



Implementasi *Data Mining* dalam Prediksi Kebutuhan Obat di Puskesmas Talang Ratu Palembang Menggunakan Regresi Linier Berganda

Muhammad Aldo Nugraha¹, Dafid²

¹Universitas Multi Data Palembang, Palembang, Indonesia, aldonugraha52@gmail.com

²Universitas Multi Data Palembang, Palembang, Indonesia, dafid@mdp.ac.id

Corresponding Author: aldonugraha52@gmail.com¹

Abstract: *Community Health Centers play an important role in providing healthcare services, particularly in managing drug inventory. One of the common problems faced is inaccurate drug demand planning due to the use of manual calculation methods that lack precision. This study aims to apply data mining techniques to predict drug requirements at Talang Ratu Community Health Center in Palembang using the Multiple Linear Regression method. The data used in this study were obtained from one year of historical records of drug demand and usage. The research methodology follows the CRISP-DM framework, which includes business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The results show that the multiple linear regression model is able to predict drug inventory needs with good accuracy. Model evaluation produced a Mean Squared Error value of 12.854 and an R-Squared value of 0.78, indicating that 78% of the variation in drug inventory can be explained by demand and usage variables. Therefore, the proposed model is considered feasible as a decision support tool for planning drug inventory to improve efficiency and the quality of healthcare services at Talang Ratu Community Health Center in Palembang.*

Keyword: *data mining, drug demand prediction, multiple linear regression, CRISP-DM, community health center.*

Abstrak: Pusat Kesehatan Masyarakat (Puskesmas) memiliki peran penting dalam pelayanan Kesehatan, salah satunya dalam pengelolaan persediaan obat. Permasalahan yang sering terjadi adalah ketidaktepatan dalam perencanaan kebutuhan obat akibat metode perhitungan manual yang kurang akurat. Penelitian ini bertujuan untuk menerapkan teknik data mining dalam memprediksi kebutuhan obat di Puskesmas Talang Ratu Palembang menggunakan metode *Regresi Linier Berganda*. Data yang digunakan berasal dari riwayat permintaan dan pemakaian obat selama satu tahun. Metode penelitian mengacu pada tahapan *CRISP-DM* yang meliputi pemahaman bisnis, pemahaman data, persiapan data, pemodelan, evaluasi, dan penerapan. Hasil penelitian menunjukkan bahwa model *regresi linier berganda* mampu memprediksi kebutuhan persediaan obat dengan tingkat akurasi yang baik. Evaluasi model menghasilkan nilai *Mean Squared Error* sebesar 12,854 dan nilai *R-Squared* sebesar 0,78,

yang menunjukkan bahwa 78% variasi persediaan obat dapat dijelaskan oleh variabel permintaan dan pemakaian. Dengan demikian, model yang dihasilkan dinilai layak digunakan sebagai alat pendukung keputusan dalam perencanaan persediaan obat guna meningkatkan efisiensi dan kualitas pelayanan kesehatan di Puskesmas Talang Ratu Palembang.

Kata Kunci: data mining, prediksi kebutuhan obat, regresi linier berganda, CRISP-DM, puskesmas.

PENDAHULUAN

Puskesmas merupakan fasilitas pelayanan kesehatan tingkat pertama yang berperan penting dalam memberikan layanan kepada masyarakat. Selain sebagai penyelenggara pelayanan publik, Puskesmas juga menjalankan fungsi pelayanan klinis. Oleh karena itu, mutu pelayanan kesehatan yang diberikan harus memenuhi standar pelayanan minimal yang telah ditetapkan. Mutu pelayanan tersebut mencakup kualitas sumber daya manusia, tata kerja, proses operasional, serta pelayanan jasa kesehatan. Pelayanan kesehatan yang diterima pasien memiliki hubungan langsung dengan tingkat kepuasan masyarakat. Semakin baik mutu pelayanan yang diberikan, maka semakin tinggi pula tingkat kepuasan masyarakat, begitu juga sebaliknya. Penelitian ini berfokus pada pembahasan mengenai ketersediaan obat bagi pasien dalam jumlah yang optimal (Suwardiyanto dkk., 2019).

Ketersediaan obat merupakan obat yang digunakan untuk pelayanan kesehatan di puskesmas minimal harus sama dengan jumlah kebutuhan obat yang seharusnya tersedia di puskesmas. Ketersediaan obat di puskesmas harus disesuaikan dengan kebutuhan untuk pelayanan pengobatan pada masyarakat di wilayah kerjanya. Tingkat ketersediaan obat adalah tingkat persediaan obat baik jenis dan jumlah obat yang diperlukan oleh pelayanan pengobatan dalam periode waktu tertentu, diukur dengan cara menghitung persediaan dan pemakaian rata-rata perbulan. Oleh karena itu, kinerja pengelola obat sangat menentukan keberhasilan dalam pengelolaan obat di Puskesmas. Kinerja pengelolaan obat meliputi tahapan perencanaan, permintaan, penerimaan, penyimpanan, pendistribusian, pengendalian, pencatatan, pelaporan dan pengarsipan, pemantauan dan evaluasi program yang terkait satu sama lain (Amiruddin dkk., 2019).

Aspek penting dalam mendukung kelancaran operasional puskesmas adalah pengelolaan persediaan obat yang efisien, perkembangan teknologi informasi dan data mining telah membuka peluang baru dalam mengelola data persediaan obat secara lebih akurat dan efisien. *Data mining*, sebagai proses analisis data besar untuk menemukan pola atau hubungan yang tidak terlihat sebelumnya, dapat dimanfaatkan untuk meningkatkan efisiensi pengelolaan persediaan obat. Dengan penerapan metode *data mining*, Puskesmas Talang Ratu dapat melakukan prediksi terhadap kebutuhan obat berdasarkan pola penggunaan historis, mengidentifikasi tren konsumsi obat tertentu, serta mengoptimalkan waktu pemesanan dan pengadaan obat.

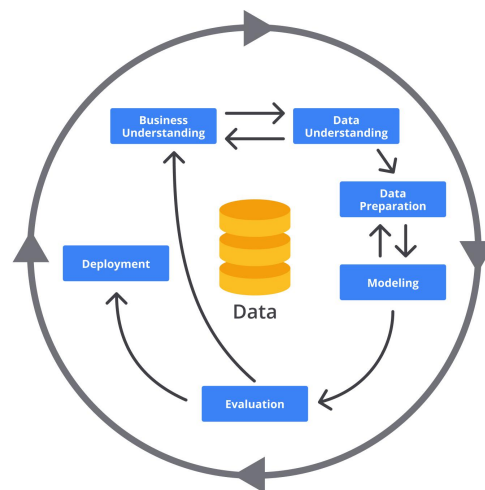
Puskesmas Talang Ratu Palembang belum bisa melakukan prediksi stok obat karena terhambat oleh beberapa faktor yaitu permintaan obat yang dibutuhkan oleh Puskesmas Talang Ratu belum tersedia di gudang farmasi. Hal tersebut sering kali membuat Puskesmas Talang Ratu kurang tepat dalam melakukan persediaan obat dikarenakan Puskesmas Talang Ratu hanya melakukan prediksi sebatas perkiraan saja dan dilakukan perhitungan secara manual dengan tidak menggunakan teknik-teknik untuk melakukan prediksi sehingga tidak mendapatkan hasil yang akurat. Dalam upaya untuk meningkatkan kinerja dari pelayanan Puskesmas Talang Ratu pada bidang farmasi dan membantu meminimalisir kesalahan yang terjadi, maka pada penelitian ini diterapkan *Data Mining* untuk memprediksi dalam menentukan jumlah stok obat untuk ke depannya dengan mengidentifikasi pola penggunaan obat berdasarkan Riwayat kunjungan pasien selama 1 tahun, hal ini memungkinkan

perencanaan pengadaan obat yang lebih akurat, efisiensi dan meningkatkan kualitas pelayanan kesehatan kepada masyarakat (Dahlia & Andri, 2021).

Data Mining yaitu proses yang dilakukan untuk menemukan sebuah pengetahuan yang ada di dalam basis data atau biasa dikenal dengan sebutan CRISP- DM (*Cross-Industry Standard Process for Data Mining*). Adapun metode yang digunakan yaitu Algoritma Regresi Linier berganda dimana algoritma ini bekerja dengan membagi dataset menjadi subset yang lebih kecil berdasarkan atribut tertentu untuk menghasilkan model yang dapat membuat keputusan atau prediksi. Regresi Linier berganda digunakan untuk memprediksi data ke dalam kategori tertentu, seperti memprediksi obat di puskesmas berdasarkan atribut persediaan stok awal, pemakaian obat dan permintaan.

METODE

Pada penelitian ini menurut (Hasanah dkk., 2021), digunakannya metodologi data mining CRISP-DM sebagai pemecah masalah yang umum untuk bisnis dan penelitian. Metodologi ini terdiri dari enam tahapan yaitu Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, dan Deployment. Proses metodologi ini terdiri dari 6 tahapan yang dapat dijelaskan sebagai berikut.



Gambar 1. Metode CRISP-DM (source : <https://dicoding-assets.sgp1.cdn.digitaloceanspaces.com/blog/wp-content/uploads/2024/06/Metodologi-data-science.png>)

CRISP-DM (*CROSS-INDUSTRY STANDARD PROCESS FOR DATA MINING*)

1. *Business Understanding*

Beberapa hal yang dilakukan pada tahap ini seperti memahami kebutuhan serta tujuan dari sudut pandang bisnis selanjutnya mengartikan pengetahuan ke dalam bentuk pendefinisian masalah pada data mining dan kemudian menentukan rencana serta strategi untuk mencapai tujuan data mining.

2. *Data Understanding*

Tahapan ini diawali dengan mengumpulkan data, mendeskripsikan data, serta mengevaluasi kualitas data.

3. *Data Preparation*

Dalam tahapan ini yaitu membangun dataset akhir dari berupa data mentah. Ada beberapa hal yang akan dilakukan mencakup melakukan pembersihan data (Data Cleaning), melakukan pemilihan data (Data Selection), record dan atribut-

atribut, dan juga melakukan transformasi terhadap data (Data Transformation) untuk dijadikan masukan dalam tahap pemodelan.

4. *Modelling*

Berbagai teknik pemodelan dipilih dan diterapkan, dan parameternya dikalibrasi dengan nilai optimal. Terdapat beberapa teknik untuk jenis masalah *data mining* yang sama. Beberapa teknik memiliki persyaratan khusus berupa data. Karena itu, tahap persiapan data sangat diperlukan.

5. *Evaluation*

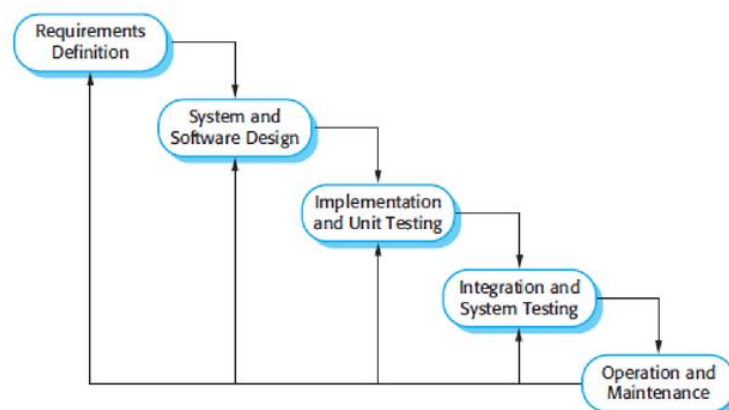
Evaluasi secara menyeluruh dan meninjau langkah-langkah yang dijalankan untuk membuatnya dilakukan, untuk memastikan model tersebut benar mencapai tujuan bisnis. Tujuan utamanya adalah untuk menentukan apakah ada beberapa masalah bisnis penting yang belum dipertimbangkan. Pada akhir tahap ini, keputusan penggunaan hasil data mining harus dicapai.

6. *Deployment*

Pembuatan model umumnya bukan akhir dari proyek. Sekalipun tujuan dari model ini adalah untuk meningkatkan pengetahuan tentang data, pengetahuan yang didapat perlu diorganisir dan disajikan sedemikian rupa sehingga pelanggan dapat menggunakannya. Pengembangan aplikasi pada tahap ini akan dilakukan dengan metodologi pengembangan sistem pada penelitian ini ialah *Waterfall*.

WATERFALL

Metode pengembangan sistem menurut (Sasmito dkk., 2017) yang diterapkan pada penelitian ini adalah dengan pengembangan sistem menggunakan metode waterfall. Metode waterfall merupakan model pengembangan sistem informasi yang sistematis dan sekuensial. Metode Waterfall memiliki tahapan-tahapan sebagai berikut.



Gambar 2. Metode *Waterfall* (source : <https://www.sekawanmedia.co.id/wp-content/uploads/2024/01/metode-watefall-menurut-sommerville.png>)

1. *Requirements analysis and definition*

Pengembang sistem memerlukan komunikasi terhadap klien untuk tujuan mencari data kebutuhan dan memahami perangkat lunak yang diharapkan oleh klien dan batasan batasan apa saja nantinya dalam pembuatan perangkat lunak. Data kebutuhan itu berupa informasi yang dilakukan dengan cara wawancara, diskusi, atau survey langsung dan kemudian akan dilakukan analisa yang menghasilkan kebutuhan yang dibutuhkan oleh klien.

2. *System and software design*

Tahap ini merupakan spesifikasi dari tahap sebelumnya yang akan dipelajari oleh pengembang dalam membuat sistem nantinya. Desain sistem membantu pengembang dalam mendefinisikan dari arsitektur dari sistem secara keseluruhan.

3. *Implementation and unit testing*

Tahap pembuatan aplikasi yang akan dibuat oleh pengembang, pengaplikasian apa yang sudah didapat dari tahap sebelumnya agar dapat membangun aplikasi yang sesuai dengan apa yang dibutuhkan.

4. *Integration and system testing*

Aplikasi yang sudah dibuat di dalam tahap sebelumnya dilakukan testing dan pengujian. Tahapan ini adalah tahapan terakhir dari pengembangan perangkat lunak, setelah tahap pengujian dan testing perangkat lunak akan dikirimkan ke klien.

5. *Operation and maintenance*

Tahap ini adalah tahapan pemeliharaan sistem setelah perangkat lunak dikirimkan kepada klien. Pemeliharaan ini termasuk dalam memperbaiki kesalahan-kesalahan yang ada pada sistem yang tidak ditemukan dalam tahap pengujian dan testing.

HASIL DAN PEMBAHASAN

Modelling

Tahap medelling dalam penelitian dilakukan menggunakan *google collab* dapat dilihat pada Gambar 3. Berdasarkan perhitungan model Algoritma Regresi Linier Berganda maka data pada setiap kelompok (*group*) terlebih dahulu di-reset indeks nya menggunakan fungsi *reset_index(drop=True)*. Proses ini dilakukan untuk memastikan bahwa indeks pada *dataframe* bersifat berurutan dan bersih dari indeks sebelumnya. Hal ini penting terutama ketika data akan divisualisasikan atau digunakan sebagai masukan ke dalam model, sehingga tidak terjadi ketidaksesuaian antara baris data dan indeksnya. Selanjutnya, variabel prediktor (fitur) dan variabel target dipisahkan. Pada kode ini, fitur yang digunakan adalah X1 dan X2, yang kemudian disimpan dalam variabel X.

```
# Reset index for clean plotting
group = group.reset_index(drop=True)
X = group[['x1', 'x2']]
y = group['Y']

# Train model
model = LinearRegression()
model.fit(X, y)
y_pred_train = model.predict(X)
```

Gambar 3. Model Data Mining

Berdasarkan hasil proses pemodelan regresi linier berganda yang diterapkan pada seluruh data obat selama periode penelitian, diperoleh persamaan regresi rata-rata dapat dilihat pada Gambar 4. Persamaan ini merupakan model estimasi yang menggambarkan hubungan antara permintaan (X_1), pemakaian (X_2), dan persediaan obat (Y) secara keseluruhan dari semua jenis obat dalam dataset.

Gambar 4. Persamaan Regresi

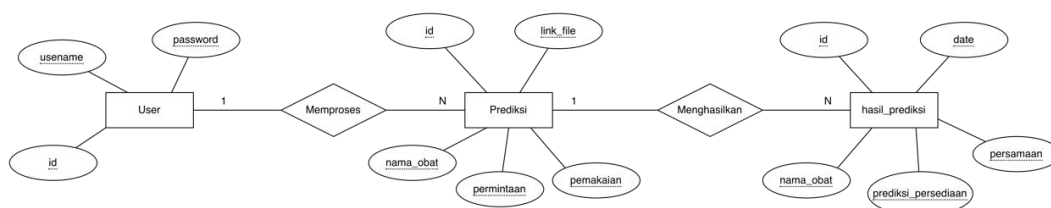
Evaluation

Tahapan evaluasi merupakan proses penting untuk memastikan bahwa model Regresi Linier Berganda yang dibangun mampu memberikan prediksi yang akurat dan sesuai dengan tujuan bisnis, yaitu mendukung perencanaan persediaan obat di Puskesmas Talang Ratu Palembang. Evaluasi dilakukan menggunakan dua metrik utama, yaitu *Mean Squared Error (MSE)* dan *R-Squared (R^2)*. *MSE* digunakan untuk mengukur rata-rata besarnya kesalahan kuadrat antara nilai aktual dan nilai hasil prediksi, sehingga semakin kecil nilai *MSE* menunjukkan semakin baik performa model. Berdasarkan hasil perhitungan, diperoleh nilai *MSE* = 12.854, yang menunjukkan bahwa selisih antara nilai aktual dan hasil prediksi cukup rendah dan masih dalam batas toleransi untuk data operasional persediaan obat.

Nilai *R-Squared (R^2)* digunakan untuk mengukur seberapa besar variasi pada variabel dependen (persediaan obat) dapat dijelaskan oleh variabel independen (permintaan dan pemakaian). Hasil evaluasi menunjukkan bahwa model memiliki nilai $R^2 = 0.78$, yang berarti bahwa 78% variasi jumlah persediaan obat dapat dijelaskan oleh variabel permintaan (X_1) dan pemakaian (X_2), sedangkan sisanya sebesar 22% dipengaruhi oleh faktor lain di luar model, seperti keterlambatan distribusi, kebijakan pengadaan, atau fluktuasi kebutuhan pasien. Dengan demikian dapat disimpulkan model dinyatakan layak digunakan sebagai alat pendukung keputusan dalam merencanakan kebutuhan stok obat. Model ini mampu membantu mengurangi risiko kekurangan ataupun kelebihan stok, serta memberikan dasar perencanaan yang lebih terukur berdasarkan data.

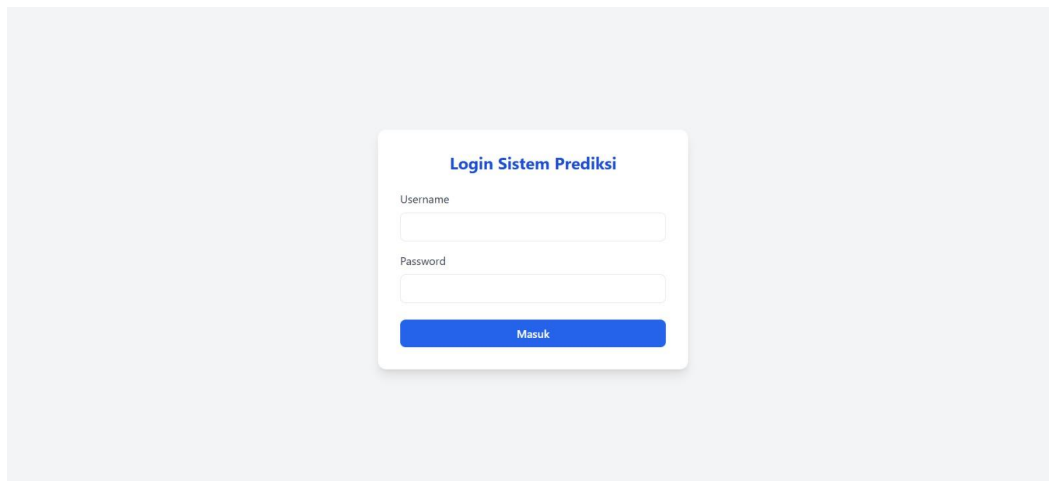
Deployment

Setelah melakukan tahap *Modelling* dan *Evaluation* dengan menyelesaikan perhitungan terhadap *dataset*, maka langkah selanjutnya adalah *deployment* yaitu membuat sebuah *output* berupa aplikasi *website* pada langkah berikutnya, di mana rancangan *Entity Relationship Diagram (ERD)* dapat dilihat pada Gambar 5. Rancangan antarmuka aplikasi data mining yang diimplementasikan ke dalam sistem berbasis framework Laravel. Model tersebut dihubungkan dengan aplikasi sehingga admin dapat memasukkan atau mengunggah data permintaan dan pemakaian obat terbaru, kemudian sistem secara otomatis melakukan perhitungan prediksi berdasarkan rumus regresi yang telah dihasilkan.

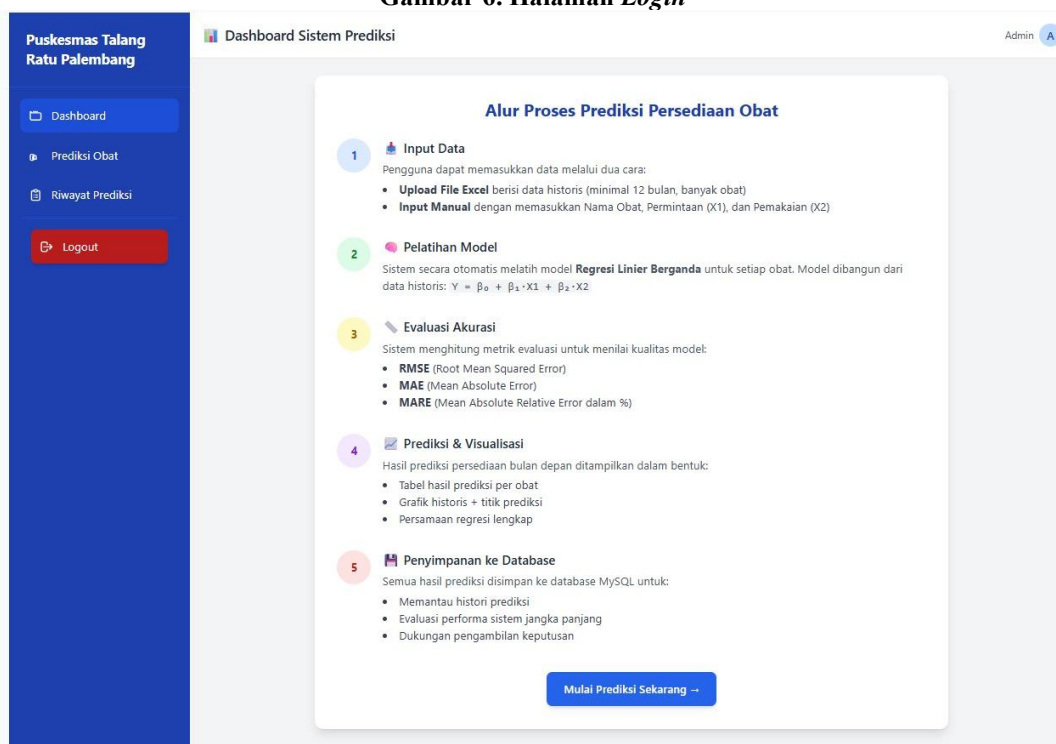


Gambar 5. Entity Relationship Diagram

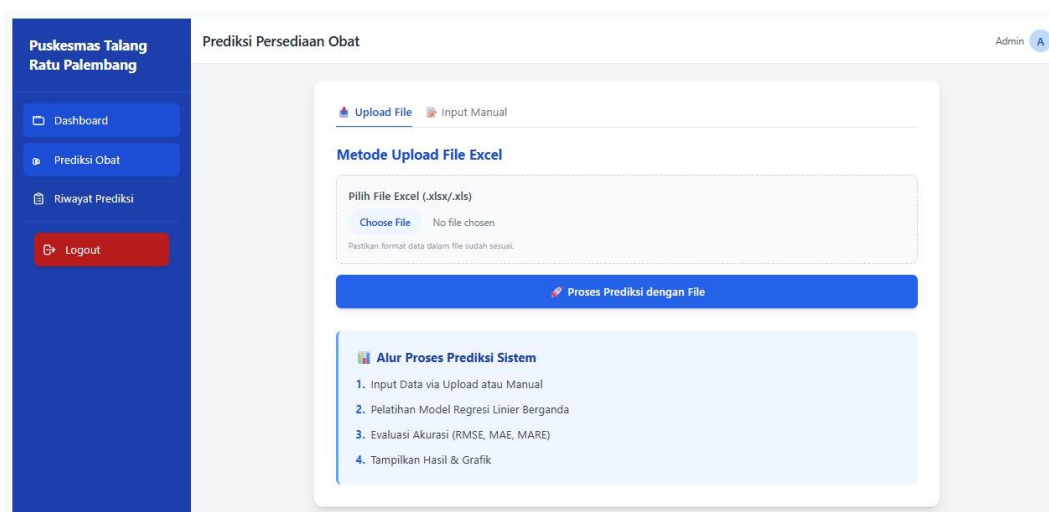
Pada Gambar 6. terdapat halaman *login*. Pada halaman login aplikasi ini pengguna harus menginput username dan password yang telah dibuat sebelumnya, setelah pengguna berhasil menginput username dan password serta mengklik tombol masuk lalu pengguna langsung dapat ke halaman dashboard.



Gambar 6. Halaman Login



Gambar 7. Halaman Dashboard



Gambar 8. Halaman Prediksi

NO	NAMA OBAT	PREDIKSI PERSEDIAAN	RMSE	MAE	MARE (%)	PERSAMAAN	WAKTU
1	Parasetamol 500mg	875.98	504.77	426.12	61.98%	$Y = 664.80 + (0.84)X1 + (0.34)X2$	2025-12-01 16:53
2	Calcium Lactat 500 mg	957.65	526.14	472.07	97.39%	$Y = 1092.57 + (-0.88)X1 + (0.80)X2$	2025-12-01 16:52
3	Chlorhexidine 2% 5 L	1002.01	463.46	379.62	63.73%	$Y = 963.54 + (-0.65)X1 + (1.61)X2$	2025-12-01 16:52
4	Antasida doen tablet	1209.89	534.00	461.29	75.59%	$Y = 1425.12 + (-1.87)X1 + (1.84)X2$	2025-12-01 16:52
5	Amlodipine 10mg	823.03	480.27	430.58	103.80%	$Y = 767.20 + (3.18)X1 + (-3.86)X2$	2025-12-01 16:52
6	Amlodipine 5mg	939.02	505.77	421.33	72.11%	$Y = 693.52 + (0.78)X1 + (0.35)X2$	2025-12-01 16:52
7	Amoxicillin 500 mg	1083.99	532.64	461.49	123.11%	$Y = 1100.11 + (-0.93)X1 + (1.45)X2$	2025-12-01 16:52
8	Allopurinol 100 mg Tab	660.39	378.66	337.27	47.06%	$Y = 20.73 + (2.44)X1 + (0.57)X2$	2025-12-01 16:52
9	Allopurinol 100mg	971.93	540.37	461.07	87.74%	$Y = 75.42 + (1.84)X1 + (3.14)X2$	2025-12-01 16:52

Gambar 9. Halaman Riwayat Prediksi

Halaman *dashboard* dapat dilihat pada Gambar 7. Pengguna dapat melihat petunjuk dari aplikasi dan alur proses data mining serta bisa memulai untuk memprediksi. Gambar 8. menunjukkan halaman prediksi. Pada halaman prediksi pengguna dapat mengupload dokumen untuk memprediksi berdasarkan historis dengan memilih model terlebih dahulu ingin memprediksi persediaan dan menampilkan hasil dari prediksi. Halaman prediksi dapat juga menginput nama obat permintaan dan pemakaian untuk diprediksi secara manual. Gambar 9 memperlihatkan halaman riwayat prediksi. Pada halaman riwayat prediksi menunjukan hasil prediksi yang berfungsi untuk menampilkan seluruh hasil prediksi yang telah dilakukan oleh admin sebelumnya, sehingga memudahkan proses penelusuran, evaluasi, dan monitoring perubahan kebutuhan obat dari waktu ke waktu.

KESIMPULAN

Pendekatan ini mampu memberikan solusi yang lebih efektif dan akurat dibandingkan proses manual yang sebelumnya digunakan oleh pihak puskesmas. Evaluasi model menunjukkan bahwa model memiliki tingkat keakuratan yang baik. *Metrik* evaluasi seperti *Mean Squared Error (MSE)*, *Mean Absolute Error (MAE)*, dan *R-Squared (R²)* memberikan gambaran bahwa model dapat memprediksi persediaan dengan tingkat kesalahan yang masih dapat diterima untuk kebutuhan operasional puskesmas.

Nilai evaluasi menunjukkan bahwa model memiliki nilai *MSE* dan *MAE* yang relatif rendah serta nilai *R²* yang cukup tinggi, menandakan bahwa sebagian besar variasi persediaan dapat dijelaskan oleh variabel permintaan dan pemakaian. Berdasarkan indikator tersebut, model dinilai layak untuk digunakan sebagai alat bantu prediksi. Sistem juga menyediakan informasi tambahan berupa *RMSE*, *MAE*, *MARE (%)*, serta persamaan regresi setiap obat, sehingga memudahkan proses evaluasi dan pengambilan keputusan.

REFERENSI

- D. Suwardiyanto, M. Nur Shodiq, D. Hidayat Kusuma, dan T. Oktalita Sari, “Sistem Prediksi Kebutuhan Obat di Puskesmas Menggunakan Metode Least Square,” *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 4, no. 1, hlm. 75–80, Jan 2019, doi: 10.30591/jpit.v4i1.1085.
- E. E. Amiruddin dan W. I. Septarani A, “Study Drug Availability at the Meo-Meo Health Center in the City of Baubau,” *Jurnal Ilmiah Kesehatan (JIKA)*, vol. 1, no. 2, hlm. 60–76, Des 2019, doi: 10.36590/jika.v1i2.11.
- Dahlia dan Andri, “Implementasi Data Mining untuk Prediksi Persediaan Obat pada Puskesmas Kertapati menggunakan Regresi Linier Berganda,” *Jurnal Sistem dan Informatika (JSI)*, vol. 15, no. 2, hlm. 95–103, Mei 2021, doi: 10.30864/jsi.v15i2.331.
- M. A. Hasanah, S. Soim, dan A. S. Handayani, “Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir,” 2021. [Daring]. Tersedia pada: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- G. W. Sasmito, J. T. Informatika, H. Bersama, J. Mataram, N. 09, dan P. Lor, “Penerapan Metode Waterfall Pada Desain Sistem Informasi Geografis Industri Kabupaten Tegal,” vol. 2, no. 1, 2017, [Daring]. Tersedia pada: <http://www.tegalkab.go.id>,