

ANALISIS PENJUALAN ROTI PADA DISTRIBUTOR MY ROTI MENGGUNAKAN METODE REGRESI LINEAR BERDASARKAN NILAI RMSE

Agus julkarnaen, Ade Irma Purnamasari, Irfan Ali

Teknik Informatika, STMIK IKMI Cirebon

Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45135

Agusjulkarnaen28@gmail.com

ABSTRAK

Pada saat ini roti menjadi salah satu kebutuhan pokok dalam kehidupan. Komposisi gizi roti bervariasi bergantung pada jenis tepung yang dipakai serta bahan tambahan lainnya. My roti menjual roti dari berbagai macam merk. Tujuan dari penelitian ini adalah untuk memprediksi penjualan roti pada 3 bulan berikutnya pada data penjualan roti merk My roti distributor Bandung timur. Regresi linear digunakan sebagai metode prediksi dengan jumlah roti yang terjual sebagai variabel Y dan periode penjualan roti sebagai variabel X. RMSE (Root Mean Squared Error) dan Relative Error digunakan untuk memutar hasil prediksi. Hasil kategori prediksi Entry pada bulan pertama terjual 564 unit, pada bulan kedua 569 unit dan bulan ketiga 575 unit dan kategori Mid pada bulan pertama terjual 63 unit, pada bulan kedua 64 unit dan bulan ketiga 65 unit. Hasil evaluasi nilai RMSE pada kategori Entry 83.318, sedangkan nilai Relative Error kategori Entry 13.25% dan Nilai RMSE kategori Mid 17.721, sedangkan Nilai Relative Error kategori Mid 17.42%. Prediksi penjualan handphone menggunakan metode regresi linier ini dapat dikatakan tergolong dalam kategori cukup atau bisa digunakan.

Kata kunci : Analisis, Penjualan, RSME, Regresi Linear

1. PENDAHULUAN

Roti memiliki peran penting dalam berbagai konteks budaya sebagai elemen fundamental dalam konsumsi pangan. Dalam konteks pola makan, roti seringkali dianggap sebagai sumber utama karbohidrat yang menyumbangkan energi yang penting bagi fungsi tubuh manusia[1].

Setiap penjualan dikumpulkan menjadi data, yang dapat diolah dengan data mining untuk menghasilkan pengetahuan baru dan memperbaiki pengambilan keputusan di masa depan[2].

Penelitian prediksi penjualan roti pada 2022-2023 untuk pengambilan keputusan. Prediksi atau Forecasting adalah ramalan terhadap kejadian masa depan yang membantu dalam pengambilan keputusan[3].

Pada penelitian ini prediksi atau Forecasting bersifat time series, data urutan waktu atau "time series data" merujuk pada data yang terkumpul, tercatat atau diamati secara berurutan sesuai dengan waktu[4].

Regresi linear, metode statistik untuk time series, mempelajari hubungan antara variabel dependent Y dan variabel independent X, digunakan untuk memprediksi nilai Y berdasarkan nilai yang diberikan pada X[5].

My roti distributor Bandung timur mengalami hambatan cashflow karena stok yang masih belum terjual. Studi data mining dengan regresi linear pada data penjualan roti di toko toko dan dievaluasi dengan 2 indikator penilaian yaitu RMSE dan Relative Error. RMSE adalah akar kuadrat rata-rata error dari suatu metode yang digunakan untuk mengevaluasi kesesuaian data sebenarnya dengan model prediksi. Semakin mendekati 0 nilai RMSE, semakin akurat prediksinya[6].

Relative error adalah perbandingan kesalahan pada nilai aktual dengan kesalahan yang diukur, yang dinyatakan dalam satuan persen. Semakin kecil persentase kesalahan, semakin akurat prediksinya. Keputusan mengenai akurasi jika Relative Error $\leq 25\%$ dianggap akurat, sedangkan nilai $> 25\%$ tetapi $\leq 50\%$ perlu dipertimbangkan jika ingin digunakan, dan estimasi dengan nilai $> 50\%$ dianggap tidak akurat [7]. diharapkan dapat memberikan prediksi penjualan dan memandu keputusan terbaik dalam menentukan stok roti.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Yubi aqsho ramadhan pada penelitian dengan judul "Prediksi Penjualan Handphone Di Toko X Menggunakan Algoritma Regresi Linear" menetapkan bahwa Tujuan utama adalah meningkatkan ketepatan prediksi penjualan roti di toko-toko, membantu dalam perencanaan persediaan, dan mengoptimalkan strategi pemasaran. Dengan demikian, penelitian ini tidak hanya bertujuan untuk mengisi kekosongan pengetahuan tentang prediksi penjualan roti, tetapi juga berusaha untuk memberikan kontribusi yang signifikan pada pengembangan metode analisis yang dapat digunakan dalam memprediksi penjualan roti menggunakan algoritma regresi linear [8].

Penelitian ini diharapkan bisa menjadi panduan praktis untuk membuat keputusan yang lebih terinformasi tentang manajemen persediaan dan strategi pemasaran. Selain itu, dengan meningkatkan efisiensi penjualan roti, penelitian ini bisa meningkatkan kesejahteraan ekonomi di tingkat lokal, memperkuat ketahanan ekonomi lokal, dan memberikan kontribusi pada literatur informatika

terkait analisis prediksi penjualan di sektor ekonomi mikro [9].

Dalam pengujian fungsional, tujuannya adalah memastikan bahwa semua fitur dari program klasifikasi dapat diimplementasikan sepenuhnya sesuai harapan.

2.2. Metode Regresi

Sistem pendukung keputusan (SPK) merupakan sebuah sistem yang bekerja berbasis komputer dimana sistem ini dapat membantu mengambil keputusan atau mengambil kebijakan [9]. Hal ini karena dalam SPK akan diberikan rekomendasi berdasarkan peringkat nilai yang telah dihitung berdasarkan pada pembobotan masing-masing variabel yang dihitung. Intinya, Sistem Pendukung Keputusan (SPK) diciptakan untuk membantu dalam seluruh proses pengambilan keputusan, mulai dari mendefinisikan masalah, memilih data yang relevan, hingga mengevaluasi pilihan alternatif. SPK dapat membantu mengorganisir informasi, menganalisis data, dan memberikan rekomendasi untuk mendukung proses pengambilan keputusan yang lebih baik dan efektif. Regresi Linier adalah teknik statistik yang digunakan untuk membangun model yang menghubungkan variabel dependen dengan satu atau lebih variabel independen. Jika hanya ada satu variabel independen, itu disebut sebagai regresi linier sederhana, sedangkan jika ada lebih dari satu variabel independen, itu disebut sebagai regresi linier berganda [10].

2.3. RMSE (Root Mean Square Error)

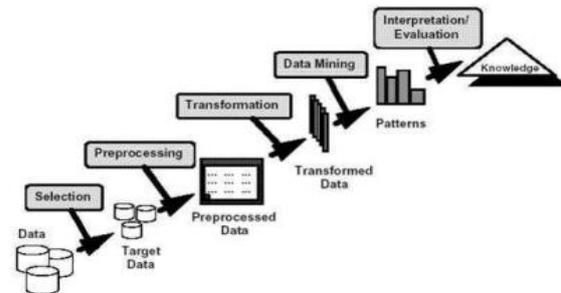
Root Mean Square Error (RMSE) adalah suatu metrik evaluasi yang digunakan untuk mengukur seberapa baik model regresi memprediksi nilai-nilai observasi yang sebenarnya. RMSE mengukur seberapa besar deviasi antara nilai-nilai yang diprediksi oleh model dengan nilai-nilai yang diamati dalam data yang sebenarnya. Secara matematis, RMSE dihitung sebagai akar dari rata-rata dari kuadrat dari selisih antara nilai-nilai prediksi dan nilai-nilai observasi. RMSE merupakan salah satu metrik evaluasi yang umum digunakan dalam analisis regresi untuk mengevaluasi kinerja model dan membantu dalam membandingkan kualitas berbagai model regresi. Semakin kecil nilai RMSE, semakin baik model regresi dalam melakukan prediksi.

3. METODE PENELITIAN

Penelitian ini menerapkan jenis penelitian deskriptif menggunakan pendekatan kuantitatif. Deskriptif kuantitatif berfungsi untuk menentukan jumlah setiap variabel dimana satu atau lebih variabel bebas, tanpa membentuk suatu ikatan atau perbedaan dengan variabel yang lain dengan tujuan memberikan penjelasan yang objektif tentang suatu kondisi [10]. Dengan menggunakan Regresi Linear. Regresi Linear merupakan metode prediksi yang mempelajari ikatan antara variabel dependent dan variabel independent.

Algoritma regresi linear berfungsi untuk mengukur kekuatan hubungan dua variabel, dan menunjukkan alur hubungan antara variabel dependent dengan variabel independent [11].

KDD (Knowledge Discovery in Database) digunakan dalam menganalisis data pada penelitian ini. Gambar tahapan KDD bisa dilihat pada Gambar 1 berikut.



Gambar 1. Tahapan KDD

3.1. KDD (Knowledge Discovery in Database)

Proses mencari hubungan antar variabel menggunakan data mining dan Knowledge Discovery in Databases (KDD) untuk menggali data tersembunyi dari informasi besar. Data mining adalah tahapan dalam KDD, yang juga terdiri dari Selection, Preprocessing, Transformation, Evaluation, dan Knowledge. KDD dan data mining berbeda konsep namun terkait [12]. Pada penelitian ini dataset tentang penjualan roti sebanyak 12989 data dari tahun 2022-2023 menggunakan 1 atribut tanggal dan 35 atribut merek roti dan 1 atribut jumlah penjualan.

3.2. Data Selection

Sebelum memasuki tahap pencarian informasi dalam proses KDD, langkah penting yang harus dilakukan adalah memilih data yang relevan dari berbagai data operasional yang tersedia. Tahap ini dikenal sebagai Data Selection [13]. Pada proses ini dilakukan proses penggalian informasi. Atribut yang dipilih dari dataset adalah tahun, periode, kategori Entry dan Mid data penjualan 2022-2023. Hasil Data Selection dapat dilihat pada Tabel 1 berikut.

Tabel 1. Hasil Data selection

Tahun	Periode	Entry	Mid
2022	Januari	460	38
	Februari	480	47
	Maret	530	45
	April	570	65
	Mei	460	76
	Juni	431	43
	Juli	462	98
	Agustus	623	62
	September	674	74
	Oktober	695	69
	November	596	59
	Desember	637	63
2023	Januari	618	61
	Februari	599	59
	Maret	550	61

Tahun	Periode	Entry	Mid
	April	461	54
	Mei	562	65
	Juni	433	75
	Juli	484	77
	Agustus	695	57
	September	665	56
	Oktober	672	61

3.3. Preprocessing

Proses preprocessing mencakup beberapa langkah, antara lain menghilangkan data yang tidak relevan atau duplikat, memeriksa data yang tidak konsisten, dan memperbaiki kesalahan dalam data seperti kesalahan penulisan atau ketik[14]. Pada tahap ini tidak dilakukan preprocessing data, karena tidak ada kesalahan pada data.

3.4. Transformation

Transformation bertujuan untuk mengurangi kerumitan data, memfasilitasi pengubahan data, serta mencegah salah satu variabel bebas mendominasi variabel bebas lainnya yang dapat berdampak pada hasil prediksi. Pada atribut periode ubah menjadi atribut nomer, untuk membantu memudahkan koordinasi data yang akan diproses. Hasil Transformation dapat dilihat pada Tabel 2 berikut.

Tabel 2. Hasil Transformasi

No	Entry	Mid
1	460	38
2	480	47
3	530	45
4	570	65
5	460	76
6	431	43
7	462	98
8	623	62
9	674	74
10	695	69
11	596	59
12	637	63
13	618	61
14	599	59
15	550	61
16	461	54
17	562	65
18	433	75
19	484	77
20	695	57
21	665	56
22	672	61

3.5. Data mining

Tahap kelima dalam proses KDD adalah langkah di mana tugas Data mining yang sesuai dipilih sesuai dengan tujuan yang telah ditetapkan pada tahap awal untuk menggali informasi dari data. Pada proses ini menggunakan aplikasi Rapidminer, dilakukan proses pemodelan pada setiap kategori Entry dan Mid untuk menentukan nilai persamaan algoritma regresi linear, setelah didapat kan hasil persamaan $Y = a+b(X)$ dari

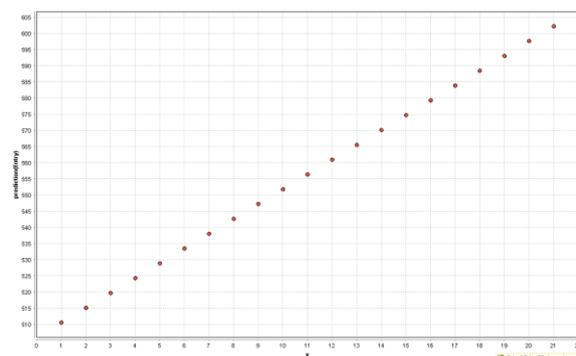
setiap kategori, selanjutnya dilakukan proses prediksi penjualan 3 bulan berikutnya dari setiap kategori, setelah mendapatkan hasil prediksi dilakukan perhitungan evaluasi dari hasil prediksi dengan menggunakan RMSE dan Relative Error untuk mengetahui akurasi hasil prediksi.

3.6. Evaluation

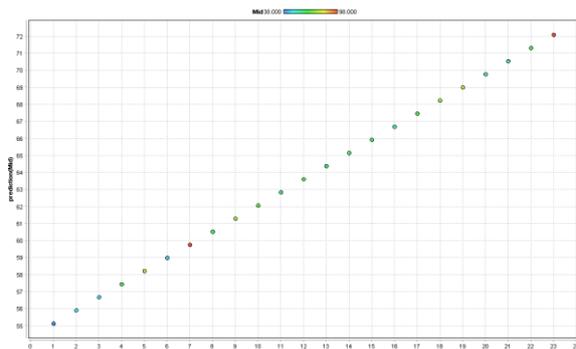
Diperlukan penyajian pola informasi hasil dari data mining dalam format yang dapat dipahami dengan mudah oleh para pihak yang berkepentingan. Evaluation merupakan tahap penting dalam proses KDD yang melibatkan pemeriksaan apakah pola atau informasi yang ditemukan sesuai dengan fakta atau hipotesis yang telah ada sebelumnya.

3.7. Knowledge

Berdasarkan hasil validitas prediksi penjualan roti pada masing-masing kategori ditampilkan plot Scatter Bubble hasil prediksi algoritma regresi linear, diketahui hasil prediksi penjualan roti memiliki nilai yang konstan. Hasil Plot Scatter dapat dilihat pada Gambar 2 dan 3 berikut.



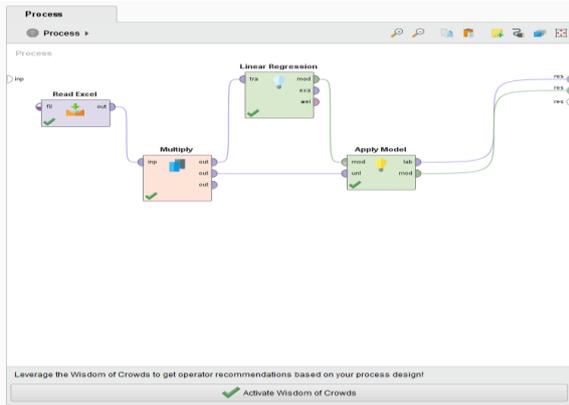
Gambar 2. Plot Scatter Entry



Gambar 3. Plot Scatter Mid

4. HASIL DAN PEMBAHASAN

Pengujian didapatkan beberapa model data mining Regresi Linear menggunakan tools Rapidminer dalam menentukan persamaan regresi linear, memprediksi penjualan periode berikutnya dan menentukan hasil evaluasi dari hasil prediksi menggunakan RMSE. Desain menentukan persamaan dapat dilihat pada Gambar 4.



Gambar 4. Dsain Menentukan persamaan

Berdasarkan proses perhitungan persamaan nilai a dan b didapatkan kategori Entry $Y = 500.08 + 5.4(X)$ dan kategori Mid $Y = 54.36 + 0.8(X)$. Hasil pemodelan dapat dilihat pada Gambar 5 dan 6 berikut.

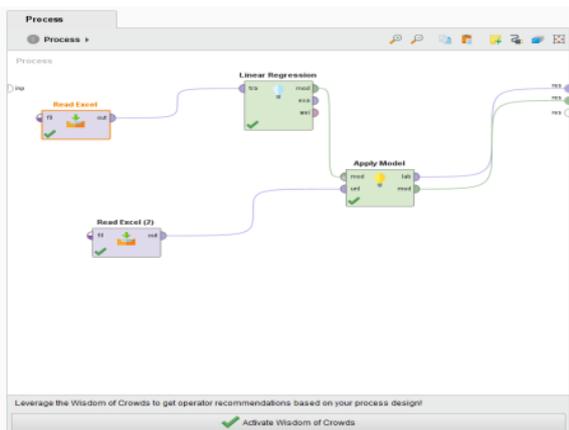
Attribute	Coefficient
Y	5.557
(Intercept)	500.078

Gambar 5. Hasil Pemodelan Entry

Attribute	Coefficient
X	0.771
(Intercept)	54.360

Gambar 6. Hasil Pemodelan Mid

Setelah persamaan didapatkan dilakukan prediksi penjualan roti pada 3 bulan dengan nilai $X = 12, 13, 14$. Desain prediksi data penjualan dapat dilihat pada gambar 3 berikut.



Gambar 7. Desain Prediksi data penjualan

Didapat kan hasil pada bulan pertama terjual 564 unit, pada bulan kedua 569 unit dan bulan ketiga 575 unit, sedangkan kategori Mid pada bulan pertama terjual 63 unit, pada bulan kedua 64 unit dan bulan ketiga 65 unit. Hasil prediksi dapat dilihat pada Gambar 8 dan 9 berikut.

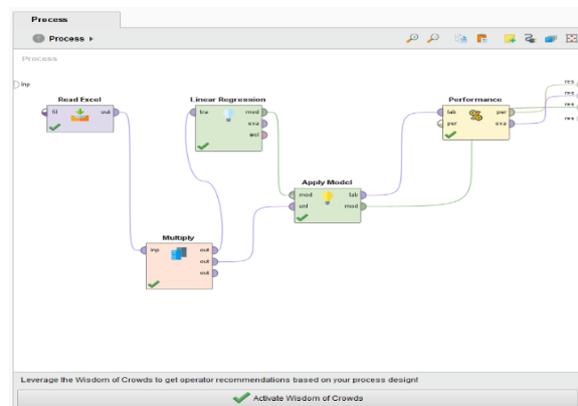
Row No.	y	prediction(P...	No	x
1	?	490.521	1	12
2	?	480.667	2	13
3	?	470.814	3	14

Gambar 8. Hasil Prediksi Entry

Row No.	y	prediction(M...	no	x
1	?	63.609	1	12
2	?	64.379	2	13
3	?	65.150	3	14

Gambar 9. Hasil Prediksi Mid

Setelah hasil prediksi didapatkan dilakukan evaluasi menggunakan 2 indikator penilaian RMSE dan Relative Error. Desain Evaluasi RMSE dan Relative Error dapat dilihat pada Gambar 10 berikut.



Gambar 10. Desain Evaluasi RMSE

Didapatkan nilai RMSE kategori Entry 83.318 dan Relative Error 13.25% sedangkan pada kategori Mid nilai RMSE 13.721 dan Relative Error 17.42%. Hasil Evaluasi RMSE dan Relative Error dapat dilihat pada Gambar 11 dan 12 berikut.

root_mean_squared_error	relative_error
root_mean_squared_error: 82.318 +/- 0.000	relative_error: 13.25% +/- 8.67%

Gambar 11. Hasil Evaluasi RMSE dan Relative error Entry

root_mean_squared_error	relative_error
root_mean_squared_error: 13.721 +/- 0.000	relative_error: 17.42% +/- 11.86%

Gambar 12. Hasil Evaluasi RMSE dan Relative error Mid

5. KESIMPULAN DAN SARAN

Pada penelitian ini diusulkan algoritma regresi linear untuk memprediksi penjualan roti di toko toko, setelah melewati proses tahapan KDD (Knowledge Discovery in Data Base) dan data mining didapatkan hasil dari prediksi penjualan 3 bulan berikutnya,

dengan nilai evaluasi yang bisa disimpulkan cukup atau bisa digunakan hasil dari persamaan regresi linear, karena hasil evaluasi RMSE dan Relative Error cukup mendekati nilai 0.

DAFTAR PUSTAKA

- [1] B. Tahun, M. J. Zaqy, R. Mentari, M. Iqbal, M. T. Informasi, and U. P. Pancabudi, "Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD Implementasi Data Mining Untuk Memprediksi Penjualan Toko Roti Mawar Menggunakan Regresi Linier Berganda Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD," vol. 6, pp. 385–392, 2023.
- [2] S. Supardi, A. K. Ajie, A. Dwiyantri, and J. Ramiaji, "Peran Data Mining dalam Memprediksi Tingkat Penjualan Sepatu Adidas Menggunakan Metode Algoritma Regresi Linear Sederhana," vol. 4, no. 5, pp. 883–890, 2023.
- [3] M. Penjualan and H. Di, *PENERAPAN ALGORITMA REGRESI LINEAR DALAM MEMPREDIKSI PENJUALAN HANDPHONE DI TOKO X (Studi Kasus: Toko Handphone X) APPLICATION OF LINEAR REGRESSION ALGORITHM IN PREDICTING HANDPHONE SALES AT STORE X (Case Study: X Handphone Store)*. 2022.
- [4] A. Citra, A. Bima, P. Susanti, and F. Nugrahanti, "Analisis Regresi Linear Untuk Memprediksi Kriteria Pengaruh Penjualan Mobil Honda Civic dan Toyota Camry pada Kota Surabaya," *Bull. Inf. Technol.*, vol. 3, no. 4, pp. 217–230, 2022.
- [5] N. B. Pamungkas, "PREDIKSI JUMLAH PENERIMAAN SISWA BARU PADA SMAN 1 PAGAR DEWA DENGAN METODE REGRESI LINIER (STUDI KASUS : SMAN 1 PAGAR DEWA)," vol. 2, no. 1, pp. 8–14, 2021.
- [6] D. S. Purnia, H. Muhajir, M. F. Adiwisastro, and D. Supriadi, "Pengukuran Kesenjangan Digital Menggunakan Metode Deskriptif Berbasis Website," vol. 8, no. 2, 2020.
- [7] M. Yasin, "Perbandingan Penerapan Estimasi Metode Regresi Linier Menggunakan Rapidminer dan Ms. Excel studi Kasus Penjualan Pada Public Dataset," *J. Lentera Ict*, vol. 8, no. 1, pp. 17–29, 2023.
- [8] R. Jumardi and S. H. Widiastuti, "Aplikasi Forecasting Penjualan Dan Persediaan Produk Usaha Mikro, Kecil Dan Menengah," *JSAI (Journal Sci. Appl. Informatics)*, vol. 4, no. 3, pp. 383–390, 2021, doi: 10.36085/jsai.v4i3.2756.
- [9] Y. A. E. Tuah and A. Anyan, "Implementasi Model Regresi Linear Sederhana Untuk Prediksi Gaji Berdasarkan Pengalaman Lama Bekerja," *JUTECH J. Educ. Technol.*, vol. 1, no. 2, pp. 56–70, 2020, doi: 10.31932/jutech.v1i2.1289.
- [10] I. Wahyuni, N. Nafi'iyah, and Masrurroh, "Sistem Peramalan Penjualan Perumahan di Kabupaten Lamongan dengan Menggunakan Metode Regresi Linier Berganda," *Semin. Nas. Sist. Inf.* 2019, no. September, pp. 1969–1973, 2019.
- [11] M. Sholeh, E. K. Nurnawati, and U. Lestari, "Penerapan Data Mining dengan Metode Regresi Linear untuk Memprediksi Data Nilai Hasil Ujian Menggunakan RapidMiner," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 8, no. 1, pp. 10–21, 2023, doi: 10.14421/jiska.2023.8.1.10-21.
- [12] 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," *J. Nas. Teknol. dan Sist. Inf.*, vol. 7, no. 1, pp. 8–17, 2021, doi: 10.25077/teknosi.v7i1.2021.8-17.
- [13] Syahril, M., Kusnasari, S., Muhazir, A., & Syahputri, A. (2023). Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD Implementasi Data Mining Untuk Rekomendasi Jurusan Menggunakan Algoritma K-Means Clustering Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD. *Teknologi Sistem Informasi Dan Sistem Komputer TGD*, 6, 235–245.
- [14] Syakir, Y., Iman Hermanto, T., Ramadhan, Y. R., Studi, P., Informatika, T., Teknologi, S. T., & Purwakarta, W. (2022). Analisis Marketplace Shopee Untuk Memprediksi Penjualan dengan Algoritma Regresi Linier. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 6(2), 904–915.