

Penerapan Regresi Linear Menggunakan RapidMiner Untuk Memprediksi Penjualan dan Persediaan

Diterima:

10 Juni 2024

Revisi:

10 Juli 2024

Terbit:

1 Agustus 2024

¹Akmal Hisyam Pradhana, ²Muhammad Irfa'udin, ³Achmad Ali Firmansyah, ⁴Aidina Ristyawan, ⁵Erna Daniati

¹Fakultas Teknik & Ilmu Komputer, ²Sistem Informasi, ³Universitas Nusantara PGRI Kediri

¹akmalhisyampradhana@gmail.com, ²irfaudin906@gmail.com,

³achmadali059@gmail.com, ⁴aidinaristi@unpkediri.ac.id,

⁵ernadaniati@unpkediri.ac.id

Abstrak—Prediksi dilakukan untuk meramalkan penjualan dan stok masa depan agar dapat memenuhi kebutuhan konsumen. Meramalkan sangat penting karena berkaitan langsung dengan jumlah barang yang terjual dan memungkinkan perencanaan strategi penjualan untuk periode mendatang. Prediksi didasarkan pada data penjualan atau permintaan dari periode sebelumnya. Penelitian bertujuan untuk menerapkan regresi linear menggunakan RapidMiner sebagai alat bantu analisis dalam memprediksi penjualan dan persediaan pada Keripik Assri. Melalui penelitian ini, UMKM dapat membuat keputusan yang lebih tepat. Metode penelitian yang digunakan adalah knowledge discovery in database (KDD) dan dilakukan menggunakan RapidMiner versi 10.3 dengan data penjualan dan stok pada tahun 2019 sampai 2024. Hasil menunjukkan bahwa nilai Root Mean Squared Error (RMSE) untuk penjualan dan stok adalah 0.000 +/- 0.000. Ini menunjukkan selisih antara hasil prediksi model dan nilai sebenarnya sangat kecil, menunjukkan tingkat akurasi yang sangat tinggi.

Kata Kunci—Rapidminer, Regresi Linear, RMSE, Penjualan

Abstract— Forecasting is done to forecast future sales and stock in order to meet consumer needs. Forecasting is very important because it is directly related to the amount of goods sold and allows planning sales strategies for future periods. Predictions are based on sales or demand data from previous periods. The research aims to apply linear regression using RapidMiner as an analytical tool in predicting sales and inventory at Assri Chips. Through this research, MSMEs can make more informed decisions. The research method used is knowledge discovery in database (KDD) and is carried out using RapidMiner version 10.3 with sales and stock data from 2019 to 2024. The results show that the Root Mean Squared Error (RMSE) value for sales and stock is 0.000 +/- 0.000. This shows that the difference between the model's predicted and actual values is very small, indicating a very high level of accuracy.

Keywords—Rapidminer, Regresi Linear, RMSE, Sale

This is an open access article under the CC BY-SA License.



Penulis Korespondensi:

Aidina Ristyawan,

Sistem Informasi,

Universitas Nusantara PGRI Kediri,

Email: aidinaristi@unpkediri.ac.id

ID Orcid: [<https://orcid.org/0009-0003-2712-1507>]

Handphone: 0812-3262-4460

I. PENDAHULUAN

UMKM sering menghadapi tantangan dalam manajemen persediaan dan penjualan. Kesalahan dalam manajemen persediaan dapat menyebabkan kerugian berdampak pada biaya operasional dan kepuasan pelanggan. Oleh karena itu, pengabdian kepada masyarakat melalui penerapan teknologi analisis data sangat penting untuk membantu UMKM dalam mengatasi masalah. Keripik Assri merupakan salah satu UMKM di Kediri yang memproduksi makanan ringan seperti keripik pisang, tempe, sukun, ubi ungu, talas, dan kue kecipir. Sebagai UMKM, Keripik Assri menghadapi masalah dalam mengelola persediaan dan penjualan. Masalah utama Keripik Assri adalah sering terjadinya kelebihan dan kekurangan persediaan keripik. Masalah dapat terpecahkan menggunakan sektor teknologi informasi yang memungkinkan manusia mendapatkan informasi secara cepat dan akurat[1]. Regresi Linear menjadi metode yang tepat guna memecahkan masalah dengan cara memprediksi persediaan dan memprediksi penjualan untuk meningkatkan efisiensi operasional dan memaksimalkan keuntungan[2]. Metode ini memungkinkan mempelajari beberapa faktor sebelumnya untuk meramalkan kejadian di masa depan. Dengan cara ini, bisa lebih siap menghadapi apa yang akan terjadi nantinya[3]. Penelitian ini bertujuan untuk menerapkan regresi linear menggunakan RapidMiner sebagai alat bantu analisis dalam memprediksi penjualan dan persediaan pada Keripik Assri. Dengan memanfaatkan RapidMiner, diharapkan UMKM ini dapat mengoptimalkan proses bisnis mereka dan membuat keputusan yang lebih akurat berdasarkan prediksi yang tepat.

UMKM dapat mengelola persediaan dengan lebih baik, mengurangi biaya penyimpanan, dan meningkatkan kepuasan pelanggan dengan memprediksi penjualan secara akurat. Dengan tujuan persediaan produk memudahkan pelaku bisnis lebih mudah memenuhi permintaan pasar[4]. Hipotesis dalam penelitian ini adalah bahwa penerapan regresi linear menggunakan RapidMiner dapat memprediksi penjualan yang tepat dan membantu dalam pengelolaan persediaan yang lebih efisien pada Keripik Assri[5]. Penelitian mengacu pada literatur dan konsep yang berkaitan dengan regresi linear dan analisis data.

Data Mining adalah proses identifikasi hubungan, pola, dan perkembangan baru dari data besar menggunakan teknologi pengenalan pola serta metode statistik dan matematika[6]. *Data Mining* melibatkan teknik statistik, matematika, kecerdasan buatan, dan machine learning untuk mengekstrak informasi dan pengetahuan bermanfaat dari kumpulan data besar [7]. *Data mining* adalah rangkaian langkah untuk menemukan nilai tambahan dari sekumpulan data dengan mengungkap kebenaran yang sebelumnya belum terdeteksi secara manual[8]. Tujuan utama data mining adalah mengungkap, menemukan, atau mengeksplorasi pemahaman dari data atau informasi yang tersedia[9].

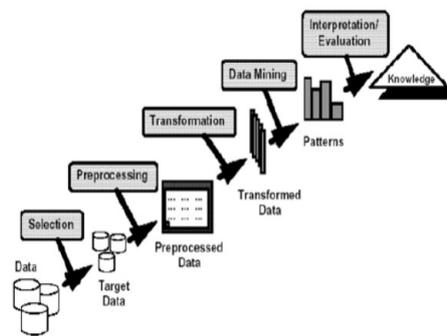
Root Mean Square Error (RMSE) adalah standar penilaian yang digunakan untuk menilai akurasi prediksi model regresi terhadap Nilai yang diamati yang sesungguhnya. RMSE kerap dipakai dalam evaluasi performa model regresi serta perbandingan mutu antara berbagai model regresi[5]. Semakin mendekati nilai RMSE yang mendekati 0, semakin tinggi akurasi prediksinya[10].

Sistem Pendukung Keputusan (SPK) adalah teknologi yang membantu dalam pengambilan keputusan atau kebijakan[11]. Regresi linear adalah algoritma yang menggambarkan interaksi antara variabel independen dan dependen, digunakan untuk menemukan pola dan tren dalam data untuk prediksi penjualan dan persediaan[12].

RapidMiner adalah platform analisis data yang memungkinkan pengguna melakukan pemodelan prediktif tanpa memerlukan keahlian pemrograman yang mendalam. *RapidMiner* dikembangkan dengan pendekatan open core, yang menggabungkan elemen-elemen open source dengan fitur-fitur tambahan yang tersedia dalam versi berbayar[13].

II. METODE

Analisis deskriptif adalah metode untuk menjelaskan objek berdasarkan data yang telah dikumpulkan dan dianalisis, untuk membuat keputusan yang berlaku secara umum. Tahapan *data mining* dilakukan dengan metode *Knowledge Discovery in Database (KDD)*, yaitu Penggalan informasi yang tersembunyi dalam dataset yang luas. istilah *Knowledge Discovery in Database/KDD* juga merujuk pada metode pencarian informasi yang belum teridentifikasi dalam sebuah basis data yang besar[14]. Tahapan prosesnya pada Gambar 1.



Gambar 1 Tahap Knowledge Discovery in Database

Penjelasan alur pada Gambar 1 :

- Seleksi data (*selection*) adalah Memilih data yang relevan dari kumpulan data untuk kebutuhan penelitian.
- Pra-pemrosesan (*Pre-processing*) adalah Membersihkan data dengan memeriksa konsistensi, menghapus data identik, dan memperbaiki kesalahan.
- Data transformation* misalnya mengubah polinomial menjadi numerik atau sebaliknya agar dapat digunakan dalam data mining. Operator digunakan untuk mengonversi yang lebih efektif untuk analisis[15].
- Data mining* adalah proses menemukan pola atau pemahaman yang sebelumnya tidak diketahui dalam kumpulan data tertentu dengan menggunakan teknik tertentu.
- Evaluasi adalah Menyajikan pola yang ditemukan dalam bentuk yang mudah dipahami oleh pengguna.

III. HASIL DAN PEMBAHASAN

3.1 Selection & Pre-processing Data

Dalam penelitian ini, dataset yang digunakan sebagai bahan prediksi meliputi informasi produksi dan penjualan yang dikumpulkan selama lima tahun terakhir. Berikut adalah data penjualan keripik seperti yang ditunjukkan pada Tabel 1.

Tabel 1 Dataset Penjualan Keripik

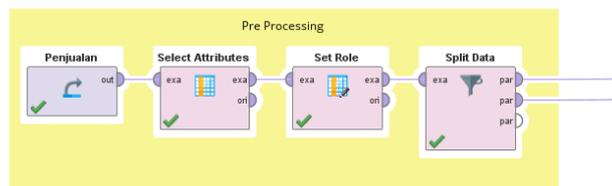
Nomor	Tanggal	Bulan	Tahun	Date	Pisang	Tempe	Sukun	Ubi Ungu	Talas	Total Penjualan	Keterangan
1	1 sd 31	1	2019	31/01/2019	99	148	49	97	99	492	Meningkat
2	1 sd 28	2	2019	28/02/2019	109	159	53	107	94	522	Meningkat
3	1 sd 31	3	2019	31/03/2019	109	159	58	112	92	530	Meningkat
.....				
62	1 sd 29	2	2024	29/02/2024	168	49	158	139	49	563	Meningkat
63	1 sd 31	3	2024	31/03/2024	168	69	178	139	48	602	Meningkat
64	1 sd 30	4	2024	30/04/2024	157	69	157	148	48	579	Menurun

Berikut adalah data persediaan keripik seperti yang ditunjukkan pada Tabel 2.

Tabel 2 Dataset Persediaan Keripik

Nomor	Tanggal	Bulan	Tahun	Date	Pisang	Tempe	Sukun	Ubi Ungu	Talas	Total Persediaan	Keterangan
1	1 sd 31	1	2019	31/01/2019	100	150	50	100	100	500	Meningkat
2	1 sd 28	2	2019	28/02/2019	110	160	55	110	95	530	Meningkat
3	1 sd 31	3	2019	31/03/2019	110	160	60	115	95	540	Meningkat
.....				
62	1 sd 29	2	2024	29/02/2024	150	70	120	120	90	550	Menurun
63	1 sd 31	3	2024	31/03/2024	170	50	160	140	50	570	Meningkat
64	1 sd 30	4	2024	30/04/2024	160	60	170	130	40	560	Menurun

Dataset penjualan dan persediaan keripik dilakukan Pre-Processing sebelum digunakan. Seperti yang ditunjukkan pada Gambar 1.

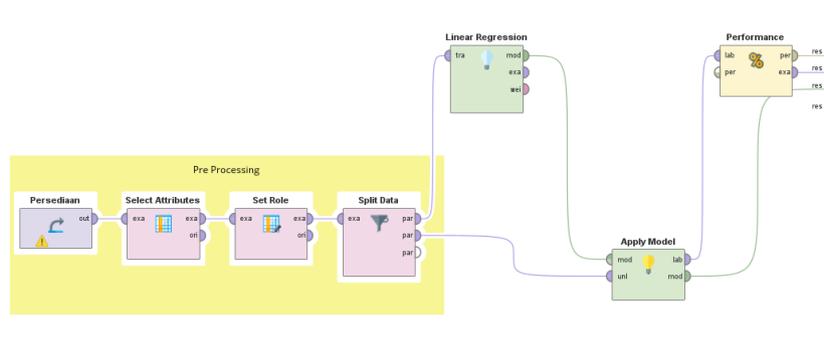


Gambar 1 PreProcessing Data Penjualan dan Persediaan

Langkah pertama memilih kolom relevan dalam dataset, yaitu "Date, Pisang, Tempe, Sukun, Ubi Ungu, Talas, dan Total Penjualan". Operator "Set Role" Menjadikan kolom Total penjualan dan Total persediaan sebagai label. Sedangkan kolom lainnya sebagai reguler. Data dibagi menjadi training dan testing dengan operator split data. dengan presentase 0.5 data training dan 0.5 data testing. Data testing yang akan diproses untuk menghasilkan hasil prediksi.

3.2 Data Mining

Proses Data Mining akan dilakukan pada data yang telah dilakukan preprocessing. Prosedur dimulai dengan preprocessing dan kemudian data mining menggunakan algoritma Linear Regression. Seperti yang ditunjukkan pada Gambar 2:



Gambar 2 Proses Data Mining

Operator Apply Model digunakan untuk menerapkan model yang telah dilatih pada data baru untuk membuat prediksi. Setelah model regresi linear dilatih, Apply Model memprediksi data uji. Evaluasi kinerja model dilakukan dengan Operator Performance (Regression) menggunakan metrik seperti Root Mean Squared Error (RMSE), yang menunjukkan seberapa besar kesalahan prediksi—nilai RMSE yang lebih kecil menunjukkan kinerja model yang lebih baik.

3.3 Evaluasi

Hasil prediksi total penjualan dan persediaan ditunjukkan pada Gambar 3.

Row No.	Total Penjualan	prediction(Total Penjualan)	Date	Pisang	Tempe	Sukun	Ubi Ungu	Telas
1	530	530.000	Mar 31, 2019	109	159	58	112	92
2	569	569.000	Jul 31, 2019	118	163	62	118	108
3	589	589.000	Sep 30, 2019	122	168	67	124	108
4	638	638.000	Oct 31, 2019	133	178	79	134	114
5	685	685.000	Feb 29, 2020	144	188	87	143	123
6	670	670.000	Mar 31, 2020	139	183	83	138	127
7	716	716.000	Apr 30, 2020	147	193	94	148	134
8	669	669.000	Jun 30, 2020	138	184	82	138	127
9	668	668.000	Aug 31, 2020	138	184	82	137	127
10	627	627.000	Sep 30, 2020	128	174	73	128	124
11	671	671.000	Oct 31, 2020	138	183	82	139	129
12	642	642.000	Nov 30, 2020	133	177	77	132	123
13	642	642.000	Jan 31, 2021	132	178	77	132	123

Gambar 3 Hasil Prediksi Total Penjualan

Hasil Root Mean Squared Error ditunjukkan pada Gambar 4



Gambar 4 Hasil Root Mean Squared Error

Pada Gambar 4 didapatkan hasil nilai Root Mean Squared Error (RMSE) penjualan dan persediaan sebesar 0.000 +/- 0.000. Hal itu menunjukkan bahwa nilai penjualan dan persediaan yang diprediksi oleh model dengan nilai sebenarnya sangat kecil dan hampir tidak ada kesalahan.

IV. KESIMPULAN

Hasil prediksi penjualan dan persediaan UMKM keripik Assri menunjukkan bahwa model regresi yang digunakan memiliki tingkat akurasi yang sangat tinggi, sebagaimana ditunjukkan oleh nilai Root Mean Squared Error (RMSE) yang mencapai 0.000 +/- 0.000. Nilai RMSE yang nol ini mengindikasikan bahwa nilai penjualan dan persediaan yang diprediksi oleh model dengan nilai sebenarnya sangat kecil, tidak ada kesalahan. Dengan demikian, model ini dapat diandalkan untuk memprediksi jumlah penjualan dan persediaan keripik secara akurat, membantu dalam perencanaan inventaris dan strategi pemasaran yang lebih efektif. Hasil ini menunjukkan potensi besar untuk mengoptimalkan manajemen stok dan meminimalkan risiko kelebihan atau kekurangan persediaan, yang pada akhirnya dapat meningkatkan efisiensi operasional dan kepuasan pelanggan.

DAFTAR PUSTAKA

- [1] E. Daniati and A. Nugroho, "K-Means clustering with Decision Support System using SAW: Determining thesis topic," in *Proceedings - 6th IEEE International Conference on Control System, Computing and Engineering, ICCSCE 2016*, 2017. doi: 10.1109/ICCSCE.2016.7893593.
- [2] R. Jumardi and S. H. Widiastuti, "Aplikasi Forecasting Penjualan Dan Persediaan Produk Usaha Mikro, Kecil Dan Menengah," *JSAI (Journal Scientific and Applied Informatics)*, vol. 4, no. 3, 2021, doi: 10.36085/jsai.v4i3.2756.
- [3] A. Ristyawan and A. C. Pradikdo, "MODEL KLASIFIKASI ABSTRAK SKRIPSI MENGGUNAKAN TEXT MINING UNTUK PENGKATEGORIAN SKRIPSI SESUAI BIDANG KAJIAN," *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, vol. 9, no. 2, 2018, doi: 10.24176/simet.v9i2.2513.
- [4] H. Jaya, R. Gunawan, and R. Kustini, "Penerapan Data Mining Untuk Memprediksi Target Produksi Berdasarkan Tingkat Penjualan Dan Banyaknya Pemesanan Produk Pada Pt. Neo National Menggunakan Metode Regresi Linier Berganda," vol. 18, no. SAINTIKOM, pp. 219–227, 2019.
- [5] A. Julkarnaen, A. Irma Purnamasari, and I. Ali, "ANALISIS PENJUALAN ROTI PADA DISTRIBUTOR MY ROTI MENGGUNAKAN METODE REGRESI LINEAR BERDASARKAN NILAI RMSE," 2024.
- [6] D. Setiawan, N. Surojudin, and W. Hadikristanto, "Prediksi Penjualan Obat Dengan Algoritma Regresi Linear," *Prosiding Sains dan Teknologi*, vol. 1, no. 1, 2022.
- [7] A. Natuzzuhriyyah and N. Nafisah, "Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Pembelajaran Secara Daring Menggunakan Algoritma Naïve Bayes," *Techno Xplore: Jurnal Ilmu Komputer dan Teknologi Informasi*, vol. 6, no. 2, 2021, doi: 10.36805/technoxplore.v6i2.1377.
- [8] R. S. Rohman, R. A. Saputra, and D. A. Firmansaha, "Komparasi Algoritma C4.5 Berbasis PSO Dan GA Untuk Diagnosa Penyakit Stroke," *CESS (Journal of Computer Engineering, System and Science)*, vol. 5, no. 1, 2020, doi: 10.24114/cess.v5i1.15225.
- [9] Miftahuljannah, Aswan Supriyadi Sunge, and Ahmad Turmudi Zy, "ANALISIS PREDIKSI PENJUALAN DENGAN METODE REGRESI LINEAR DI PT. EAGLE INDUSTRY INDONESIA," *Jurnal Informatika Teknologi dan Sains (Jinteks)*, vol. 5, no. 3, 2023, doi: 10.51401/jinteks.v5i3.3325.
- [10] D. S. Purnia, M. F. Adiwisastro, H. Muhajir, and D. Supriadi, "Pengukuran Kesenjangan Digital Menggunakan Metode Deskriptif Berbasis Website," *EVOLUSI: Jurnal Sains dan Manajemen*, vol. 8, no. 2, 2020, doi: 10.31294/evolusi.v8i2.8942.
- [11] Y. A. E. Tuah and A. Anyan, "IMPLEMENTASI MODEL REGRESI LINEAR SEDERHANA UNTUK PREDIKSI GAJI BERDASARKAN PENGALAMAN LAMA

- BEKERJA,” *JUTECH: Journal Education and Technology*, vol. 1, no. 2, 2020, doi: 10.31932/jutech.v1i2.1289.
- [12] I. Wahyuni, N. Nafi'iyah, and Masruroh, “Sistem Peramalan Penjualan Perumahan di Kabupaten Lamongan dengan Menggunakan Metode Regresi Linier Berganda,” *Seminar Nasional Sistem Informasi 2019*, no. September, 2019.
- [13] R. Nofitri and N. Irawati, “ANALISIS DATA HASIL KEUNTUNGAN MENGGUNAKAN SOFTWARE RAPIDMINER,” *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 5, no. 2, 2019, doi: 10.33330/jurteks.v5i2.365.
- [14] V. R. Prasetyo, H. Lazuardi, A. A. Mulyono, and C. Lauw, “Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Linear Regression,” *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 7, no. 1, 2021, doi: 10.25077/teknosi.v7i1.2021.8-17.
- [15] A. M. Andika, N. Suarna, R. D. Dana. "Analisa Dataset Asosiasi Penjualan Menggunakan Metode FP-Growth". *Jurnal Teknologi Ilmu Komputer, Vol 2, Nol, 2023, Hal. 80-88*, doi:DOI10.56854/jtik.v2i1.108