage: Jurnal Ilmiah Teknik Dan Ilmu Komputer*, 2(3), 100-108.
- [32] Markoulidakis, I., Kopsiaftis, G., Rallis, I., & Georgoulas, I. (2021, June). Multi-class confusion matrix reduction method and its application on net promoter score classification problem. In *Proceedings of the 14th PErvasive technologies related to assistive environments conference* (pp. 412-419).