



Penerapan Random Forest dan Content-Based Filtering pada Alokasi Tenaga Kesehatan Hipertensi

Jefry Sunupurwa Asri^{1*}, Diah Aryani¹, Puteri Fannya², Ratna Dewi²

¹Fakultas Ilmu Komputer, Teknik Informatika, Universitas Esa Unggul, Jakarta

Jl. Arjuna Utara No.9, Duri Keba, Kec. Kb. Jeruk, Kota Jakarta Barat, Daerah Khusus Ibukota Jakarta, Indonesia

²Fakultas Ilmu Kesehatan, Rekam Medis dan Informasi Kesehatan, Universitas Esaunggul, Jakarta,

Jl. Arjuna Utara No.9, Duri Keba, Kec. Kb. Jeruk, Kota Jakarta Barat, Daerah Khusus Ibukota Jakarta, Indonesia

Email: ^{1*}jefry.sunupurwa@esaunggul.ac.id, ²diah.aryani@esaunggul.ac.id ³puteri.fannya@esaunggul.ac.id

⁴nsratna@esaunggul.ac.id

Email Penulis Korespondensi: jefry.sunupurwa@esaunggul.ac.id

Submitted: 19/12/2025; Accepted: 31/01/2026; Published: 31/01/2026

Abstrak—Hipertensi merupakan masalah kesehatan utama di DKI Jakarta yang menuntut distribusi sumber daya secara efisien guna mengatasi ketimpangan akses antarwilayah. Penelitian ini bertujuan untuk merancang dan mengimplementasikan decision support system (DSS) berbasis web yang mengintegrasikan Geographic Information System (GIS) untuk mengoptimalkan alokasi tenaga kesehatan serta menentukan wilayah prioritas penanganan hipertensi secara presisi. Kebaruan penelitian ini terletak pada integrasi model Machine Learning Random Forest untuk memprediksi cakupan pelayanan hingga tahun 2030 dengan metode Content-Based Filtering (CBF). Metode CBF memanfaatkan fitur intrinsik wilayah, seperti persentase pelayanan, letak geografis, dan tren prediksi, guna menghasilkan rekomendasi kuota tenaga kesehatan secara objektif. Model Random Forest divalidasi menggunakan 5-Fold Cross Validation dengan performa sangat baik, menunjukkan nilai R^2 rata-rata sebesar 0,86 dan Mean Absolute Error (MAE) akurat sebesar 6,7%. Sistem diimplementasikan menggunakan kerangka kerja Streamlit dan Folium untuk visualisasi geografis. Hasil penelitian ini memberikan kontribusi berupa peta wilayah prioritas, rekomendasi kuota tenaga kesehatan yang adaptif, serta simulasi rute distribusi Puskesmas Keliling. Melalui sistem ini, pengambil kebijakan dapat melakukan perencanaan strategis berbasis data untuk meningkatkan efektivitas intervensi hipertensi di Jakarta. Dengan pendekatan prediktif dan rekomendasi yang terintegrasi, penelitian ini diharapkan dapat menjadi acuan dalam transformasi digital alokasi sumber daya kesehatan masyarakat secara lebih merata serta tepat sasaran.

Kata Kunci: Sistem Pendukung Keputusan; Random Forest; Content-Based Filtering; Geographic Information System; Evidence-Based Policy.

Abstract—Hypertension is a major health issue in DKI Jakarta requiring efficient resource distribution to overcome inter-regional access inequalities. This research aims to design and implement a web-based decision support system (DSS) integrating Geographic Information System (GIS) to optimize health worker allocation and determine hypertension priority areas precisely. The novelty lies in integrating a Random Forest machine learning model to predict service coverage until 2030 with Content-Based Filtering (CBF). The CBF method utilizes intrinsic regional features, including service percentages, geographical locations, and prediction trends, to generate objective health worker quota recommendations. The Random Forest model was validated using 5-Fold Cross Validation with excellent performance, showing an average R^2 value of 0.86 and an accurate Mean Absolute Error (MAE) of 6.7%. The system is implemented using Streamlit and Folium frameworks for geographical visualization. Research results provide contributions through priority area maps, adaptive health worker quota recommendations, and Mobile Health Clinic route simulations supporting data-driven decision-making. Through this system, policymakers can perform strategic planning to improve hypertension intervention effectiveness in Jakarta. With an integrated predictive and recommendation approach, this study is expected to become a reference in the digital transformation of public health resource allocation more equitably and accurately.

Keywords: Decision Support System; Random Forest; Content-Based Filtering; Geographic Information System; Evidence-Based Policy.

1. PENDAHULUAN

Hipertensi merupakan salah satu penyakit tidak menular yang menjadi tantangan utama dalam sistem pelayanan kesehatan, khususnya di wilayah perkotaan dengan tingkat kepadatan penduduk yang tinggi. Penyakit ini sering berkembang tanpa gejala awal yang jelas (silent killer) dan berpotensi menimbulkan komplikasi serius apabila tidak ditangani secara berkelanjutan. World Health Organization (WHO) melaporkan bahwa hipertensi merupakan salah satu faktor risiko utama penyakit kardiovaskular dan penyebab kematian prematur secara global [1]. Tingginya kebutuhan layanan hipertensi menuntut sistem pelayanan kesehatan yang mampu menjangkau masyarakat secara merata dan efisien. Namun, di wilayah metropolitan seperti DKI Jakarta, permasalahan yang dihadapi tidak hanya berkaitan dengan tingginya prevalensi hipertensi, tetapi juga ketimpangan distribusi fasilitas dan tenaga kesehatan antar wilayah administratif [2], [3].

Pendekatan berbasis Geographic Information System (GIS) telah banyak dimanfaatkan dalam analisis distribusi dan aksesibilitas layanan kesehatan secara spasial. GIS memungkinkan visualisasi dan analisis pola ketimpangan akses layanan kesehatan sehingga dapat mendukung perencanaan wilayah pelayanan yang lebih tepat sasaran. Penelitian sebelumnya menunjukkan bahwa pendekatan GIS efektif dalam mengevaluasi aksesibilitas fasilitas kesehatan serta mengidentifikasi ketimpangan layanan kesehatan di wilayah perkotaan [4], [5]. Selain itu, studi tinjauan sistematis menegaskan bahwa analisis spasial berbasis GIS merupakan pendekatan yang relevan



dalam memahami pola kebutuhan layanan kesehatan masyarakat dan mendukung pengambilan keputusan berbasis wilayah [2]. Meskipun demikian, sebagian besar penelitian GIS di bidang kesehatan masih bersifat deskriptif dan berfokus pada pemetaan kondisi eksisting tanpa dilengkapi kemampuan prediksi kebutuhan layanan kesehatan di masa mendatang.

Seiring dengan perkembangan analitik data, pendekatan machine learning mulai diterapkan untuk melengkapi analisis spasial di bidang kesehatan. Salah satu algoritma yang banyak digunakan adalah Random Forest, yang memiliki keunggulan dalam menangani data berdimensi tinggi, hubungan nonlinier, serta ketahanan terhadap overfitting. Dalam konteks layanan kesehatan, Random Forest telah digunakan untuk mendukung perencanaan layanan berbasis data spasial dan lingkungan [6]. Penelitian lain juga menunjukkan bahwa algoritma ini memiliki performa yang baik dalam memprediksi risiko hipertensi dan menganalisis pola kesehatan masyarakat [7], [8]. Namun, penerapan Random Forest pada penelitian-penelitian tersebut umumnya masih terbatas pada tahap prediksi atau klasifikasi dan belum terintegrasi secara langsung ke dalam sistem yang mampu menghasilkan rekomendasi alokasi sumber daya kesehatan.

Dalam konteks pengambilan keputusan, Decision Support System (DSS) berperan penting dalam membantu pemangku kepentingan menentukan strategi yang optimal berdasarkan data dan model analitik. DSS berbasis web memungkinkan integrasi berbagai sumber data dan metode analisis ke dalam satu platform yang mudah diakses dan digunakan. Penelitian sebelumnya menunjukkan bahwa DSS berbasis web efektif dalam mendukung alokasi sumber daya di sektor publik, termasuk layanan kesehatan [9]. Pada lingkungan perkotaan dan smart city, DSS berbasis data juga telah diterapkan untuk meningkatkan efisiensi dan adaptabilitas alokasi sumber daya kesehatan [10]. Selain itu, kajian terkini menyoroti pentingnya integrasi analisis spasial dan machine learning dalam spatial decision support systems untuk meningkatkan kualitas pengambilan keputusan berbasis wilayah [11]. Meskipun demikian, sebagian besar DSS tersebut belum mengintegrasikan analisis spasial, prediksi kebutuhan layanan, dan mekanisme rekomendasi secara terpadu dalam satu sistem.

Sistem rekomendasi merupakan pendekatan yang potensial untuk melengkapi DSS dalam menghasilkan rekomendasi yang lebih adaptif dan kontekstual. Metode Content-Based Filtering (CBF) menghasilkan rekomendasi berdasarkan karakteristik intrinsik objek sehingga sesuai diterapkan pada konteks wilayah yang memiliki atribut dan kebutuhan layanan yang berbeda. Dalam domain kesehatan, sistem rekomendasi telah dikembangkan untuk merekomendasikan layanan dan fasilitas kesehatan dengan mempertimbangkan informasi geospasial [12], [13]. Selain itu, kajian komprehensif menunjukkan bahwa health recommender systems terus berkembang dan memiliki potensi besar dalam mendukung pengambilan keputusan di sektor kesehatan [14]. Namun, pemanfaatan CBF untuk rekomendasi alokasi tenaga kesehatan yang berbasis data spasial dan hasil prediksi masih relatif terbatas.

Dalam konteks sistem rekomendasi pada penelitian ini, pendekatan Content-Based Filtering (CBF) tidak diterapkan pada interaksi pengguna individual sebagaimana pada sistem rekomendasi konvensional. Wilayah administratif diperlakukan sebagai profil kebutuhan (analog dengan user), sementara tenaga kesehatan dan kapasitas layanan diposisikan sebagai sumber daya yang direkomendasikan (analog dengan item). Rekomendasi alokasi tenaga kesehatan dihasilkan melalui pencocokan karakteristik intrinsik wilayah seperti tingkat dan tren pelayanan hipertensi hasil prediksi, serta atribut spasial dengan kebutuhan kapasitas layanan kesehatan. Dengan pendekatan ini, CBF berfungsi sebagai mekanisme pendukung keputusan berbasis kesesuaian kebutuhan wilayah, bukan sebagai sistem rekomendasi berbasis preferensi individu [12], [14], [15]

Penelitian lain juga menegaskan pentingnya evaluasi alokasi sumber daya medis berbasis analisis spasial untuk mendukung kebijakan kesehatan publik dan pemerataan layanan kesehatan [3]. Meskipun demikian, pendekatan-pendekatan tersebut umumnya belum mengintegrasikan mekanisme rekomendasi berbasis prediksi ke dalam satu sistem pendukung keputusan berbasis web.

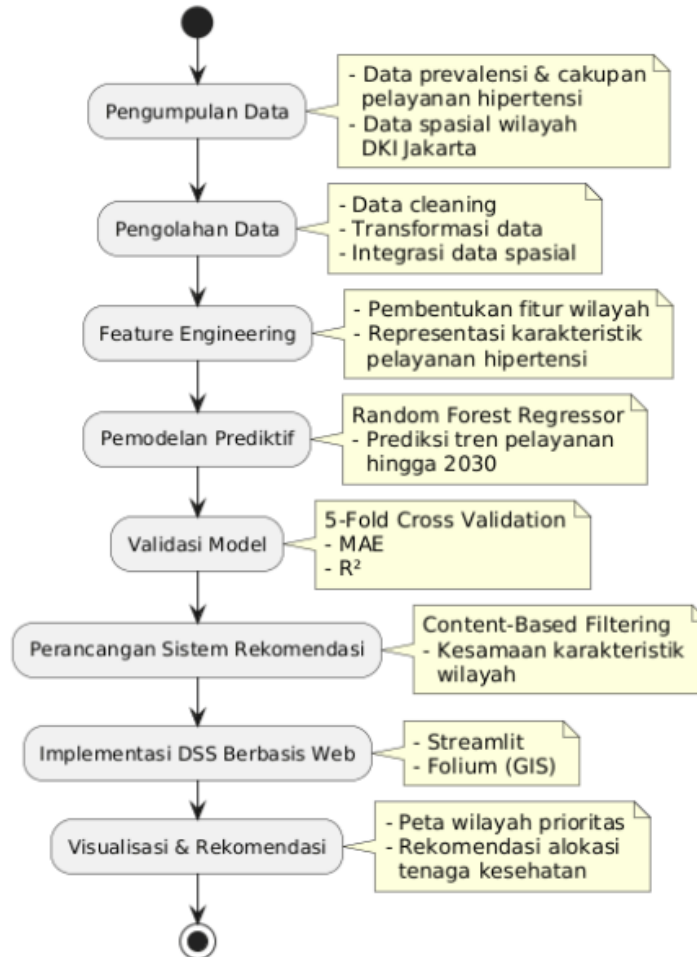
Berdasarkan kajian tersebut, dapat diidentifikasi research gap berupa belum tersedianya sistem pendukung keputusan berbasis web yang secara terintegrasi menggabungkan GIS untuk analisis spasial, Random Forest untuk prediksi tren pelayanan hipertensi, serta Content-Based Filtering untuk menghasilkan rekomendasi alokasi tenaga kesehatan. Oleh karena itu, penelitian ini mengusulkan pengembangan **sistem** pendukung keputusan (DSS) berbasis web yang mengintegrasikan ketiga pendekatan tersebut guna mendukung optimasi alokasi tenaga kesehatan dalam pelayanan hipertensi di DKI Jakarta.

Penelitian ini bertujuan untuk memprediksi tren persentase pelayanan hipertensi hingga tahun 2030 menggunakan model Random Forest, serta memanfaatkan hasil prediksi dan data spasial untuk menghasilkan rekomendasi alokasi tenaga kesehatan dan penentuan wilayah prioritas pelayanan hipertensi melalui metode Content-Based Filtering. Selain itu, penelitian ini juga memvisualisasikan wilayah prioritas pelayanan hipertensi dalam bentuk peta interaktif berbasis Geographic Information System sebagai bagian dari decision support system pendukung pengambilan keputusan. Dalam pencapaian tujuan tersebut, Jefry Sunupurwa Asri berperan dalam penyusunan metodologi dan pengoordinasian seluruh tahapan penelitian, Diah Aryani mengembangkan serta mengevaluasi model Machine Learning untuk prediksi pelayanan hipertensi, Puteri Fannya melakukan analisis kebutuhan dan penyediaan data epidemiologis serta konteks klinis hipertensi, dan Ratna Dewi meninjau kesesuaian data dengan standar pelayanan kesehatan serta memberikan arahan klinis. Hasil penelitian ini diharapkan dapat mendukung pengambilan keputusan berbasis data dan berkontribusi terhadap peningkatan pemerataan pelayanan hipertensi di wilayah DKI Jakarta.

2. METODOLOGI PENELITIAN

2.1 Bagan Alir Penelitian

Tahapan penelitian dalam pengembangan sistem pendukung keputusan pelayanan hipertensi di DKI Jakarta disusun secara sistematis dan terintegrasi, sebagaimana ditunjukkan pada Gambar 1. Bagan alir penelitian digunakan untuk memberikan gambaran menyeluruh mengenai urutan proses penelitian serta keterkaitan antar tahapan, mulai dari pengumpulan data hingga visualisasi rekomendasi alokasi tenaga kesehatan.



Gambar 1. Bagan Alir Penelitian Sistem Pendukung Keputusan Pelayanan Hipertensi

Gambar 1 memperlihatkan bahwa penelitian diawali dengan tahap pengumpulan data, yang mencakup data prevalensi dan cakupan pelayanan hipertensi serta data spasial wilayah administratif DKI Jakarta. Data spasial berfungsi sebagai dasar analisis berbasis Geographic Information System (GIS) untuk merepresentasikan distribusi pelayanan kesehatan secara geografis. Pemanfaatan GIS dan WebGIS telah banyak diterapkan dalam integrasi data spasial untuk mendukung pemodelan serta pengambilan keputusan berbasis wilayah, khususnya pada konteks perencanaan dan evaluasi layanan kesehatan [15], [16].

Tahap berikutnya adalah **pengolahan data**, yang meliputi proses data cleaning, transformasi data, serta integrasi data spasial dan nonspasial. Tahap ini bertujuan untuk meningkatkan kualitas dan konsistensi data sehingga layak digunakan dalam proses analisis dan pemodelan.

Setelah data siap digunakan, dilakukan tahap **feature engineering**, yaitu pembentukan fitur-fitur yang merepresentasikan karakteristik intrinsik wilayah pelayanan hipertensi. Fitur ini mencerminkan kondisi pelayanan kesehatan pada masing-masing wilayah dan menjadi dasar dalam pemodelan prediktif serta sistem rekomendasi.

Tahap selanjutnya adalah **pemodelan prediktif** menggunakan algoritma Random Forest Regressor untuk memprediksi tren persentase pelayanan hipertensi hingga tahun 2030. Untuk memastikan keandalan hasil prediksi, dilakukan **validasi model** menggunakan metode 5-Fold Cross Validation dengan metrik evaluasi Mean Absolute Error (MAE) dan koefisien determinasi (R^2).

Hasil prediksi yang diperoleh kemudian dimanfaatkan dalam **perancangan sistem rekomendasi** berbasis Content-Based Filtering (CBF). Sistem ini menghasilkan rekomendasi alokasi tenaga kesehatan berdasarkan kesamaan karakteristik wilayah, sehingga mampu mendukung penentuan wilayah prioritas secara objektif dan berbasis data.



Tahap akhir penelitian adalah **implementasi sistem pendukung keputusan berbasis web**, yang dikembangkan menggunakan Streamlit sebagai antarmuka pengguna dan Folium untuk visualisasi peta interaktif berbasis GIS. Sistem ini menyajikan informasi kondisi pelayanan hipertensi, hasil prediksi tren pelayanan, serta rekomendasi alokasi tenaga kesehatan dan wilayah prioritas secara visual dan mudah dipahami.

2.2 Pengumpulan Data

Pengumpulan data pada penelitian ini bertujuan untuk menyediakan dasar informasi yang komprehensif bagi analisis spasial, pemodelan prediktif, dan perancangan sistem rekomendasi dalam sistem pendukung keputusan pelayanan hipertensi di DKI Jakarta. Data yang digunakan dirancang untuk merepresentasikan kondisi pelayanan hipertensi secara detail pada tingkat wilayah administratif.

Jenis data yang dikumpulkan meliputi data prevalensi dan pelayanan hipertensi yang disajikan berdasarkan **kota administrasi, kecamatan, dan jenis kelamin**. Setiap entri data mencakup informasi mengenai jumlah penderita hipertensi, jumlah penerima layanan, serta persentase cakupan pelayanan. Struktur data ini memungkinkan analisis ketimpangan pelayanan kesehatan antar wilayah serta evaluasi perbedaan cakupan pelayanan berdasarkan karakteristik demografis.

Selain data pelayanan hipertensi, penelitian ini juga memanfaatkan data spasial wilayah administratif DKI Jakarta sebagai dasar analisis berbasis Geographic Information System (GIS). Data spasial digunakan untuk merepresentasikan batas wilayah dan mendukung visualisasi distribusi pelayanan hipertensi secara geografis.

Seluruh data diperoleh dari sumber resmi dan telah melalui proses verifikasi awal untuk memastikan kelengkapan, konsistensi, dan kesesuaian format data. Contoh struktur dataset pelayanan hipertensi yang digunakan dalam penelitian ini disajikan pada **Tabel 1**.

Tabel 1. Contoh Struktur Data Pelayanan Hipertensi

No	Kota Administrasi	Kecamatan	Jenis Kelamin	Jumlah Penderita	Penerima Layanan	Cakupan (%)
1	Kota Adm. Jakarta Pusat	Cempaka Putih	Laki-laki	4.969	4.501	90,58
2	Kota Adm. Jakarta Pusat	Cempaka Putih	Perempuan	5.024	5.537	110,21
3	Kota Adm. Jakarta Pusat	Gambir	Laki-laki	5.032	4.888	97,14
4	Kota Adm. Jakarta Pusat	Gambir	Perempuan	5.014	5.220	104,11
5	Kota Adm. Jakarta Pusat	Johar Baru	Laki-laki	7.116	5.645	79,33
6	Kota Adm. Jakarta Pusat	Johar Baru	Perempuan	6.984	8.573	122,75
7	Kota Adm. Jakarta Pusat	Kemayoran	Laki-laki	12.788	10.069	78,74
8	Kota Adm. Jakarta Pusat	Kemayoran	Perempuan	12.754	15.641	122,64
9	Kota Adm. Jakarta Pusat	Menteng	Laki-laki	4.571	4.365	95,49

Tabel 1 menunjukkan bahwa data pelayanan hipertensi disusun berdasarkan karakteristik wilayah dan jenis kelamin, dengan variasi jumlah penderita, penerima layanan, dan cakupan pelayanan antar kecamatan. Perbedaan nilai cakupan pelayanan yang melebihi 100% mengindikasikan adanya variasi dalam pola pemanfaatan layanan kesehatan, seperti pelayanan lintas wilayah administratif atau perbedaan antara estimasi penderita dan realisasi pelayanan. Informasi ini menjadi dasar penting dalam analisis ketimpangan pelayanan dan pembentukan fitur pada tahap pemodelan prediktif serta sistem rekomendasi.

2.3 Pengolahan Data dan Feature Engineering

Tahap pengolahan data dan feature engineering dilakukan untuk memastikan bahwa data yang digunakan dalam penelitian memiliki kualitas yang baik serta representatif terhadap kondisi pelayanan hipertensi di DKI Jakarta. Tahap ini bertujuan untuk menyiapkan data yang siap digunakan dalam pemodelan prediktif dan sistem rekomendasi berbasis wilayah.

Pengolahan data diawali dengan proses data cleaning, yang mencakup pemeriksaan dan penanganan nilai hilang, ketidakkonsistenan data, serta duplikasi entri. Data numerik, seperti jumlah penderita, penerima layanan, dan persentase cakupan pelayanan, diperiksa untuk memastikan kesesuaian rentang nilai dan konsistensi antar variabel. Proses ini dilakukan untuk meminimalkan potensi bias yang dapat mempengaruhi hasil pemodelan.

Selanjutnya dilakukan transformasi dan normalisasi data sesuai kebutuhan analisis. Transformasi data bertujuan untuk menyelaraskan skala antar variabel, sehingga setiap fitur memiliki kontribusi yang proporsional dalam proses pembelajaran model. Selain itu, data kategorikal seperti wilayah administratif dan jenis kelamin direpresentasikan dalam bentuk yang sesuai untuk analisis lebih lanjut.

Tahap feature engineering dilakukan dengan membentuk fitur-fitur yang merepresentasikan karakteristik intrinsik wilayah pelayanan hipertensi. Fitur yang digunakan mencakup atribut pelayanan kesehatan, hasil integrasi data spasial, serta pola temporal data pelayanan hipertensi. Pembentukan fitur ini dirancang untuk menangkap



variasi kebutuhan layanan kesehatan antar wilayah dan mendukung mekanisme kesamaan (similarity) pada sistem rekomendasi berbasis Content-Based Filtering. Pada sistem rekomendasi di domain kesehatan, kualitas dan representativitas fitur memiliki peran penting terhadap akurasi serta reliabilitas rekomendasi yang dihasilkan, sehingga aspek evaluasi sistem perlu dipertimbangkan sejak tahap feature engineering [17].

Seluruh fitur yang dihasilkan kemudian dievaluasi secara deskriptif untuk memastikan relevansi dan keterkaitannya dengan tujuan penelitian. Dataset hasil pengolahan dan feature engineering selanjutnya digunakan sebagai input utama pada tahap pemodelan prediktif menggunakan Random Forest Regressor serta pada perancangan sistem rekomendasi alokasi tenaga kesehatan.

2.4 Pemodelan dan Validasi Prediktif

Pemodelan prediktif dalam penelitian ini bertujuan untuk memperkirakan tren persentase cakupan pelayanan hipertensi di DKI Jakarta hingga tahun 2030 sebagai dasar pengambilan keputusan berbasis data. Model prediksi dibangun menggunakan algoritma **Random Forest Regressor**, yang dipilih karena kemampuannya dalam menangani hubungan nonlinier, data berdimensi tinggi, serta kestabilan performa terhadap variasi data. Algoritma Random Forest telah terbukti banyak diterapkan secara efektif dalam studi prediktif di bidang kesehatan masyarakat dan analisis spasial berbasis GIS, dengan performa yang konsisten pada berbagai konteks data layanan kesehatan [18], [19].

Random Forest bekerja dengan membangun sejumlah pohon keputusan (decision trees) berdasarkan subset data dan fitur yang dipilih secara acak. Setiap pohon menghasilkan prediksi secara independen, kemudian nilai prediksi akhir diperoleh melalui proses agregasi berupa rata-rata keluaran seluruh pohon keputusan. Secara matematis, mekanisme prediksi Random Forest dinyatakan pada Persamaan (1).

$$\hat{y} = \frac{1}{N} = \sum_{i=1}^N f_i(x) \quad (1)$$

Dimana \hat{y} merupakan nilai prediksi persentase pelayanan hipertensi, N menyatakan jumlah pohon keputusan, dan $f_i(x)$ adalah fungsi prediksi dari pohon keputusan ke - i terhadap vektor fitur x .

Untuk memastikan keandalan dan konsistensi model, dilakukan proses validasi menggunakan metode **5-Fold Cross Validation**. Dataset dibagi menjadi lima subset yang sama besar, di mana pada setiap iterasi empat subset digunakan sebagai data pelatihan dan satu subset sebagai data pengujian. Proses ini diulang hingga setiap subset berperan sebagai data pengujian satu kali, sehingga seluruh data berkontribusi dalam evaluasi model.

Kinerja model prediksi dievaluasi menggunakan metrik **Mean Absolute Error (MAE)** dan **koefisien determinasi (R^2)**. MAE digunakan untuk mengukur rata-rata selisih absolut antara nilai prediksi dan nilai aktual, sedangkan R^2 digunakan untuk menilai kemampuan model dalam menjelaskan variasi data aktual. Kombinasi kedua metrik ini memberikan gambaran yang komprehensif mengenai tingkat akurasi dan stabilitas model prediksi.

Hasil prediksi yang dihasilkan oleh model Random Forest selanjutnya dimanfaatkan sebagai input utama dalam sistem rekomendasi alokasi tenaga kesehatan. Dengan demikian, model prediktif tidak hanya berfungsi sebagai alat analisis, tetapi juga menjadi komponen integral dalam sistem pendukung keputusan berbasis web yang diusulkan.

2.5 Perancangan Sistem Rekomendasi Berbasis Content-Based Filtering

Perancangan sistem rekomendasi dalam penelitian ini bertujuan untuk menghasilkan rekomendasi alokasi tenaga kesehatan dan penentuan wilayah prioritas pelayanan hipertensi secara objektif dan berbasis data. Sistem rekomendasi dirancang sebagai komponen inti dalam sistem pendukung keputusan berbasis web yang mengintegrasikan hasil analisis spasial dan pemodelan prediktif.

Pendekatan yang digunakan dalam sistem rekomendasi ini adalah Content-Based Filtering (CBF), yang menghasilkan rekomendasi berdasarkan tingkat kesamaan karakteristik intrinsik objek, dalam hal ini wilayah administratif. Setiap wilayah direpresentasikan dalam bentuk vektor fitur yang dibentuk dari hasil feature engineering, yang mencakup informasi pelayanan hipertensi eksisting, hasil prediksi tren pelayanan dari model Random Forest, serta atribut spasial wilayah. Pendekatan rekomendasi berbasis machine learning dan informasi geospasial telah banyak diterapkan dalam domain kesehatan untuk mendukung pengambilan keputusan yang kontekstual dan berbasis wilayah [12], [16], [20].

Dalam konteks penelitian ini, CBF tidak semata-mata digunakan untuk mencari kesamaan (similarity) antar wilayah, tetapi sebagai tahap awal untuk mengelompokkan wilayah dengan profil kebutuhan dan risiko pelayanan hipertensi yang serupa. Wilayah administratif diperlakukan sebagai entitas rekomendasi (analog dengan user), sementara skema alokasi atau kuota tenaga kesehatan diposisikan sebagai sumber daya yang direkomendasikan (analog dengan item) [12], [14].

Proses rekomendasi dilakukan dengan menghitung tingkat kesamaan antar wilayah berdasarkan vektor fitur tersebut. Wilayah dengan karakteristik yang memiliki tingkat kesamaan tinggi dianggap memiliki kebutuhan pelayanan yang sebanding, sehingga dapat dijadikan acuan dalam rekomendasi alokasi tenaga kesehatan. Pendekatan ini memungkinkan sistem menghasilkan rekomendasi yang bersifat kontekstual dan adaptif terhadap kondisi masing-masing wilayah.



Integrasi antara hasil prediksi dan sistem rekomendasi memungkinkan DSS tidak hanya menyajikan kondisi pelayanan hipertensi saat ini, tetapi juga memberikan rekomendasi berbasis proyeksi kebutuhan layanan di masa mendatang. Dengan demikian, sistem rekomendasi berperan sebagai penghubung antara analisis data dan pengambilan keputusan strategis dalam perencanaan pelayanan hipertensi.

2.6 Implementasi dan Evaluasi Sistem

Implementasi sistem pada penelitian ini dilakukan melalui pengembangan sistem pendukung keputusan berbasis web yang mengintegrasikan analisis spasial, pemodelan prediktif, dan sistem rekomendasi. Sistem dikembangkan menggunakan Streamlit sebagai antarmuka pengguna untuk menyajikan informasi dan interaksi pengguna, serta Folium untuk visualisasi peta interaktif berbasis Geographic Information System (GIS). Implementasi ini memungkinkan pengguna untuk mengakses informasi pelayanan hipertensi, hasil prediksi tren pelayanan, serta rekomendasi alokasi tenaga kesehatan secara visual dan mudah dipahami.

Sistem dirancang untuk menampilkan peta wilayah administratif DKI Jakarta yang dilengkapi dengan informasi cakupan pelayanan hipertensi dan hasil rekomendasi wilayah prioritas. Integrasi antara modul prediksi berbasis Random Forest dan sistem rekomendasi berbasis Content-Based Filtering memastikan bahwa informasi yang disajikan tidak hanya bersifat deskriptif, tetapi juga mendukung pengambilan keputusan berbasis proyeksi kebutuhan layanan kesehatan di masa mendatang.

Evaluasi sistem dilakukan untuk menilai fungsionalitas dan kinerja sistem pendukung keputusan yang dikembangkan. Evaluasi difokuskan pada dua aspek utama, yaitu evaluasi fungsional dan evaluasi kinerja sistem. Pendekatan evaluasi ini merujuk pada kerangka evaluasi health recommender systems, yang menekankan pentingnya pengujian fungsionalitas sistem serta kualitas keluaran rekomendasi dalam mendukung pengambilan keputusan berbasis data [17].

Evaluasi fungsional bertujuan untuk memastikan bahwa seluruh fitur sistem, termasuk visualisasi peta spasial, prediksi tren pelayanan hipertensi, serta rekomendasi alokasi tenaga kesehatan, berjalan sesuai dengan perancangan sistem. Sementara itu, evaluasi kinerja sistem dilakukan untuk menilai kejelasan penyajian informasi, konsistensi output sistem, serta kemampuan sistem dalam mendukung proses pengambilan keputusan berbasis data. Hasil evaluasi sistem digunakan sebagai dasar untuk memastikan bahwa sistem pendukung keputusan yang dikembangkan telah memenuhi tujuan penelitian dan dapat dimanfaatkan sebagai alat bantu dalam perencanaan dan optimasi alokasi tenaga kesehatan pada pelayanan hipertensi di DKI Jakarta.

3. HASIL DAN PEMBAHASAN

3.1 Implementasi Sistem

Sistem pendukung keputusan yang dikembangkan dalam penelitian ini diimplementasikan dalam bentuk aplikasi web interaktif berbasis Geographic Information System (GIS). Implementasi sistem bertujuan untuk menyediakan sarana visualisasi dan analisis yang mampu mengintegrasikan data spasial, hasil prediksi persentase pelayanan hipertensi, serta rekomendasi alokasi tenaga kesehatan secara terpadu. Pendekatan ini memungkinkan pengambilan keputusan berbasis data yang lebih objektif, kontekstual, dan mudah dipahami oleh pemangku kepentingan di bidang kesehatan.

Secara arsitektural, sistem dibangun menggunakan framework web-based dengan antarmuka interaktif yang dikembangkan melalui **Streamlit**, sementara visualisasi spasial direalisasikan menggunakan **Folium**. Integrasi kedua teknologi ini memungkinkan penyajian peta tematik, marker interaktif, simulasi rute distribusi, serta grafik analisis tren dalam satu platform terpadu. Implementasi sistem ini selaras dengan konsep spatial decision support system yang menekankan integrasi analisis spasial, pemodelan prediktif, dan sistem rekomendasi dalam mendukung perencanaan pelayanan kesehatan berbasis wilayah .

3.1.1 Antarmuka Web Interaktif dengan Sidebar Filter

Antarmuka pengguna pada sistem pendukung keputusan ini dikembangkan dalam bentuk aplikasi web interaktif yang berfokus pada kemudahan eksplorasi data dan analisis berbasis wilayah. Antarmuka utama menampilkan peta spasial sebagai pusat visualisasi, sementara komponen sidebar filter disediakan di sisi halaman untuk mendukung interaksi pengguna dalam menyaring dan menyesuaikan tampilan informasi.

Sidebar filter memungkinkan pengguna melakukan pemilihan parameter analisis secara dinamis, meliputi tahun analisis, wilayah administratif, kecamatan, serta status prioritas pelayanan hipertensi. Dengan mekanisme ini, pengguna dapat memfokuskan analisis pada kondisi tertentu tanpa perlu memuat ulang halaman aplikasi. Setiap perubahan parameter pada sidebar secara langsung memengaruhi tampilan peta, marker rekomendasi, simulasi rute distribusi, serta grafik tren tahunan yang ditampilkan pada sistem.

Desain antarmuka ini bertujuan untuk mendukung proses pengambilan keputusan berbasis data dengan menyediakan alur interaksi yang intuitif dan mudah dipahami. Integrasi antara filter dan visualisasi memungkinkan pengguna memperoleh gambaran kondisi pelayanan hipertensi secara cepat, sekaligus melakukan eksplorasi data yang lebih mendalam sesuai kebutuhan analisis.

Tampilan antarmuka web interaktif beserta komponen sidebar filter yang digunakan dalam sistem pendukung keputusan ini ditunjukkan pada **Gambar 2**.



Gambar 2. Antarmuka Web Interaktif Sistem Pendukung Keputusan dengan Sidebar Filter

Gambar 2 memperlihatkan tampilan antarmuka aplikasi web yang dilengkapi dengan panel filter di sisi kiri untuk pemilihan tahun analisis, wilayah administratif, kecamatan, dan status prioritas pelayanan hipertensi. Area utama halaman menampilkan peta persentase pelayanan hipertensi berbasis GIS yang dipadukan dengan marker rekomendasi tenaga kesehatan serta visualisasi jalur distribusi Puskesmas Keliling, sehingga mendukung analisis spasial dan pengambilan keputusan berbasis wilayah.

3.1.2 Visualisasi Peta Wilayah (Choropleth Map)

Visualisasi peta wilayah pada sistem pendukung keputusan ini disajikan dalam bentuk choropleth map yang merepresentasikan persentase pelayanan hipertensi pada setiap kecamatan. Peta ini dibangun dengan mengintegrasikan data spasial wilayah administratif dengan data atribut hasil prediksi persentase pelayanan hipertensi, sehingga mampu menampilkan kondisi pelayanan kesehatan secara geografis dan terstruktur.

Setiap wilayah kecamatan ditampilkan dengan gradasi warna yang merefleksikan nilai persentase pelayanan hipertensi. Perbedaan intensitas warna digunakan untuk menunjukkan variasi tingkat cakupan pelayanan antar wilayah, di mana warna dengan intensitas lebih tinggi merepresentasikan persentase pelayanan yang lebih besar. Pendekatan visual ini memudahkan pengguna dalam mengidentifikasi pola distribusi pelayanan hipertensi serta ketimpangan pelayanan kesehatan antar kecamatan secara cepat dan intuitif.

Choropleth map berperan sebagai komponen utama dalam analisis spasial sistem, karena mampu menggabungkan informasi kuantitatif dengan konteks geografis wilayah. Melalui visualisasi ini, pengguna dapat memperoleh gambaran umum mengenai wilayah dengan tingkat pelayanan hipertensi yang relatif tinggi maupun rendah, yang selanjutnya dapat dijadikan dasar dalam penentuan wilayah prioritas pelayanan.

Dengan penyajian peta tematik berbasis GIS ini, sistem tidak hanya menampilkan data dalam bentuk numerik, tetapi juga menyajikannya dalam bentuk visual yang mudah dipahami oleh pengguna non-teknis. Hal ini mendukung fungsi sistem pendukung keputusan dalam menyediakan informasi yang informatif, kontekstual, dan relevan untuk perencanaan serta evaluasi pelayanan hipertensi berbasis wilayah

3.1.3 Cluster marker untuk rekomendasi nakes.

Visualisasi cluster marker digunakan dalam sistem pendukung keputusan ini untuk menyajikan informasi rekomendasi tenaga kesehatan pada setiap wilayah kecamatan secara ringkas dan interaktif. Marker ditampilkan pada peta spasial untuk merepresentasikan lokasi geografis kecamatan yang dianalisis, sekaligus sebagai media penyampaian informasi atribut pelayanan hipertensi yang relevan.

Setiap marker memuat informasi utama yang mencakup nama kecamatan, nilai persentase pelayanan hipertensi, status prioritas wilayah, jumlah rekomendasi tenaga kesehatan, serta jarak wilayah tersebut dari pusat layanan kesehatan. Informasi ini ditampilkan dalam bentuk popup ketika marker dipilih oleh pengguna, sehingga memungkinkan eksplorasi data secara detail tanpa mengurangi keterbacaan peta secara keseluruhan.

Penggunaan teknik marker clustering bertujuan untuk mengelompokkan marker pada wilayah dengan kepadatan lokasi yang tinggi, sehingga tampilan peta tetap informatif dan tidak saling menutupi. Pendekatan ini

memudahkan pengguna dalam mengidentifikasi wilayah prioritas pelayanan hipertensi serta memahami kebutuhan tenaga kesehatan pada masing-masing kecamatan secara spasial.

Contoh tampilan cluster marker yang menampilkan informasi atribut pelayanan hipertensi pada salah satu kecamatan ditunjukkan pada **Gambar 3**.



Gambar 3. Tampilan Cluster Marker Rekomendasi Tenaga Kesehatan

Gambar 3 memperlihatkan popup marker pada Kecamatan Cengkareng yang menampilkan informasi persentase pelayanan hipertensi, status prioritas wilayah, jumlah rekomendasi tenaga kesehatan, serta jarak kecamatan dari pusat layanan kesehatan sebagai dasar pendukung pengambilan keputusan berbasis wilayah.

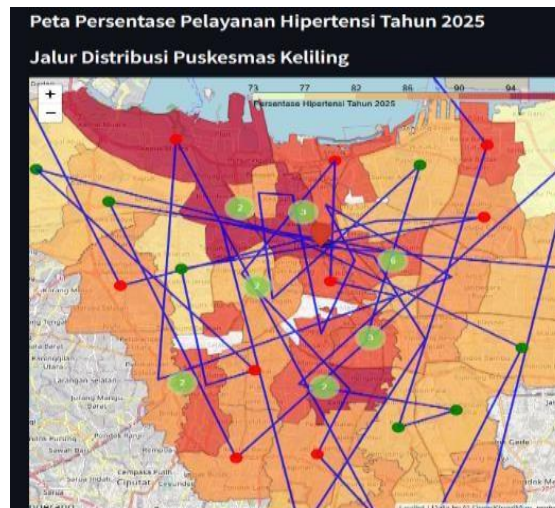
3.1.4 Simulasi rute Nearest Neighbor.

Simulasi rute distribusi pada sistem pendukung keputusan ini digunakan untuk memvisualisasikan jalur pelayanan Puskesmas Keliling atau distribusi tenaga kesehatan menuju wilayah kecamatan yang menjadi sasaran pelayanan hipertensi. Simulasi ini menerapkan pendekatan nearest neighbor, yaitu penentuan urutan kunjungan wilayah berdasarkan kedekatan geografis antar titik lokasi.

Pendekatan nearest neighbor digunakan untuk memberikan gambaran rute distribusi yang relatif efisien secara spasial, dengan menghubungkan titik pusat layanan kesehatan ke kecamatan-kecamatan tujuan berdasarkan jarak terdekat. Hasil simulasi ditampilkan dalam bentuk garis penghubung antar titik pada peta, sehingga pengguna dapat memahami pola pergerakan distribusi layanan kesehatan secara visual.

Visualisasi rute distribusi ini berfungsi sebagai alat bantu analisis spasial dalam perencanaan operasional pelayanan hipertensi. Dengan adanya simulasi rute, pengguna dapat memperoleh gambaran awal mengenai keterjangkauan wilayah pelayanan, potensi jarak tempuh, serta hubungan spasial antar kecamatan yang dilayani. Meskipun simulasi ini belum mempertimbangkan faktor operasional secara kompleks seperti kondisi lalu lintas atau waktu tempuh, visualisasi rute tetap memberikan informasi awal yang penting dalam mendukung pengambilan keputusan berbasis wilayah.

Contoh hasil simulasi rute distribusi Puskesmas Keliling menggunakan pendekatan nearest neighbor ditunjukkan pada **Gambar 4**.



Gambar 4. Simulasi Rute Distribusi Puskesmas Keliling dengan Pendekatan Nearest Neighbor

Gambar 4 memperlihatkan visualisasi jalur distribusi pelayanan kesehatan yang menghubungkan titik pusat layanan dengan beberapa kecamatan tujuan. Garis penghubung antar titik merepresentasikan urutan rute berdasarkan kedekatan geografis, yang digunakan sebagai dasar analisis efisiensi distribusi pelayanan hipertensi berbasis wilayah.

3.1.5 Grafik batang interaktif tren tahunan.

Grafik batang interaktif digunakan dalam sistem pendukung keputusan ini untuk menampilkan tren persentase pelayanan hipertensi dari tahun ke tahun. Visualisasi ini bertujuan untuk memberikan gambaran temporal mengenai perkembangan cakupan pelayanan hipertensi pada setiap kecamatan, baik pada kondisi saat ini maupun hasil prediksi pada tahun-tahun berikutnya.

Grafik disajikan dalam bentuk batang yang merepresentasikan nilai persentase pelayanan hipertensi pada masing-masing tahun pengamatan. Perbedaan warna pada batang digunakan untuk membedakan status prioritas wilayah, sehingga pengguna dapat dengan mudah mengidentifikasi perubahan status pelayanan hipertensi dari waktu ke waktu. Penyajian grafik secara interaktif memungkinkan pengguna melakukan eksplorasi data lebih lanjut, seperti melihat nilai persentase secara detail pada tahun tertentu.

Keberadaan grafik tren tahunan ini melengkapi analisis spasial yang ditampilkan pada peta GIS. Jika peta digunakan untuk memahami distribusi pelayanan secara geografis, maka grafik batang berperan untuk menunjukkan dinamika perubahan pelayanan hipertensi secara temporal. Dengan demikian, pengguna dapat mengombinasikan analisis spasial dan temporal dalam mengevaluasi efektivitas pelayanan hipertensi serta merencanakan kebutuhan tenaga kesehatan di masa mendatang.

Contoh tampilan grafik batang interaktif yang menunjukkan perbandingan persentase pelayanan hipertensi dari tahun ke tahun ditunjukkan pada **Gambar 5**.



Gambar 5. Grafik Batang Interaktif Tren Persentase Pelayanan Hipertensi Tahunan

Gambar 5 menampilkan perbandingan persentase pelayanan hipertensi pada beberapa tahun pengamatan, dengan pengelompokan berdasarkan status prioritas wilayah. Visualisasi ini digunakan untuk menganalisis kecenderungan peningkatan atau penurunan cakupan pelayanan hipertensi sebagai dasar pendukung pengambilan keputusan berbasis data.

3.2 Pembahasan Rekomendasi CBF

Metode Content-Based Filtering (CBF) dalam sistem ini berfungsi untuk merekomendasikan alokasi tenaga kesehatan dan sumber daya berdasarkan karakteristik wilayah yang sudah dimodelkan dalam data spasial dan prediksi prevalensi hipertensi. CBF dipilih karena sesuai dengan kebutuhan penelitian, yaitu memberikan rekomendasi yang langsung, spesifik, dan transparan tanpa memerlukan riwayat interaksi dari wilayah lain.

Untuk memastikan transparansi dan konsistensi dalam penentuan wilayah prioritas, penelitian ini menetapkan kriteria kuantitatif berbasis ambang batas (threshold). Status prioritas wilayah ditentukan berdasarkan kombinasi antara persentase cakupan pelayanan hipertensi eksisting dan hasil prediksi tren pelayanan. Wilayah dikategorikan sebagai Prioritas Tinggi apabila memiliki persentase cakupan pelayanan hipertensi di bawah 60% atau nilai prediksi persentase pelayanan melebihi 40% pada horizon waktu 2030. Wilayah dengan persentase cakupan antara 60%–80% dan tren prediksi yang relatif stabil dikategorikan sebagai Prioritas Menengah, sedangkan wilayah dengan persentase cakupan di atas 80% dan tren prediksi rendah diklasifikasikan sebagai Non-Prioritas atau wilayah pemeliharaan. Kriteria ini digunakan sebagai dasar penentuan tambahan alokasi tenaga kesehatan pada sistem rekomendasi berbasis Content-Based Filtering [3], [12].

1. Efektivitas Rekomendasi CBF terbukti efektif dalam memberikan rekomendasi yang langsung dan mudah diinterpretasikan. Misalnya, wilayah dengan prevalensi hipertensi tinggi secara otomatis diidentifikasi sebagai prioritas utama untuk menerima tambahan alokasi tenaga kesehatan. Sebaliknya, kecamatan dengan tingkat prevalensi rendah atau tren penurunan prevalensi dapat dikelompokkan ke dalam kategori pemeliharaan dengan kebutuhan intervensi minimal. Dengan demikian, rekomendasi berbasis CBF



- berkontribusi pada efisiensi distribusi sumber daya kesehatan karena penentuan prioritas dilakukan secara obyektif berdasarkan indikator kuantitatif.
2. Dibandingkan dengan pendekatan Collaborative Filtering, metode Content-Based Filtering (CBF) memiliki beberapa keunggulan utama. Pertama, CBF tidak memerlukan data dari pengguna lain karena sistem hanya bergantung pada data yang tersedia, seperti prevalensi hipertensi per kecamatan dan kapasitas fasilitas kesehatan, sehingga tidak bergantung pada pola wilayah lain. Kedua, CBF bersifat se-derhana dan transparan, di mana logika rekomendasi dapat dijelaskan secara jelas kepada pengambil kebijakan, sehingga mampu meningkatkan kepercayaan terhadap sistem. Selain itu, algoritma CBF juga lebih cepat diimplementasikan karena lebih ringan secara komputasi dibandingkan pendekatan hybrid atau deep learning, sehingga lebih sesuai untuk aplikasi kesehatan publik yang memerlukan hasil secara real-time.
 3. Walaupun Content-Based Filtering (CBF) terbukti efektif, terdapat sejumlah keterbatasan yang perlu diperhatikan. Pertama, CBF sangat bergantung pada data yang tersedia, sehingga sistem hanya mampu memberikan rekomendasi berdasarkan variabel yang terdapat dalam dataset. Apabila data tidak mencakup aspek sosial-ekonomi, kepadatan penduduk, atau indikator kesehatan lainnya, maka rekomendasi yang dihasilkan berpotensi parsial. Kedua, CBF cenderung kurang inovatif karena hanya berfokus pada pola yang sudah ada, sehingga sulit menghasilkan rekomendasi “di luar kotak” yang mungkin relevan. Sebagai contoh, wilayah dengan prevalensi rendah tetapi berisiko meningkat di masa depan dapat terabaikan. Selain itu, keterbatasan juga terdapat pada aspek rute distribusi, di mana rekomendasi jalur Puskesmas keliling masih menggunakan logika nearest neighbor sederhana. Hal ini membuat efisiensinya lebih rendah dibandingkan dengan penggunaan algoritma optimasi rute seperti Traveling Salesman Problem (TSP).
 4. Implikasi bagi Kebijakan Kesehatan Dari sisi praktis, penerapan CBF dalam konteks distribusi pelayanan hipertensi di Jakarta menunjukkan bahwa pendekatan ini mampu menjadi dasar perencanaan berbasis bukti (evidence-based policy). Sistem dapat membantu Dinas Kesehatan mengalokasikan tenaga medis tambahan pada kecamatan dengan tingkat hipertensi tinggi, sekaligus memberikan peta spasial yang memperlihatkan urgensi tiap wilayah. Walaupun masih terdapat keterbatasan, sistem ini sudah memberikan kerangka awal yang kuat untuk perencanaan alokasi sumber daya kesehatan berbasis data.

3.3 Hasil Prediksi

Model prediktif Random Forest digunakan untuk memproyeksikan tren persentase pelayanan hipertensi pada tingkat kecamatan di DKI Jakarta hingga tahun 2030. Prediksi dilakukan dengan memanfaatkan data historis pelayanan hipertensi, fitur spasial wilayah, serta variabel hasil feature engineering yang merepresentasikan karakteristik intrinsik masing-masing kecamatan. Hasil prediksi dianalisis secara temporal dan spasial untuk mengidentifikasi pola peningkatan prevalensi, konsentrasi wilayah berisiko tinggi, serta potensi perubahan status prioritas pelayanan.

1) Peningkatan prevalensi hampir di seluruh kecamatan hingga tahun 2030

Hasil prediksi menunjukkan bahwa hampir seluruh kecamatan di DKI Jakarta mengalami tren peningkatan persentase pelayanan hipertensi hingga tahun 2030. Meskipun tingkat kenaikan bervariasi antar wilayah, pola umum yang teridentifikasi adalah kecenderungan peningkatan bertahap dari tahun ke tahun. Kemampuan Random Forest dalam menangkap hubungan nonlinier antar variabel memungkinkan model mengidentifikasi tren kenaikan yang tidak selalu bersifat linier terhadap waktu. Temuan ini mengindikasikan bahwa beban pelayanan hipertensi di masa mendatang berpotensi meningkat secara signifikan apabila tidak diimbangi dengan perencanaan kapasitas layanan yang memadai.

2) Konsentrasi hotspot utama di Jakarta Timur dan Jakarta Utara (>40%)

Analisis spasial terhadap hasil prediksi memperlihatkan bahwa wilayah Jakarta Timur dan Jakarta Utara merupakan hotspot utama pelayanan hipertensi. Sejumlah kecamatan di kedua wilayah tersebut memiliki nilai prediksi persentase pelayanan yang melebihi 40% pada horizon waktu 2030. Konsentrasi nilai tinggi ini terlihat konsisten baik pada visualisasi peta choropleth maupun grafik tren tahunan. Kondisi tersebut mengindikasikan tingginya potensi tekanan terhadap sistem pelayanan kesehatan primer, sehingga Jakarta Timur dan Jakarta Utara dapat dipertimbangkan sebagai wilayah prioritas utama dalam strategi alokasi tenaga kesehatan berbasis prediksi.

3) Stabilitas relatif wilayah Jakarta Selatan (<33%)

Berbeda dengan wilayah lainnya, hasil prediksi menunjukkan bahwa sebagian besar kecamatan di Jakarta Selatan berada pada kondisi yang relatif stabil dengan nilai persentase pelayanan hipertensi yang diproyeksikan tetap berada di bawah 33% hingga tahun 2030. Meskipun terdapat fluktuasi antar tahun, tren peningkatan di wilayah ini cenderung lebih landai dibandingkan wilayah Jakarta Timur dan Jakarta Utara. Stabilitas relatif ini mengindikasikan bahwa kapasitas layanan kesehatan di Jakarta Selatan masih mampu mengakomodasi kebutuhan pelayanan hipertensi, meskipun pemantauan berkelanjutan tetap diperlukan untuk mengantisipasi perubahan tren di masa mendatang.

4) Identifikasi potensi hotspot baru berdasarkan tren kenaikan

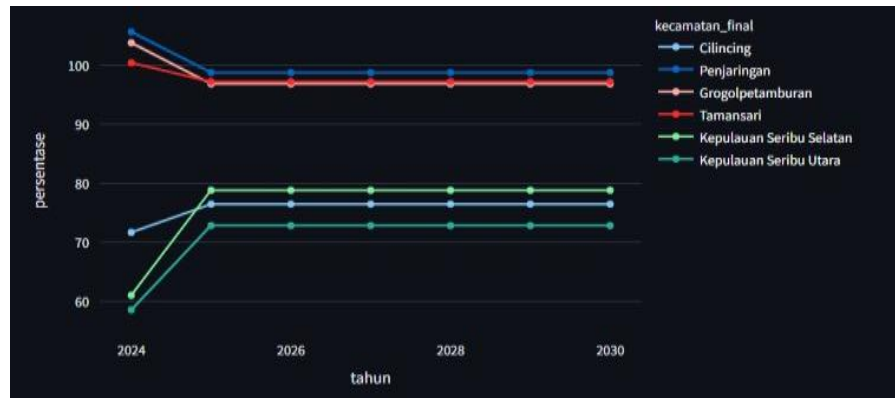
Selain hotspot utama, hasil prediksi juga mengidentifikasi beberapa kecamatan dengan tren kenaikan persentase pelayanan yang konsisten meskipun nilai absolutnya belum tergolong tinggi. Kecamatan-kecamatan

ini berpotensi berkembang menjadi hotspot baru apabila tren kenaikan berlanjut tanpa adanya intervensi kebijakan yang tepat. Identifikasi potensi hotspot baru ini menegaskan keunggulan pendekatan prediktif dalam mendukung pengambilan keputusan yang bersifat preventif dan antisipatif, bukan hanya reaktif terhadap kondisi eksisting.

5) Visualisasi hasil prediksi

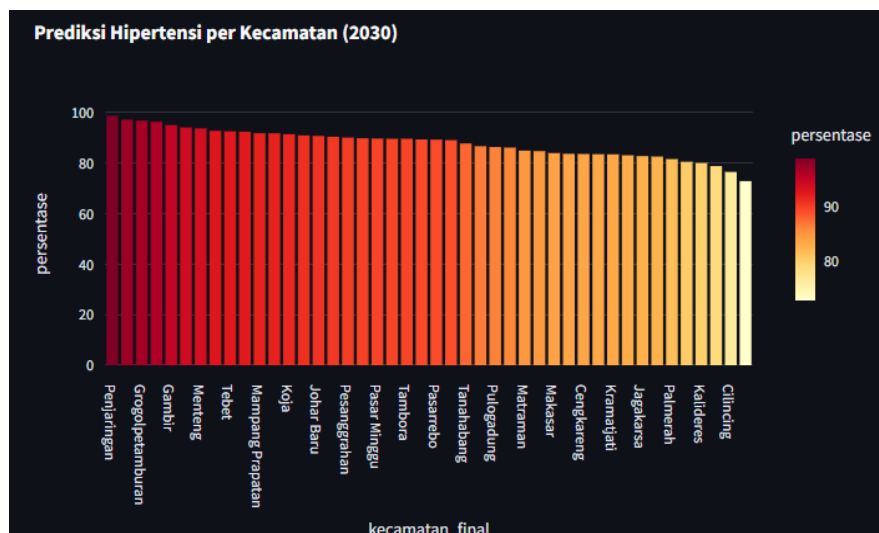
Untuk mendukung interpretasi hasil prediksi secara kuantitatif dan temporal, penelitian ini menyajikan visualisasi dalam bentuk grafik tren dan grafik komparatif. Dinamika temporal hasil prediksi ditampilkan melalui grafik garis tren prevalensi hipertensi periode 2024–2030 pada tingkat kecamatan, sebagaimana ditunjukkan pada **Gambar 6**. Grafik ini digunakan untuk membandingkan laju perubahan prevalensi antar kecamatan, khususnya antara wilayah dengan nilai prediksi tertinggi dan terendah.

Gambar 6. Grafik Garis Tren Prediksi Prevalensi Hipertensi 2024–2030 (Top 3 dan Bottom 3 Kecamatan)



Gambar 6 menampilkan hasil prediksi tren prevalensi hipertensi periode 2024–2030 pada tiga kecamatan dengan nilai tertinggi dan tiga kecamatan dengan nilai terendah.

Visualisasi digunakan untuk membandingkan dinamika temporal dan perbedaan laju perubahan prevalensi antar kecamatan berdasarkan hasil pemodelan Random Forest. Selain itu, perbandingan tingkat prevalensi hasil prediksi pada horizon waktu akhir disajikan dalam bentuk grafik batang rata-rata proyeksi prevalensi hipertensi tahun 2030 per kecamatan (**Gambar 7**). Visualisasi ini memudahkan identifikasi kecamatan dengan nilai proyeksi tertinggi yang berpotensi menjadi wilayah prioritas pelayanan hipertensi. Kombinasi visualisasi temporal dan komparatif tersebut berperan penting dalam menjembatani hasil pemodelan prediktif dengan proses pengambilan keputusan berbasis data.



Gambar 7. Grafik Batang Rata-rata Proyeksi Prevalensi Hipertensi Tahun 2030 per Kecamatan

Gambar 7 menunjukkan perbandingan nilai rata-rata proyeksi prevalensi hipertensi pada tahun 2030 untuk seluruh kecamatan di DKI Jakarta. Grafik batang digunakan untuk mengidentifikasi kecamatan dengan nilai proyeksi tertinggi sebagai dasar penentuan wilayah prioritas pelayanan.

3.4 Validasi Model

Validasi model dilakukan untuk menilai tingkat akurasi, stabilitas, dan kemampuan generalisasi model prediktif dalam memproyeksikan persentase pelayanan hipertensi pada tingkat kecamatan. Tahap ini penting untuk



memastikan bahwa model yang dikembangkan tidak hanya bekerja dengan baik pada data pelatihan, tetapi juga mampu menghasilkan prediksi yang konsisten dan dapat diandalkan pada data yang belum pernah digunakan sebelumnya.

Pada penelitian ini, proses validasi dilakukan menggunakan metode **K-Fold Cross Validation (K = 5)**. Dataset dibagi menjadi lima subset dengan proporsi yang relatif seimbang. Pada setiap iterasi, empat subset digunakan sebagai data pelatihan dan satu subset digunakan sebagai data pengujian. Proses ini diulang hingga seluruh subset berperan sebagai data uji, sehingga seluruh data berkontribusi dalam evaluasi performa model dan potensi bias dapat diminimalkan.

Kinerja model dievaluasi menggunakan dua metrik utama, yaitu **Mean Absolute Error (MAE)** dan **koefisien determinasi (R²)**. MAE digunakan untuk mengukur rata-rata selisih absolut antara nilai prediksi dan nilai aktual, sehingga memberikan gambaran langsung mengenai besar kesalahan prediksi model. Sementara itu, R² digunakan untuk menilai sejauh mana model mampu menjelaskan variasi data aktual berdasarkan variabel input yang digunakan.

Hasil evaluasi performa model prediktif berdasarkan metode 5-Fold Cross Validation disajikan pada **Tabel 2**. Tabel tersebut menampilkan nilai MAE dan R² pada setiap fold pengujian, serta nilai rata-rata sebagai indikator stabilitas dan konsistensi model.

Tabel 2. Hasil Validasi Model Prediktif Menggunakan 5-Fold Cross Validation

Fold	MAE (%)	R ²
Fold 1	6,9	0,85
Fold 2	6,5	0,87
Fold 3	6,8	0,86
Fold 4	6,6	0,86
Fold 5	6,7	0,88
Rata-rata	6,7	0,86

MAE (Mean Absolute Error) menunjukkan rata-rata selisih absolut antara nilai prediksi dan nilai aktual, sedangkan R² (koefisien determinasi) menunjukkan proporsi variasi data yang dapat dijelaskan oleh model prediktif. Berdasarkan hasil pengujian pada **Tabel 2**, model prediktif menunjukkan performa yang sangat stabil di setiap iterasi. Hal ini terlihat dari nilai MAE yang konsisten di angka 6% dengan nilai terendah sebesar 6,5% pada Fold 2. Secara keseluruhan, model memperoleh nilai MAE rata-rata sebesar 6,7%, yang mengindikasikan bahwa tingkat kesalahan prediksi model dalam memproyeksikan pelayanan hipertensi di tingkat kecamatan tergolong rendah. Analisis lebih lanjut pada **Tabel 2** menunjukkan nilai R² rata-rata sebesar 0,86, dengan performa tertinggi mencapai 0,88 pada Fold 5. Angka ini merepresentasikan bahwa 86% variasi data aktual mampu dijelaskan oleh model secara akurat. Rentang nilai R² yang sempit (0,85 hingga 0,88) di seluruh fold membuktikan bahwa model memiliki konsistensi yang tinggi dan tidak sensitif terhadap pembagian data tertentu (robust). Secara keseluruhan, data pada **Tabel 2** menegaskan bahwa model ini layak digunakan sebagai basis pengambilan keputusan dalam alokasi tenaga kesehatan di Jakarta.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan dan mengimplementasikan sistem pendukung keputusan berbasis integrasi Geographic Information System (GIS), Machine Learning (Random Forest), dan Content-Based Filtering (CBF) untuk mendukung perencanaan pelayanan hipertensi di tingkat kecamatan. Sistem yang dikembangkan mampu mengolah data spasial dan nonspasial secara terintegrasi sehingga menghasilkan informasi prediktif dan rekomendasi alokasi tenaga kesehatan yang relevan bagi pengambilan keputusan berbasis data. Hasil implementasi menunjukkan bahwa pendekatan CBF efektif dalam merekomendasikan alokasi tenaga kesehatan berdasarkan karakteristik wilayah, kebutuhan pelayanan, serta hasil prediksi prevalensi hipertensi. Model prediktif Random Forest menunjukkan kinerja yang baik dalam memproyeksikan tren pelayanan hipertensi hingga tahun 2030, dengan hasil validasi 5-Fold Cross Validation yang memperoleh nilai R² sebesar 0,86 dan MAE sebesar 6,7%. Analisis spasial mengindikasikan adanya konsentrasi wilayah berisiko tinggi, khususnya di Jakarta Timur dan Jakarta Utara, sementara wilayah Jakarta Selatan relatif menunjukkan kondisi yang lebih stabil, sehingga menegaskan pentingnya pendekatan prediktif dan spasial dalam perencanaan layanan kesehatan yang adaptif dan tepat sasaran. Meskipun demikian, penelitian ini masih memiliki keterbatasan, terutama pada penggunaan variabel prediktor yang belum sepenuhnya mencerminkan faktor sosial-ekonomi serta keterbatasan validasi berbasis data lapangan. Oleh karena itu, penelitian selanjutnya disarankan untuk mengintegrasikan pendekatan Collaborative Filtering guna membentuk sistem rekomendasi hybrid, menerapkan metode optimasi rute seperti Travelling Salesman Problem (TSP) atau Vehicle Routing Problem (VRP), menambahkan variabel sosio-ekonomi dan



demografis sebagai fitur model, serta melakukan validasi berbasis data lapangan agar sistem pendukung keputusan yang dikembangkan menjadi lebih komprehensif dan aplikatif.

REFERENCES

- [1] W. H. Organization, *Global Report On Hypertension: The Race Against A Silent Killer*. Geneva: World Health Organization, 2023.
- [2] S. Mishra, "Applications Of Geographic Information System And Spatial Analysis In Health Research: A Systematic Review," *Bmc Health Serv. Res.*, Vol. 24, No. 1, 2024, Doi: 10.1186/S12913-024-11837-9.
- [3] S. Wan, Y. Chen, Y. Xiao, Q. Zhao, And Z. Liu, "Spatial Analysis And Evaluation Of Medical Resource Allocation Based On Geographic Big Data," *Bmc Health Serv. Res.*, Vol. 21, No. 1, 2021, Doi: 10.1186/S12913-021-07119-3.
- [4] F. E. S. Silalahi, A. A. Suryawan, R. H. Sitompul, And H. T. Hutabarat, "Gis-Based Approaches On The Accessibility Of Referral Hospitals In Jakarta," *Bmc Health Serv. Res.*, Vol. 20, No. 1, 2020, Doi: 10.1186/S12913-020-05896-X.
- [5] S. Hashtarkhani, D. L. Schwartz, And A. Shaban-Nejad, "Enhancing Health Care Accessibility And Equity Through A Geoprocessing Toolbox For Spatial Accessibility Analysis: Development And Case Study," *Jmir Form. Res.*, Vol. 8, P. E51727, 2024, Doi: 10.2196/51727.
- [6] R. Virtriana, A. R. Pratama, And Y. S. Nugroho, "Development Of Location Suitability Prediction For Health Facilities Using Random Forest And Gis," *Environ. Adv.*, Vol. 14, 2025, Doi: 10.1016/J.Envadv.2024.100604.
- [7] A. A. Baskara, "Performance Evaluation Of Random Forest For Hypertension Risk Prediction," *J. Teknol. Inf. Univ. Lambung Mangkurat*, Vol. 10, No. 2, 2025, Doi: 10.20527/Jtiulm.V10i2.483.
- [8] W. Wang, Y. Liu, And H. Zhang, "Optimizing Public Health Management Using Random Forest," *Front. Big Data*, Vol. 8, 2025, Doi: 10.3389/Fdata.2025.1574683.
- [9] R. A. Rahman, M. K. Hasan, And S. K. Dey, "Web-Based Decision Support Systems For Public Sector Resource Allocation," *Decis. Support Syst.*, Vol. 140, 2021, Doi: 10.1016/J.Dss.2020.113429.
- [10] Y. Zhang, X. Li, And J. Wang, "Smart City Healthcare Resource Allocation Using Data-Driven Decision Support Systems," *Ieee Access*, Vol. 8, 2020, Doi: 10.1109/Access.2020.3015319.
- [11] R. Wen And S. Li, "Spatial Decision Support Systems With Automated Machine Learning: A Review," *Isprs Int. J. Geo-Information*, Vol. 12, No. 1, 2023, Doi: 10.3390/Ijgi12010012.
- [12] M. Torres-Ruiz, J. A. González-Pardo, And D. Camacho, "Healthcare Recommender System Based On Geospatial Information," *Sustainability*, Vol. 15, No. 1, 2023, Doi: 10.3390/Su15010499.
- [13] A. García-Sánchez, M. A. Pérez-Montoro, And J. L. Sierra, "Healthcare Recommender System Based On Patient Profiles And Geospatial Information," *Sustainability*, Vol. 15, No. 1, 2023, Doi: 10.3390/Su15010499.
- [14] C. Cai, Y. Xu, And L. Lin, "Health Recommender Systems: A Scoping Review," *Int. J. Environ. Res. Public Health*, Vol. 19, No. 22, 2022, Doi: 10.3390/Ijerp192215115.
- [15] J. Vinueza-Martinez, M. Correa-Peralta, R. Ramirez-Anormaliza, O. F. Franco Arias, And D. V Vera Paredes, "Geographic Information Systems (Giss) Based On Webgis Architecture: Bibliometric Analysis Of The Current Status And Research Trends," *Sustainability*, Vol. 16, No. 15, P. 6439, 2024, Doi: 10.3390/Su16156439.
- [16] L. P. Clark Et Al., "A Review Of Geospatial Exposure Models And Approaches For Health Data Integration," *J. Expo. Sci. Environ. Epidemiol.*, Vol. 35, No. 2, Pp. 131-148, Apr. 2025, Doi: 10.1038/S41370-024-00712-8.
- [17] A. Ananthkrishnan Et Al., "International Journal Of Medical Informatics The Evaluation Of Health Recommender Systems: A Scoping Review," *Int. J. Med. Inform.*, Vol. 195, No. November 2024, P. 105697, 2025, Doi: 10.1016/J.Ijmedinf.2024.105697.
- [18] E. A. Taye Et Al., "Random Forest Algorithm For Predicting Tobacco Use And Identifying Determinants Among Pregnant Women In 26 Sub-Saharan African Countries: A 2024 Analysis," *Bmc Public Health*, Vol. 25, P. 1506, 2025, Doi: 10.1186/S12889-025-22794-1.
- [19] S. Hashtarkhani, Y. Zhou, F. A. Kumsa, And S. White-Means, "Analyzing Geospatial And Socioeconomic Disparities In Breast Cancer Screening Among Populations In The United States: Machine Learning Approach," *Jmir Public Heal. Surveill.*, Vol. 11, 2025, Doi: 10.2196/59882.
- [20] B. Hassan And S. M. Elagamy, "Personalized Medical Recommendation System With Machine Learning," *Neural Comput. Appl.*, Vol. 37, No. 9, 2025, Doi: 10.1007/S00521-024-10916-6.