

PENERAPAN ALGORITMA REGRESI LINEAR UNTUK PREDIKSI PENJUALAN ALAT TULIS KANTOR (ATK) DI BUMDES

Alfia Hurifiani ¹, Ade Irma Purnamasari ², Irfan Ali ³

^{1,2} Teknik Informatika, STMIK IKMI Cirebon

³ Rekayasa Perangkat Lunak, STMIK IKMI Cirebon

Jl. Perjuangan No. 10 B Majasem Kec. Kesambi Kota Cirebon Tlp. 0231) 490480 - 490481

hurifianialfia@gmail.com

ABSTRAK

Pertumbuhan Badan Usaha Milik Desa (BUMDes) sebagai pilar pemberdayaan ekonomi di tingkat desa semakin berkembang pesat. Di antara berbagai kegiatan usaha BUMDes, penjualan Alat Tulis Kantor (ATK) muncul sebagai potensi yang signifikan. Untuk meningkatkan efisiensi dan produktivitas penjualan ATK, penerapan metode analisis prediksi menjadi esensial, terutama dengan menggunakan algoritma regresi linear. Penelitian ini dilakukan untuk menganalisis dan memprediksi penjualan ATK di BUMDes Desa Cintarasa melalui pendekatan kuantitatif. Data penjualan ATK selama beberapa bulan terakhir dihimpun, melibatkan variabel-variabel seperti jumlah produk terjual, harga ATK, dan faktor-faktor lain yang mempengaruhi penjualan. Analisis hasil menunjukkan adanya hubungan signifikan antara variabel yang diamati, memberikan gambaran akurat dampak masing-masing variabel terhadap penjualan ATK. Pengujian keakuratan prediksi menggunakan *Root Mean Square Error* (RMSE), *Absolute Error*, dan *Relative Error* memvalidasi hasil penelitian. Hasil prediksi total penjualan didapatkan pada hari pertama total penjualan adalah 55.435, pada hari kedua 28.272, dan hari ketiga 31.667. Hasil evaluasi nilai RMSE pada total penjualan 7802.802, nilai *absolute error* 6894.646, dan nilai *Relative Error* 21.05%. Prediksi penjualan alat tulis kantor (ATK) dengan menggunakan metode regresi linear dapat dikatakan cukup baik dengan Pengujian akurasi terhadap semua atribut yang menunjukkan bahwa nilai *Root Mean Square Error* (RMSE), *Absolute Error*, dan *Relative Error* memenuhi standar yang ditetapkan.

Kata kunci : BUMDes, Penjualan ATK, Regresi Linear, Analisis Prediksi Penjualan, RMSE (*Root Mean Squared Error*)

1. PENDAHULUAN

Di era pesatnya perkembangan bidang Informatika, dampak revolusi teknologi telah menciptakan paradigma baru di berbagai sektor kehidupan, meluas dari domain teknologi hingga bisnis, pendidikan, dan aspek lainnya. Penelitian ini bertujuan untuk menjelajahi penggunaan algoritma regresi linear dalam konteks prediksi penjualan Alat Tulis Kantor (ATK) di Badan Usaha Milik Desa (BUMDes), khususnya di Desa Cintarasa sebagai studi kasus. Tujuan penelitian ini mencakup pemahaman dan penerapan metode regresi linear sebagai instrumen efektif dalam meningkatkan manajemen penjualan ATK di tingkat desa. Dengan menghadapi tantangan dinamika ekonomi dan perubahan kebutuhan masyarakat lokal, penerapan regresi linear diharapkan dapat memberikan kontribusi positif pada pertumbuhan ekonomi dan keberlanjutan BUMDes.

Pola data pada BUMDes sangat kompleks, melibatkan berbagai faktor seperti fluktuasi musiman, ketidakpastian permintaan pasar, dan adopsi teknologi informasi. Penggunaan metode regresi linear menjadi relevan dalam mengatasi permasalahan ini, dengan fokus pada analisis data historis penjualan, pengumpulan data ekonomi desa, dan faktor-faktor sosial. Kendala akses data, fluktuasi musiman, faktor eksternal, dan kompleksitas analisis prediksi menjadi bagian integral dari pola data yang dihadapi. Pemilihan regresi linear sebagai metode penelitian dipertegas oleh kebutuhan untuk menangani tantangan tersebut,

mengoptimalkan akurasi prediksi penjualan ATK, dan memberikan panduan praktis bagi BUMDes dalam meningkatkan efisiensi penjualan. Hasil penelitian ini diharapkan memberikan kontribusi positif pada pengembangan metode analisis prediksi yang dapat diterapkan dalam konteks ekonomi lokal, menciptakan dasar untuk penelitian lanjutan, dan memperkaya pemahaman terhadap pengelolaan data di tingkat desa. Dengan demikian, penelitian ini tidak hanya relevan untuk pemahaman analisis prediksi penjualan ATK di BUMDes tetapi juga memiliki dampak lebih luas pada perkembangan teknologi informatika secara umum. Penelitian ini menjadi relevan dengan temuan penelitian sebelumnya yang menggunakan algoritma regresi linear dalam konteks prediksi penjualan, seperti pada produk elektronik, impor dan ekspor, penjualan teh, dan manajemen PT. *Eagle Industry* Indonesia. Tantangan yang dihadapi dalam penelitian ini mencakup kendala akses data, fluktuasi musiman, faktor eksternal, kompleksitas analisis prediksi, ketidakpastian permintaan pasar, dan adopsi teknologi informasi. Hasil penelitian ini dapat berkontribusi secara positif pada pengembangan metode analisis prediksi yang dapat digunakan dalam konteks ekonomi lokal, memberikan panduan praktis bagi BUMDes, meningkatkan efisiensi penjualan ATK, dan menguatkan ketahanan ekonomi desa. Proses pendekatan penelitian melibatkan analisis data historis penjualan, pengumpulan data ekonomi desa, dan faktor-faktor sosial, serta integrasi teknologi informasi

dan komputasi untuk mengelola data heterogen dan kompleks. Hasil temuan dapat menjadi dasar untuk penelitian lanjutan dan memberikan kontribusi nyata pada perkembangan teknologi informatika, membuka peluang baru dalam pengelolaan dan analisis data di tingkat lokal. Dengan demikian, penelitian ini tidak hanya memiliki dampak pada pemahaman analisis prediksi penjualan ATK di BUMDes tetapi juga memberikan sumbangan pada kemajuan teknologi informatika secara lebih luas.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Berdasarkan penelitian sebelumnya yang berjudul Prediksi Penjualan produk elektronik yang terlaris pada CV. ISTANA KOMPUTER Palembang dengan Menggunakan Algoritma Regresi Linear Sederhana, penelitian ini memanfaatkan Metodologi Regresi Linier untuk menerapkan Data Mining. Penghitungan dilakukan dengan menggunakan bantuan *tool* dari Rapidminer [1]. Menurut Penelitian sebelumnya dengan judul *Implementation of data mining to predict the value of indonesian oil and non-oil and gas import exports using the linear regression method*. Impor yang diprediksi data nilai nilai migas dan nonmigas bulan Juni sebesar 209.162.268, dan perkiraan data nilai ekspor migas dan nonmigas bulan Juni sebesar 349.285.781 dan nonmigas gas yang lebih banyak diperkirakan memiliki nilai tertinggi dibandingkan dengan nilai minyak dan gas yang masuk setiap bulan [2].

Selain itu, penelitian sebelumnya yang berjudul Analisa Dataset Penjualan Teh Menggunakan Algoritma Linear Regresi. Hasil analisis ini menunjukkan bahwa terdapat hubungan yang signifikan antara penjualan teh dan faktor-faktor seperti harga teh dan kegiatan promosi. Penurunan harga teh dan peningkatan upaya promosi tampaknya berdampak positif terhadap peningkatan volume penjualan. Namun, dalam dataset yang kami gunakan, faktor cuaca tidak dapat dianggap sebagai pengaruh yang signifikan terhadap penjualan teh. Selain itu, penerapan model regresi linear pada variabel-variabel yang dianalisis menunjukkan tingkat keakuratan yang baik dalam memprediksi penjualan teh. [3].

Penelitian sebelumnya dengan judul Analisis prediksi penjualan dengan metode regresi linear di PT. *Eagle Industry* Indonesia. Dengan menggunakan algoritma regresi linear, hasil tersebut menunjukkan bahwa variabel atau atribut yang digunakan dalam penelitian ini (produk, aktual, dan plan) tidak berdampak pada hasilnya. *Root Mean Squared Error* adalah 36241.241 +/- 0.000, dan *Squared Error* adalah 1313427569.481 +/- 5882150128.134. Hal Ini menunjukkan bahwa *squared Error* menghasilkan nilai yang tinggi dari *Root Mean Squared Error* [4]. Penelitian sebelumnya dengan Judul Implementasi Algoritma Regresi linear berganda untuk prediksi penjualan. Hasil yang didapatkan dari penelitian ini

menghasilkan jumlah prediksi penjualan barang untuk tahun 2020 sebesar 169715 barang [5].

Penelitian yang membahas mengenai Analisis Regresi Linear Sederhana dalam memprediksi tingkat penjualan album *KPOP*. Tujuan penelitian ini adalah untuk menyelidiki sejauh mana penjualan album K-pop memengaruhi masyarakat dari berbagai lapisan. Peneliti menggunakan metode data mining dengan menerapkan algoritma estimasi, khususnya Regresi Linear, untuk mengevaluasi dan meramalkan tingkat penjualan album *K-pop*. Pendekatan ini melibatkan perhitungan nilai X , Y , X^2 , Y^2 , dan XY , yang kemudian digunakan untuk menentukan koefisien dan variabel terikat dari setiap atribut atau label. Hasil penelitian menunjukkan persamaan regresi $Y = 8.08 + 5.15X$, dengan akurasi sebesar 10.86%. Penggunaan RapidMiner dalam analisis data menghasilkan performa vektor sebesar 0.089% untuk tingkat penjualan album "NCT127 *집주 2 Baddies*" [6].

Berdasarkan penelitian yang membahas mengenai Analisis *Marketplace Shopee* untuk memprediksi penjualan dengan algoritma Regresi Linier. Dengan menggunakan RapidMiner Studio 9.10, hasil pengujian menunjukkan bahwa nilai *Mean Square Error (MSE)* mencapai 5.172.628.212.404, *Root Mean Square Error (RMSE)* mencapai 2.274.341, dan *Mean Absolute Percentage Error (MAPE)* sebesar 4,34%. Berdasarkan MAPE, dapat disimpulkan bahwa akurasi algoritma regresi linier dalam memprediksi penjualan *Butik Ariqa Collection* di *platform Shopee* sangat tinggi [7].

Menurut Penelitian yang membahas mengenai Penerapan Algoritma Regresi Linier Berganda pada estimasi penjualan sepeda motor. Metode ini dipilih karena mampu membuat suatu estimasi dengan memanfaatkan data-data mengenai penjualan. Hasil dari estimasi (perkiraan) penjualan sepeda motor manual tahun 2021 dengan bulan januari adalah sebanyak 56,941 atau 57 sepeda motor dengan kategori manual. Artinya terdapat penambahan jumlah sepeda motor manual sebanyak 5 motor, sedangkan hasil sampai bulan Mei 2021 berjumlah 65,710 motor. Sehingga dapat disimpulkan bahwa penjualan sepeda motor di CV. Kerinci Motor mengalami kenaikan penjualan pada 5 bulan kedepan [8].

Penelitian yang membahas mengenai Penerapan metode linear regresi dalam prediksi produksi pia. tujuan dari penelitian ini untuk melakukan prediksi penjualan pia dengan teknik data mining dan untuk mengetahui tingkat keakuratan metode linier regresi dalam memprediksi penjualan pia dengan menggunakan metode linier regresi. Hasil dari penelitian ini Dalam melakukan prediksi produksi pia dengan menggunakan metode linier regresi dan Tingkat keakuratan metode linier regresi dalam memprediksi produksi pia sebesar 5.775% [9].

Penelitian yang dilakukan oleh (Ramadhan et al., 2023) membahas mengenai Prediksi penjualan *handphone* di toko X menggunakan algoritma regresi linear menunjukkan hasil yang positif. Pada kategori

Entry, terdapat penjualan sebanyak 84 unit pada bulan pertama, 86 unit pada bulan kedua, dan 88 unit pada bulan ketiga. Sementara itu, pada kategori *Mid*, terjual 28 unit pada bulan pertama, 29 unit pada bulan kedua, dan 30 unit pada bulan ketiga. Evaluasi hasil menunjukkan nilai RMSE untuk kategori *Entry* sebesar 10.36, dengan *Relative Error* sebesar 19.11%, sedangkan untuk kategori *Mid*, nilai RMSE mencapai 7.50, dengan *Relative Error* sebesar 32.97%. Dengan demikian, dapat disimpulkan bahwa prediksi penjualan *handphone* menggunakan regresi linear ini dapat dianggap cukup baik dan layak digunakan. [10].

Berdasarkan Penelitian yang membahas mengenai Analisis metode Regresi Linear Berganda dan *ordinary Least Squared* dalam mengelola manajemen operasional penjualan MPM motor dealer di Lombok timur. Menanggapi hal tersebut, dilakukan penelitian dengan menggunakan metode Regresi Linier Berganda dan *Ordinary Least Squared* untuk membantu tim perusahaan dalam memprediksi pendapatan penjualan berdasarkan fluktuasi rata-rata transaksi barang dalam suatu periode tertentu. Metode Regresi Linier Berganda digunakan untuk menganalisis hubungan linier antara dua atau lebih variabel independen, sementara *Ordinary Least Squared* digunakan sebagai pendekatan dalam analisis koefisien regresi. Hasil penelitian ini mencakup pembentukan model persamaan regresi untuk proses estimasi dan koefisien regresi yang diperoleh selama uji coba. Dari hasil penelitian ini dapat disimpulkan bahwa model regresi yang dihasilkan melalui kombinasi metode Regresi Linier Berganda dan *Ordinary Least Squared* dapat efektif digunakan untuk mengestimasi jumlah penjualan [11].

Menurut Penelitian yang membahas mengenai Analisis Performansi Algoritma Linear *Regression* dengan *Generalized* linear model untuk prediksi penjualan pada usaha mikro, kecil, dan menengah. Dari hasil perhitungan, terlihat bahwa algoritma Regresi Linier memiliki nilai RSME, MSE, dan MAPE berturut-turut sebesar 1,983; 3,933; dan 1,518. Sementara itu, algoritma *Generalized* Linear Model memberikan nilai RSME, MSE, dan MAPE berturut-turut sebesar 4,827; 23,295; dan 3,882. Berdasarkan prediksi yang dihasilkan oleh algoritma Regresi Linier dan *Generalized* Linear Model, dapat disimpulkan bahwa nilai RSME pada algoritma Regresi Linier menunjukkan perhitungan yang paling baik, karena memiliki nilai RSME yang paling rendah [12].

Menurut penelitian yang membahas mengenai Analisis Peramalan Penjualan Produk Pada PT. Enseval Putera Megatrading TBK Menggunakan Metode Regresi Linear Sederhana. Penelitian ini bertujuan untuk melakukan prediksi produk farmasi di PT. Enseval Putera Megatrading Tbk Palembang, yang nantinya dapat digunakan sebagai panduan dalam pengambilan keputusan terkait perencanaan dan pengendalian persediaan produk farmasi di perusahaan tersebut. [13].

Menurut penelitian yang membahas mengenai Penerapan data mining untuk memprediksi penjualan kain tenun menggunakan regresi linear studi kasus: UD. Bintang Remawe Sukarare. Penelitian ini akan memprediksi penjualan produk kain tenun dengan menggunakan data penjualan masa lalu dan menerapkan model Regresi Linier. Untuk menguji kinerja algoritma, digunakan metrik evaluasi seperti *Mean Square Error (MSE)*, *Mean Square Error (MSE)*, dan *Root Mean Square Error (RMSE)*. Hasil dari pemodelan regresi linear menunjukkan skor sebesar 0.8041320270845731. Sedangkan hasil pengujian menunjukkan bahwa nilai *Mean Square Error (MSE)* mencapai 47.377, dan nilai *Root Mean Square Error (RMSE)* sebesar 6.883125. Sementara itu, skor MAE (*Mean Absolute Error*) diperoleh sebesar 3.373572 [14].

Menurut penelitian yang membahas mengenai Perbandingan penerapan estimasi metode regresi linier menggunakan rapidminer dan Ms. Excel studi kasus penjualan pada *public* dataset. Tahap estimasi naik volume penjualan dimulai dengan menentukan dataset yang diambil dari sumber data publik di <https://www.kaggle.com/>. Selanjutnya, menggunakan metode data mining untuk estimasi, algoritma regresi linier diterapkan dengan menggunakan perangkat lunak MS Excel dan RapidMiner, masing-masing menghasilkan persamaan regresi linear ($0,053 * tv + 0,224 * radio + 0,023 * newspaper$) dari proses RapidMiner dan persamaan ($3,314 + (0,043 * tv) + (0,191 * radio) + (0,002 * newspaper)$) dari proses MS Excel. Penerapan pada data uji menghasilkan nilai RMSE (*root mean squared error*) sebesar 1,744 pada RapidMiner dan 3,277 pada MS Excel, dengan menggunakan data pelatihan sebanyak 160 dan data uji sebanyak 40, atau perbandingan 8:2. Dengan demikian, karena nilai RMSE yang dihasilkan oleh RapidMiner lebih kecil dibandingkan dengan hasil perhitungan MS Excel, dapat disimpulkan bahwa penggunaan perangkat lunak RapidMiner lebih cocok daripada MS Excel untuk tujuan ini. [15].

Berdasarkan penelitian yang membahas mengenai implementasi algoritma regresi linier berganda untuk prediksi penjualan. Minimarket merupakan tempat atau toko yang menyediakan berbagai kebutuhan pokok masyarakat dan umumnya menggunakan teknologi atau mesin untuk melakukan transaksi pembayaran. Salah satu tantangan yang sering dihadapi oleh minimarket adalah ketidakpastian terkait jumlah penjualan, dan hal tersebut juga menjadi permasalahan yang dihadapi oleh Minimarket Sr. Penelitian ini memiliki tujuan untuk memproyeksikan jumlah penjualan barang dengan menggunakan algoritma regresi linier berganda sebagai metode prediksi yang diterapkan. Hasil dari penelitian ini menunjukkan bahwa prediksi jumlah penjualan barang untuk tahun 2020 adalah sebanyak 169.715 barang. [16].

2.2. Rumus Regresi Linear Sederhana

Analisis regresi linear sederhana bergantung pada hubungan kausal atau fungsional antara satu variabel independen dengan variabel dependen. Koefisien regresi bertujuan untuk menentukan apakah variabel independen yang terdapat dalam persamaan regresi secara individual memiliki pengaruh terhadap nilai variabel dependen. Analisis regresi linear sederhana berperan dalam menguji hubungan sebab-akibat antara variabel yang menjadi faktor penyebab dengan variabel yang menjadi akibatnya.

Analisis regresi sederhana merupakan suatu metode pendekatan untuk memodelkan hubungan antara satu variabel dependen dan satu variabel independen. Dalam kerangka regresi, variabel independen digunakan untuk menjelaskan variasi dalam variabel dependen. Dalam analisis regresi sederhana, terdapat asumsi bahwa perubahan dalam variabel independen (X) akan mengakibatkan perubahan pada variabel dependen (Y) secara konsisten. Namun, pada hubungan yang bersifat non-linier, perubahan dalam variabel X tidak diikuti oleh perubahan variabel Y secara proporsional. Persamaan regresi linier sederhana ditunjukkan pada rumus berikut:

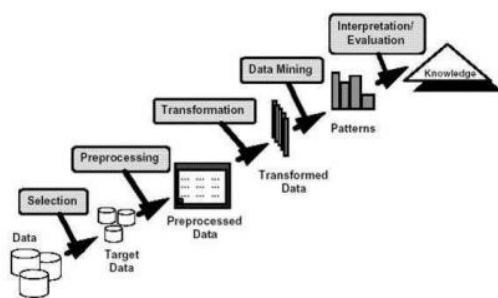
$$Y = a + bX$$

Dimana:

- a : konstanta (titik potong Y)
- b : koefisien dari variabel X
- Y : variabel dependen
- X : variabel independen

3. METODE PENELITIAN

Pada penelitian ini, analisis data dilakukan menggunakan tahapan proses KDD (*Knowledge Discovery In Database*) dan Algoritma Regresi Linear. Penerapan proses KDD bertujuan untuk memastikan bahwa langkah-langkah yang dilakukan berjalan secara terstruktur dan sistematis.



Gambar 1. Alur KDD

Berikut tahapan penelitian yang dilakukan menggunakan proses KDD:

3.1. Selection

Selection adalah proses analisis data yang relevan dari database karena seringkali ditemukan bahwa tidak semua data diperlukan dalam tahap data mining. Pada tahap ini, data yang digunakan sebagai penelitian

adalah data yang diambil terbaru selama 5 bulan penjualan lalu dipilih atribut sesuai yang dibutuhkan.

3.2. Preprocessing

Preprocessing merupakan bagian dari persiapan data sebelum memulai proses data mining. Data yang mengalami *preprocessing* biasanya melibatkan langkah-langkah seperti pembersihan, reduksi, dan integrasi. Sebelum melibatkan proses prediksi, langkah *preprocessing* diperlukan terlebih dahulu, termasuk penghapusan data yang tidak diperlukan dan penanganan data yang hilang atau missing.

3.3. Transformation

Pada tahap ini, terjadi transformasi data ke format yang tepat untuk kemudian diproses oleh algoritma data mining. Hal ini bertujuan untuk mempermudah konsistensi data yang diolah oleh algoritma dan alat yang digunakan dalam penelitian.

3.4. Data Mining

Data mining adalah suatu proses pengolahan data yang mengikuti algoritma yang sesuai dengan langkah-langkah pada proses data mining. Dalam penelitian ini, algoritma yang diterapkan adalah Regresi Linear untuk melakukan prediksi terhadap penjualan Alat Tulis Kantor (ATK).

3.5. Evaluation

Evaluasi digunakan dalam tahap ini untuk mengevaluasi hasil yang diperoleh dari penerapan Algoritma Regresi Linear. Tujuannya adalah untuk menarik kesimpulan dari proses data mining serta mengidentifikasi pola berdasarkan tindakan yang telah diambil.

3.6. Knowledge

Tahap knowledge ini menampilkan informasi yang dihasilkan dari proses prediksi, mempresentasikan hasil sehingga data tersebut menjadi lebih mudah untuk dipahami.

4. HASIL DAN PEMBAHASAN

4.1. Data Selection

Pada tahap seleksi data, data yang akan digunakan untuk prediksi dipilih menggunakan *tools* RapidMiner. Proses seleksi ini dilakukan melalui *Microsoft Excel*, di mana dalam penelitian ini, data penjualan bulan Mei hingga September 2023 dipilih untuk digunakan dalam proses data mining guna memprediksi penjualan alat tulis kantor (ATK). Hasil dari seleksi data diketahui pada data penjualan terdapat 5 atribut yaitu atribut Hari dan tanggal, nama barang, harga, total penjualan, dan total pembelian.

4.2. Data Preprocessing

Pada tahap Data *Preprocessing*, dilakukan proses Pembersihan Data yang mencakup identifikasi dan penanganan data yang hilang atau kosong sebelum memasuki tahap selanjutnya. Karena dataset yang

digunakan tidak mengandung data yang hilang atau kosong, maka tidak perlu dilakukan penghapusan data yang kosong.

4.3. Data Transformation

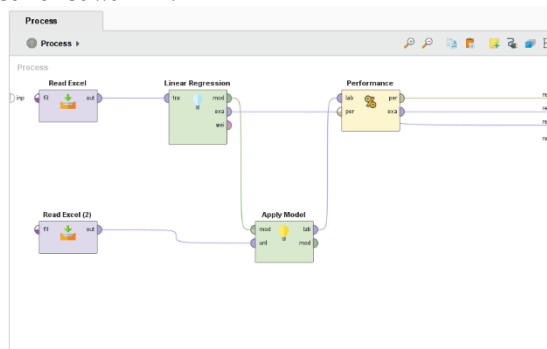
Pada tahap ini, dilakukan proses transformasi data menjadi format yang tepat untuk dapat diproses dalam pemodelan data mining. Bertujuan untuk memudahkan koordinasi data yang diproses oleh algoritma dan *tools* yang digunakan dalam penelitian yaitu rapidminer. Pada tahap transformasi pada data penjualan terdapat 4 atribut yaitu no, tanggal, total penjualan seperti yang terlihat pada tabel sebagai berikut.

Tabel 1. Data Transformation

No	TANGGAL	TOTAL PENJUALAN (Y)
1	01-May-23	Rp 57.500,00
2	02-May-23	Rp 33.500,00
3	03-May-23	Rp 171.000,00
4	05-May-23	Rp 26.000,00
5	08-May-23	Rp 38.000,00
6	10-May-23	Rp 72.000,00
7	11-May-23	Rp 40.000,00
8	15-May-23	Rp 16.000,00
9	16-May-23	Rp 29.500,00
10	17-May-23	Rp 21.000,00
11	19-May-23	Rp 15.000,00
12	22-May-23	Rp 19.000,00
13	23-May-23	Rp 14.000,00
...		
88	26-Sep-23	Rp 23.000,00
89	27-Sep-23	Rp 45.000,00
90	28-Sep-23	Rp 60.000,00

4.4. Data mining (Pengolahan Data dengan RapidMiner)

Proses data mining yang diterapkan dalam penelitian ini yaitu prediksi menggunakan regresi linear. Pada penelitian ini, akan dimuat perhitungan prediksi dari total penjualan per hari dan nilai evaluasi hasil prediksi diukur dengan RMSE, *Absolute Error*, dan *Relative Error*. Setelah data diolah, perhitungan akan dilakukan menggunakan RapidMiner. Proses dimulai dengan mengonfigurasi data ke dalam RapidMiner, yang melibatkan penyesuaian dataset dan pembentukan model, sebagaimana terlihat pada gambar di bawah ini.



Gambar 2. Data mining

Dataset dibagi menjadi dua, yaitu *training*, dan *testing*. Dataset yang ada dalam *testing* inilah akan diolah sehingga menghasilkan keluaran yang bisa dijadikan sebagai hasil prediksi. Pada proses prediksi yang akan ditunjukkan pada gambar 4.1 akan dilakukan pengujian data *testing* dan *training* 90% dan 10% dengan penambahan 5 tahapan proses seperti berikut :

1. Dataset Input Training 90%

Pada dataset *input training* merupakan 90% dari data yang telah diolah dan dilakukan transformasi, dataset *training* dapat dilihat pada gambar dibawah ini.

Open in Turbo Prep Auto Model

Row No.	TOTAL PENJ...	TANGGAL	TOTAL PEM...
1	57500	May 1, 2023	9
2	33500	May 2, 2023	10
3	171000	May 3, 2023	13
4	26000	May 5, 2023	13
5	38000	May 8, 2023	7
6	72000	May 10, 2023	25
7	40000	May 11, 2023	11
8	16000	May 15, 2023	4
9	29500	May 16, 2023	5
10	21000	May 17, 2023	5
11	15000	May 19, 2023	5
12	19000	May 22, 2023	3
13	14000	May 23, 2023	3
14	25000	May 24, 2023	3
15	17000	May 25, 2023	6

ExampleSet (81 examples, 1 special attribute, 2 regular attributes)

Gambar 3. Dataset Input Training 90%

2. Dataset Input Testing 10%

Pada dataset *input testing* merupakan 10% dari data yang telah diolah dan dilakukan transformasi, dataset *testing* dapat dilihat pada gambar dibawah ini:

Open in Turbo Prep Auto Model

Row No.	TOTAL PENJ...	TANGGAL	TOTAL PEM...
1	44500	Sep 18, 2023	15
2	40000	Sep 19, 2023	7
3	22000	Sep 20, 2023	8
4	35000	Sep 21, 2023	7
5	33000	Sep 22, 2023	7
6	32000	Sep 25, 2023	8
7	23000	Sep 26, 2023	8
8	45000	Sep 27, 2023	14
9	60000	Sep 28, 2023	17

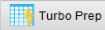
Gambar 4. Dataset Input Testing 10%

3. Linear Regrssion

Pada tahap *linear regression*, data tersebut akan diproses melalui metode regresi linear.

4. Apply model

Apply model merupakan tahap terakhir di mana model data mining regresi linear diterapkan berdasarkan proses-proses sebelumnya, termasuk penggunaan *data training*, *data testing*, dan regresi linear. Setelah melalui seluruh proses, hasil model prediksi dari data testing akan diperoleh, mencapai sekitar 10%, seperti terlihat pada gambar di bawah ini.

Open in  Turbo Prep  Auto Model

Row No.	TOTAL PENJ...	prediction(T...	TANGGAL	TOTAL PEM...
1	44500	55435.439	Sep 18, 2023	15
2	40000	28272.607	Sep 19, 2023	7
3	22000	31667.961	Sep 20, 2023	8
4	35000	28272.607	Sep 21, 2023	7
5	33000	28272.607	Sep 22, 2023	7
6	32000	31667.961	Sep 25, 2023	8
7	23000	31667.961	Sep 26, 2023	8
8	45000	52040.085	Sep 27, 2023	14
9	60000	62226.147	Sep 28, 2023	17

Gambar 5. Apply model

5. Performance

Performance vector hasil dari prediksi total alat tulis kantor (ATK) menggunakan algoritma regresi linear dengan memberikan gambaran tentang seberapa baik model tersebut dapat memprediksi nilai penjualan. Dalam konteks ini, dapat diasumsikan bahwa hasil prediksi dievaluasi menggunakan beberapa metrik kinerja.

PerformanceVector

```
PerformanceVector:
root_mean_squared_error: 7802.802 +/- 0.000
absolute_error: 6894.646 +/- 3653.433
relative_error: 21.05% +/- 13.58%
```

Gambar 6. Hasil PerformanceVector

Dalam evaluasi kinerja model, beberapa metrik penting digunakan, seperti *Root Mean Squared Error (RMSE)*, *Absolute Error*, dan *Relative Error*. Untuk RMSE, nilai yang dicapai adalah sebesar 7802.802 dengan deviasi standar 0.000, mengindikasikan tingkat konsistensi yang tinggi dan potensi kinerja model yang baik. *Absolute Error*, yang mengukur selisih *absolute* antara prediksi dan nilai seharusnya, mencapai nilai 6894.646 dengan deviasi standar 3653.433, menunjukkan performa yang lebih baik ketika nilai error semakin rendah. Sementara itu, *Relative Error*, yang diukur dalam *persentase*, menghasilkan persentase *relative error* sebesar 21.05% dengan deviasi standar 13.58%, memberikan gambaran tentang seberapa besar model dapat bervariasi dalam memprediksi nilai dalam bentuk *persentase* dari nilai sebenarnya.

4.5. Evaluation

Hasil evaluasi nilai RMSE pada pada kategori Total Penjualan 7802.802, nilai *absolute Error* 6874.646, dan nilai *Relative Error* 21,05%

4.6. Knowledge

Pada tahapan yang telah dilakukan bahwa sesuai prediksi data penjualan alat tulis kantor (ATK) di BUMDes menggunakan algoritma regresi linear dengan proses tahapan KDD. Berdasarkan hasil validasi prediksi penjualan alat tulis kantor (ATK) dalam setiap kategori, *Plot Scatter Bubble* dan *plot line* dari prediksi algoritma regresi linear menunjukkan bahwa prediksi penjualan ATK memiliki nilai yang tidak stabil. Bisa dilihat di gambar dibawah ini:



Gambar 7. Plot Line Prediksi Total penjualan ATK

Pada data penjualan alat tulis kantor (ATK) yang disajikan di atas, tercatat bahwa pada tanggal 18 September 2023, total penjualan mencapai 44,500, sedangkan pada tanggal 19, 20, 21, 22, 25, 26, 27, dan 28 September 2023, total penjualan berturut-turut adalah 40,000, 22,000, 35,000, 33,000, 32,000, 23,000, 45,000, dan 6,000. Selanjutnya, berdasarkan prediksi, penjualan diperkirakan akan mencapai 55,435, 28,272, 31,667, 28,272, 28,272, 31,667, 31,667, 52,040, dan 62,226 dalam 12 hari mendatang. Prediksi ini memberikan gambaran perkiraan tingkat penjualan harian yang dapat digunakan sebagai panduan untuk perencanaan dan analisis lebih lanjut terkait kinerja penjualan ATK dalam periode tersebut.

4.7. Penerapan prediksi penjualan alat tulis kantor (ATK) menggunakan algoritma regresi linear

Pada penelitian ini dilakukan penerapan prediksi penjualan alat tulis kantor (ATK) di BUMDes di Desa Citarasa menggunakan algoritma regresi linear, metode analisa yang digunakan adalah KDD (*Knowledge Discovery in Database*) dan tools yang digunakan adalah RapidMiner. Tahapan analisa terdiri dari beberapa tahap yaitu *Selection*, *Preprocessing*, *Transformation*, *Data mining*, *Evaluation* dan *Knowledge*. Pada tahap *data mining* diterapkan algoritma regresi linear untuk pemodelan prediksi penjualan alat tulis kantor (ATK) secara spesifik atau satu persatu disetiap kategori ATK yang sudah diseleksi pada tahap *selection*, yaitu menjadi kategori

no, total penjualan (dan total pembelian, setelah dilakukan prediksi dari setiap atribut, maka didapatkan hasil prediksi penjualan. $Y = 34.246,52 + (-81,20)(4005)$ atau $Y = 34.246,52 - 81,20(4005)$.

4.8. Hasil evaluasi RMSE, Absolute Error, dan Relative Error pada prediksi menggunakan algoritma regresi linear

Hasil dari proses menggunakan algoritma regresi linear di peroleh model prediksi Total penjualan $Y = 34.246,52 + (-81,20)(X)$ atau $Y = 34.246,52 - 81,20(X)$. Model prediksi akan dinilai menggunakan tiga indikator evaluasi, yaitu *Root Mean Square Error* (RMSE), *Absolute Error* dan *Relative Error*. Hasil yang diperoleh dari model tersebut, pada kategori total penjualan nilai RMSE yang diperoleh adalah 7802.802, nilai absolute error 6894.646, dan nilai *Relative Error* 21,05 %. Berdasarkan penelitian, Apabila *Relative Error* $\leq 25\%$, dianggap sebagai tingkat akurasi yang baik. Dalam hal nilai yang $> 25\%$ tetapi $\leq 50\%$, mungkin dapat dipertimbangkan untuk digunakan, sedangkan perkiraan dengan *Relative Error* $> 50\%$ dianggap tidak akurat [10]. Berdasarkan penelitian terdahulu hasil RMSE dan *Relative Error* prediksi penjualan alat tulis kantor (ATK) di BUMDes termasuk model cukup, artinya adalah model yang dapat digunakan untuk memprediksi penjualan di hari berikutnya.

4.9. Hasil evaluasi RMSE, Absolute Error, dan Relative Error pada prediksi menggunakan algoritma regresi linear untuk menentukan perhitungan hari berikutnya

Setelah mengetahui nilai evaluasi menunjukkan bahwa model termasuk dalam kategori cukup, hal ini menandakan bahwa model tersebut bisa digunakan untuk memprediksi penjualan Alat tulis kantor (ATK) pada periode selanjutnya. Pada penelitian ini akan di prediksi hari berikutnya. Pada beberapa hari berikutnya penjualan kategori total penjualan nilai $X = 80, 81, 82$, dari setiap nilai X yang diketahui, dilakukan perhitungan dengan model yang dihasilkan pada total penjualan adalah $Y = 34.246,52 + (-81,20)(X)$ atau $Y = 34.246,52 - 81,20(X)$, didapatkan hasil pada hari pertama total penjualan adalah 55.435, pada hari kedua 28.272, dan hari ketiga 31.667.

5. KESIMPULAN DAN SARAN

Dari hasil penelitian yang telah dilakukan, melalui proses prediksi menggunakan atribut total penjualan, diperoleh hasil prediksi untuk beberapa hari berikutnya berdasarkan data penjualan Alat Tulis Kantor (ATK) di Badan Usaha Milik Desa (BUMDes) menggunakan algoritma Regresi Linear. Dalam rangkaian hasil tersebut, beberapa simpulan dapat diambil, yaitu: Pada hari pertama penjualan, total penjualan mencapai 55.435, diikuti oleh 28.272 pada hari kedua, dan 31.667 pada hari ketiga. Dalam evaluasi kategori total penjualan, akurasi error diukur menggunakan RMSE, *Absolute Error*, dan *Relative*

Error. Hasilnya menunjukkan RMSE sebesar 7,802.802 +/- 0.000, *Absolute Error* sebesar 6,894.646 +/- 3,653.433, dan *Relative Error* sebesar 21.05% +/- 13.58%. Dalam konteks ini, nilai hasil prediksi dianggap baik. Ini mengindikasikan bahwa prediksi total penjualan alat tulis kantor (ATK) di BUMDes cukup akurat dan dapat diandalkan.

Berdasarkan hasil penelitian, terdapat beberapa saran sebagai berikut: Untuk meningkatkan penjualan alat tulis kantor (ATK) di BUMDes Cintarasa, diperlukan strategi seperti analisis pasar, pengembangan promosi, manajemen stok efektif, pemanfaatan teknologi, dan peningkatan kualitas layanan. Penelitian dapat di kembangkan lagi dengan menggunakan algoritma *forecasting* yang lainnya, seperti *Neural Network*, *Naïve Bayes*, *Decision Tree*, dll. Pada penelitian selanjutnya bisa menggunakan dataset beberapa tahun sebelumnya .

DAFTAR PUSTAKA

- [1] R. D. Putri, "PREDIKSI PENJUALAN PRODUK ELEKTRONIK YANG TERLARIS PADA CV. ISTANA KOMPUTER PALEMBANG MENGGUNAKAN ALGORITMA REGRESI LINEAR ...," *J. Mantik*, 2022, [Online]. Available: <http://iocscience.org/ejournal/index.php/mantik/article/view/2798>
- [2] E. S. Ompusunggu, W. Sinaga, and ..., "IMPLEMENTATION OF DATA MINING TO PREDICT THE VALUE OF INDONESIAN OIL AND NON-OIL AND GAS IMPORT EXPORTS USING THE LINEAR ...," ... *Sist. Inf. dan ...*, 2023, [Online]. Available: <http://jurnal.unprimdn.ac.id/index.php/JUSIKO M/article/view/4081>
- [3] I. Riyana, N. Suarna, and G. Dwilestari, "Analisa Dataset Penjualan Teh Menggunakan Algoritma Linear Regresi," *J. ...*, 2023, [Online]. Available: <https://ejurnal.bangunharapanbangsa.com/index.php/jutek/article/view/194>
- [4] A. S. Sunge and A. T. Zy, "ANALISIS PREDIKSI PENJUALAN DENGAN METODE REGRESI LINEAR DI PT. EAGLE INDUSTRY INDONESIA," *J. Inform. Teknol. dan Sains (Jinteks ...)*, 2023, [Online]. Available: <http://www.jurnal.uts.ac.id/index.php/JINTEKS/article/view/3325>
- [5] D. Tampubolon, D. Saripurna, and A. Azlan, "Implementasi Regresi Linier Berganda Untuk Memprediksi Tingkat Penjualan Alat Kelistrikan," *J. Cyber Tech*, 2020, [Online]. Available: <http://ojs.trigunadharma.ac.id/index.php/jct/article/view/3070>
- [6] S. Lestari, "Analisis Algoritma Regresi Linear Sederhana dalam Memprediksi Tingkat Penjualan Album KPOP," *INSOLOGI J. Sains dan Teknol.*, 2023, [Online]. Available: <https://journal.literasisains.id/index.php/insologi>

- /article/view/1692
- [7] 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>
- [8] E. Rahayu, I. Parlina, and Z. A. Siregar, "Penerapan Algoritma Regresi Linier Berganda Pada Estimasi Penjualan Sepeda Motor Application of Multiple Linear Regression Algorithm for Motorcycle..., " *Jupiter. download.garuda.kemdikbud.go.id*, 2022, [Online]. Available: <http://download.garuda.kemdikbud.go.id/article.php?article=2998591&val=27036&title=Application of Multiple Linear Regression Algorithm for Motorcycle Sales Estimation>
- [9] A. R. K. Haba, "Penerapan Metode Linier Regresi dalam Prediksi Produksi Pia," *J. Sist. dan Tek. Komput.*, 2021, [Online]. Available: <http://ejournal.caturakti.ac.id/index.php/simtek/article/view/112>
- [10] Y. A. Ramadhan, A. Faqih, and ..., "Prediksi Penjualan Handphone di Toko X menggunakan Algoritma Regresi Linear," *J. Inform. ...*, 2023, [Online]. Available: <https://journal.nurulfikri.ac.id/index.php/jit/article/view/692>
- [11] L. Erwin, K. Kartarina, R. Anggriani, and ..., "Analisis Metode Regresi Linier Berganda Dan Ordinary Least Squared Dalam Mengelola Manajemen Operasional Penjualan MPM MOTOR Dealer di Lombok Timur," *JISAMAR (Journal ...)*, 2023, [Online]. Available: <http://journal.stmikjayakarta.ac.id/index.php/jisamar/article/view/1016>
- [12] F. H. Hamdanah and D. Fitrianah, "Analisis Performansi Algoritma Linear Regression dengan Generalized Linear Model untuk Prediksi Penjualan pada Usaha Mikra, Kecil, dan Menengah," *J. Nas. Pendidik. ...*, 2021, [Online]. Available: <https://ejournal.undiksha.ac.id/index.php/janapati/article/view/31035>
- [13] A. Arisandi and U. Ependi, "Analisis Peramalan Penjualan Produk Pada PT. Enseval Putera Megatrading TBK Menggunakan Metode Regresi Linear Sederhana," *JUPITER (Jurnal Penelit. Ilmu dan ...)*, 2023, [Online]. Available: <https://jurnal.polsri.ac.id/index.php/jupiter/article/view/5318>
- [14] F. Widia and W. Murniati, "Penerapan Data Mining Untuk Memprediksi Penjualan Kain Tenun Mnggunakan Regresi Linear," *J. Ilm. Tek. Mesin ...*, 2022, [Online]. Available: <https://ejurnal.stie-trianandra.ac.id/index.php/JURITEK/article/view/284>
- [15] M. Yasin, "... PENERAPAN ESTIMASI METODE REGRESI LINIER MENGGUNAKAN RAPIDMINER DAN MS. EXCEL STUDI KASUS PENJUALAN PADA PUBLIC DATASET," *J. LENTERA ICT*, 2023, [Online]. Available: <https://www.plj.ac.id/ojs/index.php/jrict/article/view/874>
- [16] Y. Asohi and A. Andri, "Impelementasi Algoritma Regresi Linier Berganda Untuk Prediksi Penjualan," *J. Nas. Ilmu Komput.*, 2020, [Online]. Available: <http://journal.jis-institute.org/index.php/jnik/article/view/161>