

Implementasi Metode Regresi Linier Berganda untuk Prediksi Jumlah Kunjungan Pasien di Puskesmas Simpang Kawat Menggunakan RapidMiner

Rosma Windah¹, Reny Wahyuning Astuti², Sri Mulyati³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Nurdin Hamzah,
Jambi, Indonesia

e-mail: windahrosma@gmail.com, r3ny4stuti@gmail.com, mulyati.sri52@gmail.com

Abstrak

Puskesmas merupakan fasilitas kesehatan tingkat pertama yang memiliki peran penting dalam memberikan pelayanan kepada masyarakat. Namun, jumlah kunjungan pasien yang fluktuatif sering kali menjadi tantangan dalam perencanaan pelayanan, seperti pengaturan tenaga medis dan ketersediaan obat-obatan. Oleh karena itu, diperlukan suatu analisis prediksi guna membantu pihak Puskesmas dalam mengambil keputusan yang lebih tepat. Penelitian ini dilakukan untuk memprediksi jumlah kunjungan pasien pada Puskesmas Simpang Kawat menggunakan metode Regresi Linier Berganda. Alasan dipilihnya metode Regresi Linier Berganda adalah karena mampu menganalisis hubungan lebih dari satu variabel independen terhadap variabel dependen, sehingga dapat memberikan gambaran yang lebih akurat terhadap jumlah kunjungan berdasarkan beberapa faktor. Variabel independen yang digunakan dalam prediksi kunjungan pasien berdasarkan layanan pada puskesmas adalah Klaster 2, Klaster 3, Klaster 4, Gigi dan Mulut, serta UGD. Variabel independen dalam memprediksi jumlah kunjungan pasien hipertensi adalah jumlah pasien dewasa dan lansia. Data yang digunakan merupakan data kunjungan pasien dari tahun 2019 hingga 2024 untuk memprediksi tahun 2025, yang kemudian dianalisis menggunakan perhitungan manual dan dibantu dengan perangkat lunak RapidMiner. Hasil penelitian menunjukkan bahwa metode Regresi Linier Berganda mampu memprediksi jumlah kunjungan pasien dengan tingkat error yang dapat diterima, sehingga dapat dijadikan acuan awal dalam perencanaan pelayanan dan pengelolaan sumber daya di puskesmas.

Kata Kunci: Jumlah Kunjungan Pasien, Prediksi, Puskesmas, Regresi Linier Berganda.

Abstract

Community health centers are primary health facilities that play an important role in providing services to the community. However, fluctuating patient visits often pose challenges in service planning, such as the allocation of medical personnel and the availability of medicines. Therefore, a predictive analysis is needed to help community health centers make more informed decisions. This study was conducted to predict the number of patient visits at the Simpang Kawat Puskesmas using the Multiple Linear Regression method. The Multiple Linear Regression method was chosen because it can analyze the relationship between more than one independent variable and the dependent variable, thereby providing a more accurate picture of the number of visits based on several factors. The independent variables used in predicting patient visits based on services at the health center are Cluster 2, Cluster 3, Cluster 4, Dental and Oral Health, and Emergency Room. The independent variables in predicting the number of hypertensive patient visits are the number of adult and elderly patients. The data used are patient visit data from 2019 to 2024 to predict 2025, which were then analyzed using manual

calculations and assisted by RapidMiner software. The results of the study indicate that the Multiple Linear Regression method is capable of predicting the number of patient visits with an acceptable error rate, making it a useful initial reference for service planning and resource management at community health centers.

Keywords: *Community Health Center, Multiple Linear Regression, Number of Patient Visits, Prediction.*

1. Pendahuluan

Puskesmas merupakan fasilitas atau tempat yang menyediakan layanan terkait kesehatan untuk masyarakat yang dibiayai oleh pemerintah [1]. Seiring dengan pertumbuhan populasi manusia dan meningkatnya kesadaran masyarakat tentang kesehatan juga semakin meningkat. Sehingga pasien yang datang ke Puskesmas akan terus meningkat, tidak hanya datang dalam kondisi sedang sakit saja tetapi banyak juga masyarakat yang datang dalam keadaan sehat untuk mengecek kesehatannya. Dalam dunia kesehatan lonjakan jumlah pasien di fasilitas pelayanan kesehatan, seperti Puskesmas menjadi tantangan yang harus dihadapi oleh pengelola layanan kesehatan.

Puskesmas Simpang Kawat merupakan salah satu layanan kesehatan yang berlokasi di Jl. Buton No.16, Payo Lebar, Kec. Jelutung, Kota Jambi. Setiap hari, fasilitas kesehatan masyarakat ini menerima kunjungan pasien untuk mengatasi permasalahan kesehatan mereka. Lonjakan jumlah pasien yang tidak menentu di Puskesmas Simpang Kawat mengakibatkan kurangnya pelayanan secara optimal apabila tidak di persiapkan dengan baik. Seperti terjadi antrian panjang pasien pada administrasi maupun di layanan kesehatan tertentu yang cukup ramai. Jika jumlah pasien meningkat tidak diiringi dengan kesiapan layanan untuk menghadapi peningkatan pasien, stok obat yang tersedia bisa tidak mencukupi, sehingga menghambat kenyamanan dan efektivitas pelayanan.

Prediksi jumlah kunjungan Puskesmas adalah memperkirakan berapa banyak orang yang akan mengunjungi Puskesmas dalam waktu tertentu. Dengan memprediksi berdasarkan data historis seperti jumlah total pasien berkunjung ke Puskesmas, jumlah pasien Klaster 2 (Ibu dan Anak), jumlah pasien Klaster 3 (Dewasa dan Lansia), jumlah pasien klaster 4 (penyakit menular), jumlah pasien layanan gigi dan mulut, dan jumlah pasien layanan UGD, dan prediksi kunjungan pasien Hipertensi dengan variabel independent jumlah pasien dewasa dan lansia. Pasien hipertensi dipilih karena merupakan penyakit yang sering di tangani oleh pihak puskesmas. Prediksi jumlah pasien ini berdasarkan klaster yang ada di puskesmas, dan prediksi penyakit hipertensi. Klaster di Puskesmas adalah pengelompokan pelayanan kesehatan berdasarkan kelompok usia dan jenis penyakit.

Penelitian terdahulu yang menggunakan metode regresi linier berganda seperti Penelitian Simanungkalit et al., 2020, berjudul “*Implentasi Data Mining untuk Mengestimasi Jumlah Kunjungan Pasien RS Vita Insani Pematang Siantar dengan Menggunakan Metode Regresi Linier Berganda*” pada penelitian tersebut regresi linier berganda mampu memprediksi jumlah kunjungan pasien BPJS dengan akurat berdasarkan data-data yang mempengaruhi tinggi rendahnya jumlah kunjungan pasien BPJS di rumah sakit Vita Insani yaitu jumlah poli UGD dan jumlah Termasuk ICU.

Penelitian Kurniawan, 2023, berjudul “*Penerapan Metode Regresi Linier Sederhana untuk Prediksi Penyebaran Vaksin Covid-19 di Kabupaten Cilacap*” Penelitian ini menganalisis hubungan antara jumlah kasus positif (variabel x) dan umlah vaksinasi dosis 1 (variabel y). Metode regresi linier berhasil memprediksi penyebaran vaksin di kabupaten Cilacap. Metode prediksi regresi linier dapat menghasilkan prediksi dengan beberapa kriteria, dimana terdapat model prediksi regresi linier yang memiliki nilai *error* kurang dari 18% artinya mempunyai akurasi sebesar 82% atau termasuk dalam kategori baik

Penelitian Baihaqi et al., 2019 berjudul, “*Regresi Linier Sederhana untuk Memprediksi Kunjungan Pasien di Rumah Sakit Berdasarkan Jenis Layanan dan Umur Pasien*”, Metode

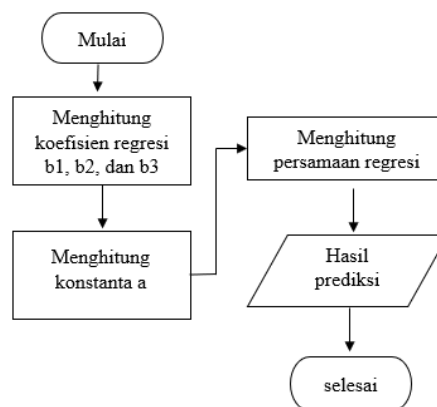
regresi linier sederhana digunakan untuk menganalisis hubungan antara variabel independen (jenis layanan dan umur pasien) dan variabel dependen (jumlah kunjungan pasien). Berdasarkan hasil prediksi yang diperoleh kunjungan pasien baik dari pasien rawat inap, rawat jalan, dan gawat darurat pada tahun 2019 cenderung stabil setiap bulannya, tidak ada penurunan dan kenaikan secara signifikan

Data Mining adalah penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar [2]. Salah satu metode analisis dalam data mining adalah metode Regrsi Linier, terdapat dua jenis regresi dalam regresi linier yaitu regresi linier sederhana dan regresi linier berganda. Pada penelitian ini digunakan metode regresi linier berganda. Regresi Linier Berganda merupakan model persamaan yang menjelaskan hubungan satu variabel tak bebas/response (Y) dengan dua atau lebih variabel bebas/predictor (X1, X2, ..., Xn). memprediksi nilai variabel tak bebas/response (Y) apabila nilai-nilai variabel bebasnya/predictor (X1, X2, ..., Xn) diketahui [3]. Metode regresi linier berganda dirancang untuk memprediksi nilai numerik (continuous), seperti jumlah pasien. Sedangkan banyak metode data mining, misalnya klasifikasi, lebih cocok untuk memprediksi kategori atau kelompok, bukan angka pasti.

2. Metode Penelitian

Regresi linier berganda merupakan model persamaan yang menggunakan satu variabel tak bebas/ *response* (Y) dengan dua atau lebih variabel bebas/*predictor* (X1, X2...Xn). memprediksi nilai variable tak bebas/*response* (Y) apabila nilai-nilai variabel bebasnya/*predictor* (X1, X2,..., Xn) diketahui [3]. Regresi linier berganda digunakan untuk memodelkan hubungan antara satu variabel dependen (Y) dengan dua atau lebih variabel independen (X1, X2...Xn). Adapun rumus perhitungan metode regresi linier berganda sebagai berikut:

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n \quad (1)$$



Gambar 1. Flowchart Algoritma Regresi Linier Berganda

Diawali dengan menghitung koefisien regresi b1, b2, dan b3 Kemudian dilanjutkan menghitung konstanta a dengan mengambil hasil dari perhitungan konstanta. Dan untuk nilai variable x1, x2, dan x3 menggunakan data sebelumnya. Lalu muncullah hasil prediksinya [4]. Pada penelitian ini digunakan data jumlah pasien dari tahun 2019 hingga 2024 untuk memprediksi jumlah pasien pada tahun 2025. Data tersebut diperoleh langsung dari sistem informasi Puskesmas Simpang Kawat, sehingga memiliki tingkat akurasi dan konsistensi yang lebih tinggi. Jumlah data yang digunakan sebanyak 6 data berdasarkan data tahunan selama enam tahun.

Adapun langkah-langkah Regresi Linier Berganda adalah sebagai berikut:

1. Langkah pertama dengan memulai atau start.

2. Input data jumlah kunjungan pasien klaster 2 (layanan ibu dan anak), klaster 3 (layanan usia dewasa dan lansia), klaster 4 (layanan penyakit menular), Gigi dan Mulut, UGD, pasien dewasa, lansia dan hipertensi dan total pasien 2019-2024
3. Menghitung koefisien regresi b_1 , b_2 , b_3 , b_4 , b_5
4. Menghitung konstanta a
5. Menentukan nilai X pada tahun 2025
6. Menghitung persamaan regresi
7. Hasil prediksi

3. Hasil dan Pembahasan

Dengan menerapkan metode Regresi Linier Berganda pada pemrograman menggunakan RapidMiner untuk memprediksi pasien hipertensi dan prediksi jumlah pasien keseluruhan. Data yang digunakan dari tahun 2019 hingga 2024. Pada perhitungan awal data dibagi menjadi 3 bagian yaitu data training 2019-2022 untuk memprediksi data testing 2023, data training 2019-2023 digunakan untuk prediksi data testing 2024, dan data 2019-2024 digunakan sebagai data prediksi 2025. Perhitungan Menggunakan bantuan *Microsoft Excel* serta *tools* RapidMiner sebagai perbandingan hasil.

1. Langkah pertama dalam mencari nilai koefisien regresi (b) dari data training adalah menjumlahkan seluruh data sesuai variabelnya
2. Langkah kedua adalah mengalikan masing-masing nilai variabel independen (X) dengan variabel dependen (Y), yaitu X_1Y , X_2Y , X_3Y , dan seterusnya, kemudian menjumlahkan hasil perkalian tersebut.
3. Langkah ketiga adalah mengalikan setiap pasangan variabel independen, seperti X_1X_2 , X_1X_3 , hingga X_4X_5 , kemudian menjumlahkan hasilnya.
4. Langkah keempat adalah mengkuadratkan setiap nilai pada variabel independen (X_1^2 , X_2^2 , X_3^2 , X_4^2 , X_5^2), kemudian menjumlahkan hasil kuadrat masing-masing variabel.
5. Langkah kelima adalah melakukan perhitungan menggunakan rumus matriks untuk mendapatkan nilai intercept (a) dan koefisien regresi b_1 hingga b_5 .

Tabel 1. Hasil koefisien dan konstanta 2023

a	45481,16
b1	0,12
b2	0,69
b3	0,07
b4	-2,19
b5	19,39

Tabel 2. Hasil data testing 2023

Tahun	X1	X2	X3	X4	X5	Y aktual	Y prediksi
2023	19466	34469	12757	1908	239	68839	72907

Tabel 3. Hasil koefisien dan konstanta 2024

a	48126,32
b1	-0,49
b2	0,90
b3	0,07

b4	-3,07
b5	28,44

Tabel 4. Hasil data testing 2024

Tahun	X1	X2	X3	X4	X5	Y aktual	Y prediksi
2024	19098	36924	12973	1608	297	70900	76227

3.1. Perhitungan Error

Untuk mengetahui kesalahan antara hasil prediksi dan data aktual, digunakan rumus *Mean Absolute Error* MAE dan *Mean Absolute Percentage Error* MAPE sebagai berikut:

$$MAE = Y \text{ aktual} - Y \text{ prediksi}$$

$$MAE \text{ 2023} = 68839 - 72907 = 4068$$

$$MAE \text{ 2024} = 70900 - 76227 = 5327$$

$$MAPE = Y \text{ aktual} - Y \text{ prediksi} / Y \text{ actual} \times 100\%$$

$$MAPE \text{ 2023} = 4068 / 68839 \times 100 \\ = -5,91 \text{ atau } 5,91\%$$

$$MAPE \text{ 2024} = 5327 / 70900 \times 100 \\ = -7,51 \text{ atau } 7,51\%$$

Tabel 5. Persentase keakurasian MAPE

Sumber [5]

Nilai MAPE	Kategori Akurasi	Keterangan
$MAPE \leq 10\%$	Sangat Baik (Excellent)	Prediksi sangat akurat
$10\% < MAPE \leq 20\%$	Baik (Good)	Prediksi cukup akurat
$20\% < MAPE \leq 50\%$	Cukup (Reasonable)	Prediksi dapat diterima, tapi perlu perbaikan
$MAPE > 50\%$	Kurang (Poor)	Prediksi tidak akurat

Dari hasil yang di dapat disimpulkan bahwa hasil tersebut masuk dalam kategori prediksi sangat baik, karena MAPE kurang dari 10%, pada Tahun 2023 MAPE yang dihasilkan sebesar 5,91% pada Tahun 2024 MAPE sebesar 7,51%.

Tabel 6. Hasil prediksi layanan 2025

Tahun	X1	X2	X3	X4	X5	Y
2019	22703	41791	15270	2052	381	82197
2020	22042	42021	14291	1795	293	80442
2021	20189	39721	13210	1760	241	75121
2022	18313	44727	12697	2207	185	78129
2023	19466	34469	12757	1908	239	68839
2024	19098	36924	12973	1608	297	70900
2025	20302	39942	13533	1888	273	75938

Hasil prediksi menunjukkan bahwa jumlah kunjungan pasien pada tahun 2025 mencapai 75.938 pasien, yang berarti mengalami peningkatan dibandingkan tahun sebelumnya. Peningkatan ini dapat menjadi indikator meningkatnya kebutuhan layanan kesehatan masyarakat, sekaligus menjadi peringatan bagi pihak Puskesmas Simpang Kawat untuk menyiapkan strategi pelayanan, sumber daya manusia, serta logistik kesehatan yang memadai agar mampu mengantisipasi lonjakan pasien di tahun mendatang.

Tabel 7. Hasil prediksi Hipertensi 2025

Tahun	X1 Dewasa	X2 Lansia	Y Hipertensi
2019	10832	11521	4027
2020	11552	11920	3820
2021	10987	10907	3052
2022	11052	10529	2983
2023	10791	10105	2701
2024	10537	10368	3096
2025	11043	10996	3279,83

Selain peningkatan jumlah pasien secara keseluruhan, hasil prediksi juga menunjukkan bahwa jumlah pasien hipertensi pada tahun 2025 diperkirakan meningkat menjadi 3.279 pasien, dibandingkan dengan tahun 2024 yang berjumlah 3.096 pasien. Hal ini menandakan adanya kecenderungan peningkatan kasus hipertensi di wilayah kerja Puskesmas Simpang Kawat.

Row No.	total pasien ...	prediction(L...	Tahun	klaster 2	klster 3	klster 4	gigi	ugd
1	82197	82197.000	2019	22703	41791	15270	2052	381
2	80442	80442.000	2020	22042	42021	14291	1795	293
3	75121	75121.000	2021	20189	39721	13210	1760	241
4	78129	78129.000	2022	18313	44727	12697	2207	185
5	68839	68839.000	2023	19466	34469	12757	1908	239
6	70900	70900.000	2024	19098	36924	12973	1608	297
7	?	75937.018	2025	20302	39942	13533	1888	273

Gambar 2. Tampilan Hasil Prediksi RapidMiner 2025

Row No.	prediction(Y ...	Tahun	X1 Dewasa	X2 Lansia	Y Hipertensi
1	3955.037	2019	10832	11521	4027
2	3802.376	2020	11552	11920	3820
3	3274.585	2021	10987	10907	3052
4	2878.055	2022	11052	10529	2983
5	2673.443	2023	10791	10105	2701
6	3095.504	2024	10537	10368	3096
7	3271.881	2025	10959	10892	?

Gambar 3. Tampilan Hasil Prediksi Hipertensi RapidMiner 2025

Hasil prediksi menggunakan RapidMiner menunjukkan jumlah pasien tahun 2025 sebesar 75.938 pasien dan 3271.881 pasien hipertensi, sama dengan hasil perhitungan manual menggunakan Excel. Hal ini menunjukkan bahwa model regresi linier berganda memiliki tingkat akurasi yang sangat tinggi dan dapat diandalkan untuk melakukan prediksi jumlah kunjungan pasien di tahun berikutnya.

4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai prediksi jumlah kunjungan pasien pada Puskesmas Simpang Kawat menggunakan metode Regresi Linier Berganda, dapat disimpulkan bahwa:

1. Metode Regresi Linier Berganda mampu digunakan untuk memprediksi jumlah kunjungan pasien berdasarkan variabel independen yaitu jumlah kunjungan per klaster layanan (Klaster 2, Klaster 3, Klaster 4, Gigi dan Mulut, dan UGD) dan prediksi jumlah pasien hipertensi berdasarkan variabel independen yaitu pasien dewasa usia 18 – 59 tahun dan lansia usia 60 ke atas.
2. Proses prediksi dilakukan dengan menggunakan perhitungan manual Excel dan juga pengolahan data menggunakan aplikasi RapidMiner. Data tahun 2019 hingga 2022 digunakan sebagai data *training*. Data tahun 2023 dan 2024 digunakan sebagai data *testing* untuk mengukur tingkat akurasi model dengan membandingkan dengan data aktual.
3. Hasil prediksi tahun 2023, nilai MAPE (*Mean Absolute Percentage Error*) dari hasil prediksi menggunakan RapidMiner adalah sebesar 0,22%, sedangkan hasil dari Excel adalah 5,91%. Sementara pada data tahun 2024, nilai MAPE dari RapidMiner sebesar 0,17%, sedangkan Excel menghasilkan MAPE sebesar 7,51%. Ini menunjukkan bahwa akurasi model dari RapidMiner lebih tepat dibandingkan metode perhitungan manual.
4. Prediksi jumlah kunjungan pasien 2025 digunakan data 2019 hingga 2024. Pada prediksi jumlah kunjungan pasien tahun 2025, pada Excel menghasilkan prediksi yaitu 75938, pada RapidMiner hasil prediksi yaitu 75937,018 pasien atau dibulatkan menjadi 75938 pasien.
5. Prediksi pasien hipertensi dengan variabel independen jumlah pasien dewasa dan lansia di hasilkan prediksi sebesar 3271.881 hasil RapidMiner dan 3279.83 hasil Microsoft Excel.
6. Untuk penelitian selanjutnya, disarankan agar dilakukan pengembangan model prediksi menggunakan metode lain seperti ARIMA, Random Forest, atau Artificial Neural Network, sehingga hasil prediksi dapat dibandingkan dan diverifikasi dengan berbagai pendekatan. Selain itu, disarankan menggunakan rentang data waktu yang lebih panjang agar pola tren dan variasi musiman dapat teridentifikasi dengan lebih akurat.
7. Hasil penelitian ini dapat dijadikan bahan pertimbangan bagi manajemen Puskesmas Simpang Kawat dalam perencanaan sumber daya manusia, ketersediaan obat, serta strategi pelayanan kesehatan agar dapat mengantisipasi potensi peningkatan jumlah kunjungan pasien di tahun-tahun mendatang. Selain itu, hasil prediksi dapat digunakan sebagai dasar pengambilan keputusan kebijakan operasional, misalnya penjadwalan tenaga medis, pembagian beban kerja, dan penyediaan fasilitas pelayanan.

Daftar Pustaka

- [1] S. S. Alamsyah, Maimunah, and P. Sukmasetya, "Prediksi Jumlah Kedatangan Pasien Puskesmas Menggunakan Metode Backpropagation Artificial Neural Network," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 4, no. 6, pp. 2842–2849, 2024, doi: 10.30865/klik.v4i6.1922.

- [2] S. Puspitorini, R. Wahyuning Astuti, and M. Jannah Vemi Putri, "Association Rule Mining Untuk Promo Whatsapp Blast Pada Produk Ace Hardware Jambi Prima Mall," *J. Akad.*, vol. 15, no. 1, pp. 90–96, 2022, doi: 10.53564/akademika.v15i1.848.
- [3] R. Zana Nainggolan, K. Ibnutama, and M. Gilang Suryanata, "Implementasi Data Mining Dengan Metode Regresi Linier Berganda Dalam Estimasi Mahasiswa Baru Pada Sekolah Tinggi Agama Islam Raudhatul Akmal Batang Kuis," *J. CyberTech*, vol. 1, no. 1, pp. 13–20, 2021, [Online]. Available: <https://ojs.trigunadharma.ac.id/index.php/jct>
- [4] S. Adiguno, Y. Syahra, and M. Yetri, "Prediksi Peningkatan Omset Penjualan Menggunakan Metode Regresi Linier Berganda," *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 1, no. 4, p. 275, 2022, doi: 10.53513/jursi.v1i4.5331.
- [5] A. Anggara, K. Auliasari, and Y. Agus Pranoto, "Metode Regresi Linier Berganda Untuk Prediksi Omset Penyewaan Kamera Di Joe Kamera," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 1, pp. 852–858, 2023, doi: 10.36040/jati.v7i1.6158.