

ANALISIS PENJUALAN PRODUK UMKM DI SHOPEE PADA TOKO AGUNG0NA9 MENGGUNAKAN MODEL ALGORITMA REGRESI LINEAR

Adi Supriyatna, Ade Irma Purnamasari, Irfan Ali

Teknik Informatika, STMIK IKMI Cirebon
Rekayasa Perangkat Lunak, STMIK IKMI Cirebon
Jl. Perjuangan No. 10 B Majasem Kota Cirebon
Adisupriatna652@gmail.com,

ABSTRAK

Shopee merupakan platform e-commerce yang beroperasi secara online dan hadir di berbagai negara di Asia Tenggara, Shopee telah berkembang pesat dalam beberapa tahun terakhir memungkinkan konsumen membeli produk secara online, Toko umkm Agung0na9 Menjual produk kategori musik dan juga kerajinan. Tujuan dari penelitian ini adalah menerapkan algoritma regresi linier untuk memprediksi produk dengan Kategori musik dan kerajinan yang akan terjual dalam waktu 3 bulan berikutnya. Regresi linear digunakan sebagai metode prediksi dengan jumlah produk yang terjual sebagai variabel Y dan periode sebagai variabel X. *relative Error* digunakan untuk mengevaluasi hasil prediksi. Hasil prediksi kategori Musik pada bulan pertama terjual 68 pcs, pada bulan kedua 69 pcs dan bulan ketiga 70 pcs dan kategori Kerajinan pada bulan pertama terjual 1078 pcs, pada bulan kedua 1029 pcs dan bulan ketiga 1066 pcs. Hasil evaluasi nilai Nilai *Relative Error* pada kategori Musik 13.64%, sedangkan, Nilai *Relative Error* kategori Kerajinan 22.65%. Prediksi penjualan handphone menggunakan metode regresi linear ini dapat dikatakan tergolong dalam kategori cukup atau bisa digunakan.

Kata kunci: *Prediksi, Relative error, Regresi linear, Penjualan*

1. PENDAHULUAN

Di era globalisasi dan pesatnya perkembangan teknologi informasi, dunia bisnis mengalami perubahan besar [1]. Usaha Mikro, Kecil, dan Menengah (UMKM) merupakan salah satu penopang utama perekonomian negara dan memberikan kontribusi penting dalam penciptaan lapangan kerja dan peningkatan pendapatan masyarakat. Dalam konteks ini, *platform e-commerce* seperti Shopee telah menjadi sarana penting bagi UMKM untuk menjual produknya secara luas [2]. Toko Agung0na9 mewakili UMKM yang beroperasi di platform Shopee, menunjukkan semangat beradaptasi dengan tren teknologi untuk mendapatkan keunggulan kompetitif [3]

Ketika persaingan meningkat, perkiraan penjualan telah menjadi elemen penting dalam perencanaan strategis [4]. Prediksi atau *Forecasting* adalah Ramalan terhadap kejadian masa depan yang membantu dalam pengambilan keputusan [5]. 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 [6]. *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 [7]Toko agung0na9 mengalami permasalahan cashflow karena stok produk kategori x yang belum terjual. Studi data mining dengan *regresi linear* pada data penjualan produk di evaluasi dengan menggunakan indikator *relative error*. *Relative error* adalah akar kuadrat rata-rata *error* dari suatu metode yang digunakan untuk mengevaluasi kesesuaian data

sebenarnya dengan model prediksinya [8]. 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 [9]. Diharapkan bisa memberikan prediksi penjualan dan memberikan keputusan terbaik dalam menentukan stok produk.

2. TINJAUAN PUSTAKA

2.1. Penelitian terdahulu

Salah satu penelitian terkait yang sudah bahasil dilakukan diantaranya penelitian oleh yubi Aqsho Ramadhan , Ahmad Faqih dan Gifthera Dwilestari 2023 berjudul "PREDIKSI PENJUALAN HANDPHONE DI TOKO X MENGGUNAKAN ALGORITMA REGRESI LINEAR" Kesimpulan dari jurnal ini adalah bahwa penggunaan regresi linear dalam memprediksi penjualan handphone di toko X telah memberikan hasil yang cukup baik [10]. Evaluasi menggunakan RMSE dan Relative Error menunjukkan hasil yang mendekati nilai 0, menunjukkan tingkat akurasi yang tinggi dalam prediksi penjualan. Selain itu, artikel ini juga memberikan gambaran metodologi penelitian yang digunakan, termasuk pemilihan data, preprocessing, transformasi, data mining, evaluasi, dan pengetahuan. Hasilnya disajikan dalam scatter plot untuk menunjukkan konsistensi prediksi. Dengan demikian, jurnal ini memberikan kontribusi dalam memahami dan memprediksi penjualan handphone berdasarkan data historis menggunakan regresi linear [11].

Kemudian Hasil penelitian dari Riyana, Suarna and Dwilestari, 2023 Yang Berjudul "Analisa Dataset

Penjualan Teh Menggunakan Algoritma Linear Regresi” Hasil penelitian ini membahas tentang analisis penjualan teh dengan menggunakan algoritma regresi linier. Hal ini bertujuan untuk mengidentifikasi pola dan tren terkait penjualan serta mengetahui hubungan antara variabel penjualan dengan variabel yang mempengaruhinya[12]. Selain itu, makalah ini membahas strategi pemasaran produk teh dan dampak berbagai faktor terhadap penjualan. Metodologi penelitian meliputi pengumpulan data, pembersihan data, transformasi data, penambangan data, dan interpretasi & evaluasi. Penelitian ini menggunakan proses Knowledge Discovery in Databases (KDD) dan alat RapidMiner untuk analisis data[13].

2.2. Landasan Teori

2.2.1. Data mining

Data mining adalah proses yang menggunakan metode statistik, matematika, kecerdasan buatan dan pembelajaran mesin untuk mengidentifikasi informasi yang berguna dan temukan pengetahuan tersembunyi dari berbagai database [13]

2.2.2. Regresi Linear

Algoritma regresi Linear adalah sebuah metode statistik yang digunakan untuk memodelkan hubungan linear antara satu atau lebih variabel independen (predisiktor) dan variabel dependen (hasil atau respons). Dalam konteks prediksi penjualan produk UMKM di Shopee, algoritma Regresi Linear digunakan untuk menemukan garis linear terbaik yang merepresentasikan hubungan antara faktor-faktor tertentu dengan jumlah penjualan produk.[14]

2.2.3. Relative error

Relative error adalah akar kuadrat rata-rata error dari suatu metode yang digunakan untuk mengevaluasi kesesuaian data sebenarnya dengan model prediksinya, Relative Error berguna dalam banyak konteks, termasuk pengukuran akurasi model prediksi [14]

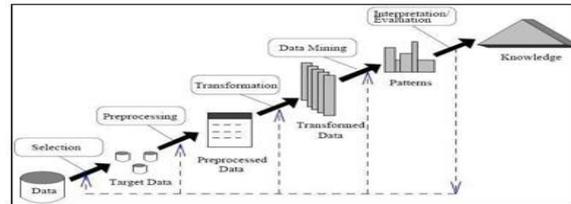
2.2.4. Knowledge Discovery in Database (KDD)

Knowledge Discovery in Database (KDD) adalah pendekatan yang diterapkan untuk mendapatkan informasi yang berasal dari basis data yang tersedia. Hasil temuan dari proses ini dapat dimanfaatkan untuk membentuk basis pengetahuan yang digunakan dalam konteks pengambilan keputusan[5]

3. METODE PENELITIAN

Metode Penelitian ini menggunakan jenis penelitian deskriptif dengan menggunakan pendekatan kuantitatif. Fungsi deskriptif kuantitatif dirancang untuk menentukan jumlah setiap variabel, dengan tujuan memberikan gambaran obyektif tentang suatu kondisi, dan satu atau lebih variabel dihubungkan atau dibedakan dengan variabel lain, bersifat independen tanpa menimbulkan apa pun[10]Gunakan regresi

linier. Regresi linier merupakan metode prediksi yang menguji hubungan antara variabel terikat dan bebas. Algoritma regresi linier mengukur kekuatan hubungan antara dua variabel dan menunjukkan aliran hubungan antara variabel terikat dan bebas [10]KDD (Knowledge Discovery in Database) digunakan untuk analisis data dalam penelitian ini. Gambar stage KDD ditunjukkan pada Gambar 1 di bawah ini.



Gambar 1. Tahapan Proses KDD

Proses *Knowledge Discovery in Database (KDD)* dapat diuraikan sebagai berikut :

a. Selection

Data penjualan pada toko Agung0na9 merupakan data tahun 2022-2023.. Pada tahapan ini data dari penjualan dari bulan oktober 2022- oktober 2023 , terjadi penyeleksian atribut untuk menentukan atribut mana yang akan digunakan dalam proses data mining.

b. Pre-processing

Pada tahapan ini, hanya fokus pada data yang akan diolah. Ini juga termasuk membersihkan data untuk memperbaiki duplikat, nilai yang hilang, dan memperbaiki kesalahan seperti kesalahan ketik.

c. Transformation

Merupakan proses transformasi untuk data terpilih, sehingga data tersebut layak dan sesuai untuk diolah dalam data mining.

d. Data Mining

Proses data mining ini melibatkan penerapan algoritma atau pencarian pengetahuan, dan dalam penelitian ini, digunakan algoritma *Regresi Linear* pada *tools rapidminer* versi 10.2.

e. Interpretation / Evaluation

Tahapan ini merupakan tahap akhir dari proses data mining. Tahap ini dilakukan setelah data diolah dengan algoritma yang dipilih menggunakan *tools rapidminer* dan tujuannya adalah untuk melakukan evaluasi terhadap hasil proses data mining.

4. HASIL DAN PEMBAHASAN

Analisis data yang digunakan Knowledge Discovery in Database (KDD) digunakan untuk analisis data.

4.1. Selection

Sebelum memasuki tahap pencarian Informasi dalam proses KDD, Langkah yang penting yang harus dilakukan yaitu memilih data yang relevan dari berbagai data yang tersedia. Tahap ini dikenal

dengan Data Selection[10]. pada proses ini dilakukan penggalian informasi atribut yang dipilih untuk dari *dataset* adalah periode, Kategori Musik dan Kerajinan dapat dilihat pada tabel 1 berikut

Tabel 1. Hasil Data Selection

Periode	Musik	Kerajinan
Okt-22	63	1029
Nov-22	51	964
Des-22	46	752
Jan-23	58	530
Feb-23	64	1015
Mar-23	76	860
Apr-23	69	974
Mei-23	52	1298
Jun-23	71	1587
Jul-23	83	955
Agt-23	69	1158
Sep-23	57	1280
Oktr-23	66	1320

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 [12]. Pada tahap ini tidak dilakukan preprocessing data, karena tidak ada kesalahan pada data

4.2. Preprocessing

Dalam tahap ini dilakukan *Preprocessing* data yaitu proses *cleaning* data yang hilang atau kosong sebelum ke proses selanjutnya, karena data yang peneliti gunakan tidak terdapat data yang missing atau kosong maka tidak dilakukan penghapusan data yang kosong.

4.3. Transformation

Transformation bertujuan untuk mengurangi kerumitan data, memfasilitasi perubahan data, serta mencegah salah satu variabel bebas mendominasi variabel bebas lainnya yang dapat berdampak pada hasil prediksi [10]. 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 Transformation

No	Musik	Kerajinan
1	63	1029
2	51	964
3	46	752
4	58	530
5	64	1015
6	76	860
7	69	974
8	52	1298
9	71	1587
10	83	955
11	69	1158
12	57	1280
13	66	1320

4.4. Data Mining

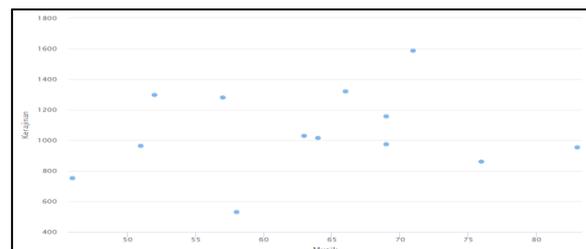
Dalam proses langkah di mana tugas Data mining yang sesuai dipilih sesuai dengan tujuan yang telah ditetapkan pada tahap awal untuk menggali informasi dari data [3]. Pada proses ini menggunakan aplikasi Rapidminer, dilakukan proses pemodelan pada setiap kategori Musik dan Kerajinan untuk menentukan nilai 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 Relative Error untuk mengetahui akurasi hasil prediksi. untuk memprediksi bulan kedepan. Pada proses ini dilakukan pemodelan algoritma regresi linear dengan mengubah nama atribut pada data. Beberapa operator yang digunakan adalah *Read Excel* berfungsi untuk membaca isi dari file *Microsoft Excel* yaitu data penjualan Musik lalu operator *Multiply* untuk mengambil objek dari *port input* atau objek dari operator *Read Excel* dan mengirimkan salinannya ke *port output* dihubungkan pada operator *Linear Regression*, pada tahap ini dilakukan perhitungan model regresi linear dari *input ExampleSet*, lalu setelah itu dihubungkan pada operator *Apply Model*, berfungsi untuk menerapkan model yang telah dilatih dari operator *Linear Regression* pada data *testing* dari output operator *Multiply*, lalu terakhir dihubungkan pada *port Result*.

4.5. Evaluation

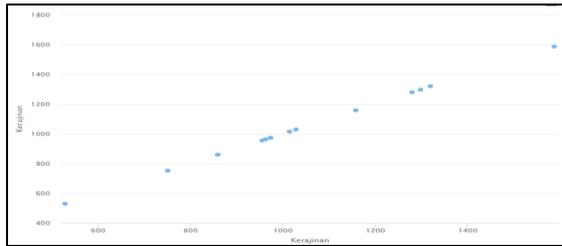
Diperlukannya informasi hasil dari data mining dalam format yang dapat dipahami dengan mudah oleh para pihak yang berkepentingan. Evaluation merupakan tahap penting dalam proses KDD yang melibatkan pemeriksaan apakah pola atau informasi yang ditemukan sesuai dengan fakta atau hipotesis yang telah ada sebelumnya ditetapkan dari hasil yang sudah diprediksi. Nilai Relative error bisa dikatakan baik apabila $Relative Error \leq 25\%$ dianggap akurat.

4.6. Knowledge

Berdasarkan prediksi penjualan produk dengan kategori musik dan kerajinan masing-masing kategori ditampilkan *plot scatter bubble* hasil prediksi algoritma regresi linear, berikut dapat dilihat dari gambar diberikut.

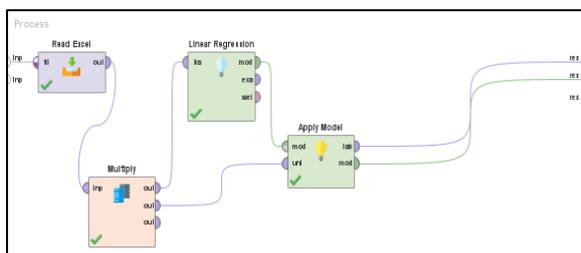


Gambar 2. Plot Scatter Musik



Gambar 3. Plot Scatter Kerajinan

Pengujian dari model algoritma Regresi Linear menggunakan tool Rapidminer dalam menentukan persamaan regresi linear, prediksi penjualan periode berikutnya dan menentukan hasil evaluasi dari hasil prediksi menggunakan *relative error*. Desainnya dapat dilihat dari gambar 4 berikut.



Gambar 5 Desain Menentukan persamaan

Berdasarkan proses perhitungan persamaan nilai a dan b didapatkan kategori musik $Y=55.58+ 1.13(X)$ dan kategori kerajinan $Y= 4.11+796 (X)$. hasil permodelan dapat dilihat pada gambar berikut.

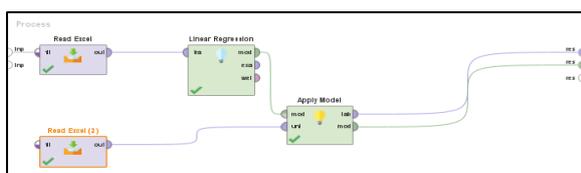
Attribute	Coefficient
x	1.126
(Intercept)	55.577

Gambar 6. Hasil Pemodelan Musik

Attribute	Coefficient
y	4.106
(Intercept)	794.958

Gambar 7. Hasil Pemodelan Kerajinan

Setelah persamaan didapatkan dilakukan prediksi penjualan produk pada 3 bulan berikutnya dengan nilai $X=12,13,14$. Desain prediksi data penjualan dapat dilihat dari gambar 8 berikut.



Gambar 8. Desain Prediksi Penjualan

Didapatkan kategori produk musik pada bulan pertama terjual 68, pada bulan kedua 69 dan bulan ketiga 70 sedangkan kategori produk kerajinan bulan pertama terjual 1078, pada bulan kedua 1029 dan bulan ketiga 1066. Hasil prediksi dapat dilihat pada gambar 9 dan 10 berikut.

Row No.	Musik	prediction(Musik)	x
1	69	67.967	11
2	57	69.093	12
3	66	70.220	13

Gambar 9. Hasil Prediksi Musik

Row No.	Kerajinan	prediction(K...	y
1	1158	1078.280	11
2	1280	1029.007	12
3	1320	1065.962	13

Gambar 10. Hasil Prediksi Kerajinan

Setelah didapatkan hasil prediksi maka dilakukan proses selanjutnya dengan evaluasi menggunakan indikator Penilaian *Relative Error* Hasil evaluasi *Relative Error* pada kategori Musik dan Kerajinan, pada nilai akurasi error pada kategori musik didapatkan *Relative Error* 13.6%. sedangkan Kerajinan nilai *Relative Error* 22.65%. Desain bisa dilihat pada gambar 11 dan 12 berikut.

```
relative_error
relative_error: 13.64% +/- 10.40%
```

Gambar 11. Nilai Relative Error Musik

```
relative_error
relative_error: 22.65% +/- 23.08%
```

Gambar 12. Nilai Relative Error Kerajinan

5. KESIMPULAN DAN SARAN

Pada Penelitian yang dilakukan menggunakan algoritma regresi linear untuk memprediksi penjualan produk kategori musik dan kerajinan di toko AgungOna9, dilakukan beberapa proses tahapan KDD (*Knowledge Discovery In Data Base*) dan didapatkan hasil untuk memprediksi 3 bulan selanjutnya yaitu bulan pertama musik 68 pcs, bulan dua 69 dan bulan tiga 70 pcs untuk kategori kerajinan bulan pertama 1078, bulan kedua 1029 dan bulan ketiga 1066. dengan nilai evaluasi untuk kategori musik 13% dan kerajinan 22,65% yang bisa dikatakan cukup atau bisa digunakan hasil dari persamaan regresi linear, karena hasil evaluasi *Relative Error* cukup mendekati.

DAFTAR PUSTAKA

[1] A. Azzanuddin and A. Calam, "Data Mining Untuk Memprediksi Penjualan Buah Sawit Menggunakan Metode Multi Regresi Pada PT. Karya Hevea Indonesia," *Jurnal Cyber Tech*, 2022, [Online]. Available:

- <https://ojs.trigunadharma.ac.id/index.php/jct/article/view/1923>
- [2] A. Arisandi and U. Ependi, "Analisis Peramalan Penjualan Produk Pada PT. Enseval Putera Megatrading TBK Menggunakan Metode Regresi Linear Sederhana," *JUPITER (Jurnal Penelitian Ilmu dan ...*, 2023, [Online]. Available: <https://jurnal.polsri.ac.id/index.php/jupiter/article/view/5318>
- [3] S. Lestari, "Analisis Algoritma Regresi Linear Sederhana dalam Memprediksi Tingkat Penjualan Album KPOP," *INSOLOGI: Jurnal Sains dan Teknologi*, 2023, [Online]. Available: <https://journal.literasisains.id/index.php/insologi/article/view/1692>
- [4] Y. Syakir, T. I. Hermanto, and ..., "Analisis Marketplace Shopee Untuk Memprediksi Penjualan dengan Algoritma Regresi Linier," *J-SAKTI (Jurnal Sains ...*, 2022, [Online]. Available: <http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/501>
- [5] I. Riyana, N. Suarna, and G. Dwilestari, "Analisa Dataset Penjualan Teh Menggunakan Algoritma Linear Regresi," *Jurnal ...*, 2023, [Online]. Available: <https://ejurnal.bangunharapanbangsa.com/index.php/jutek/article/view/194>
- [6] A. S. Sunge and A. T. Zy, "ANALISIS PREDIKSI PENJUALAN DENGAN METODE REGRESI LINEAR DI PT. EAGLE INDUSTRY INDONESIA," *Jurnal Informatika Teknologi dan Sains (Jinteks ...*, 2023, [Online]. Available: <http://www.jurnal.uts.ac.id/index.php/JINTEKS/article/view/3325>
- [7] V. R. Prasetyo, H. Lazuardi, and ..., "Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Regresi Linier," ... *Teknologi dan Sistem ...*, 2021, [Online]. Available: <http://repository.ubaya.ac.id/39606/>
- [8] M. L. Mu'tashim, T. Muhayat, and ..., "Analisis prediksi harga rumah sesuai spesifikasi menggunakan multiple linear regression," *Informatik: Jurnal ...*, 2021, [Online]. Available: <https://ejournal.upnvj.ac.id/informatik/article/view/3635>
- [9] M. Galih and P. D. Atika, "Prediksi Penjualan Menggunakan Algoritma Regresi Linear Di Koperasi Karyawan" Usaha Bersama," *Journal of Informatic and Information ...*, 2022, [Online]. Available: <https://ejurnal.ubharajaya.ac.id/index.php/jiforty/article/view/1354>
- [10] Y. A. Ramadhan, A. Faqih, and ..., "Prediksi Penjualan Handphone di Toko X menggunakan Algoritma Regresi Linear," *Jurnal Informatika ...*, 2023, [Online]. Available: <https://journal.nurulfikri.ac.id/index.php/jit/article/view/692>
- [11] F. H. Hamdanah and D. Fitrihanah, "Analisis Performansi Algoritma Linear Regression dengan Generalized Linear Model untuk Prediksi Penjualan pada Usaha Mikra, Kecil, dan Menengah," *Jurnal Nasional Pendidikan ...*, 2021, [Online]. Available: <https://ejournal.undiksha.ac.id/index.php/janapati/article/view/31035>
- [12] A. Anggrawan, H. Hairani, and ..., "Prediksi Penjualan Produk Unilever Menggunakan Metode Regresi Linear," *Jurnal ...*, 2022, [Online]. Available: <https://journal.universitasbumigora.ac.id/index.php/bite/article/view/2416>
- [13] E. P. A. Akhmad, "Data Mining Menggunakan Regresi Linear untuk Prediksi Harga Saham Perusahaan Pelayaran," *Jurnal Aplikasi Pelayaran dan ...*, 2020, [Online]. Available: <https://pdp-journal.hangtuah.ac.id/index.php/jurnal/article/view/13>
- [14] W. J. PUTRA, *Analisa Algoritma Regresi Linear dan Decision Tree Dalam Prediksi Penjualan Produk (Studi Kasus: Lookma Boutique)*. repository.mercubuana.ac.id, 2022. [Online]. Available: <https://repository.mercubuana.ac.id/70498/>