

Prediksi Penjualan Produk Menggunakan Algoritma *Xtreme Gradient Boosting*

Zhehan Gustiyandi¹, Aradea², Heni Sulastr³

^{1,2}Program Studi Informatika, Fakultas Teknik, Universitas Siliwangi, Tasikmalaya

³Program Studi Sistem Informasi, Fakultas Teknik, Universitas Siliwangi, Tasikmalaya

e-mail: ¹zehan265@gmail.com, ²aradea@unsil.ac.id, ³henisulastr@unsil.ac.id

Abstrak

Pesatnya perkembangan era digital membuat persaingan di sektor ritel semakin ketat, termasuk di gerai SRC Pak Didin. Salah satu tantangan yang dihadapi gerai tersebut adalah pengelolaan produk yang belum optimal sehingga berdampak pada efisiensi operasional. Data Science menawarkan solusi untuk meningkatkan kinerja bisnis, seperti meningkatkan efisiensi operasional dan mengoptimalkan strategi pemasaran atau penjualan. Penelitian ini bertujuan untuk memprediksi penjualan produk di gerai SRC Pak Didin menggunakan algoritma XGBoost dan mengusulkan strategi penjualan yang dapat diterapkan untuk meningkatkan operasional gerai. Dataset yang digunakan dalam penelitian ini terdiri dari 24 jenis produk selama tiga bulan terakhir, terhitung sejak September 2024. Produk-produk tersebut antara lain mi instan (berbagai rasa seperti gorengan, kuah kaldu ayam, Aceh, dan soto), teh botol, minuman ringan, minuman herbal, kopi instan, makanan ringan, biskuit, air minum dalam kemasan, roti, tepung, gula, bumbu penyedap, es krim, susu kotak, dan makanan ringan lainnya. Penelitian ini menggunakan algoritma XGBoost untuk menganalisis data penjualan selama tiga bulan terakhir dan memprediksi penjualan pada bulan berikutnya. Metrik evaluasi yang digunakan meliputi Mean Squared Error (MSE) dan R-squared (R^2). Model XGBoost diuji dalam tiga skenario: XGBoost Regression, XGBoost Regression Linear (variabel tunggal x), dan XGBoost Regression Linear (dua variabel x), dengan tujuan mengidentifikasi model dengan kinerja terbaik. Hasil akurasi menunjukkan bahwa model XGBoost Regression mencapai 96,56%, model XGBoost Regression Linear dengan satu variabel x mencapai 99,22%, dan model XGBoost Regression Linear dengan dua variabel x mencapai 99,59%. Model XGBoost Regression Linear dengan dua variabel dipilih sebagai model terbaik karena memiliki skor akurasi tertinggi. Model ini dapat secara efektif memprediksi penjualan produk dan memberikan wawasan yang dapat ditindaklanjuti untuk mengembangkan strategi penjualan, yang menguntungkan operasi toko SRC Pak Didin.

Kata kunci— XGBoost Regression Linear, Prediction, Mean Squad Error, R-squared, Sales Strategy

1. PENDAHULUAN

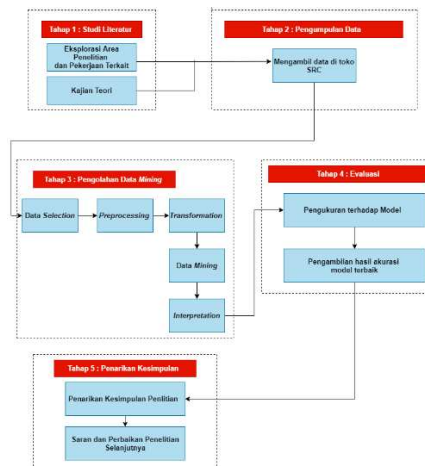
Seiring dengan perkembangan teknologi, maka dunia usaha khususnya dibidang retail mengalami persaingan yang sangat ketat dalam menjalankan bisnisnya [1]. Dampak dari berkembang pesatnya era digital, membuat data menjadi aset penting bagi sebuah toko dalam menjalankan bisnisnya [2]. Data yang terkait dengan penjualan seperti volume transaksi, tren musiman, harga, promosi, dan faktor eksternal lainnya menyediakan peluang besar untuk menganalisis dalam penjualan produk [3]. Namun, Toko SRC Pak Didin menghadapi beberapa permasalahan dalam pengelolaan produk yang dijual, terutama terkait manajemen stok. Dalam praktiknya, sering kali terjadi kendala, seperti penurunan penjualan pada beberapa produk atau peningkatan permintaan yang tidak diimbangi dengan ketersediaan stok. Hal ini mengakibatkan ketidakseimbangan antara permintaan dan pasokan produk, yang dapat memengaruhi kelancaran operasional toko [4]. Permasalahan tersebut tidak hanya berdampak pada operasional toko, tetapi juga memengaruhi tingkat kepuasan pelanggan. Ketidaksesuaian ketersediaan produk dengan kebutuhan pelanggan dapat mengakibatkan hilangnya peluang penjualan sekaligus

meningkatkan biaya penyimpanan untuk produk yang tidak terjual. Oleh karena itu, diperlukan solusi yang efektif untuk memprediksi penjualan secara akurat guna mengoptimalkan distribusi, pengelolaan stok, dan penjualan produk [5]. Penggunaan *Data Science* dapat membantu bisnis dalam berbagai hal, seperti mengoptimalkan strategi pemasaran atau penjualan [2]. Salah satu penerapannya adalah menggunakan metode prediksi atau peramalan [6]. Metode prediksi digunakan dengan menganalisis data penjualan produk yang pernah terjadi dalam kurun waktu tertentu [7]. Misalnya data penjualan produk yang terjual setiap bulan, tahun untuk menganalisis pola penjualan untuk melakukan prediksi pada bulan atau tahun selanjutnya [8].

Algoritma *XGBoost* dapat digunakan untuk melakukan prediksi [9]. Prediksi atau *forecasting* merupakan suatu perhitungan untuk meramalkan keadaan di masa mendatang dengan melakukan pengujian terhadap keadaan dimasa lalu [6], [10]. *XGBoost* menggunakan model yang lebih teratur untuk membangun struktur pohon regresi, sehingga dapat memberikan kinerja yang lebih baik dan mampu mengurangi kompleksitas model untuk menghindari overfitting [11]. Model *XGBoost* berbentuk pohon keputusan atau pohon regresi [5]. Penelitian [8] melakukan prediksi dengan menggunakan algoritma *XGBoost* terhadap dataset sales data big mart 2013 menjelaskan bahwa algoritma *XGBoost* dapat diterapkan terhadap dataset penjualan produk, buktinya adalah membandingkan model *Random Forest*, *LightBM*, dan *XGBoost* dengan melakukan pengukuran menggunakan metode *MAE* dan *RMSE* hasil akurasi dari setiap model yaitu pada *Random Forest* sebesar 0.32, *LightBM* 0.60, dan *XGBoost* 0.61. Penelitian [12] menjelaskan bahwa penggunaan optimasi dengan menambahkan hyperparameter dapat meningkatkan hasil akurasi *XGBoost* dengan melakukan pengukuran menggunakan *confusion matrix* hasil akurasi yang didapatkan sebelum dan sesudah yaitu 80.02% dan 83.42%. Penelitian [13] dalam penelitiannya menjelaskan bahwa algoritma *XGBoost* dapat diterapkan dengan baik pada dataset penjualan produk dari tahun 2013 -2017 dengan hasil pengukuran menggunakan metode *R-squared* akurasi yang didapatkan sebesar 0.92. Oleh karena itu, dalam penelitian ini dilakukan beberapa eksperimen dari model *XGBoost* yang digunakan yaitu model *XGBoost Regression* dan *XGBoost Regression Linear*. Metode evaluasi yang digunakan untuk mengukur hasil akurasi yang dihasilkan oleh model menggunakan metode *Mean Squad Error (MSE)* dan *R-squared (R^2)*. Tujuannya yaitu melakukan prediksi penjualan produk yang dapat dijadikan sebagai strategi dalam melakukan penjualan [14].

2. METODE PENELITIAN

Metode penelitian kuantitatif dilakukan dengan pendekatan eksperimen, penelitian ini menghasilkan nilai dari setiap variable, sehingga nilai tersebut perlu diukur pengaruh sebab-akibat antara variabel *independen* dan *dependen* [15]. Pada penelitian ini, objek yang dikembangkan adalah data penjualan produk pada toko SRC pak didin. Pengolahan data mining yang dilakukan menggunakan *KDD (Knowledge Discovery in Database)* [16]. Gambar 1 merupakan tahapan penelitian terdiri dari studi literatur, pengumpulan data, pengolahan data mining, evaluasi dan penarikan kesimpulan.



Gambar 1. Tahapan Penelitian

2.1 Studi Literatur

Studi literatur dilakukan dengan mencari informasi yang bersumber dari jurnal, buku mengenai topik penelitian yang terkait [17] dengan penerapan algoritma XGboost untuk melakukan prediksi terhadap data penjualan produk [18].

2.2 Pengumpulan Data

Data produk yang dikumpulkan sebanyak 24 dimana pengumpulan data tersebut dilakukan dengan metode observasi langsung dengan teknik wawancara terhadap pemilik toko SRC. Data yang direkap merupakan data penjualan produk selama tiga bulan terakhir terhitung sejak bulan September 2024.

2.3 Pengolahan Data Mining

Pengolahan data mining dilakukan dengan tahapan Knowledge Discovery In Database [16], diantaranya:

1. *Data Selection*, yaitu Pada tahapan ini dilakukan teknik perolehan sebuah pengurangan representasi dari data dan meminimalkan hilangnya informasi data.
2. *Preprocessing*, yaitu pada tahapan ini dilakukan pembersihan data atau penghapusan atribut yang tidak digunakan pada dataset. Metode yang digunakan yaitu data *cleaning*.
3. *Transformation*, yaitu Pada tahapan ini dilakukan proses merubah data yang telah dipilih agar sesuai untuk proses data *mining*.
4. *Data Mining*, yaitu Pada tahapan ini dilakukan proses mencari informasi dalam data terpilih dengan menggunakan algoritma *XGBoost Regression* dan *XGBoost Regression Linear* untuk menentukan berapa nilai dari prediksi penjualan dari masing-masing produk.
5. *Interpretation*, yaitu Pada tahapan ini menampilkan grafik hasil prediksi penjualan produk dan dibandingkan dengan data penjualan produk tiga bulan terakhir dengan tujuan membuat strategi penjualan.

2.4 Evaluasi

Tahapan evaluasi menggunakan metode pengukuran performa model yaitu *Mean Squared Error (MSE)* dan *R-squared (R2)* untuk menentukan model yang paling baik dengan hasil akurasi yang didapatkan.

2.5 Penarikan Kesimpulan

Tahapan ini merupakan tahapan akhir dari penelitian yang menyimpulkan hasil prediksi penjualan produk dan membuat strategi penjualan.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Data yang digunakan merupakan data penjualan produk pada toko SRC pak didin yang diambil dari penjualan periode tiga bulan terakhir terhitung dari bulan September 2024, tertera pada Tabel 1.

Tabel 1. Data Penjualan Produk Tiga Bulan Terakhir

No	Nama Produk	Kuantitas Produk	Kuantitas Terjual	No	Nama Produk	Kuantitas Produk	Kuantitas Terjual
1	Mie Indomie Goreng	1460	1430	13	Larutan Cap Kaki Tiga	346	280
2	Mie Indomie Ayam Bawang	1460	1435	14	Teh Kotak	316	276
3	Roti Aoka	1630	1610	15	Teh Botol Sosro	310	288
4	Es Krim Alice	1995	1975	16	Gula	220	170
5	Nabati Sip	865	825	17	Tepung Terigu	210	175
6	Kacang Sukro	805	780	18	Pop Mie	176	130
7	Susu Kotak Ultra Milk	570	530	19	Biskuit Roma	174	144
8	Susu Kotak Indomilk	570	540	20	Aqua Botol	161	135
9	Mie Indomie Soto	505	475	21	Makanan Ringan (Snack Taro)	162	148
10	Mie Indomie Aceh	502	478	22	Biskuit Gery	152	131
11	Kopi Kapal Api	382	375	23	Bumbu Racik	154	129
12	Kopi Good Day	350	345	24	Shampo Botol	92	40

3.2 Data Selection

Data *selection* dilakukan untuk menentukan data variabel pada dataset. Pada dataset penjualan produk, variabel yang digunakan adalah kolom nama produk, kuantitas produk, kuantitas terjual, dan satuan.

3.3 Data Preprocessing

Pada tahapan ini dilakukan proses data *cleaning* dan data *transformation*. Proses data *cleaning* dilakukan agar data yang memiliki subjek kosong, duplikasi data dapat dihapus dan memperbaiki format penulisan data. Hasil sebelum dan sesudah data *cleaning* tertera pada Gambar 2 dan Gambar 3.

	No	Nama Produk	Unnamed: 2	Kuantitas Produk	Unnamed: 4	Kuantitas Terjual	Unnamed: 6	Satuan
0	1	Mie Indomie Goreng	NaN	1460	NaN	1430	NaN	pcs
1	2	Mie Indomie Ayam Bawang	NaN	1460	NaN	1435	NaN	pcs
2	3	Roti Aoka	NaN	1630	NaN	1610	NaN	pcs

Gambar 2. Dataset Sebelum Data Cleaning

Pada Gambar 2. merupakan dataset yang belum dilakukan proses data *cleaning*, masih terdapat beberapa kolom yang memiliki subjek kosong, contohnya kolom Unnamed: 2, Unnamed:4, Unnamed: 6, dan NaN. Selanjutnya pada kolom satuan dihapus, karena kolom tersebut tidak mempengaruhi untuk pengolahan data *mining*.

	Nama Produk	Kuantitas Produk	Kuantitas Terjual
0	Mie Indomie Goreng	1460	1430
1	Mie Indomie Ayam Bawang	1460	1435
2	Roti Aoka	1630	1610

Gambar 3. Dataset Sesudah Data Cleaning

Pada Gambar 3. merupakan dataset yang sudah dilakukan proses data *cleaning*. Beberapa kolom yang sebelumnya ada pada Gambar 2. dihapus, sehingga menyisakan tiga kolom yang akan digunakan pada pengolahan data *mining* yaitu kolom nama produk, kuantitas produk, dan kuantitas terjual. Sementara itu, proses data *transformation* dilakukan untuk melakukan penggabungan atau pemilihan terhadap variabel dan atribut pada dataset yang akan dipilih. Pada tahapan ini, variabel x yang digunakan adalah kolom kuantitas produk dan variabel y yang digunakan adalah kolom kuantitas terjual.

3.4 Exploring Data Analysis

Exploring Data Analysis dilakukan untuk lebih memahami lagi mengenai struktur dan karakteristik dari dataset yang digunakan. Pada tahapan ini bertujuan untuk mendeteksi *outlier*, menemukan hubungan antar variabel, dan menghasilkan hipotesis sementara [19]. Pada penelitian ini tahapan *EDA* yang dilakukan yaitu melihat struktur dataset, menampilkan lima produk paling banyak dan paling sedikit terjual, boxplot kuantitas produk dan kuantitas terjual, menampilkan produk berdasarkan kategori makanan, kategori minuman, kategori bahan dapur, dan kategori lainnya.

3.5 Pengembangan Model

Pengembangan model dilakukan menggunakan bantuan *google colab* dan beberapa *library python* untuk membantu pengembangannya. Dimulai dari

1). Pembagian Data

Pembagian dataset dibagi menjadi dua yaitu data pelatihan dan data pengujian. Rasio untuk data pelatihan sebesar 70% dan data pengujian sebesar 30%.

2). Modeling *XGBoost Regression*

Tahapan modeling *XGBoost Regression* yaitu dataset yang digunakan dilakukan dulu serangkaian proses mulai dari data *selection*, data *preprocessing*, kemudian masuk ke pembagian data untuk diproses ke tahapan prediksi. Konsep utama dari model *XGBoost Regression* merupakan sebuah pohon keputusan, dimana terdapat beberapa proses iterasi yang menghasilkan pohon keputusan baru untuk memperoleh nilai prediksi yang baik. pembuatan model *XGBoost Regression* dilakukan dengan menggunakan *Google Colab* dengan *library* yang digunakan yaitu *XGBRegressor*.

3). Modeling *XGBoost Regression Linear (satu variabel x)*

Tahapan modeling *XGBoost Regression Linear* (satu variabel x) yaitu dataset yang digunakan dilakukan dulu serangkaian proses mulai dari data *selection*, data *preprocessing*, kemudian masuk ke pembagian data untuk diproses ke tahapan prediksi. Pada model ini ditambahkan beberapa parameter yaitu 'booster': 'gblinear', 'objective': 'reg:squarederror', 'learning_rate': 0.1, 'alpha': 0.1, 'lambda': 0.1. Tujuan dari penambahan parameter ini adalah untuk meningkatkan performa model. Pembuatan model *XGBoost Regression Linear* (satu variabel x) dilakukan dengan menggunakan *Google Colab* dengan *library* yang digunakan yaitu *XGBRegressor*. Hasil akhirnya merupakan proses untuk melakukan prediksi tertera pada Gambar 4.

```
[ ] # Input data baru
new_data = pd.DataFrame({'Kuantitas Produk': []}) # Ganti nilai sesuai kebutuhan

# Melakukan prediksi pada data baru
predicted_sales_xg_linear = model_xg_linear.predict(new_data[['Kuantitas Produk']])

# Menampilkan hasil prediksi sebelum post-processing
print("Prediksi Kuantitas Terjual (Sebelum Post-processing):", predicted_sales_xg_linear[0])
```

Gambar 4. Model Prediksi XGBoost Regression Linear (satu variabel x)

4). Modeling *XGBoost Regression Linear* (dua variabel x)

Tahapan modeling *XGBoost Regression Linear* (satu variabel x) yaitu dataset yang digunakan dilakukan dulu serangkaian proses mulai dari data *selection*, data *preprocessing*, kemudian masuk ke pembagian data untuk diproses ke tahapan prediksi. Pada model ini ditambahkan beberapa parameter yaitu 'booster': 'gblinear', 'objective': 'reg:squarederror', 'learning_rate': 0.1, 'alpha': 0.1, 'lambda': 0.1. Pada model ini dilakukan terlebih dahulu proses encoding menggunakan *Leave One Out Encoder* [20]. Kolom yang di encoding yaitu nama produk. Tujuan dari penambahan parameter dan penambahan variabel x adalah untuk meningkatkan performa model pada tahapan modeling sebelumnya yaitu model *XGBoost Regression Linear* (satu variabel x). Pada pembuatan model *XGBoost Regression Linear* (dua variabel x) dilakukan dengan menggunakan *Google Colab* dengan library yang digunakan yaitu *XGBRegressor*. Hasil akhirnya merupakan proses untuk melakukan prediksi tertera pada Gambar 5.

```
[ ] # Input data baru
new_data = pd.DataFrame({'Kuantitas Produk': [], # Ganti nilai sesuai kebutuhan
                        'Nama Produk Encoded': []}) # Tambahkan kolom 'Nama Produk Encoded'

# Melakukan prediksi pada data baru
predicted_sales_xg_linear1 = model_xg_linear1.predict(new_data[['Kuantitas Produk',
                                                                'Nama Produk Encoded']]) # Sertakan kolom 'Nama Produk Encoded' saat prediksi

# Menampilkan hasil prediksi sebelum post-processing
print("Prediksi Kuantitas Terjual (Sebelum Post-processing):", predicted_sales_xg_linear1[0])
```

Gambar 5. Model Prediksi XGBoost Regression Linear (dua variabel x)

5). Tujuan Pengembangan Tiga Model Eksperimen

Tujuan dilakukan eksperimen model XGBoost adalah untuk mencari model dengan performa yang baik untuk melakukan prediksi terhadap dataset penjualan produk yang digunakan. Hasil prediksi data penjualan produk berdasarkan beberapa eksperimen yang dilakukan tertera pada Tabel 2. Model 1 adalah XGBoost Regression, model 2 adalah XGBoost Regression Linear (satu variabel x), dan model 3 adalah XGBoost Regression Linear (dua variabel x).

Tabel 2. Hasil Prediksi Ketiga Experimen Model

Nama Produk	Model 1	Model 2	Model 3	Nama Produk	Model 1	Model 2	Model 3
Mie Indomie Goreng	180	423	474	Biskuit Roma	26	100	54
Mie Indomie Ayam Bawang	175	414	463	Biskuit Gery	-1	82	56
Mie Indomie Aceh	57	167	155	Shampo Botol	71	104	92
Mie Indomie Soto	60	175	185	Nabati Sip	0	286	297
Teh Kotak	4	166	174	Kacang Sukro	40	242	278
The Botol Sosro	8	114	104	Tepung Terigu	-5	104	96
Larutan Cap Kaki Tiga	0	164	132	Gula	0	123	75
Kopi Good Day	-65	113	54	Bumbu Racik	1	84	76
Kopi Kapal Api	100	122	75	Es Krim Alice	0	546	552
Pop Mie	40	114	105	Susu Kotak Ultra Milk	5	205	208

Nama Produk	Model 1	Model 2	Model 3	Nama Produk	Model 1	Model 2	Model 3
Aqua Botol	40	88	50	Susu Kotak Indomilk	-5	185	172
Roti Aoka	365	460	535	Makanan Ringan (Snack Taro)	27	73	66

3.6. Evaluasi Model

Evaluasi model dilakukan untuk mengukur kinerja algoritma *XGBoost Regression*, algoritma *XGBoost Regression Linear* (satu variabel x), dan algoritma *XGBoost Regression Linear* (dua variabel x). Evaluasi model yang digunakan untuk mengukur hasil kinerja dalam melakukan prediksi adalah metode *Mean Squad Error (MSE)* [21] dan *R-squared (R²)* [22].

1) Perbandingan Model

Perbandingan Model dapat dilihat Tabel 3.

Tabel 3. Perbandingan Hasil Pengukuran Model

Score	<i>XGBoost Regression</i>	<i>XGBoost Regression Linear</i> (satu variabel x)	<i>XGBoost Regression Linear</i> (dua variabel x)
<i>MSE</i>	5861.11351	1320.36282	683.74895
<i>R-squared</i>	0.96567	0.99226	0.99599

Hasil dari pengukuran performa model yaitu model *XGBoost Regression* memiliki nilai akurasi sebesar 0.96567, model *XGBoost Regression Linear* (satu variabel x) memiliki nilai akurasi sebesar 0.99226, dan model *XGBoost Regression Linear* (dua variabel x) memiliki nilai akurasi sebesar 0.99599. Model dengan akurasi paling baik adalah model *XGBoost Regression Linear* (dua variabel x).

2) Analisis Hasil Prediksi

Pada tahapan ini melakukan analisis terhadap hasil prediksi dari ketiga percobaan model yang dilakukan yang tertera pada Tabel 2. Kemudian hasil prediksi tersebut dibandingkan dengan penjualan produk selama tiga bulan terakhir yang tertera pada Tabel 4.

Tabel 4. Data Penjualan Produk Selama Tiga Bulan

Nama Produk	Bulan 1	Bulan 2	Bulan 3	Nama Produk	Bulan 1	Bulan 2	Bulan 3
Mie Indomie Goreng	460	500	470	Biskuit Roma	40	58	46
Mie Indomie Ayam Bawang	465	490	480	Biskuit Gery	38	48	45
Mie Indomie Aceh	145	168	165	Shampo Botol	8	12	20
Mie Indomie Soto	140	175	160	Nabati Sip	260	295	270
Teh Kotak	70	108	98	Kacang Sukro	235	280	265
The Botol Sosro	75	116	97	Tepung Terigu	40	70	65
Larutan Cap Kaki Tiga	65	100	115	Gula	40	60	70
. Kopi Good Day	115	110	120	Bumbu Racik	35	46	48
Kopi Kapal Api	125	118	132	Es Krim Alice	650	665	660
Pop Mie	26	60	44	Susu Kotak Ultra Milk	165	180	185
Aqua Botol	40	45	50	Susu Kotak Indomilk	160	190	190
Roti Aoka	530	550	530	Makanan Ringan (Snack Taro)	40	58	50

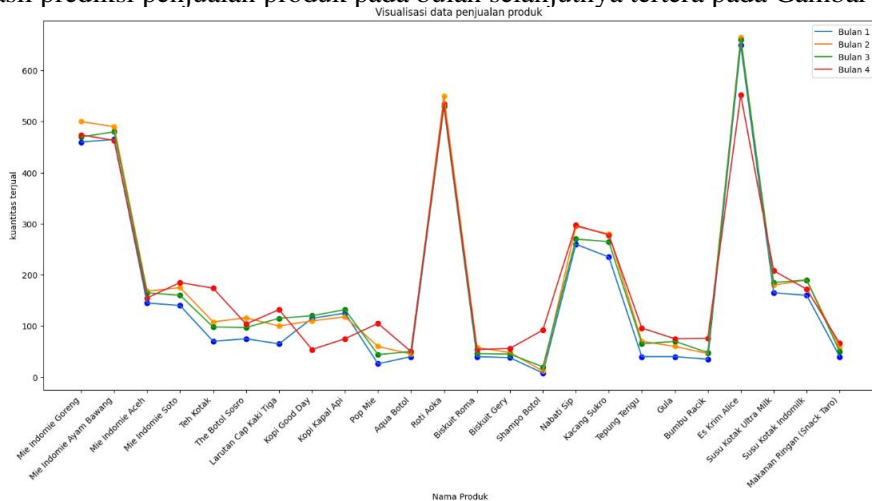
Pada model pertama dengan menggunakan model *XGBoost Regression*, hasil prediksi yang diperoleh tidak sesuai dengan pola penjualan produk pada bulan-bulan sebelumnya. Salah satu anomali yang muncul adalah adanya nilai prediksi penjualan yang negatif, sementara pada data historis penjualan, tidak pernah tercatat nilai penjualan negatif. Hal ini menunjukkan bahwa pada percobaan pertama terdapat keterbatasan dalam model untuk melakukan prediksi, terutama dalam mempertimbangkan faktor kuantitas produk yang memengaruhi jumlah produk terjual.

Pada model kedua, menggunakan model *XGBoost Regression Linear* (satu variabel x), hasil prediksi yang diperoleh menunjukkan kesesuaian yang lebih baik dengan pola penjualan produk pada bulan-bulan sebelumnya. Hasil prediksi penjualan menunjukkan bahwa jumlah kuantitas produk dan kuantitas terjual sudah cukup akurat. Dibandingkan dengan percobaan pertama, batasan nilai dalam memprediksi kuantitas terjual dari kuantitas produk mengalami perbaikan. Namun, masih terdapat beberapa produk dengan nilai prediksi penjualan yang tidak sepenuhnya sesuai dengan kuantitas produk yang tersedia.

Pada model ketiga, menggunakan model *XGBoost Regression Linear* (dua variabel x), hasil prediksi yang diperoleh hampir sebaik percobaan kedua, namun menunjukkan tingkat keoptimalan yang lebih tinggi dibandingkan dua percobaan sebelumnya. Pada percobaan ini, batasan nilai untuk memprediksi penjualan berdasarkan data kuantitas produk telah sesuai, di mana kuantitas terjual tidak melebihi kuantitas produk yang tersedia. Pola prediksi penjualan yang dihasilkan dalam percobaan ini cukup konsisten dengan pola penjualan pada bulan-bulan sebelumnya, karena hasil prediksinya mendekati nilai aktual dan tidak menunjukkan perbedaan yang signifikan antara jumlah produk yang terjual dan kuantitas produk yang tersedia.

3.7. Strategi Penjualan Produk

Strategi penjualan produk dibuat berdasarkan grafik hasil penjualan produk tiga bulan terakhir dan hasil prediksi penjualan produk pada bulan selanjutnya tertera pada Gambar 7.



Gambar 4. Grafik Penjualan Produk

Bulan 1 sampai 3 merupakan data penjualan produk selama tiga bulan terakhir, sementara itu Bulan 4 adalah hasil prediksi penjualan produk pada bulan selanjutnya. Strategi penjualan produk yang dapat dilakukan berdasarkan Gambar 7 sebagai berikut.

1. Produk varian mie memiliki pola penjualan yang cukup stabil pada setiap bulannya. Untuk itu bisa dilakukan penambahan stok pada produk tersebut untuk meminimalisir apabila

terjadi kelonjakan dalam penjualan. Contohnya pada produk mie indomie aceh dan mie indomie soto mengalami kenaikan dalam penjualan.

2. Pada beberapa produk yang mengalami penurunan penjualan, sebaiknya dilakukan promosi atau paket bundling untuk mengatasi permasalahan dalam penjualan.
3. Produk aqua botol, biskuit roma, biskuit Gerry, tepung terigu, gula dan bumbu racik berdasarkan hasil prediksi pada Bulan 4, mengalami penjualan yang stabil. Untuk itu stok produk yang masuk bisa dilakukan seperti biasanya dan tidak perlu menambah jumlah stok produk yang masuk. Jika ingin meningkatkan penjualan, bisa dilakukan penjualan berupa promosi atau paket bundling, dengan catatan stok produk yang masuk ditambah lagi kuantitasnya.
4. Dalam mengatasi penurunan penjualan produk, bisa dilakukan strategi *marketing* berupa diskon, promosi, paket bundling.
5. Dalam mengatasi kelonjakan penjualan produk, bisa dilakukan penambahan kuantitas stok produk yang masuk. Tetapi harus melihat terlebih dahulu pola penjualan pada bulan-bulan sebelumnya terhadap produk tersebut.

4. KESIMPULAN

Prediksi penjualan produk menggunakan algoritma *XGBoost* dengan tiga eksperimen yaitu *XGBoost Regression*, *XGBoost Regression Linear* (satu variabel x), dan *XGBoost Regression Linear* (dua variabel x). Berdasarkan ketiga eksperimen model tersebut, model *XGBoost Regression Linear* (dua variabel x) merupakan model yang paling baik dengan akurasi sebesar 0.99599. Hasil prediksi data penjualan produk yang didapatkan sangat sesuai jika dilihat dari pola penjualan tiga bulan sebelumnya, sehingga berdasarkan data tersebut dibuat strategi penjualan produk yang dapat diterapkan oleh pihak toko untuk bulan-bulan selanjutnya. Oleh karena itu, untuk pengembangan selanjutnya dataset yang digunakan ditambah lagi dengan menggunakan data penjualan produk di beberapa toko SRC berbeda untuk menghasilkan strategi penjualan produk yang lebih baik lagi.

DAFTAR PUSTAKA

- [1] D. Suparman, "Pengaruh Harga Dan Kualitas Pelayanan Terhadap Penjualan Spare Part Motor Di Pt. Slm (Selamat Lestari Mandiri)," *J. Ekon.*, vol. Vol. 07 No, no. 2, p. 2, 2018.
- [2] Hartatik *et al.*, *Data Science for Business (Pengantar dan Penerapan Berbagai Sektor)*, no. May. 2023.
- [3] P. Dankorpho, "Sales Forecasting for Retail Business using XGBoost Algorithm," *J. Comput. Sci. Technol. Stud.*, vol. 6, no. 2, pp. 136–141, 2024, doi: 10.32996/jcsts.2024.6.2.15.
- [4] D. Y. Y. Sim and Z. Wei, "XGBoost Regression Algorithms for Efficient Predictions on Inventory Sales and Management," no. September, pp. 66–71, 2024, doi: 10.1109/eeet61723.2023.00033.
- [5] R. Siringoringo, R. Perangin-angin, and M. J. Purba, "Segmentasi Dan Peramalan Pasar Retail Menggunakan Xgboost Dan Principal Component Analysis," *METHOMIKA J. Manaj. Inform. dan Komputerisasi Akunt.*, vol. 5, no. 1, pp. 42–47, 2021, doi: 10.46880/jmika.vol5no1.pp42-47.
- [6] A. Alfani W.P.R., F. Rozi, and F. Sukmana, "Prediksi Penjualan Produk Unilever Menggunakan Metode K-Nearest Neighbor," *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 6, no. 1, pp. 155–160, 2021, doi: 10.29100/jipi.v6i1.1910.

- [7] Ab. Bin Tayyab and M. F. Nasim, "Walmart Sales Prediction Using Machine Learning Algorithms," *Ann. Rom. Soc. Cell Biol.*, vol. 25, pp. 7872-7882–7872 – 7882, 2024.
- [8] F. Riza, "Analisis dan Prediksi Data Penjualan Menggunakan Machine Learning dengan Pendekatan Ilmu Data," *Data Sci. Indones.*, vol. 1, no. 2, pp. 62–68, 2022, doi: 10.47709/dsi.v1i2.1308.
- [9] M. Riyyasy, Azfa, Rasikh, W. Aghniya, Nouval, and H. Tantyoko, "Penerapan Algoritma Machine Learning Untuk Memprediksi Term Deposit Nasabah Perbankan," *J. Inform. Inf. Technol.*, vol. 2, no. 3, pp. 145–156, 2023.
- [10] N. Ambari, N. Puspitasari, and A. Septiarini, "Prediction of Budget Planning Using The Long Short Term Memory," *J. Artif. Intell. Softw. Eng.*, vol. 5, no. 1, pp. 116–125, 2025.
- [11] D. L. Yuliani, N. Mulyatini, and E. Herlina, "ANALISIS MATERIAL REQUIREMENT PLANNING DAN MANAJEMEN RANTAI PASOKAN DALAM MENINGKATKAN KEUNGGULAN BERSAING (Suatu Studi Pada ...," ... *Entrep. J.*, vol. 1, no. September, pp. 32–46, 2019.
- [12] S. E. Herni Yulianti, Oni Soesanto, and Yuana Sukmawaty, "Penerapan Metode Extreme Gradient Boosting (XGBOOST) pada Klasifikasi Nasabah Kartu Kredit," *J. Math. Theory Appl.*, vol. 4, no. 1, pp. 21–26, 2022, doi: 10.31605/jomta.v4i1.1792.
- [13] K. Li, "A Sales Prediction Method Based on XGBoost Algorithm Model," *BCP Bus. Manag.*, vol. 36, pp. 367–371, 2023, doi: 10.54691/bcpbm.v36i.3487.
- [14] D. Riando and A. Afiyati, "IMPLEMENTATION OF XGBOOST ALGORITHM TO PREDICT THE SELLING PRICE OF CAYENNE PEPPERS IN," vol. 4, no. 09, pp. 741–749, 2024.
- [15] S. Priadana and D. Sunarsi, *METODE PENELITIAN KUANTITATIF*. 2021.
- [16] D. T. Larose and C. D. Larose, *Discovering Knowledge in Data: An Introduction to Data Mining: Second Edition*, vol. 9780470908. 2014.
- [17] K. Abdullah *et al.*, *Metodologi Penelitian Kuantitatif*, vol. 3, no. 2. 2021.
- [18] X. Dairu and Z. Shilong, "Machine Learning Model for Sales Forecasting by Using XGBoost," *2021 IEEE Int. Conf. Consum. Electron. Comput. Eng. ICCECE 2021*, no. Iccece, pp. 480–483, 2021, doi: 10.1109/ICCECE51280.2021.9342304.
- [19] A. Wibowo, "Analisa Dan Visualisasi Data Penjualan Menggunakan Exploratory Data Analysis Pada PT. Telkominfra," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 9, no. 3, pp. 2292–2304, 2022, doi: 10.35957/jatisi.v9i3.2737.
- [20] F. Bolikulov, R. Nasimov, A. Rashidov, F. Akhmedov, and Y. I. Cho, "Effective Methods of Categorical Data Encoding for Artificial Intelligence Algorithms," *Mathematics*, vol. 12, no. 16, 2024, doi: 10.3390/math12162553.
- [21] D. Arisandi, S. Salamun, and A. R. Putra, "Prediksi Penerimaan Siswa Baru dengan Metode Single Exponential Smoothing Melalui Metrik Evaluasi MAD, MSE dan MAPE," *J. Inf. Syst. Res.*, vol. 4, no. 4, pp. 1197–1204, 2023, doi: 10.47065/josh.v4i4.3658.
- [22] A. M. M. Fattah, A. Voutama, N. Heryana, and N. Sulistiyowati, "Pengembangan Model Machine Learning Regresi sebagai Web Service untuk Prediksi Harga Pembelian Mobil dengan Metode CRISP-DM," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 5, p. 1669, 2022, doi: 10.30865/jurikom.v9i5.5021.