

Analisis dan Perancangan Sistem Informasi Prediksi Data Penjualan pada *E-commerce* Khumaira Cookies & Snack

Nayla Andina Oktavia¹, Tri Wahyu Widyaningsih²

^{1,2} Program Studi Teknik Informatika, Fakultas Teknik dan Teknologi, Universitas Tanri Abeng
nayla.andina@student.tau.ac.id¹, tri.widyaningsih@tau.ac.id²

Diterima : 30 Agustus 2025

Disetujui : 20 September 2025

Abstract—Perkembangan teknologi digital telah mendorong perubahan signifikan dalam dunia bisnis, termasuk cara konsumen berbelanja secara online. Penelitian ini bertujuan mengimplementasikan metode data mining menggunakan pendekatan regresi linear sederhana dalam memprediksi penjualan pada e-commerce Khumaira Cookies & Snack. Permasalahan yang diangkat meliputi sistem penjualan yang masih dilakukan secara manual serta kurangnya pemanfaatan teknologi dalam proses prediksi permintaan pasar. Metode penelitian menggunakan pendekatan Research and Development (R&D) dengan tahapan perencanaan, analisis data, perancangan sistem e-commerce, pengembangan fitur prediksi penjualan, hingga evaluasi sistem. Kebaruan penelitian ini terletak pada integrasi sistem e-commerce dengan fitur prediksi penjualan berbasis data mining yang ditujukan khusus bagi UMKM, yang masih jarang dikaji. Hasil dari analisis regresi menunjukkan bahwa metode ini mampu memberikan gambaran tren penjualan lima tahun ke depan dan membantu bisnis dalam menyusun strategi produksi serta pemasaran yang lebih tepat sasaran. Sistem e-commerce yang dibangun juga berhasil meningkatkan efisiensi transaksi dan pengalaman pengguna. Keterbatasan penelitian ini adalah penggunaan satu variabel independen, sehingga faktor eksternal belum diperhitungkan. Dengan demikian, integrasi sistem digital berbasis prediksi dapat meningkatkan daya saing UMKM dalam era transformasi digital.

Keywords —Data Mining, Regresi Linear, Prediksi Penjualan, *E-Commerce*, R&D, Pengembangan Sistem Informasi.

I. PENDAHULUAN

Perkembangan teknologi informasi telah membawa dampak besar terhadap berbagai aspek kehidupan, salah satunya adalah perubahan dalam pola konsumsi masyarakat dalam berbelanja. Akses internet yang semakin luas dan beragamnya perangkat digital telah mendorong pergeseran perilaku konsumen dari belanja konvensional ke platform e-commerce yang menawarkan kemudahan, kecepatan, dan efisiensi dalam transaksi.

Khumaira Cookies & Snack merupakan usaha mikro yang bergerak di bidang penjualan makanan ringan. Saat ini, sistem operasional yang digunakan masih bersifat manual,

mencakup pencatatan pesanan, stok barang, transaksi, serta pelaporan yang dilakukan secara tertulis dan diolah kembali menggunakan Excel. Ketergantungan pada interaksi manual melalui media sosial dalam proses pemesanan sering kali menyulitkan konsumen dalam memperoleh informasi penting seperti ketersediaan stok dan total pembayaran.

Untuk menjawab tantangan tersebut, dibutuhkan sistem e-commerce yang mampu mengotomatisasi proses bisnis dan menyajikan informasi secara real-time. Sistem ini diharapkan dapat meningkatkan efisiensi pelayanan sekaligus mendukung pengelolaan bisnis secara terstruktur [1]. Selain itu, analisis terhadap data historis

penjualan dapat memberikan kontribusi yang signifikan dalam proses pengambilan keputusan. Salah satu metode yang tepat untuk analisis prediktif adalah regresi linier, yang mampu memodelkan hubungan antara variabel waktu dan penjualan untuk menghasilkan estimasi nilai penjualan di masa depan secara numerik.

Berbagai penelitian telah membuktikan bahwa penerapan teknik data mining dapat meningkatkan efisiensi operasional dan mendukung pengambilan keputusan strategis. Apriana and Yuliansyah [2], mengindikasikan bahwa penggunaan data historis dan analisis perilaku pelanggan dapat mengoptimalkan strategi pemasaran dan berkontribusi pada peningkatan konversi penjualan, meskipun belum dibahas secara rinci penerapannya dalam skala UMKM. Pendekatan clustering digunakan oleh Nugraha et al. [3], metode K-Means diterapkan untuk mengelompokkan produk berdasarkan tingkat penjualannya, dengan hasil yang mendukung pengembangan strategi pembelian ulang. Namun masih belum mempertimbangkan faktor eksternal seperti tren musiman dan kondisi pasar. Sementara itu, Asih dan Martanto [4], menerapkan algoritma FP-Growth dalam mengidentifikasi pola pembelian produk dengan tingkat kepercayaan yang tinggi, namun belum dijelaskan bagaimana pola tersebut digunakan dalam praktik pemasaran.

Beberapa penelitian lainnya menggunakan regresi linier sebagai metode prediktif. Inawati et al. [5], membuktikan bahwa regresi linier mampu memberikan hasil prediksi penjualan album K-Pop yang akurat berdasarkan nilai error yang rendah. Halim [6], menggabungkan dashboard monitoring penjualan dengan regresi linier untuk mendukung pengambilan keputusan bisnis, meskipun belum mengulas strategi jangka panjang. Syahril et al. [7], menggunakan algoritma Apriori untuk mengidentifikasi pola penjualan peralatan sekolah, namun penelitian ini juga belum menjelaskan penerapannya dalam strategi perusahaan secara menyeluruh. Athallah dan Rozi [8], mengindikasikan regresi linear cocok digunakan dalam memprediksikan penjualan dengan tingkat kesalahan yang kecil, namun keterbatasan pada data historis menjadi

tantangan dalam pengembangan model. Jaya et al. [9], menerapkan regresi linier berganda dalam memprediksi produksi berdasarkan data penjualan dan jumlah pemesanan, dan menunjukkan efektivitas metode ini dalam mendukung perencanaan produksi.

Di sisi lain, metode klasifikasi seperti Naive Bayes juga banyak digunakan. Hutahaean dan Handoko [10], menerapkan algoritma ini untuk memprediksi penjualan obat di klinik dan membantu pengelolaan stok. Harahap et al. [11], memanfaatkannya dalam memprediksi produk AC terlaris, menghasilkan tingkat akurasi 75% dengan precision dan recall yang cukup tinggi. Namun, kedua penelitian tersebut merekomendasikan penambahan variabel dan data yang lebih besar untuk meningkatkan akurasi model.

Penelitian ini dilatarbelakangi oleh tantangan UMKM yang masih banyak mengandalkan proses manual dalam operasional bisnisnya. Khumaira Cookies & Snack merupakan salah satu contoh UMKM yang menghadapi masalah pada pencatatan, stok, dan transaksi. Beberapa penelitian sebelumnya telah menggunakan data mining untuk mendukung pengambilan keputusan penjualan, namun penerapan spesifik pada konteks UMKM yang terintegrasi langsung dengan sistem e-commerce masih sangat terbatas. Hal ini menjadi gap penelitian yang ingin dijawab melalui pengembangan sistem prediksi penjualan berbasis regresi linear sederhana dalam e-commerce.

Kebaruan penelitian ini bukan pada metode regresi linear itu sendiri, melainkan pada implementasi sistem informasi prediksi penjualan yang terintegrasi langsung dengan platform e-commerce UMKM. Integrasi ini dinilai penting karena dapat menggabungkan otomatisasi transaksi dengan kemampuan prediktif yang aplikatif bagi pemilik usaha kecil.

II. TINJAUAN PUSTAKA

A. E-Commerce

E-commerce adalah sistem perdagangan yang dilakukan melalui jaringan elektronik, terutama internet. Sistem e-commerce modern tidak hanya

menyediakan informasi produk dan memfasilitasi transaksi, tetapi juga mengintegrasikan fitur-fitur seperti manajemen stok, pelaporan penjualan, dan analisis data. Dengan demikian, e-commerce dapat menjadi solusi strategis bagi pelaku usaha skala kecil dan menengah (UMKM) dalam meningkatkan efisiensi bisnis mereka [12].

B. Data Mining

Data mining adalah proses untuk mengidentifikasi pola dan informasi penting dari kumpulan data yang besar. Tujuan utamanya adalah untuk mengungkapkan pengetahuan tersembunyi yang dapat memberikan manfaat dalam pengambilan keputusan. Dalam konteks bisnis, data mining sering digunakan untuk analisis perilaku konsumen, deteksi tren penjualan, dan peramalan (forecasting). Salah satu metode yang banyak digunakan dalam data mining untuk tujuan prediksi adalah regresi [13].

C. Regresi Linear

Dalam penelitian ini, regresi linear sederhana diterapkan untuk menganalisis hubungan antara variabel waktu (tahun) sebagai variabel independen (X) dan total pendapatan penjualan sebagai variabel dependen (Y). Metode ini bertujuan untuk mengukur pengaruh waktu terhadap pertumbuhan pendapatan serta menghasilkan prediksi penjualan lima tahun ke depan (2025–2029). Pendekatan ini dinilai efektif karena kesederhanaannya serta kemampuannya dalam memberikan estimasi kuantitatif yang dapat mendukung pengambilan keputusan berbasis data [14].

D. R-Squared

Koefisien determinasi (R^2) merupakan ukuran yang menggambarkan sejauh mana variasi dalam variabel dependen (Y) dapat dijelaskan oleh variabel independen (X) dalam model regresi. Nilainya berada antara 0 hingga 1; semakin mendekati 1, semakin efektif model dalam menjelaskan data [15].

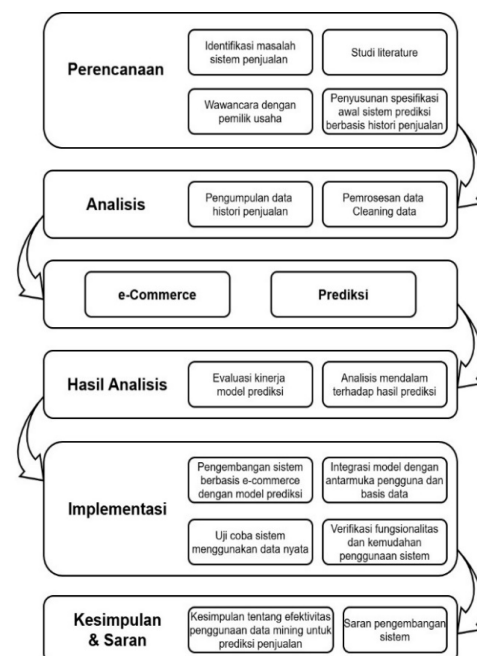
Beberapa penelitian sebelumnya telah membahas data mining untuk prediksi penjualan. Nugraha et al. menggunakan K-Means untuk pengelompokan produk, namun belum

mempertimbangkan tren musiman. Asih dan Martanto menerapkan FP-Growth untuk pola pembelian, tetapi belum dikaitkan dengan strategi penjualan jangka panjang. Penelitian Inawati et al. dan Halim membuktikan regresi linear akurat untuk prediksi penjualan, namun penerapannya lebih pada kasus menengah hingga besar, bukan pada UMKM.

Dari analisis tersebut terlihat bahwa masih terdapat kekosongan penelitian (research gap) dalam penerapan regresi linear sederhana untuk UMKM dengan integrasi penuh pada sistem e-commerce. Penelitian ini berusaha menutup gap tersebut dengan menawarkan sistem prediksi penjualan yang praktis, terjangkau, dan sesuai konteks bisnis kecil.

III. METODE PENELITIAN

A. Tahapan Penelitian



Gambar 1. Tahapan Penelitian

Tahapan penelitian ini disusun secara sistematis guna mendukung implementasi metode data mining dalam memprediksi data penjualan pada sistem e-commerce Khumaira Cookies & Snack. Adapun tahapan yang dilakukan meliputi perencanaan, analisis, pengembangan e-commerce & prediksi, hasil analisis, implementasi, serta kesimpulan dan saran. Berikut adalah penjelasan dari tahapan-tahapan dalam penelitian:

A.1 Perencanaan

Pada tahap ini dilakukan identifikasi awal terhadap permasalahan sistem penjualan yang masih dilakukan secara manual. Proses ini dilengkapi dengan studi literatur dan wawancara langsung dengan pemilik usaha untuk memahami kebutuhan sistem dan menyusun spesifikasi awal sistem prediksi berdasarkan data historis penjualan.

A.2 Analisis

Setelah kebutuhan sistem teridentifikasi, dilakukan analisis data untuk memastikan kualitas dan kelayakannya dalam proses pemodelan. Tahap ini mencakup pengumpulan, pemetaan struktur, serta pembersihan data (data cleaning) guna menghilangkan nilai kosong, duplikat, dan ketidaksesuaian. Data penjualan tahunan Khumaira Cookies & Snack dari tahun 2020 hingga 2024 diperiksa secara visual dan statistik untuk mendeteksi potensi anomali. Hasil observasi menunjukkan tidak terdapat missing value maupun outlier yang signifikan, sehingga seluruh data digunakan secara utuh tanpa proses imputasi, trimming, atau winsorizing, karena dianggap merepresentasikan kondisi penjualan yang aktual.

Tahun 2020									
No.	Hari/Tanggal	Nama Customer	Nama Barang	Jumlah Barang	Total Jumlah Barang	Harga Barang	Total Harga Barang	Jumlah Total	
1.	Sabtu 16 Mei 2020	Rynn	Nutella	27	168	Rp. 25.000	Rp. 675.000	Rp. 3.495.000	
			Katering?	23		Rp. 20.000	Rp. 460.000		
			Samper?	22		Rp. 20.000	Rp. 440.000		
			Puri Salju	27		Rp. 20.000	Rp. 540.000		
			Coklat Oles	22		Rp. 20.000	Rp. 440.000		
			Susu Kacang	22		Rp. 20.000	Rp. 440.000		
			Cheese Biskuit	13		Rp. 20.000	Rp. 260.000		
			Triple Cheese Almond	18		Rp. 20.000	Rp. 360.000		
			Nutella	14		Rp. 25.000	Rp. 350.000		
			Katering?	12		Rp. 20.000	Rp. 240.000		
2.	Sabtu 16 Mei 2020	Yenni	Samper?	12	84	Rp. 20.000	Rp. 240.000	Rp. 1.780.000	
			Puri Salju	17		Rp. 20.000	Rp. 340.000		
			Coklat Oles	12		Rp. 20.000	Rp. 240.000		
			Susu Kacang	12		Rp. 20.000	Rp. 240.000		
			Triple Cheese Almond	3		Rp. 20.000	Rp. 60.000		
			Nutella	25		Rp. 25.000	Rp. 625.000		
			Katering?	19		Rp. 20.000	Rp. 380.000		
			Samper?	16		Rp. 20.000	Rp. 320.000		
			Puri Salju	16		Rp. 20.000	Rp. 320.000		
			Coklat Oles	19		Rp. 20.000	Rp. 380.000		
3.	Sabtu 16 Mei 2020	Yohis	Susu Kacang	16	110	Rp. 20.000	Rp. 320.000	Rp. 2.320.000	
			Triple Cheese Almond	3		Rp. 20.000	Rp. 60.000		
			Nutella	3		Rp. 25.000	Rp. 75.000		
			Katering?	2		Rp. 20.000	Rp. 40.000		
			Samper?	2		Rp. 20.000	Rp. 40.000		
			Puri Salju	2		Rp. 20.000	Rp. 40.000		
			Coklat Oles	2		Rp. 20.000	Rp. 40.000		
			Susu Kacang	2		Rp. 20.000	Rp. 40.000		
			Cheese Biskuit	1		Rp. 20.000	Rp. 20.000		
			Nutella	3		Rp. 25.000	Rp. 75.000		
4.	Jumat 13 Mei 2020	Rialta	Katering?	1	21	Rp. 20.000	Rp. 20.000	Rp. 445.000	
			Samper?	1		Rp. 20.000	Rp. 20.000		
			Puri Salju	2		Rp. 20.000	Rp. 40.000		
			Coklat Oles	2		Rp. 20.000	Rp. 40.000		
			Susu Kacang	2		Rp. 20.000	Rp. 40.000		
			Cheese Biskuit	1		Rp. 20.000	Rp. 20.000		
			Nutella	3		Rp. 25.000	Rp. 75.000		
			Katering?	1		Rp. 20.000	Rp. 20.000		
			Samper?	1		Rp. 20.000	Rp. 20.000		
			Puri Salju	1		Rp. 20.000	Rp. 20.000		
5.	Minggu 17 Mei 2020	Feni	Coklat Oles	1	10	Rp. 20.000	Rp. 20.000	Rp. 215.000	
			Susu Kacang	1		Rp. 20.000	Rp. 20.000		
			Nutella	2		Rp. 25.000	Rp. 50.000		
			Katering?	1		Rp. 20.000	Rp. 20.000		
			Samper?	1		Rp. 20.000	Rp. 20.000		
			Puri Salju	1		Rp. 20.000	Rp. 20.000		
			Coklat Oles	1		Rp. 20.000	Rp. 20.000		
			Susu Kacang	1		Rp. 20.000	Rp. 20.000		
			Nutella	2		Rp. 25.000	Rp. 50.000		
			Samper?	3		Rp. 20.000	Rp. 60.000		
6.	Sabtu 8 Mei 2020	Nina	Katering?	3	17	Rp. 25.000	Rp. 75.000	Rp. 355.000	
			Samper?	2		Rp. 20.000	Rp. 40.000		
			Puri Salju	5		Rp. 20.000	Rp. 100.000		
			Coklat Oles	2		Rp. 20.000	Rp. 40.000		
			Susu Kacang	2		Rp. 20.000	Rp. 40.000		
			Cheese Biskuit	1		Rp. 20.000	Rp. 20.000		
			Nutella	10		Rp. 25.000	Rp. 250.000		
			Katering?	7		Rp. 20.000	Rp. 140.000		
			Samper?	2		Rp. 20.000	Rp. 40.000		
			Puri Salju	2		Rp. 20.000	Rp. 40.000		
7.	Minggu 17 Mei 2020	Ida	Cheese Biskuit	7	20	Rp. 20.000	Rp. 140.000	Rp. 450.000	
			Nutella	2		Rp. 25.000	Rp. 50.000		
			Katering?	2		Rp. 20.000	Rp. 40.000		
			Samper?	2		Rp. 20.000	Rp. 40.000		
			Puri Salju	2		Rp. 20.000	Rp. 40.000		
			Coklat Oles	2		Rp. 20.000	Rp. 40.000		
			Susu Kacang	2		Rp. 20.000	Rp. 40.000		
			Cheese Biskuit	3		Rp. 20.000	Rp. 60.000		
			Nutella	2		Rp. 25.000	Rp. 50.000		
			Katering?	1		Rp. 20.000	Rp. 20.000		
8.	Sabtu 9 Mei 2020	Irena	Cheese Biskuit	3	15	Rp. 20.000	Rp. 60.000	Rp. 325.000	
			Nutella	2		Rp. 25.000	Rp. 50.000		
			Katering?	1		Rp. 20.000	Rp. 20.000		
			Samper?	1		Rp. 20.000	Rp. 20.000		
			Puri Salju	2		Rp. 20.000	Rp. 40.000		
			Coklat Oles	2		Rp. 20.000	Rp. 40.000		
			Susu Kacang	2		Rp. 20.000	Rp. 40.000		
			Cheese Biskuit	3		Rp. 20.000	Rp. 60.000		
			Nutella	2		Rp. 25.000	Rp. 50.000		
			Katering?	1		Rp. 20.000	Rp. 20.000		
9.	Sabtu 9 Mei 2020	Biqis	Puri Salju	1	6	Rp. 20.000	Rp. 20.000	Rp. 125.000	
			Cheese Biskuit	1		Rp. 20.000	Rp. 20.000		
			Nutella	1		Rp. 25.000	Rp. 25.000		
			Katering?	1		Rp. 20.000	Rp. 20.000		
			Samper?	1		Rp. 20.000	Rp. 20.000		
			Puri Salju	1		Rp. 20.000	Rp. 20.000		
			Coklat Oles	1		Rp. 20.000	Rp. 20.000		
			Susu Kacang	1		Rp. 20.000	Rp. 20.000		
			Cheese Biskuit	1		Rp. 20.000	Rp. 20.000		
			Nutella	1		Rp. 25.000	Rp. 25.000		
10.	Minggu 17 Mei 2019	Ratih	Cheese Biskuit	1	3	Rp. 20.000	Rp. 20.000	Rp. 65.000	
			Nutella	1		Rp. 25.000	Rp. 25.000		
			Katering?	1		Rp. 20.000	Rp. 20.000		
			Samper?	1		Rp. 20.000	Rp. 20.000		
			Puri Salju	1		Rp. 20.000	Rp. 20.000		
			Coklat Oles	1		Rp. 20.000	Rp. 20.000		
			Susu Kacang	1		Rp. 20.000	Rp. 20.000		
			Cheese Biskuit	1		Rp. 20.000	Rp. 20.000		
			Nutella	1		Rp. 25.000	Rp. 25.000		
			Katering?	1		Rp. 20.000	Rp. 20.000		
11.	Minggu 17 Mei 2019	Egi	Samper?	2	44	Rp. 20.000	Rp. 40.000	Rp. 965.000	
			Puri Salju	2		Rp. 20.000	Rp. 40.000		
			Coklat Oles	2		Rp. 20.000	Rp. 40.000		
			Susu Kacang	2		Rp. 20.000	Rp. 40.000		
			Cheese Biskuit	2		Rp. 20.000	Rp. 40.000		
			Nutella	2		Rp. 25.000	Rp. 50.000		
			Katering?	2		Rp. 20.000	Rp. 40.000		
			Samper?	2		Rp. 20.000	Rp. 40.000		
			Puri Salju	2		Rp. 20.000	Rp. 40.000		
			Coklat Oles	2		Rp. 20.000	Rp. 40.000		
12.	Minggu 17 Mei 2019	Sams	Nutella	3	5	Rp. 25.000	Rp. 75.000	Rp. 125.000	
			Katering?	1		Rp. 20.000	Rp. 20.000		
			Samper?	1		Rp. 20.000	Rp. 20.000		
			Puri Salju	1		Rp. 20.000	Rp. 20.000		
			Coklat Oles	1		Rp. 20.000	Rp. 20.000		
			Susu Kacang	1		Rp. 20.000	Rp. 20.000		
13.	Minggu 17 Mei 2020	Yenni	Nutella	1	1	Rp. 25.000	Rp. 25.000	Rp. 25.000	
			Katering?	1		Rp. 20.000	Rp. 20.000		
			Samper?	1		Rp. 20.000	Rp. 20.000		
			Puri Salju	1		Rp. 20.000	Rp. 20.000		
			Coklat Oles	1		Rp. 20.000	Rp. 20.000		
			Susu Kacang	1		Rp. 20.000	Rp. 20.000		
14.	Minggu 17 Mei 2020	Fahri	Nutella	1	2	Rp. 25.000	Rp. 25.000	Rp. 65.000	
			Katering?	1		Rp. 20.000	Rp. 20.000		
			Samper?	1		Rp. 20.000	Rp. 20.000		
			Puri Salju	1		Rp. 20.000	Rp. 20.000		
			Coklat Oles	1		Rp. 20.000	Rp. 20.000		
			Susu Kacang	1		Rp. 20.000	Rp. 20.000		
15.	Sabtu 16 Mei 2020	Total Seluruh Jumlah Kos	Nutella	3	3	Rp. 25.000	Rp. 75.000	Rp. 10.385.000	
			Katering?	1		Rp. 20.000	Rp. 20.000		
						889	Total Seluruh Harga		

Perhitungan Koefisien:

Untuk mendapatkan nilai a dan b, dapat dihitung menggunakan rumus:

$$a = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{n \sum X^2 - (\sum X)^2} \quad (2)$$

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n \sum X^2 - (\sum X)^2} \quad (3)$$

Di mana:

- n merupakan jumlah data
- $\sum XY$ merupakan jumlah perkalian X dan Y
- $\sum X$ merupakan jumlah seluruh nilai X
- $\sum Y$ merupakan jumlah seluruh nilai Y
- $\sum X^2$ merupakan jumlah kuadrat nilai X

Tabel 1. Data Pendapatan Tahun 2020 - 2024

Tahun	X (Tahun ke)	Y (Pendapatan)
2020	1	Rp. 10.785.000
2021	2	Rp. 12.765.000
2022	3	Rp. 21.015.000
2023	4	Rp. 23.865.000
2024	5	Rp. 34.750.000

Untuk menghitung rumus regresi, perlu menghitung nilai tambahan, X^2 dan XY :

Tabel 2. Nilai X^2 dan XY

X	Y	X^2	XY
1	Rp10.785.000	1	Rp. 10.785.000
2	Rp12.765.000	4	Rp. 25.530.000
3	Rp21.015.000	9	Rp. 63.045.000
4	Rp23.865.000	16	Rp. 95.460.000
5	Rp34.750.000	25	Rp. 173.750.000

Lalu menentukan nilai $\sum X$, $\sum Y$, $\sum XY$, $\sum X^2$:

Tabel 3. Nilai $\sum X$

Tahun	X
2020	1
2021	2
2022	3
2023	4
2024	5

Cara menentukan $\sum X$ dengan menjumlahkan seluruh data X (Tahun).

$$\sum X = 1 + 2 + 3 + 4 + 5 = 15$$

Tabel 4. Nilai $\sum Y$

Tahun	Y (Pendapatan)
2020	Rp. 10.785.000
2021	Rp. 12.765.000
2022	Rp. 21.015.000
2023	Rp. 23.865.000
2024	Rp. 34.750.000

Cara menentukan $\sum Y$ dengan menjumlahkan seluruh data Y (Pendapatan).

$$\sum Y = 10.785.000 + 12.765.000 + 21.015.000 + 23.865.000 + 34.750.000 = 103.180.000$$

Tabel 5. Nilai $\sum XY$

Tahun	X	Y	XY
2020	1	Rp10.785.000	Rp. 10.785.000
2021	2	Rp12.765.000	Rp. 25.530.000
2022	3	Rp21.015.000	Rp. 63.045.000
2023	4	Rp23.865.000	Rp. 95.460.000
2024	5	Rp34.750.000	Rp. 173.750.000

Cara menentukan $\sum XY$ dengan menjumlahkan perkalian seluruh data X dan Y (Pendapatan).

$$\sum XY = 10.785.000 + 25.530.000 + 63.045.000 + 95.460.000 + 173.750.000 = 368.570.000$$

Tabel 6. Nilai $\sum X^2$

X	X^2
1	1
2	4
3	9
4	16
5	25

Cara menentukan $\sum X^2$ dengan menjumlahkan seluruh data X^2 .

$$\sum X^2 = 1 + 4 + 9 + 16 + 25 = 55$$

Setelah mendapatkan seluruh nilai:

$$\sum X = 15$$

$$\sum Y = 103.180.000$$

$$\sum XY = 368.570.000$$

$$\sum X^2 = 55$$

$$n = 5 \text{ (jumlah tahun)}$$

Hitung Koefisiensi Regresi (a dan b)

Rumus a:

$$a = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{n \sum X^2 - (\sum X)^2}$$

$$a = \frac{103.180.000(55) - 15(368.570.000)}{5(55) - 15^2}$$

$$a = \frac{5.674.900.000 - 5.528.550.000}{275 - 225}$$

$$a = \frac{146.350.000}{50}$$

$$a = 2.927.000$$

Rumus b:

$$b = \frac{n(\sum XY) - (\sum X)(\sum Y)}{n\sum X^2 - (\sum X)^2}$$

$$b = \frac{5(368.570.000) - (15)(103.180.000)}{5(55) - (15)^2}$$

$$b = \frac{1.842.850.000 - 1.547.700.000}{275 - 225}$$

$$b = \frac{295.150.000}{50}$$

$$b = 5.903.000$$

Persamaan Regresi Linear:

$$Y = a + b(X)$$

$$Y = 2.927.000 + 5.903.000(X)$$

Prediksi Pendapatan 5 Tahun ke Depan (2025-2029) X = 6 sampai 10:

Tabel 7. Prediksi Pendapatan 5 Tahun Kedepan

Tahun	X	Prediksi Y (Pendapatan)
2025	6	$2.927.000 + 5.903.000(6) = 38.345.000$
2026	7	$2.927.000 + 5.903.000(7) = 44.248.000$
2027	8	$2.927.000 + 5.903.000(8) = 50.151.000$
2028	9	$2.927.000 + 5.903.000(9) = 56.054.000$
2029	10	$2.927.000 + 5.903.000(10) = 61.957.000$

Berdasarkan hasil persamaan regresi yang diperoleh, dilakukan proyeksi pendapatan penjualan Khumaira Cookies & Snack untuk lima tahun mendatang (2025–2029) dengan rentang nilai X = 6 hingga X = 10. Hasil prediksi menunjukkan tren kenaikan yang konsisten dari tahun ke tahun. Model regresi linear dinilai mampu memberikan estimasi tren penjualan yang cukup akurat. Meskipun model ini hanya menggunakan satu variabel independen, yaitu waktu, pendekatan ini tetap relevan dan efektif dalam konteks bisnis berskala kecil.

A.5 Akurasi R-Squared

Model prediksi yang telah dibangun dievaluasi menggunakan metrik koefisien determinasi (R^2) untuk mengetahui tingkat akurasi model. Analisis lanjutan dilakukan terhadap hasil prediksi untuk melihat pola

pertumbuhan dan potensi peningkatan penjualan di masa mendatang.

Rumus perhitungan R^2 :

$$R^2 = 1 - \frac{SSE}{SST} \quad (4)$$

Dengan:

Sum of Squared Error, merupakan jumlah kuadrat dari selisih antara nilai aktual dan nilai yang diprediksi.

$$SSE = \sum (Y_i - \hat{Y}_i)^2 \quad (5)$$

Total Sum of Squares, merupakan jumlah kuadrat dari selisih antara nilai aktual dan rata-rata nilai aktual.

$$SSE = \sum (Y_i - \bar{Y})^2 \quad (6)$$

Dimana:

- Y_i : nilai aktual
- \hat{Y}_i : nilai prediksi
- \bar{Y} : rata-rata Y

A.6 Validasi Model Prediksi

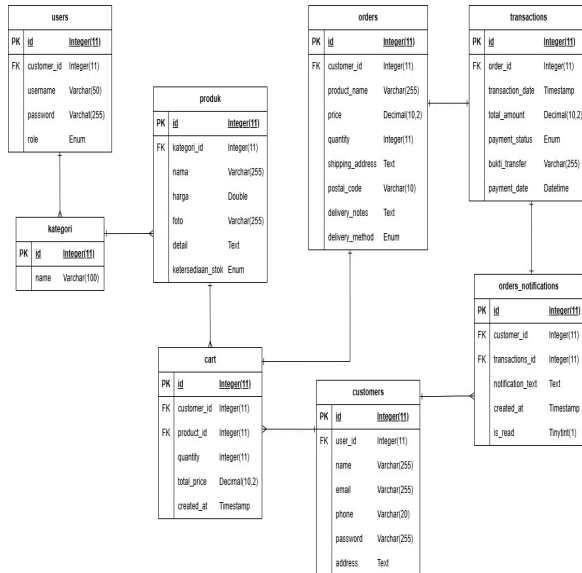
Validasi dilakukan dengan membagi data penjualan tahunan menjadi data latih (2020–2023) dan data uji (2024). Hasil evaluasi menunjukkan nilai R^2 sebesar 0,9449, yang berarti model mampu menjelaskan 94,49% variasi data penjualan. Nilai MAE Rp1.365.000, MSE Rp2.076.450.000.000, dan RMSE sebesar Rp1.441.670 menunjukkan tingkat kesalahan prediksi masih dalam batas wajar. Model juga dinilai tidak mengalami overfitting maupun underfitting berdasarkan hasil prediksi terhadap data uji.

B. Perancangan Sistem

Pada bagian ini, menjelaskan perancangan sistem prediksi penjualan berbasis data mining yang akan diterapkan pada sistem e-commerce Khumaira Cookies & Snack. Perancangan dilakukan berdasarkan hasil analisis kebutuhan serta hasil evaluasi dari tahapan sebelumnya, dengan tujuan untuk membangun sistem yang mampu membantu manajemen dalam mengambil keputusan berdasarkan prediksi data penjualan.

B.1 Entity Relationship Diagram (ERD)

Dalam sistem prediksi penjualan Khumaira Cookies & Snack, terdapat 8 tabel yang dirancang untuk mengelola data yang dibutuhkan dalam proses prediksi penjualan, digambarkan pada gambar berikut:

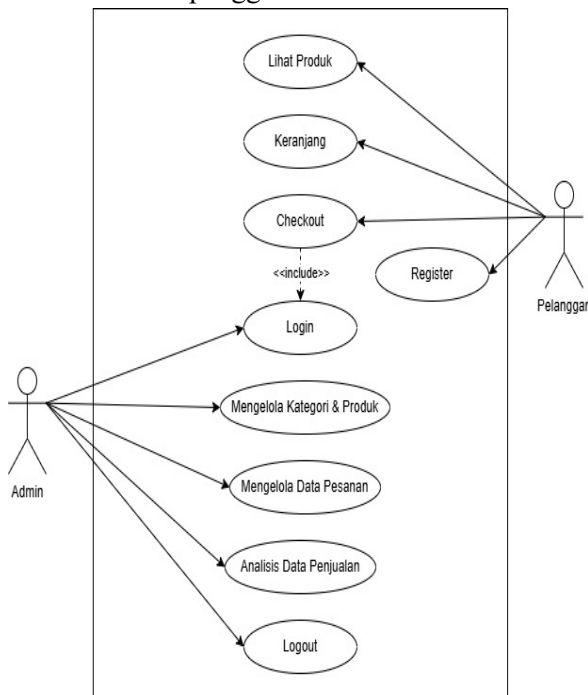


Gambar 3. ERD

Pada sistem prediksi penjualan Khumaira Cookies & Snack, terdapat 8 tabel yang dirancang untuk mengelola data yang dibutuhkan dalam proses prediksi penjualan.

B.2 Use Case Diagram

Use Case Diagram digunakan untuk menggambarkan interaksi antara pengguna dan sistem dalam konteks fungsionalitas yang disediakan. Diagram ini memvisualisasikan aktor yang terlibat, serta use case yang merepresentasikan berbagai fungsi yang dapat dilakukan oleh pengguna.

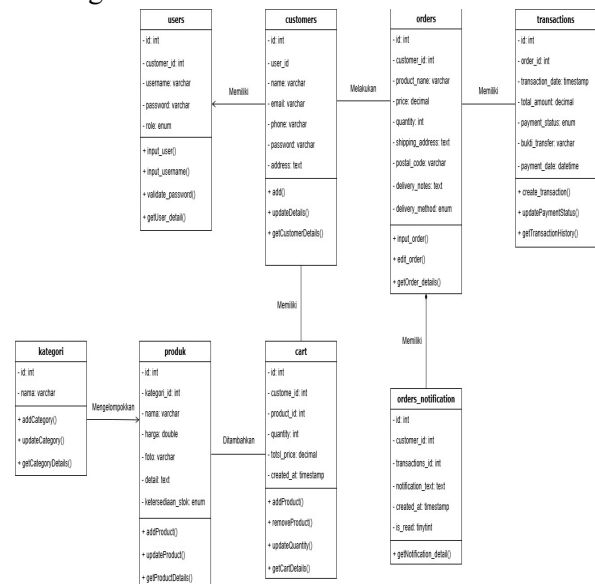


Gambar 4. Use Case Diagram

Dalam sistem manajemen pemesanan Khumaira Cookies & Snack, Use Case Diagram mencakup dua aktor yaitu admin dan Pelanggan, yang berinteraksi dengan sistem untuk melakukan berbagai aktivitas. Diagram ini memberikan gambaran tentang bagaimana pengguna berinteraksi dengan sistem.

B.3 Class Diagram

Class Diagram digunakan untuk menggambarkan struktur dari sistem dengan menunjukkan kelas-kelas yang ada, atribut yang dimiliki oleh masing-masing kelas, serta hubungan antar kelas tersebut.



Gambar 5. Clas Diagram

Dalam sistem e-commerce website Khumaira Cookies & Snack, Class Diagram mencakup beberapa kelas yang masing-masing memiliki atribut dan metode yang relevan.

IV. HASIL DAN PEMBAHASAN

A. Pengembangan Sistem

Metode Rapid Application Development (RAD) digunakan dalam penelitian ini karena mampu mempersingkat proses pengembangan sistem dengan pendekatan iteratif dan keterlibatan langsung pengguna. Metode ini dinilai efektif untuk membangun sistem e-commerce yang dilengkapi fitur prediksi penjualan, karena memungkinkan penyesuaian cepat terhadap kebutuhan pengguna UMKM yang dinamis [16].

Tahapan utama dalam RAD meliputi:

- **Perencanaan Awal:** Identifikasi kebutuhan sistem dan ekspektasi pengguna dilakukan melalui diskusi langsung, agar sistem sesuai tujuan bisnis.
- **Desain Prototipe:** Dibuat rancangan awal antarmuka dan alur sistem untuk divisualisasikan dan dikaji bersama pengguna.
- **Pengembangan:** Sistem dikembangkan secara bertahap dan diuji secara berulang untuk menerima masukan langsung dari pengguna.
- **Implementasi:** Sistem diintegrasikan secara penuh, termasuk modul e-commerce dan prediksi penjualan, dengan uji coba langsung oleh pengguna.
- **Uji Coba dan Validasi:** Seluruh fitur diuji secara fungsional dan hasil prediksi divalidasi menggunakan data penjualan aktual untuk menjamin keakuratan sistem.

Penggunaan RAD memungkinkan sistem yang dibangun menjadi lebih fleksibel, responsif, dan sesuai dengan kebutuhan operasional Khumaira Cookies & Snack.

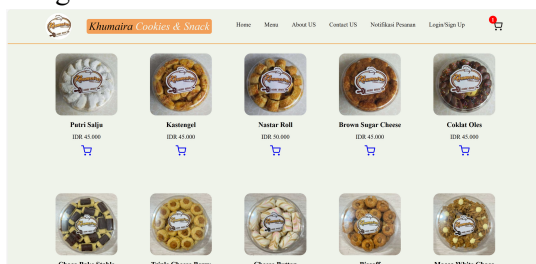
B. Implementasi Sistem E-Commerce

Sistem e-commerce Khumaira Cookies & Snack dikembangkan sebagai solusi digital untuk mendukung proses penjualan dan pengambilan keputusan berbasis data. PHP digunakan sebagai bahasa pemrograman utama dalam pembangunan sistem ini, dengan HTML dan CSS untuk antarmuka pengguna, serta MySQL sebagai basis data.

Fitur-fitur utama sistem meliputi:

1. Halaman Menu

Menampilkan daftar produk yang dijual secara online, lengkap dengan nama produk, harga, dan gambar.



Gambar 6. Halaman Menu Pelanggan

2. Halaman Admin

Digunakan oleh pemilik usaha untuk mengelola data produk, data transaksi, serta

melihat laporan penjualan dan hasil prediksi penjualan.



Gambar 7. Halaman Admin

3. Keranjang dan Checkout

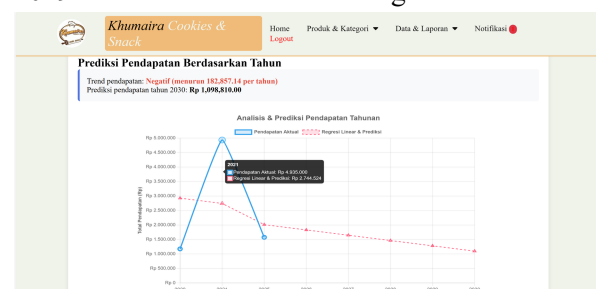
Pengguna dapat menambahkan produk ke keranjang, menentukan jumlah, dan melanjutkan ke proses checkout. Sistem secara otomatis menghitung total harga dan menyimpan data transaksi.



Gambar 8. Halaman Keranjang dan Checkout

4. Dashboard Prediksi Pendapatan

Pada halaman admin, tersedia fitur grafik prediksi pendapatan dari tahun 2025 hingga 2029 berdasarkan hasil model regresi linear.



Gambar 9. Dashboard Prediksi Pendapatan

C. Hasil Prediksi Penjualan

Berdasarkan hasil perhitungan menggunakan model regresi linier sederhana yang telah dibangun, diperoleh prediksi pendapatan penjualan Khumaira Cookies & Snack dalam lima tahun ke depan (2025 hingga 2029). Prediksi ini dihitung menggunakan persamaan regresi yang telah ditentukan sebelumnya, dengan asumsi bahwa tren pertumbuhan penjualan akan mengikuti pola linear yang konsisten seperti tahun-tahun sebelumnya. Tabel

berikut menyajikan hasil prediksi pendapatan penjualan tahunan.

Tabel 8. Hasil Prediksi

Tahun	X	Prediksi Y (Pendapatan)
2025	6	$2.927.000 + 5.903.000(6) = 38.345.000$
2026	7	$2.927.000 + 5.903.000(7) = 44.248.000$
2027	8	$2.927.000 + 5.903.000(8) = 50.151.000$
2028	9	$2.927.000 + 5.903.000(9) = 56.054.000$
2029	10	$2.927.000 + 5.903.000(10) = 61.957.000$

D. Validasi Model Prediksi

Validasi dilakukan dengan membagi data penjualan tahunan menjadi data latih (2020–2023) dan data uji (2024). Hasil evaluasi menunjukkan nilai R^2 sebesar 0,9449, yang berarti model mampu menjelaskan 94,49% variasi data penjualan. Nilai MAE Rp1.365.000, MSE Rp2.076.450.000.000, dan RMSE sebesar Rp1.441.670 menunjukkan tingkat kesalahan prediksi masih dalam batas wajar. Model juga dinilai tidak mengalami overfitting maupun underfitting berdasarkan hasil prediksi terhadap data uji.

E. Pengujian Sistem

Pada penelitian ini, menggunakan metode blackbox sebagai pengujian sistem yang berfokus pada pengujian fungsionalitas tiap fitur berdasarkan respons terhadap input yang diberikan. Metode blackbox digunakan untuk memastikan agar setiap proses, seperti login, checkout, manajemen produk, dan analisis penjualan, berjalan sesuai dengan kebutuhan pengguna dan spesifikasi sistem [17].

Tabel 7. Blackbox Pengujian Sistem

Fungsi yang diuji	Input	Output	Hasil
Login	Username dan Password valid	Pengguna berhasil login	Berhasil
Login	Username dan/atau Password salah	Muncul pesan "Login gagal"	Berhasil
Regist er	Data lengkap dan valid	Akun berhasil dibuat	Berhasil
Regist er	Email sudah terdaftar	Muncul pesan "Email sudah digunakan"	Berhasil

Lihat Produk	Akses halaman katalog produk	Daftar produk tampil	Berhasil
Pilih Produk	Klik tombol "Tambah ke Keranjang"	Produk masuk ke keranjang	Berhasil
Check out	Isi data pengiriman dan pembayaran	Transaksi berhasil dan pesanan tercatat	Berhasil
Check out	Data pengiriman tidak lengkap	Muncul pesan "Data tidak boleh kosong"	Berhasil
Manajemen Produk	Tambah produk dengan data lengkap	Produk berhasil ditambahkan ke sistem	Berhasil
Manajemen Produk	Hapus produk tertentu	Produk berhasil dihapus	Berhasil
Analisis Data Penjualan	Akses halaman analisis	Grafik dan data penjualan tampil	Berhasil
Logout	Klik tombol logout	Pengguna keluar dari sistem	Berhasil

F. Evaluasi Kepuasan Pengguna

Tabel 8. Evaluasi Kepuasan Pengguna

Respon	Jumlah	Skor	Jumlah X Skor
Sangat setuju	133	5	665
Setuju	145	4	580
Netral	40	3	120
Tidak setuju	12	2	24
Sangat tidak setuju	8	1	8
Total			1.397

Pada penelitian ini, jumlah responden (R) yang terlibat adalah sebanyak 30 orang, dan jumlah pertanyaan (S) dalam kuesioner adalah sebanyak 11 butir. Maka, skor maksimal (Q) dihitung sebagai:

$$Q = R \times S \times 5 = 30 \times 11 \times 5 = 1.650$$

Dengan total skor aktual (P) yang diperoleh dari kuesioner adalah sebesar 1.397, maka nilai usability (U) dapat dihitung sebagai:

$$U = P/Q = 1.397/1.650 = 0,85$$

Nilai $U = 0,85$ menunjukkan bahwa sistem memiliki tingkat usability sebesar 85%, nilai tersebut masuk dalam kategori sangat baik atau sangat memuaskan. Hasil ini mengindikasikan bahwa sistem e-commerce yang dibangun telah memberikan pengalaman penggunaan yang efektif, efisien, dan menyenangkan bagi pengguna.

G. Akurasi R-Squared

Akurasi model diukur menggunakan nilai R-Squared (R^2), yang menunjukkan seberapa besar variasi pendapatan penjualan dapat dijelaskan oleh variabel tahun. Berdasarkan perhitungan:

$$SSE = 20.312.470.000.000$$

$$SST = 368.789.920.000.000$$

$$R^2 = 0,9449$$

Nilai R^2 sebesar 0,9449 mengindikasikan bahwa 94,49% variasi data dapat dijelaskan oleh model. Ini menunjukkan bahwa model regresi linear memiliki tingkat akurasi tinggi dan layak digunakan untuk prediksi penjualan dan perencanaan bisnis di masa mendatang.

H. Evaluasi Prediksi dan Akurasi

Evaluasi terhadap hasil prediksi penjualan dilakukan untuk menilai seberapa baik model regresi linier sederhana dalam menggambarkan pola pertumbuhan pendapatan tahunan Khumaira Cookies & Snack. Beberapa metrik evaluasi digunakan dalam analisis ini, yaitu koefisien determinasi (R^2) serta nilai-nilai kesalahan prediksi seperti Mean Absolute Error (MAE), Mean Squared Error (MSE), dan Root Mean Squared Error (RMSE).

Hasil prediksi menunjukkan tren kenaikan konsisten dalam lima tahun ke depan, dengan nilai R^2 sebesar 0,9449 yang menandakan akurasi tinggi. Namun, model ini memiliki keterbatasan karena hanya menggunakan variabel waktu sebagai faktor prediktor. Hal ini membuat model belum mampu menangkap pengaruh faktor eksternal seperti tren musiman, promosi, atau kondisi pasar. Oleh karena itu, meskipun hasilnya valid untuk kondisi data

terbatas, model ini masih terbatas sehingga diperlukan pengembangan lebih lanjut untuk meningkatkan ketahanan dan konsistensi hasil prediksi.

Selain itu, nilai kesalahan prediksi yang diperoleh adalah sebagai berikut:

$$\text{Mean Absolute Error: Rp } 1.365.000$$

$$\text{Mean Squared Error: Rp } 2.076.450.000.000$$

$$\text{Root Mean Squared Error: Rp } 1.441.670$$

Nilai MAE dan RMSE yang relatif rendah menunjukkan bahwa deviasi antara nilai aktual dan nilai prediksi masih berada dalam batas yang wajar untuk skala data penjualan tahunan. Hal ini memperkuat validitas model sebagai alat prediksi yang dapat diandalkan.

Model juga diuji terhadap potensi overfitting dan underfitting. Hasil pengujian menunjukkan bahwa model mampu memprediksi data uji (tahun 2024) dengan akurasi yang cukup tinggi dan nilai yang mendekati realisasi aktual, sehingga dapat disimpulkan bahwa model tidak mengalami overfitting maupun underfitting.

Secara keseluruhan, hasil evaluasi menunjukkan bahwa model regresi linier sederhana yang dibangun memiliki performa yang baik dan layak digunakan sebagai dasar dalam perencanaan bisnis, khususnya dalam menyusun strategi produksi dan penjualan yang lebih efektif dan berbasis data.

V. KESIMPULAN

Penelitian ini berhasil mengintegrasikan sistem e-commerce dengan fitur prediksi penjualan berbasis data mining menggunakan regresi linear sederhana. Kebaruan terletak pada penerapan sistem prediksi yang ditujukan khusus untuk UMKM, yang selama ini masih jarang diimplementasikan. Hasil validasi menunjukkan akurasi prediksi tinggi dengan nilai R^2 sebesar 0,9449. Sistem juga mendapat respons positif dari pengguna dengan tingkat usability 85%.

Keterbatasan penelitian adalah model prediksi hanya menggunakan satu variabel (tahun) sehingga tidak mempertimbangkan faktor eksternal. Untuk penelitian selanjutnya, disarankan penggunaan metode yang lebih kompleks seperti regresi berganda, atau algoritma machine learning agar prediksi lebih

akurat dan aplikatif. Dengan demikian, sistem ini dapat menjadi dasar bagi UMKM dalam pengambilan keputusan bisnis yang lebih strategis dan berbasis data.

DAFTAR PUSTAKA

- [1] R. A. Putri, "Buku Ajar Basis Data," Jun. 2022. [Online]. Available: www.medsan.co.id
- [2] D. Apriana and C. Yuliansyah, "Optimalisasi Penjualan Online Melalui Teknik Data Mining (Studi Kasus E-Commerce)," *AL MIKRAJ J. Stud. Islam dan Hum.*, vol. 4, no. 2, pp. 514–527, 2024, [Online]. Available: <https://ejournal.insuriponorogo.ac.id/index.php/almikraj/article/view/4774>
- [3] A. Nugraha, O. Nurdiawan, and G. Dwilestari, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Yana Sport," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 6, no. 2, pp. 849–855, 2022, doi: 10.36040/jati.v6i2.5755.
- [4] N. Asih and M. Martanto, "Penerapan Data Mining Pada Transaksi Penjualan Untuk Menentukan Pola Pembelian Produk Menggunakan Algoritma Fp-Growth," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 2, pp. 1425–1431, 2024, doi: 10.36040/jati.v8i2.8961.
- [5] W. Inawati, R. Kurniawan, and Y. A. Wijaya, "Penerapan Aplikasi Rapidminer pada Prediksi Data Penjualan Album K-Pop Menggunakan Metode Regresi Linear," *J. Ris. dan Apl. Mhs. Inform.*, vol. 5, no. 3, pp. 585–592, 2024, doi: 10.30998/jrami.v5i3.10648.
- [6] N. Halim, "Perancangan Dashboard Dan Prediksi," *J. Ilmu Komput. dan Sist. Inf.*, vol. 11, no. 1, pp. 1–6, 2023.
- [7] M. Syahril, K. Erwansyah, and M. Yetri, "Penerapan Data Mining Untuk Menentukan Pola Penjualan Peralatan Sekolah Pada Brand Wigglo Dengan Menggunakan Algoritma Apriori," *J-SISKO TECH (Jurnal Teknol. Sist. Inf. dan Sist. Komput. TGD)*, vol. 3, no. 1, p. 118, 2020, doi: 10.53513/jsk.v3i1.202.
- [8] M. R. Athallah and A. F. Rozi, "Implementasi Data Mining Untuk Prediksi Peramalan Penjualan Produk Hj Karpet Menggunakan Metode Linear Regression," *J. Sains dan Teknol.*, vol. 2, no. 3, pp. 180–187, 2022, doi: 10.47233/jsit.v2i3.550.
- [9] 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," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 18, no. 2, p. 219, 2019, doi: 10.53513/jis.v18i2.162.
- [10] M. Hutahaean and K. Handoko, "Penerapan Data Mining Untuk Memprediksi Penjualan Obat Di Klinik Harapan Kita Batam," *J. Comasie*, vol. 6, no. 5, pp. 52–60, 2022.
- [11] F. Harahap, W. Fahrozi, R. Adawiyah, E. T. Siregar, and A. Y. N. Harahap, "Implementasi Data Mining dalam Memprediksi Produk AC Terlaris untuk Meningkatkan Penjualan Menggunakan Metode Naive Bayes," *J. Unitek*, vol. 16, no. 1, pp. 41–51, 2023, doi: 10.52072/unitek.v16i1.541.
- [12] A. Sadali, "Buku Ajar e-Commerce," 2022.
- [13] P. W. Rahayu *et al.*, "Buku Ajar Data Mining," 2024. [Online]. Available: <https://www.researchgate.net/publication/377415198>
- [14] N. L. W. S. R. Ginantra *et al.*, "Data Mining dan Penerapan Algoritma," 2021.
- [15] A. N. Iman, A. G. Putrada, S. Prabowo, and D. Perdana, "Peningkatan Kinerja AMG8833 sebagai Thermocam dengan Metode Regresi AdaBoost untuk Pelaksanaan Protokol COVID-19," *J. Elektro dan Telekomun. Terap.*, vol. 8, no. 1, p. 978, 2021, doi: 10.25124/jett.v8i1.3894.
- [16] Sugiyono, *Metode penelitian kuantitatif kualitatif dan R&D*. 2013.
- [17] S. R. Wicaksono, *Black Box Testing Teori Dan Studi Kasus*, no. February. 2022. doi: 10.5281/zenodo.7659674.