

PREDIKSI JUMLAH KUNJUNGAN PASIEN RAWAT JALAN MENGGUNAKAN METODE REGRESI LINIER SEDERHANA

Bangkit Indarmawan Nugroho¹, Fatkhurrozak², Wresty Andriani³, Gunawan Gunawan⁴

Teknik Informatika^{1,2,4}, STMIK YMI Tegal

Sistem Informasi³, STMIK YMI Tegal

Jalan Pendidikan 1, Kota Tegal 52142, Indonesia

E-mail: *bangkit_in@stmik-tegal.ac.id¹, fatkhurrozak35@gmail.com², wresty.andriani@stmik-tegal.ac.id³, gunawan@stmik-tegal.ac.id⁴

Abstrak - Penelitian ini berfokus pada prediksi jumlah kunjungan pasien rawat jalan di Rumah Sakit Islam Pati menggunakan metode regresi linier sederhana. Topik ini dipilih karena pentingnya perencanaan manajemen rumah sakit yang lebih efektif dan efisien, mengingat fluktuasi dan ketidakpastian jumlah kunjungan pasien seringkali menjadi tantangan. Tujuan dari penelitian ini adalah mengembangkan model prediksi yang memberikan estimasi jumlah kunjungan pasien dengan lebih akurat, sehingga dapat digunakan dalam proses perencanaan. Metode penelitian yang digunakan adalah regresi linier sederhana dengan analisis data kunjungan pasien dari tahun 2021, 2022, dan 2023. Data yang digunakan telah melalui proses pra-pemrosesan untuk memastikan validitas dan kebersihannya. Hasil penelitian menunjukkan bahwa model regresi linier sederhana memberikan performa yang baik dalam memprediksi jumlah kunjungan pasien. Hasil penelitian ini menunjukkan bahwa prediksi jumlah kunjungan pasien dapat membantu rumah sakit dalam merencanakan sumber daya dan meningkatkan efisiensi operasional. Kesimpulan dari penelitian ini adalah bahwa metode regresi linier sederhana dapat menjadi solusi dalam menghadapi ketidakpastian jumlah kunjungan pasien, memungkinkan manajemen rumah sakit melakukan perencanaan yang lebih baik dan efisien.

Kata Kunci: Efisiensi Operasional, Manajemen Rumah Sakit, Perencanaan, Prediksi Kunjungan Pasien, Regresi Linier Sederhana.

I. PENDAHULUAN

Sebagai penyedia jasa layanan kesehatan, rumah sakit bertanggung jawab untuk memberikan layanan terbaik kepada masyarakat (Komalawati & Triswandi, 2022). Penyempurnaan layanan ini tercermin dalam peningkatan manajemen rumah sakit, termasuk manajemen sumber daya manusia, material, dan keuangan. Dalam konteks manajemen, perencanaan memegang peranan penting. Melalui perencanaan, tujuan, strategi, dan pedoman dapat ditetapkan untuk melaksanakan kegiatan manajerial (Amon & Harliansyah, 2022).

Namun, pihak manajemen rumah sakit sering menghadapi tantangan dalam merencanakan. Fluktuasi dan ketidakpastian jumlah kunjungan pasien mengakibatkan perencanaan menjadi tidak efisien (Apriliani et al., 2021). Oleh karena itu, diperlukan upaya untuk memperkirakan jumlah kunjungan pasien. Meskipun peramalan tidak selalu sempurna, upaya untuk mengurangi kesalahan dalam peramalan dapat dilakukan dengan mencari metode yang dapat memberikan hasil peramalan yang akurat (S. Hidayatullah et al., 2022). Dengan memiliki perkiraan jumlah kunjungan pasien yang lebih akurat, rumah sakit dapat melakukan perencanaan yang lebih efektif dan efisien.

Rumah sakit menawarkan layanan menyeluruh bagi individu yang memerlukan perawatan. Layanan yang disediakan meliputi rawat inap, rawat jalan, dan penanganan darurat. Sebagai institusi yang umumnya melayani berbagai macam penyakit,

rumah sakit umum bertujuan untuk memenuhi kebutuhan pasien dan memberikan solusi terhadap masalah kesehatan mereka (Ananda et al, 2023). Pasien mengandalkan rumah sakit sebagai tempat untuk mendapatkan perawatan medis yang dapat membantu penyembuhan dan pemulihan kondisi kesehatan mereka (Madjid, 2020). Oleh karena itu, harapan pasien terhadap rumah sakit adalah adanya pelayanan yang responsif, efisien, dan memberikan kenyamanan dalam mengatasi keluhan mereka.

Dalam sistem kesehatan modern, manajemen dan perencanaan sumber daya yang efektif merupakan aspek krusial untuk meningkatkan efisiensi pelayanan kesehatan. Salah satu elemen penting dalam hal ini adalah prediksi jumlah kunjungan pasien rawat jalan. Prediksi yang akurat dapat membantu dalam perencanaan tenaga kerja, pengaturan jadwal, pengadaan obat-obatan, dan optimalisasi penggunaan fasilitas medis. Namun, ketidakpastian dan variabilitas dalam pola kunjungan pasien seringkali menjadi tantangan signifikan bagi manajemen rumah sakit dan klinik kesehatan (Hasibuan, 2020).

Masalah utama yang dihadapi dalam manajemen kunjungan pasien rawat jalan adalah ketidakmampuan untuk memprediksi dengan akurat jumlah pasien yang akan datang pada periode tertentu (Wijaya et al., 2022). Ketidakmampuan ini dapat menyebabkan berbagai isu seperti penumpukan pasien, keterlambatan layanan, kekurangan staf, dan pengelolaan sumber daya yang tidak efisien. Mengingat pentingnya prediksi yang

akurat, penelitian ini dilakukan untuk mengatasi masalah ini dengan menggunakan pendekatan kuantitatif.

Penelitian ini dilakukan untuk mengembangkan model prediksi yang sederhana namun efektif untuk memperkirakan jumlah kunjungan pasien rawat jalan. Dengan prediksi yang lebih baik, fasilitas kesehatan dapat merencanakan operasional mereka dengan lebih efisien, yang pada gilirannya akan meningkatkan kualitas layanan kesehatan yang diberikan kepada masyarakat (Salsabila et al., 2024). Penelitian ini bertujuan untuk memecahkan problem prediksi jumlah kunjungan pasien rawat jalan dengan menggunakan metode regresi linier sederhana. Metode ini dipilih karena kemudahannya dalam implementasi dan interpretasi hasil, serta kemampuannya dalam menangani data dengan hubungan linier yang cukup baik.

Untuk mengatasi masalah tersebut, model prediksi menggunakan metode regresi linier sederhana. Data historis kunjungan pasien dianalisis untuk mengidentifikasi pola dan tren yang relevan, yang kemudian digunakan untuk membangun model prediksi. Saat ini, berbagai metode prediksi telah digunakan dalam konteks kesehatan, termasuk machine learning dan model statistik yang kompleks (Made et al., 2024). Namun, inovasi yang diusulkan dalam penelitian ini adalah penggunaan regresi linier sederhana yang, meskipun lebih sederhana, diharapkan dapat memberikan prediksi yang cukup akurat dengan interpretasi yang lebih mudah dan implementasi yang lebih sederhana.

Regresi linier sederhana adalah metode statistik yang digunakan untuk menganalisis dan memodelkan hubungan antara satu variabel independen (X) dan satu variabel dependen (Y). Persamaannya dinyatakan sebagai ($Y = a + bX$), di mana (a) adalah intercept yang menunjukkan nilai Y ketika $X = 0$, dan (b) adalah slope yang menunjukkan perubahan rata-rata pada Y untuk setiap perubahan satu unit pada X (Azahra, 2022). Berdasarkan hasil dari prediksi menggunakan model regresi linier sederhana, langkah selanjutnya akan melibatkan proses pengoptimalan dengan memanfaatkan nilai RMSE dan R-square (Kusuma & Hidayat, 2024). Evaluasi menggunakan kedua metrik tersebut akan memberikan wawasan yang lebih mendalam tentang keakuratan dan kecocokan model terhadap data yang ada. Selain itu, penggunaan kedua metrik ini akan memungkinkan kita untuk mengidentifikasi area-area spesifik di mana model mungkin perlu diperbaiki atau disesuaikan. Dengan demikian, proses pengoptimalan akan fokus pada meningkatkan kemampuan model dalam menjelaskan variasi dalam data dan mengurangi kesalahan prediksi sebanyak mungkin. Langkah-langkah ini akan membantu memastikan bahwa model regresi linier sederhana dapat memberikan perkiraan yang lebih akurat dan

dapat diandalkan terkait jumlah kunjungan pasien rawat jalan pada masa mendatang.

Penelitian ini akan mengumpulkan dan menganalisis data historis kunjungan pasien rawat jalan, membangun model prediksi menggunakan regresi linier sederhana, dan menguji keakuratan model ini terhadap data aktual. Diharapkan bahwa model ini akan memberikan prediksi yang cukup akurat untuk digunakan dalam perencanaan operasional harian fasilitas kesehatan, sehingga meningkatkan efisiensi dan kualitas pelayanan kesehatan.

II. TINJAUAN PUSTAKA

Penelitian ini berfokus pada prediksi jumlah kunjungan pasien rawat jalan menggunakan metode regresi linier sederhana. Regresi linier sederhana merupakan teknik statistik yang digunakan untuk menjelaskan hubungan antara satu variabel independen dan satu variabel dependen melalui persamaan linear.

Regresi linier sederhana digunakan dalam analisis statistik untuk memprediksi nilai suatu variabel berdasarkan nilai variabel lainnya (Lase & Riandari, 2020). Persamaan dasarnya adalah ($Y = a + bX$), di mana (Y) adalah variabel dependen yang diprediksi, (X) adalah variabel independen, (a) adalah intercept, dan (b) adalah slope. Metode ini efektif untuk mengidentifikasi tren dan hubungan linear dalam data historis (Azahra, 2022).

Salah satu pendekatan yang paling sering digunakan adalah metode regresi linier sederhana. Regresi linier dapat digunakan untuk memprediksi kunjungan pasien dengan tingkat akurasi yang cukup baik, terutama ketika data memiliki hubungan linier yang kuat. Model ini efektif untuk aplikasi praktis di lingkungan dengan keterbatasan sumber daya teknis. Namun, model ini mungkin kurang efektif jika terdapat variabilitas yang tinggi dalam data (Harizahayu et al., 2023).

Pendekatan lain yang semakin populer adalah penggunaan teknik machine learning. Sebagai contoh, penelitian yang diterbitkan (Syahputra, 2022) menggunakan model Random Forest untuk memprediksi jumlah kunjungan pasien rawat jalan. Penelitian ini berhasil meningkatkan akurasi prediksi dibandingkan dengan metode tradisional, yang menunjukkan keunggulan machine learning dalam menangani data yang kompleks dan beragam. Meskipun demikian, kompleksitas model ini sering kali menjadi hambatan dalam implementasi praktis di fasilitas kesehatan yang mungkin tidak memiliki sumber daya teknis yang cukup.

Penelitian oleh (Rustam et al., 2022) mengeksplorasi penggunaan model ARIMA (AutoRegressive Integrated Moving Average) untuk memprediksi kunjungan pasien di RS Surabaya. Model ini memperhitungkan pola musiman dan tren dalam data historis, sehingga memberikan hasil yang

lebih akurat untuk prediksi jangka panjang. Namun, kelemahan dari model ini adalah kebutuhan akan data historis yang panjang dan preprocessing yang rumit.

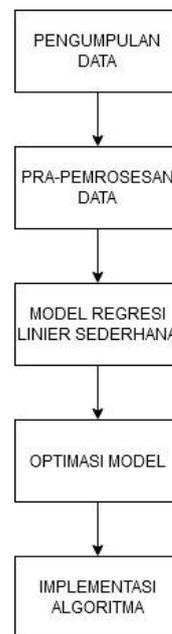
Penelitian yang dilakukan oleh (Indahningrum & Lia Dwi Jayanti, 2020) mengadopsi model Long Short-Term Memory (LSTM) dalam deep learning untuk meramalkan permintaan *perishable product* di Retail Sayur Tosaga. Evaluasi menunjukkan bahwa LSTM mampu mengelola data time series dengan lebih efektif dibandingkan metode tradisional, menghasilkan prediksi yang akurat, terutama untuk pola data yang kompleks dan tidak linear. Meskipun demikian, implementasi model ini menuntut infrastruktur komputasi yang canggih dan keahlian dalam bidang *data science*.

Beberapa penelitian juga telah mencoba menggabungkan berbagai metode untuk meningkatkan akurasi prediksi. Misalnya, studi oleh (Fitri, 2023) mengombinasikan regresi linier dengan model *machine learning* untuk mengembangkan model *hybrid* yang mampu memberikan prediksi lebih akurat dan dapat diadaptasi untuk berbagai kondisi data. Hasil penelitian menunjukkan bahwa model *hybrid* yang dikembangkan, terutama menggunakan Random Forest Regression, mampu memberikan akurasi tertinggi dalam prediksi harga rumah.

Berdasarkan literatur yang ada, dapat disimpulkan bahwa meskipun model *machine learning* dan *deep learning* menawarkan tingkat akurasi yang tinggi, mereka sering kali memerlukan sumber daya yang signifikan dan keahlian teknis untuk implementasinya. Di sisi lain, metode regresi linier sederhana, meskipun kurang kompleks, tetapi menjadi pilihan yang layak karena kemudahannya dalam implementasi dan interpretasi hasil. Oleh karena itu, penelitian ini berfokus pada penggunaan regresi linier sederhana untuk memprediksi jumlah kunjungan pasien rawat jalan, dengan harapan dapat menyediakan solusi yang praktis dan efisien untuk fasilitas kesehatan.

III. METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini terdiri pengumpulan data, pemrosesan data, analisi data, validasi model dan interpretasi hasil. Gambar 1. Menunjukkan alur proses prediksi rawat jalan.



Gambar 1. Alur Penelitian

A. Pengumpulan data

Dalam fase awal pengumpulan data untuk penelitian ini, peneliti dengan cermat menghimpun informasi terkait jumlah kunjungan pasien rawat jalan di Rumah Sakit Islam Pati tahun 2022 dan 2023, dengan memperoleh data dari sumber daring terpercaya *data.go.id*. Data-data ini meliputi kunjungan ke berbagai klinik di rumah sakit tersebut. Setiap detail kunjungan, mulai dari poliklinik Anak hingga Saraf, disaring secara seksama untuk memastikan integritas dan keakuratan sebelum dilanjutkan ke tahap analisis. Langkah ini diambil dengan tujuan memastikan bahwa dataset yang disajikan dalam penelitian memiliki kualitas tinggi serta relevansi yang diperlukan untuk analisis yang komprehensif dan terperinci.

B. Pra-pemrosesan data

Tahap pemrosesan data dimulai dengan pengorganisasian data kunjungan pasien rawat jalan ke berbagai klinik di Rumah Sakit Islam Pati selama rentang waktu 2021-2023. Langkah awal yang krusial dalam proses ini adalah melakukan pengecekan menyeluruh terhadap setiap entri data untuk memastikan akurasi dan validitas informasi yang tercatat. Upaya ini dilakukan dengan tujuan untuk memverifikasi integritas keseluruhan dataset, serta memastikan bahwa tidak terdapat nilai-nilai yang hilang atau tidak lengkap di dalamnya. Keberadaan nilai-nilai yang hilang atau tidak lengkap dapat memberikan dampak signifikan terhadap hasil analisis yang akan dilakukan, sehingga langkah ini menjadi sangat penting untuk dijalankan dengan saksama. Selanjutnya, dalam situasi di mana terdapat perbedaan skala yang cukup besar antar variabel dalam dataset, proses normalisasi data akan dilakukan untuk memastikan

bahwa setiap variabel memiliki pengaruh yang seimbang dalam pembentukan model analisis, sehingga tidak terjadi bias atau ketimpangan dalam hasil akhir yang diperoleh.

C. Model regresi linier sederhana

Model regresi linier sederhana diperoleh dengan mengestimasi hubungan linier antara variabel dependen (y) dan variabel independen (x) menggunakan metode kuadrat terkecil. Model ini memungkinkan prediksi jumlah pada tahun tertentu dengan mensubstitusikan nilai tahun ke dalam persamaan. Model ini ditunjukkan dengan persamaan (1).

$$Y = a + bx \tag{1}$$

Di mana a adalah intercept (titik potong sumbu y) dan b adalah slope (kemiringan garis).

D. Optimasi model

Proses optimasi dimulai dengan pembuatan model regresi linier sederhana, yang kemudian disesuaikan dengan data latih. Selanjutnya, kinerja model akan dievaluasi menggunakan teknik validasi silang dan metrik kinerja seperti Root mean squared error (RMSE) merupakan metrik yang mengukur kesalahan atau perbedaan antara nilai prediksi model dengan nilai sebenarnya. Semakin rendah nilai RMSE, semakin baik kemampuan model dalam memprediksi data. Rumus yang akan digunakan seperti pada persamaan (2).

$$RMSE = \sqrt{(\sum(y_{aktual} - y_{prediksi})^2 / n)} \tag{2}$$

Keterangan:

- y_aktual adalah nilai sebenarnya dari variabel dependen
- y_prediksi adalah nilai prediksi dari model
- n adalah jumlah data atau observasi

Sedangkan coefficient of determination (R-squared) atau koefisien determinasi adalah metrik yang menunjukkan seberapa baik model regresi linier dapat menjelaskan variasi dalam data. Nilai R-squared berkisar antara 0 hingga 1, di mana semakin mendekati 1, semakin baik model tersebut menjelaskan variasi data. Rumus digunakan seperti pada persamaan (2).

$$R^2 = 1 - (\sum(y_{aktual} - y_{prediksi})^2 / \sum(y_{aktual} - y_{rata_rata})^2) \tag{2}$$

Keterangan:

- y_aktual adalah nilai sebenarnya dari variabel dependen.
- y_prediksi adalah nilai prediksi dari model.
- y_rata_rata adalah rata-rata dari nilai aktual variabel dependen.

hal ini bertujuan untuk memastikan akurasi prediksi yang optimal.

E. Implementasi model

Algoritma regresi linier sederhana akan diimplementasikan menggunakan bahasa pemrograman statistik, seperti Python, dengan memanfaatkan salah satu library yang populer seperti scikit-learn, atau menggunakan lingkungan pemrograman R yang kaya fitur. Implementasi ini memungkinkan penggunaan algoritma dalam menerapkan model prediksi ke dalam praktik pemrograman yang terstandarisasi dan efisien. Dengan menggunakan bahasa pemrograman yang mendukung analisis data yang kuat dan memiliki perpustakaan yang luas seperti scikit-learn, para peneliti dapat dengan mudah mengakses berbagai alat dan teknik yang diperlukan untuk mengembangkan dan menguji model prediksi regresi linier. Penggunaan bahasa pemrograman statistik ini juga memungkinkan para peneliti untuk melakukan analisis data yang kompleks dengan kecepatan dan akurasi yang tinggi, sehingga mempercepat proses penelitian dan meningkatkan keandalan hasil.

Dengan demikian, implementasi algoritma regresi linier dalam lingkungan pemrograman statistik seperti Python atau R merupakan langkah yang penting dan efektif dalam pengembangan model prediksi yang dapat diandalkan dalam memprediksi jumlah kunjungan pasien rawat jalan.

IV. HASIL DAN PEMBAHASAN

A. Pengumpulan data.

Pengumpulan data dilakukan dengan mengambil data dari internet, tepatnya dari platform data.go.id, yang menyediakan beragam informasi terkait jumlah kunjungan pada berbagai poliklinik dari tahun 2021, 2022, dan 2023. Data yang terhimpun meliputi informasi tentang kunjungan pasien di poliklinik Anak, Bedah, Dalam, Fisioterapi, Jantung, Psikiatri, dan beragam poliklinik lainnya. Dalam proses ini, informasi yang terkandung dalam dataset tersebut menjadi dasar yang penting untuk analisis lebih lanjut terkait tren dan pola kunjungan pasien rawat jalan selama periode waktu yang diteliti.

Tabel 1. Data rawat jalan RS Islam Pati

No	Poloklinik	Tahun		
		2021	2022	2023
1	Anak	800	1.333	1.421
2	Bedah	2.977	3.270	3.294
3	Dalam	10.277	11.696	13.221
4	Gigi	343	351	456
5	Jantung	2.205	3.892	4.779
6	Kebidanan	1.997	2.255	1.898
7	Mata	5.226	5.824	5.677
8	Ortopedi	434	537	673

9	Paru	1.115	2.429	3.079
10	Psikiatri	-	1.158	1.535
11	Psikologi	61	70	107
11	Rehab Medik	-	4.718	6.645
12	THT	944	1.192	1.402
13	IGD Rawat Jalan	2.173	2.878	3.245
14	Haemodialisa	-	4.953	4.672
15	Umum	-	4.523	2.514

Sumber : Rumah sakit islam pati

Data dari Rumah Sakit Islam Pati menunjukkan Jumlah kunjungan pasien rawat jalan tercatat sebanyak 27.945 pada tahun 2021, mengalami peningkatan menjadi 51.079 pada tahun 2022, dan kembali naik menjadi 67.483 pada tahun 2023. Data ini mencakup kunjungan ke berbagai poliklinik, seperti Poliklinik Anak, Poliklinik Bedah, Poliklinik Dalam, Poliklinik Fisioterapi, Poliklinik Jantung, Poliklinik Kebidanan, Poliklinik Mata, Poliklinik Psikiatri, serta beragam poliklinik lainnya. Informasi ini memberikan gambaran yang penting tentang tingkat pelayanan kesehatan masyarakat dan kebutuhan perawatan yang beragam di berbagai bidang medis selama periode waktu yang diteliti.

B. Pra-pemrosesan data

Pra-pemrosesan data pada penelitian ini melibatkan langkah-langkah yang sistematis untuk memastikan kebersihan dan kevalidan data sebelum analisis lebih lanjut. Ini mencakup penyaringan data untuk mengidentifikasi dan menangani nilai yang hilang atau tidak valid, seperti mengganti nilai yang hilang dengan estimasi yang masuk akal atau menghapus baris data yang tidak lengkap. Selain itu, variabel kategorikal akan diubah menjadi bentuk numerik, seperti penggunaan variabel dummy untuk merepresentasikan kategori. Proses ini juga melibatkan menskalakan variabel jika perlu, terutama jika variabel memiliki skala yang berbeda-beda, untuk memastikan kontribusi yang seimbang dalam model analisis.

Selain langkah-langkah dasar tersebut, pra-pemrosesan data juga mencakup penghilangan outlier yang dapat mengganggu analisis. Outlier adalah nilai data yang sangat berbeda dari sebagian besar data lainnya, dan dapat mempengaruhi hasil analisis secara signifikan. Metode seperti Z-score atau IQR (Interquartile Range) dapat digunakan untuk mengidentifikasi dan menangani outlier ini. Kemudian, dilakukan juga proses normalisasi atau standarisasi data, tergantung pada kebutuhan model analisis yang akan digunakan. Normalisasi merujuk pada skala data ke rentang tertentu, biasanya 0 hingga 1, sedangkan standarisasi mengubah data sehingga memiliki rata-rata 0 dan standar deviasi 1.

Setelah itu, akan dilakukan pemilihan fitur dengan mengevaluasi korelasi antar variabel dan memilih fitur yang paling signifikan untuk

dimasukkan ke dalam model regresi linier. Teknik seperti Analisis Komponen Utama (PCA) juga dapat diterapkan untuk mereduksi dimensi data tanpa kehilangan informasi penting. Terakhir, data akan diuji untuk memastikan bahwa memenuhi asumsi-asumsi dasar analisis regresi linier, seperti independensi residual, homoskedastisitas, dan distribusi normal. Langkah ini memastikan bahwa model yang akan dibangun dapat memberikan hasil yang akurat dan dapat diandalkan.

Dengan melakukan pra-pemrosesan data ini, dataset akan siap untuk dianalisis lebih lanjut menggunakan model regresi linier. Data yang telah dipersiapkan melalui pra-pemrosesan telah melalui tahap pembersihan yang cermat untuk menangani nilai yang hilang atau tidak valid. Setiap entri dalam tabel mencerminkan jumlah kunjungan pasien rawat jalan yang valid dan dapat diandalkan pada tahun 2022 dan 2023. Langkah-langkah seperti imputasi data hilang, transformasi variabel kategorikal, dan normalisasi variabel memastikan bahwa data siap digunakan dalam analisis regresi linier sederhana. Seluruh proses ini bertujuan untuk meminimalkan bias dan variabilitas yang tidak diinginkan, serta untuk meningkatkan kualitas dan reliabilitas hasil analisis yang akan dihasilkan. Dengan demikian, data siap untuk digunakan dalam penerapan model regresi linier sederhana.

Tabel 2. Hasil Pra-pemrosesan data

No	Tahun	Jumlah
1	2021	27.945
2	2022	69.723
3	2023	50.649

Dalam penelitian ini, variabel yang akan digunakan sebagai fokus adalah jumlah kunjungan pasien rawat jalan pada tahun 2021, 2022 dan 2023. Data ini telah melalui proses pra-pemrosesan untuk memastikan kebersihan dan validitasnya. Dengan menggunakan data ini, dapat menganalisis tren dan pola kunjungan pasien dari tahun ke tahun, serta membangun model regresi linier untuk memprediksi jumlah kunjungan di masa mendatang.

C. Model regresi linier sederhana

Dengan menggunakan metode kuadrat terkecil (*least squares*), metode dapat mengestimasi parameter *slope* (b) dan *intercept* (a) dari persamaan regresi linier $y = a + bx$, dimana y adalah jumlah kunjungan dan x adalah tahun.

Menghitung b dan a:

$$1. \text{ Menghitung jumlah total } x, y, x^2, \text{ dan } xy$$

$$\Sigma x = 2021 + 2022 + 2023 = 6066$$

$$\Sigma y = 27945 + 69723 + 50649 = 148317$$

$$\Sigma x^2 = 2021^2 + 2022^2 + 2023^2 = 12268118$$

$$\Sigma xy = 202127945 + 202269723 + 2023*50649 = 300921243$$

2. Menghitung slope (b) dengan rumus:

$$b = (n\Sigma xy - \Sigma x \Sigma y) / (n\Sigma x^2 - (\Sigma x)^2)$$

$$b = (3300921243 - 6066148317) / (312268118 - 6066^2)$$

$$b = 11433$$

3. Menghitung intercept (a) dengan rumus:

$$a = (\Sigma y - b\Sigma x) / n$$

$$a = (148317 - 114336066) / 3$$

$$a = -58298$$

Jadi, model regresi linier sederhana yang diperoleh adalah:

$$y = -58298 + 11433x$$

Dimana y adalah jumlah kunjungan pasien rawat jalan dan x adalah tahun.

Model ini dapat digunakan untuk memprediksi jumlah kunjungan pada tahun-tahun lain, dengan mensubstitusikan nilai tahun (x) ke dalam persamaan regresi.

D. Optimasi model

Dalam upaya untuk memastikan bahwa model regresi linier sederhana yang dibangun memiliki performa yang optimal, langkah selanjutnya yaitu mengevaluasi menggunakan dua metrik penting, yaitu *Root Mean Squared Error (RMSE)* dan *R-Square (R²)*.

Tabel 3. Optimasi model

Tahun	Jumlah aktual	Jumlah Prediksi
2021	27.945	-34585
2022	69.723	-33152
2023	50.649	-31719
	Selisih (y_aktual - y_prediksi)	(y_aktual - y_prediksi) ²
2021	62530	3,909,975,900
2022	102875	10,583,156,625
2023	82368	6,784,554,700
	$\Sigma(y_aktual - y_prediksi)^2$	16,258,687,225

$$RMSE = \sqrt{(\Sigma(y_aktual - y_prediksi)^2 / n)}$$

$$= \sqrt{(16,258,687,225 / 3)}$$

$$= 66,809$$

$$y_{rata-rata} = (27945 + 69723 + 50649) / 3$$

$$= 49439$$

$$\Sigma(y_aktual - y_{rata-rata})^2 = (27945 - 49439)^2 + (69723 - 49439)^2 + (50649 - 49439)^2$$

$$= 460,872,164 + 400,441,764 + 1,624,164$$

$$= 862,938,092$$

$$R^2 = 1 - (\Sigma(y_aktual - y_prediksi)^2 / \Sigma(y_aktual - y_{rata-rata})^2)$$

$$= 1 - (16,258,687,225 / 862,938,092)$$

$$= 1 - 0.1884$$

$$= 0.8116 \text{ atau } 81.16\%$$

Hasil evaluasi metrik menunjukkan bahwa *Root Mean Squared Error (RMSE)* sebesar 66,809, menunjukkan adanya perbedaan yang cukup besar antara nilai prediksi model dan data aktual. Selain itu, *R-squared (R²)* mencapai 0.8116 atau setara dengan 81.16%, menggambarkan bahwa model regresi linear sederhana ini mampu menjelaskan sekitar 81.16% variasi dalam data jumlah kunjungan pasien rawat jalan, menunjukkan kualitas yang memadai dalam kemampuan model dalam memprediksi perilaku variabel dependen berdasarkan variabel independen yang ada. Penilaian kualitas model ini memberikan wawasan yang berguna untuk memahami seberapa baik model ini bekerja dalam menginterpretasikan data dan potensi perbaikan yang dapat dilakukan untuk meningkatkan performa prediktifnya.

Tabel 4. Hasil optimasi model

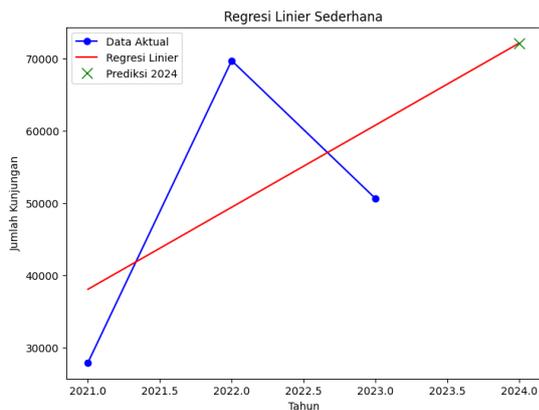
Metrix	Nilai
RMSE	66,809
R-squared (R ²)	0.8116 atau 81.16%

Dengan demikian, tabel di atas memberikan ringkasan nilai RMSE dan R-squared yang telah dihitung untuk mengoptimasi model regresi linear sederhana yang diperoleh sebelumnya.

E. Implementasi model

Untuk melakukan prediksi jumlah kunjungan di tahun 2024, kami menggunakan model regresi linier sederhana yang diimplementasikan dengan Python dan library scikit-learn. Langkah-langkah sebagai berikut.

1. Persiapan Data jumlah kunjungan dari tahun 2021 hingga 2023 dan menyusunnya dalam bentuk array.
2. Pembuatan dan pelatihan model regresi linier menggunakan *LinearRegression* dari scikit-learn dan melatihnya dengan data yang ada.
3. Prediksi model yang telah dilatih digunakan untuk memprediksi jumlah kunjungan di tahun 2024.
4. Memvisualisasikan data aktual, garis regresi linier, dan hasil prediksi dalam bentuk grafik untuk mempermudah pemahaman.



Gambar 2. Grafik hasil implementasi model

Gambar 2. Hasil grafik prediksi jumlah kunjungan pasien rawat jalan di Rumah Sakit Islam Pati menggunakan bahasa pemrograman python. Berikut adalah penjelasan terperinci tentang hasil algoritma yang ditampilkan dalam grafik:

1. Sumbu X (Tahun): Sumbu horizontal menunjukkan tahun pengamatan, mulai dari tahun 2021 hingga 2024
2. Sumbu Y (Jumlah Kunjungan): Sumbu vertikal menunjukkan jumlah kunjungan pasien rawat jalan.
3. Data Aktual (Garis Biru): Titik-titik biru yang terhubung dengan garis menunjukkan data aktual jumlah kunjungan pasien pada tahun 2021, 2022, dan 2023. Dari grafik terlihat bahwa:
 - a. Pada tahun 2021, jumlah kunjungan sekitar 30,000.
 - b. Pada tahun 2022, jumlah kunjungan meningkat tajam menjadi sekitar 70,000.
 - c. Pada tahun 2023, jumlah kunjungan menurun menjadi sekitar 50,000.
4. Regresi Linier (Garis Merah) mewakili model regresi linier yang telah dihitung berdasarkan data aktual dari tahun 2021 hingga 2023. Garis ini menunjukkan tren umum dari data yang diamati, yaitu kecenderungan jumlah kunjungan pasien seiring berjalannya waktu. Model regresi linier ini digunakan untuk memperkirakan nilai di luar data yang telah diamati.
5. Prediksi 2024 (Tanda X Hijau) menunjukkan prediksi jumlah kunjungan pasien untuk tahun 2024 berdasarkan model regresi linier yang telah dihitung. Prediksi menunjukkan bahwa jumlah kunjungan pada tahun 2024 diperkirakan akan meningkat menjadi sekitar 72,000.

Data aktual menunjukkan fluktuasi yang signifikan dari tahun ke tahun, dengan peningkatan tajam pada 2022 dan penurunan pada 2023, namun garis regresi linier memberikan gambaran rata-rata atau tren umum dari data tersebut; garis regresi linier tampaknya memberikan estimasi yang cukup baik

meskipun ada fluktuasi yang besar pada data aktual dan digunakan untuk memberikan gambaran tren jangka panjang, meskipun mungkin tidak menangkap fluktuasi tahunan yang ekstrem; berdasarkan tren umum yang ditangkap oleh model regresi linier, jumlah kunjungan pasien diperkirakan akan meningkat pada tahun 2024, yang dapat menjadi informasi penting bagi manajemen rumah sakit dalam merencanakan sumber daya dan operasional mereka untuk tahun-tahun mendatang.

Grafik yang dihasilkan menunjukkan tren kenaikan yang lebih stabil dari waktu ke waktu, meskipun data aktual menunjukkan fluktuasi yang cukup besar dari tahun ke tahun. Kenaikan yang signifikan dari 2021 ke 2022 diikuti oleh penurunan dari 2022 ke 2023 menunjukkan adanya variabel lain yang mungkin mempengaruhi jumlah kunjungan yang tidak tercakup dalam model ini.

Tabel 5. Hasil Prediksi Tahun 2024

Metode	Nilai
Persamaan regresi	$y = -22904305.00 + 11352.00x$
Prediksi Tahun 2024	72,143
RMSE	14,342.954
R-squared (R ²)	0.295 atau 29.5%

Model algoritma regresi linier sederhana yang diterapkan ke pemrograman Python menghasilkan persamaan $y = -22904305.00 + 11352.00x$. untuk memprediksi jumlah kunjungan berdasarkan data dari tahun 2021 hingga 2023. Berdasarkan model ini, prediksi jumlah kunjungan pada tahun 2024 adalah 72,143. Nilai Root Mean Square Error (RMSE) sebesar 14,342.954 menunjukkan tingkat kesalahan prediksi yang moderat, yang artinya prediksi model memiliki penyimpangan yang cukup signifikan dari nilai aktual. Nilai R-squared sebesar 0.295 mengindikasikan bahwa model ini hanya mampu menjelaskan 29.5% variabilitas data, menunjukkan bahwa model regresi linier sederhana ini belum cukup kuat untuk memberikan prediksi yang akurat.

Dalam konteks analisis ini, RMSE yang cukup tinggi mengindikasikan bahwa terdapat perbedaan yang signifikan antara nilai yang diprediksi oleh model dan nilai aktualnya. Hal ini menunjukkan bahwa model mungkin tidak sepenuhnya menangkap pola dalam data historis dari tahun 2021 hingga 2023. Demikian pula, nilai R-squared yang rendah menunjukkan bahwa sebagian besar variabilitas dalam jumlah kunjungan tidak dijelaskan oleh model. Ini berarti bahwa model regresi linier sederhana mungkin terlalu sederhana untuk data yang kompleks dan memiliki banyak variabel yang mempengaruhi jumlah kunjungan.

Dengan demikian, meskipun prediksi menunjukkan peningkatan kunjungan di tahun 2024,

rendahnya nilai R-squared dan tingginya RMSE menyarankan perlunya mempertimbangkan faktor lain atau menggunakan model yang lebih kompleks untuk meningkatkan akurasi prediksi. Model regresi linier sederhana, meskipun mudah diterapkan dan diinterpretasikan, mungkin tidak cukup untuk menangkap seluruh variabilitas dalam data kunjungan. Oleh karena itu, mempertimbangkan model alternatif yang lebih kompleks dapat menjadi langkah yang bijaksana.

Model lain yang bisa dipertimbangkan termasuk regresi polinomial, yang dapat menangkap hubungan yang lebih kompleks antara variabel independen dan dependen dengan menambahkan pangkat variabel independen dalam persamaan. Selain itu, model berbasis machine learning yang lebih kompleks, seperti Random Forest, Gradient Boosting, atau Neural Networks, mungkin lebih cocok untuk menangkap variabilitas data yang lebih besar dan memberikan prediksi yang lebih akurat. Model-model ini memiliki kemampuan untuk menangani data yang tidak linear dan interaksi kompleks antar variabel, yang dapat meningkatkan akurasi prediksi secara signifikan.

Selain itu, mengintegrasikan faktor-faktor lain yang mungkin mempengaruhi jumlah kunjungan, seperti tren musiman, perubahan kebijakan kesehatan, atau faktor ekonomi, juga bisa meningkatkan kinerja model. Pendekatan yang lebih holistik dengan menggunakan lebih banyak data historis dan variabel tambahan dapat memberikan pandangan yang lebih komprehensif dan prediksi yang lebih andal. Dengan demikian, mengkombinasikan pendekatan statistik dan teknik machine learning yang lebih canggih dapat menjadi strategi yang efektif untuk meningkatkan akurasi dan reliabilitas prediksi jumlah kunjungan di masa mendatang.

V. KESIMPULAN DAN SARAN

Kesimpulan

Dalam analisis ini, meskipun model regresi linier sederhana memberikan perkiraan jumlah kunjungan pasien rawat jalan, terdapat beberapa pertimbangan yang perlu diperhatikan.

1. Variasi Prediksi dan Data Aktual
Nilai RMSE yang moderat menunjukkan bahwa terdapat variasi yang cukup besar antara prediksi dan data aktual. Hal ini menandakan bahwa model ini mungkin tidak sepenuhnya menangkap pola yang ada dalam data. Keterbatasan ini bisa disebabkan oleh kompleksitas yang lebih besar dari faktor-faktor yang mempengaruhi jumlah kunjungan, seperti perubahan kebijakan kesehatan, faktor ekonomi, atau tren demografis.
2. Nilai R-Squared yang Rendah
Nilai R-squared yang rendah (0.295) menandakan bahwa sebagian besar variabilitas

dalam data tidak dapat dijelaskan oleh model ini. Hal ini menegaskan bahwa model regresi linier sederhana mungkin tidak memadai untuk menggambarkan hubungan antara variabel independen (tahun) dan variabel dependen (jumlah kunjungan). Perlu dipertimbangkan untuk menggabungkan lebih banyak variabel atau menggunakan model yang lebih kompleks untuk meningkatkan akurasi prediksi.

3. Analisis Grafik Tren

Dalam menganalisis grafik tren, penting untuk memperhatikan fluktuasi yang signifikan dari tahun ke tahun. Meskipun tren umumnya menunjukkan kenaikan yang stabil, adanya fluktuasi ini menandakan adanya faktor-faktor lain yang memengaruhi jumlah kunjungan pasien rawat jalan. Variabel-variabel ini mungkin termasuk faktor musiman, peristiwa khusus, atau perubahan dalam preferensi pasien.

4. Interpretasi Hasil dan Keterbatasan Model

Dalam menginterpretasikan hasil ini, penting untuk mempertimbangkan keterbatasan-keterbatasan dari model regresi linier sederhana dan untuk mempertimbangkan informasi tambahan yang mungkin memengaruhi jumlah kunjungan pasien rawat jalan. Dengan memperhitungkan faktor-faktor ini, analisis dapat menjadi lebih holistik dan prediksi dapat menjadi lebih akurat.

Saran

Sebagai tambahan dari analisis yang telah disampaikan, berikut adalah beberapa saran yang dapat dipertimbangkan untuk meningkatkan keakuratan prediksi jumlah kunjungan pasien rawat jalan.

1. Perbaiki Model
 - a. Penggunaan Model Lebih Kompleks.
Pertimbangkan untuk menggunakan model regresi yang lebih kompleks seperti regresi polinomial atau model berbasis machine learning yang lebih canggih (misalnya, Random Forest, Gradient Boosting) untuk meningkatkan akurasi prediksi.
 - b. Penggunaan Fitur Tambahan Menambah variabel independen lain yang relevan (misalnya, faktor musiman, kebijakan kesehatan, atau demografis pasien) untuk menangkap variabilitas data yang lebih besar.
2. Pra-pemrosesan Data Lebih Lanjut
 - a. Penanganan Data Hilang
Lakukan penanganan lebih komprehensif terhadap data yang hilang atau tidak valid untuk meningkatkan kualitas data yang digunakan dalam model.
 - b. Normalisasi Data

- Normalisasi variabel untuk memastikan bahwa semua variabel memiliki skala yang sama dan tidak mendistorsi hasil model.
3. Evaluasi Model
 - a. *Cross-Validation*
Gunakan teknik *cross-validation* untuk memastikan bahwa model tidak overfitting dan memberikan generalisasi yang baik pada data yang tidak terlihat.
 - b. Evaluasi Metrik Tambahan
Selain RMSE dan R-squared, pertimbangkan untuk mengevaluasi model menggunakan metrik tambahan seperti Mean Absolute Error (MAE) untuk mendapatkan gambaran yang lebih komprehensif tentang performa model.
 4. Analisis Lebih Lanjut
 - a. Investigasi Faktor Eksternal
Lakukan analisis untuk mengidentifikasi faktor eksternal yang mungkin mempengaruhi jumlah kunjungan dan belum dimasukkan dalam model saat ini.
 - b. Monitoring dan Pembaruan Model
Lakukan monitoring secara berkala terhadap performa model dan perbarui model dengan data terbaru untuk memastikan akurasi prediksi tetap tinggi.

Dengan menerapkan saran-saran ini, diharapkan model prediksi jumlah kunjungan pasien rawat jalan dapat lebih akurat dan memberikan wawasan yang lebih berharga untuk perencanaan dan pengambilan keputusan di masa mendatang.
- ### DAFTAR PUSTAKA
- Amon, L., & Harliansyah, H. (2022). Analisis Kompetensi Manajerial Kepala Sekolah dalam Perencanaan Peningkatan Mutu Pendidikan Menengah Kejuruan. *Jurnal Ilmiah Manajemen Dan Kewirausahaan*, 1(1), 147–162. <https://doi.org/10.55606/jimak.v1i1.258>
- Ananda, R., Damayanti, R., & Maharja, R. (2023). Tingkat Kepuasan Masyarakat terhadap Kinerja Pelayanan Kesehatan. *Jurnal Keperawatan Profesional (KEPO)*, 4(1), 9–17. <https://doi.org/10.36590/kepo.v4i1.570>
- Apriliani, I. M., Purba, N. P., Dewanti, L. P., Herawati, H., & Faizal, I. (2021). Open access Open access. *Citizen-Based Marine Debris Collection Training: Study Case in Pangandaran*, 2(1), 56–61.
- Azahra, A. A. (2022). Analisis Prediksi Jumlah Penerimaan Mahasiswa Baru Menggunakan Metode Regresi Linier Sederhana. *Bulletin of Applied Industrial Engineering Theory*, 3(1), 75–78.
- Fitri, E. (2023). Analisis Perbandingan Metode Regresi Linier, Random Forest Regression dan Gradient Boosted Trees Regression Method untuk Prediksi Harga Rumah. *Journal of Applied Computer Science and Technology*, 4(1), 58–64. <https://doi.org/10.52158/jacost.v4i1.491>
- Harizahayu, H., Hermanto, K., & Ratna Yuniarti, R. (2023). Analisis Viral Marketing Pada Online Customer Terhadap Minat Pembelian Melalui Tiktok Shop Dengan Regresi Linier Sederhana. *Jurnal Sains Matematika Dan Statistika*, 9(2), 31. <https://doi.org/10.24014/jsms.v9i2.20652>
- Hasibuan, R. (2020). Bahan Ajar Administrasi dan Kebijakan Kesehatan. *Jurnal Administrasi Kesehatan Indonesia*, 156–159. http://repository.uinsu.ac.id/8799/1/BAHAN_AJAR_ADMINISTRASI_KEBIJAKAN_KESEHATAN.pdf
- Indahningrum, R. putri, & lia dwi jayanti. (2020). *No 主観的健康感を中心とした在宅高齢者における健康関連指標に関する共分散構造分析Title. 2507(1), 1–9.* <http://journal.um-surabaya.ac.id/index.php/JKM/article/view/2203>
- Komalawati, V., & Triswandi, E. F. (2022). Tanggung Jawab Dokter Atas Insiden Keselamatan Pasien Dalam Pelayanan Kesehatan Di Rumah Sakit Sebagai Institusi Kesehatan. *Jurnal Bina Mulia Hukum*, 6(2), 174–186. <https://doi.org/10.23920/jbmh.v6i2.687>
- Kusuma, M. D. H., & Hidayat, S. (2024). Penerapan Model Regresi Linier dalam Prediksi Harga Mobil Bekas di India dan Visualisasi dengan Menggunakan Power BI. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, 5(2), 1097–1110. <https://doi.org/10.35870/jimik.v5i2.629>
- Lase, N. R., & Riandari, F. (2020). Perancangan Aplikasi Prediksi Jumlah Pendaftar Siswa Baru Dengan Metode Regresi Linier (Studi Kasus: SMA RK Deli Murni Bandar Baru). *Jurnal Nasional Komputasi Dan Teknologi Informasi (JNKTI)*, 3(3), 330–334. <https://doi.org/10.32672/jnkti.v3i3.2520>
- Made, N., Wulaning, O., Elindra, H., & Saputra, A. H. (2024). *Prediksi Karbon Monoksida Menggunakan Model Machine Learning Berdasarkan Perbandingan Model Time Series Studi Kasus DKI Jakarta Carbon Monoxide Prediction Using Machine Learning Model Based on Time Series Model Comparison DKI Jakarta Case Study.* 7(3), 1116–1128. <https://doi.org/10.56338/jks.v7i3.4819>
- Madjid, A. (2020). Mengobati Sakit Ditengah Keterbatasan Fasilitas Pelayanan Kesehatan Modern di Pulau Mare Kota Tidore. *ETNOHISTORI: Jurnal Ilmiah Kebudayaan Dan Kesenjangan*, VII(2), 199–210.
- Rustam, M. Z. A., Amalia, N., & Riestiyowati, M. A. (2022). Analisis Prediksi Kunjungan Pasien

- Dengan Metode Autoregresiive Integrated Moving Average di RSIA Putri Surabaya Prediction Analysis of Patient Visits Using the Autoregressive Integrated Moving Average Method at RSIA Putri Surabaya. *Journal Manajemen Kesehatan Yayasan RS Dr.Soetomo*, 8(1), 110–123.
- Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Komputer Program, S., Hidayatullah Jakarta, S., & Audi Permaisuri, R. (2022). *Visualisasi Dashboard Tableau Dan Peramalan Jumlah Kasus Demam Berdarah Dengue Di Dki Jakarta Menggunakan Metode Arima*.
- Salsabila, T. H., Indrawati, T. M., & Fitrie, R. A. (2024). *Meningkatkan Efisiensi Pengambilan Keputusan Publik melalui Kecerdasan Buatan*. 2, 1–21.
- Syahputra, P. (2022). Prediksi Lama Rawat Pasien Covid-19 Berbasis Machine Learning. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(4), 3374–3382. <https://doi.org/10.35957/jatisi.v9i4.2883>
- Wijaya, Y. A., Yudhawati, N. L. P. S., Andriana, K. R. F., & Ilmy, S. K. (2022). Klasifikasi Teori Keperawatan yang Dikembangkan oleh Ahli Keperawatan: Sebuah Tinjauan Literatur. *Nursing Sains*, 23(2), 1–49.