

PENERAPAN ALGORITMA REGRESI LINEAR PADA DATA HARGA CABAI RAWIT DI PASAR INDIHIANG

Yusuf Ikhwan Fahrudin, Rudi Kurniawan, Yudhistira Arie Wijaya

¹Teknik Informatika, STMIK IKMI Cirebon

²Rekayasa Perangkat Lunak, STMIK IKMI Cirebon

³Sistem Informasi, STMIK IKMI Cirebon

Jl. Perjuangan No.10B, Karyamulya, Kesambi, Kota Cirebon, Jawa Barat, 45131, Indonesia

yusufikhwanfahrudin@gmail.com

ABSTRAK

Sayuran, termasuk cabai rawit, memegang peran vital dalam perekonomian nasional. Meningkatnya harga cabai rawit tidak hanya dipengaruhi oleh peningkatan permintaan selama hari-hari besar keagamaan seperti Bulan Ramadhan, Hari Raya Idul Fitri, Hari Raya Idul Adha, khususnya jika terjadi gangguan cuaca atau pasokan. Penelitian ini bertujuan untuk memprediksi harga cabai rawit di Pasar Indihiang dan mencari atribut yang mempengaruhi nilai *Root Mean Squared Error* (RMSE) terbaik. Algoritma regresi linear terkenal karena kemampuannya menganalisis korelasi data dan memberikan prediksi yang akurat. Data harga cabai rawit dari 2021 hingga 2022 mencakup informasi harga, bulan, cuaca, pasokan, supplier, dan nama pedagang. Melalui analisis Regresi Linear menggunakan *Cross-Industry Standard Process for Data Mining* (CRISP-DM) dan RapidMiner versi 10.1, penelitian ini membangun model matematis untuk memprediksi harga cabai rawit berdasarkan faktor-faktor tersebut. Hasil penelitian ini menunjukkan atribut yang mempengaruhi nilai RMSE adalah Bulan dan Stok Cabai/KG terhadap label Harga Cabai dengan nilai RMSE sebesar 14975.095. Model prediksi harga diharapkan dapat membantu pedagang merancang strategi penjualan yang efektif, memungkinkan konsumen merencanakan belanja dengan lebih tepat, dan membantu petani menyesuaikan waktu tanam dan panen cabai rawit.

Kata kunci : *Data Mining, Regresi Linear, Cross-Industry Standard for Data Mining (CRISP-DM), Harga Cabai Rawit, Pasar Indihiang*

1. PENDAHULUAN

Sayur mayur merupakan salah satu subsektor unggulan yang memiliki peranan dalam menunjang perekonomian nasional [1]. Komoditas hortikultura merupakan komoditas potensial yang memiliki nilai ekonomi tinggi [2]. Selain itu komoditas hortikultura dapat menjadi sumber pendapatan bagi masyarakat dan petani baik berskala kecil, menengah maupun besar [3]. Meningkatnya harga cabai rawit tidak hanya dipengaruhi oleh peningkatan permintaan selama hari-hari besar keagamaan seperti Bulan Ramadhan, Hari Raya Idul Fitri, serta Hari Raya Idul Adha, khususnya jika terjadi gangguan cuaca atau pasokan [4].

Proses pembentukan harga suatu komoditas umumnya dipengaruhi oleh sejumlah faktor, termasuk ketersediaan terbatas pasokan cabai rawit. Faktor-faktor seperti kondisi cuaca, musim tanam, masalah spekulasi harga, biaya transportasi, dan peningkatan dalam distribusi dapat berkontribusi pada fluktuasi harga cabai rawit [5]. Fluktuasi harga yang tidak stabil sering tidak diinginkan oleh masyarakat karena dapat menimbulkan risiko dan ketidakpastian yang harus dihadapi oleh mereka [6].

Penelitian ini bertujuan untuk mengidentifikasi dan memahami harga cabai rawit di Pasar Indihiang. Analisis regresi adalah teknik statistik yang digunakan untuk memodelkan hubungan antara variabel dependen dan satu set independen [7]. Hasil dari pengolahan data penjualan tersebut menjadi sebuah informasi dan pengetahuan yang diharapkan, sehingga

dapat digali suatu potensi atau pengetahuan yang lebih baik atau akurat dalam pembacaan datanya, tepat dan cepat dari data tersebut sehingga dapat menganalisa prediksi penjualan dan menemukan peluang-peluang yang baru serta menemukan rencana strategis dan untuk analisa stok persediaan barang, selain itu bisa digunakan sebagai sarana untuk mengambil keputusan [8].

Dalam konteks ini, *Data Mining* memiliki peran signifikan dalam menggali informasi dari suatu kumpulan data yang berupa pengetahuan yang sebelumnya tidak diketahui secara manual [9]. Terdapat beberapa kelompok tugas dalam *Data Mining*, termasuk Deskripsi, Estimasi, Prediksi, Klasifikasi, Pengklusteran, dan Asosiasi [10]. Prediksi, sebagai salah satu tugas *Data Mining*, adalah suatu proses sistematis untuk memperkirakan peristiwa di masa depan berdasarkan informasi masa lalu dan saat ini dengan tujuan mengurangi kesalahan prediksi [11]. Tujuan prediksi melibatkan pemahaman dan estimasi prospek ekonomi atau kegiatan usaha [12]. Penting dicatat bahwa prediksi tidak harus memberikan jawaban pasti terkait peristiwa yang akan terjadi; sebaliknya, tujuannya adalah mencari jawaban seakurat mungkin terhadap perkiraan yang diberikan [13].

Metode yang digunakan untuk mengolah dataset tersebut adalah Regresi Linear [14]. Pengujian akurasi dilakukan dengan mengukur *Root Mean Squared Error* (RMSE) [15]. RMSE berfungsi untuk

mengevaluasi nilai yang diprediksi oleh suatu model perhitungan dengan nilai dari hasil pengamatan. Dengan kata lain, RMSE mengukur kualitas kesesuaian antara data yang sebenarnya dan model prediksi. Tingkat kesalahan dari hasil prediksi jika semakin kecil atau mendekati 0 nilai RMSE maka semakin akurat nilai prediksinya [16]. Pada data supervised algoritma Regresi Linear *data mining* bekerja dengan menggunakan label yang telah diketahui nilainya. Biasanya nilai label - label tersebut dapat berupa nilai numerik maupun nilai kategorikal [17]. Tahapan-tahapan yang dilakukan pada penelitian kali ini mengikuti proses standar dari data mining yaitu *Cross-Industry Standard for Data Mining* (CRISP-DM) dimana ada enam fase *Business Understanding* (pemahaman bisnis), *Data Understanding* (pemahaman data), *Data Preparation* (persiapan data), *Modeling* (pemodelan data), *Evaluation*, *Deployment* [18].

Penelitian ini diharapkan memberikan kontribusi yang signifikan dalam memahami dinamika harga cabai rawit dan dampaknya pada berbagai aspek kehidupan [19]. Informasi yang diperoleh dari analisis data mining dan Algoritma Regresi Linear diharapkan dapat mendalami pemahaman tentang faktor-faktor yang mempengaruhi harga cabai rawit. Ini akan mendukung pedagang dalam perencanaan stok dan pengaturan harga yang lebih efisien, serta membantu petani dalam mengoptimalkan musim tanam mereka. Terlebih lagi, penelitian ini memiliki potensi untuk memberikan manfaat praktis bagi konsumen, membantu mereka merencanakan anggaran rumah tangga dengan lebih efektif saat harga cabai rawit naik. Lebih dari itu, dampaknya juga dapat dirasakan dalam pengelolaan pasar dan perumusan kebijakan ekonomi lokal yang lebih baik. Penelitian ini juga membuka jalan bagi penelitian lebih lanjut di bidang pertanian, ekonomi, dan *data mining*, dengan harapan dapat memajukan teknik analisis data dalam mendukung stabilitas pasar serta kesejahteraan ekonomi dan sosial yang lebih luas di Pasar Indihiang dan masyarakat. Seiring dengan itu, hasil penelitian ini diharapkan dapat menjadi acuan berharga bagi penelitian-penelitian mendatang yang berfokus pada dinamika harga komoditas pertanian di konteks pasar tradisional.

2. TINJAUAN PUSTAKA

Penelitian yang berjudul “Prediksi Peningkatan Omset Penjualan Menggunakan Metode Regresi Linier Berganda” yang dilakukan oleh Adiguno, Syahra, dan Yetri. Hasil dari penelitian tersebut menunjukkan bahwa penggunaan analisis data mining dengan dukungan regresi linear berganda dapat membantu memprediksi pendapatan penjualan di PT. Makmur Jaya dengan lebih akurat dan efisien. Penelitian ini menunjukkan bahwa implementasi analisis regresi linear berganda memberikan solusi untuk memprediksi pendapatan penjualan di masa depan dengan tepat. [13]

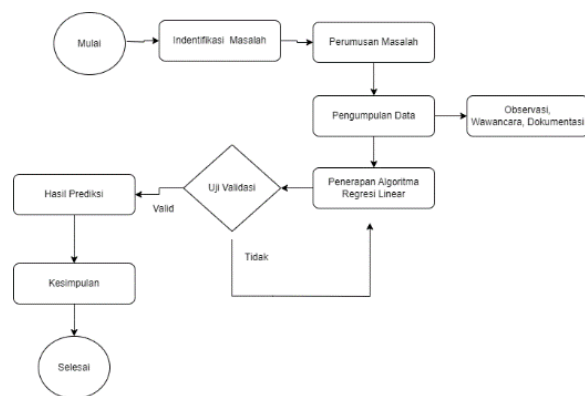
Penelitian yang berjudul “Machine Learning Untuk Model Prediksi Harga Sembako Dengan Metode Regresi Linier Berganda” yang dilakukan oleh Puteri dan Silvanie. Hasil dari penelitian tersebut menunjukkan bahwa variabel tanggal, komoditas, dan pasar secara parsial maupun simultan berpengaruh secara signifikan terhadap harga sembako, dengan persentase sumbangan pengaruh variabel bebas terhadap variabel terikat sebesar 84,2%, sedangkan sisanya sebesar 15,8% dipengaruhi oleh variabel yang tidak dimasukkan ke dalam penelitian ini.[20]

Penelitian yang berjudul “Analisis Pembentukan Harga Komoditas Cabai Rawit dan Bawang Merah Pada Tingkat Eceran di Kota Ambon” yang dilakukan oleh Timisela, Salampey, dan Apituley. Kesimpulan dari penelitian ini adalah bahwa faktor-faktor seperti transportasi, kemasan, selisih pasokan dan permintaan, produksi, daya beli konsumen, dan barang substitusi mempengaruhi pembentukan harga cabai rawit dan bawang merah di tingkat eceran di Kota Ambon. Selain itu, distribusi dan gangguan cuaca juga dapat mempengaruhi harga cabai rawit.[5]

3. METODE PENELITIAN

3.1. Metode Penelitian

Metode yang digunakan untuk mengolah dataset tersebut adalah Algoritma Regresi Linear. Dalam melakukan prediksi data harga cabai yang akan diuji sesuai pemodelan data yang akan digunakan agar mempermudah penelitian dan berjalan sesuai diinginkan maka dibuat alur atau tahapan dalam penelitian seperti terlihat pada Gambar 1 di bawah ini.



Gambar 1. Metode Penelitian

3.2. Sumber Data

Dalam penelitian ini sumber data yang digunakan adalah data primer, sumber data penelitian yang diperoleh secara langsung dengan melalui metode pengumpulan data seperti Observasi, wawancara. Data harga cabai rawit yang terkumpul dari tahun 2021 hingga 2022 mencakup informasi tentang harga, bulan, cuaca, pasokan, supplier, dan nama pedagang.

3.3. Data Testing dan Data Training

Pada bagian ini, akan diuraikan mengenai data testing yang menjadi fokus penelitian, yaitu seluruh kelompok yang ingin diteliti, serta data training, yang

merupakan subset dari data testing yang akan dikaji lebih rinci.

3.3.1. Data Testing

Data testing adalah bagian terpisah dari dataset yang tidak digunakan selama pelatihan model. Setelah model dilatih dengan data training, data testing digunakan untuk menguji kemampuan model dalam menerapkan pengetahuan yang telah dipelajarinya pada data yang belum pernah ditemuinya sebelumnya. Proses pengujian data ini membantu mengevaluasi sejauh mana model dapat menghasilkan prediksi yang akurat.

3.3.2. Data Training

Data training adalah subset dari data yang digunakan untuk melatih model atau algoritma. Ini adalah data yang diberikan ke dalam model untuk memungkinkannya belajar pola atau hubungan antara variabel-variabel yang ada. Proses pelatihan melibatkan pengoptimalan parameter model sehingga dapat membuat prediksi yang akurat ketika dihadapkan pada data yang belum pernah dilihat sebelumnya.

3.4. Teknik Pengumpulan Data

Teknik pengumpulan data adalah metode yang digunakan untuk memperoleh dan mengumpulkan informasi yang relevan terhadap masalah yang diangkat. Dalam proses pengumpulan data, berbagai teknik dapat diterapkan, seperti: berikut:

3.4.1. Observasi

Observasi merupakan teknik pengumpulan data yang dilakukan dengan mengamati langsung objek yang diteliti. Parameter yang diamati yaitu kenaikan harga cabai rawit dan jumlah responden untuk di wawancara.

3.4.2. Wawancara

Tahapan wawancara adalah proses tanya jawab yang bertujuan untuk mengumpulkan data/informasi antara pewawancara (peneliti) dengan responden narasumber. Wawancara yang digunakan dalam penelitian ini adalah wawancara yang terstruktur yaitu wawancara yang berhubungan dengan harga penjualan Cabai Rawit (*Capsicum frutescens* L.) di Pasar Indihiang yang disusun dan terstruktur oleh peneliti.

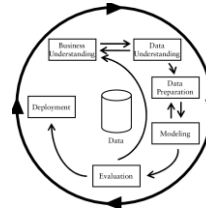
3.4.3. Dokumentasi

Dokumentasi merupakan teknik pengumpulan data berupa foto, dokumen, maupun tulisan. Dokumentasi dilakukan untuk memenuhi suatu data atau informasi untuk kepentingan variabel peneliti yang telah didesain sebelumnya.

3.5. Teknik Analisis Data

Teknik analisis data yang digunakan yaitu teknik data mining dengan metode *Cross-Industry Standard Process for Data Mining* (CRISP-DM) yang

digunakan sebagai strategi pemecahan masalah secara umum dari bisnis atau unit penelitian. Berikut pada Gambar 2 adalah proses pada *Cross-Industry Standard Process for Data Mining* (CRISP-DM).



Gambar 2. Cross-Industry Standard Process for Data Mining (CRISP-DM)

Metode CRISP-DM terdiri dari 6 tahapan, berikut adalah penjelasan dari setiap tahapannya:

1) Tahap Pemahaman Bisnis (*Business Understanding*)

Merupakan langkah awal CRISP-DM dan merupakan bagian penting. Pada titik ini, pengetahuan tentang objek bisnis, cara menghasilkan atau mengambil data, dan cara mencocokkan tujuan model dengan tujuan bisnis diperlukan agar model terbaik dapat dibangun. Kegiatan yang dilakukan meliputi: Mendefinisikan dengan jelas tujuan dan kebutuhan secara keseluruhan, menerjemahkan tujuan tersebut dan menentukan batasan dalam perumusan masalah data mining, kemudian menyiapkan strategi awal untuk mencapai tujuan yang dibelanjakan.

2) Tahap Pemahaman Data (*Data Understanding*)

Fase ini dimulai dengan pengumpulan data, dilanjutkan dengan proses untuk mendapatkan wawasan tentang data, mengidentifikasi masalah kualitas data, atau mengungkap bagian menarik dari data yang dapat digunakan untuk menghipotesiskan informasi tersembunyi.

3) Tahap Persiapan Data (*Data Preparation*)

Langkah ini mencakup semua kegiatan membangun dataset final yaitu data yang akan diproses pada langkah pemodelan, dataset akhir dibangun dari data asli/mentah yang sudah dibersihkan. data clean atau bersih untuk membersihkan nilai-nilai yang kosong, dataset yang tidak konsisten atau kosong (missing value dan noise). Langkah ini dapat diulang berkali-kali, dan juga mencakup pemilihan tabel, rekaman, dan atribut data.

4) Tahap *Modelling*

Pada langkah ini, teknik pemodelan yang berbeda diidentifikasi dan diterapkan, dan beberapa parameter disesuaikan untuk mencapai nilai maksimum.

5) Tahap Evaluasi (*Evaluation*)

Pada titik ini, model telah dilatih dan harus berkualitas baik dari perspektif analisis data. Pada tahap ini akan dilakukan evaluasi terhadap keefektifan dan kualitas suatu model sebelum digunakan dan akan menentukan apakah model

yang digunakan dapat mencapai tujuan yang telah ditetapkan dalam tahap Business Understanding, career atau tidak

6) Tahap Penyebaran (*Deployment*)

Sebagai langkah terakhir dari CRISP-DM, pengetahuan atau informasi yang diperoleh disebarluaskan dalam bentuk khusus yang dapat digunakan oleh pengguna. Tahap implementasi dapat berupa laporan atau melakukan proses penambangan data berulang-ulang di dalam perusahaan.

4. HASIL DAN PEMBAHASAN

Pada tahap ini dilakukan pengolahan data berdasarkan tahapan-tahapan yang ada pada metode CRISP-DM yang sudah dijelaskan sebelumnya.

4.1. Tahap Pemahaman Bisnis (*Business Understanding*)

Analisis data harga cabai rawit dengan menggunakan algoritma regresi linear bertujuan untuk memahami faktor-faktor yang mempengaruhi perubahan harga cabai rawit. Dengan fokus pada variabel seperti bulan, iklim, supplier dan penawaran pasokan, tujuan utama adalah mengidentifikasi korelasi antara faktor-faktor tersebut dengan perubahan harga. Melalui pemahaman yang mendalam terhadap hubungan ini, analisis ini bertujuan untuk mengembangkan model regresi linear yang dapat memberikan prediksi harga yang akurat.

4.2. Tahap Pemahaman Data (*Data Understanding*)

Data yang akan diproses dengan data mining merupakan data kuantitatif, karena kumpulan data yang akan digunakan berbentuk nominal dan numerik. Adapun kumpulan data tersebut berisi Nama Pedagang, Tahun, Bulan, Cuaca, Supplier, Harga Cabai Rawit, Stok Cabai/Kg. Adapun sumber data didapatkan berdasarkan hasil observasi dan wawancara di Indihiang yang berlokasi di Jl. Brigjen Wasita Kusuma (Komplek Pasar Indihiang) (kantor Pasar Indihiang), Kec. Indihiang, Kota Tasikmalaya, Jawa Barat, ID 46151. Data penelitian berasal dari data penjualan selama kurun waktu 2 tahun dari Januari 2021 sampai Desember 2022.

4.3. Tahap Persiapan Data (*Data Preparation*)

4.3.1. Data

Data dalam bentuk file excel yang bernama Dataset Update Kenaikan Harga Cabai 2 diperoleh dari hasil wawancara di Pasar Indihiang. Jumlah record sebanyak 720 record, terdiri dari 8 atribut seperti tampak pada Tabel dibawah ini.

Tabel 1. Dataset

No.	Nama Pedagang	Tahun		Harga Cabai	Stok Cabai/KG
1	Ibu Ade	2021		80.000	450
2	Ibu Ade	2021		90.000	450

No.	Nama Pedagang	Tahun		Harga Cabai	Stok Cabai/KG
3	Ibu Ade	2021		70.000	450
4	Ibu Ade	2021		60.000	450
.....					
719	Bpk Irwan	2022		40.000	450
720	Bpk Irwan	2022		61.000	300

4.3.2. Import Data

Agar dataset tersebut ada di repository Rapidminer, maka perlu dilakukan import data Dataset Update Harga Cabai 2.xlsx menjadi *operator Retrieve* seperti tampak pada gambar dibawah ini.

Retrieve Dataset Update_Harga Cabai2



Gambar 3. Operator Retrieve

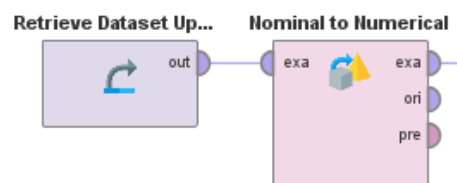
Adapun hasil import data sebagai berikut.

Tabel 2. Hasil Import Data

Row No.	No	Nama Pedagang		Harga Cabai	Stok Cabai/KG
1	?	Ibu Ade		80.000	450
2	?	Ibu Ade		90.000	450
3	?	Ibu Ade		70.000	450
4	?	Ibu Ade		60.000	450
.....					
719	?	Bpk Irwan		40.000	450
720	?	Bpk Irwan		61.000	300

4.3.3. Mengubah tipe data nominal menjadi numerik

Karena tipe data pada dataset ada yang berjenis nominal sedangkan Algoritma Regresi Linear harus bertipe numeric, maka diperlukan transformation untuk mengubah data bertipe nominal menjadi numeric menggunakan operator Nominal to Numerical seperti pada gambar dibawah ini.



Gambar 4. Operator Nominal to Numerical

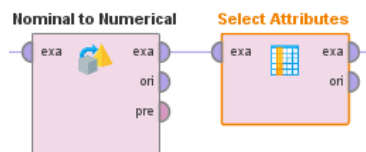
Hasil pembacaan operator Nominal to Numerical pada tabel berikut.

Tabel 3. Hasil operator Nominal to Numerical

No.	Nama Pedagang	Tahun		Harga Cabai	Stok Cabai/KG
1	0	0		80000	450
2	0	1		90000	450
3	0	2		70000	450
4	0	3		60000	450
....					
719	29	10		40000	450
720	29	11		61000	300

4.3.4. Memilih atribut yang akan digunakan

Untuk menentukan atribut yang akan digunakan dan tidak, digunakan operator *Select Atribut* untuk memilih data atribut.



Gambar 5. Select Attributes

Hasil setelah menjalankan operator select atribut sebagai berikut.

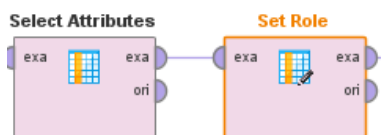
Tabel 4. Hasil Select Attribute

Atribut yang tidak digunakan	Atribut yang digunakan
No	Bulan
Nama Pedagang	Stok Cabai/KG
Supplier	Harga Cabai
Tahun	
Cuaca	

4.4. Tahap Modelling

1) Menentukan atribut label

Untuk menentukan id dan label pada dataset, digunakan operator *Set Role* seperti tampak pada gambar dibawah ini.



Gambar 6. Operator Set Role

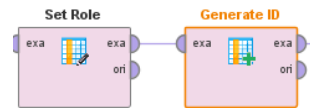
Dari hasil pembacaan operator *Set Role* didapat informasi Harga Cabai sebagai label.

Tabel 5. Penentuan label pada Set Role

Label
Harga Cabai

2) Membuat ID

Karena pada dataset data harga cabai rawit di pasar indihiang tidak memiliki id maka menggunakan operator *Generate ID*



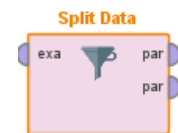
Gambar 7. Operator Generate ID

Dari hasil pembacaan operator *Generate ID* di dapatkan id seperti pada tabel berikut.

Tabel 6 Hasil Generate ID

ID
3
4
7
....
719

- Membagi data testing dan data training
Pada tahap ini akan menggunakan data testing dan data training dengan ratio 0.8 untuk data training dan 0.2 untuk data testing untuk dilakukan penelitian.



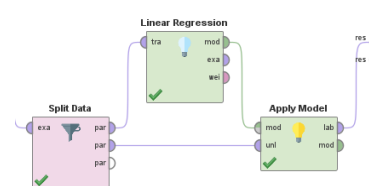
Gambar 9. Split Data

Hasil dari operator split data dapat dilihat pada tabel dibawah ini.

Tabel 7. Pembagian data testing dan training

0.8	Data Training
0.2	Data Testing

- Implementasi Algoritma Regresi Linear
Setelah melakukan pembagian data, selanjutnya mencari nilai prediksi dengan operator *Linear Regression* seperti tampak pada gambar dibawah ini.



Gambar 10. Implementasi Algoritma Regresi Linear

Pada tahap implementasi Algoritma Regresi Linear pengujian terhadap *prediction* harga cabai rawit, dapat dilihat pada tabel dibawah ini.

Tabel 8. Hasil Implementasi Algoritma Regresi Linear

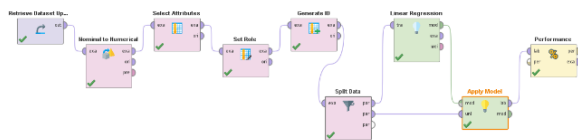
id	Harga Cabai	Prediction (Harga Cabai)
3	70000	62718.016
4	60000	60565.379

id	Harga Cabai	Prediction (Harga Cabai)
7	40000	54107.467
13	70000	72547.294
20	60000	57478.833
26	95000	64870.653
31	41000	54017.467

Dari tabel di atas, didapatkan informasi bahwa harga cabai sebesar Rp. 70.000 diprediksi akan mengalami penurunan menjadi Rp. 62.718 pada bulan selanjutnya, harga cabai Rp. 60.000 diprediksi akan mengalami kenaikan menjadi Rp. 60.565, harga cabai Rp. 40.000 diprediksi akan mengalami kenaikan menjadi Rp. 54.107, harga cabai Rp. 70.000 diprediksi akan mengalami kenaikan menjadi Rp. 72.547, harga cabai Rp. 60.000 diprediksi akan mengalami penurunan menjadi Rp. 57.478, harga cabai Rp. 95.000 diprediksi akan mengalami penurunan menjadi Rp. 64.870, dan harga cabai Rp. 41.000 diprediksi akan mengalami kenaikan menjadi Rp. 54.017 pada bulan selanjutnya.

4.5. Tahap Evaluasi (Evaluation)

Setelah menyelesaikan tahap pemodelan, langkah berikutnya adalah tahap evaluasi. Evaluasi dilakukan untuk menilai kinerja model. Proses evaluasi menggunakan data uji yang telah ditetapkan sebesar 20% dari seluruh data. Pengujian dilakukan dengan operator kinerja, sebagaimana terlihat pada Gambar 11, yang menggambarkan langkah pengujian dengan operator kinerja.



Gambar 8. Proses pengujian dengan melihat *performance*

Pada tahap implementasi Algoritma Regresi Linear pengujian terhadap *Parameter feature selection* dan *Performance* untuk mendapatkan nilai *Root Mean Squared Error* (RMSE), dapat dilihat pada tabel dibawah ini.

Tabel 9. Hasil eksperimen *Parameter feature selection* dan *Performance*

Atribut	Label	Parameter feature selection	Nilai RMSE
Stok Cabai/KG		M5 Prime	14975.095
		Greedy	14975.095
		T-Test	14975.095
		Iterative T-Test	14975.095
Bulan	Harga Cabai	M5 Prime	15307.903
		Greedy	15307.903
		T-Test	15307.903
		Iterative T-Test	15307.903

Setelah menjalankan eksperimen, hasil dari tabel di atas menunjukkan bahwa atribut yang signifikan memengaruhi nilai RMSE adalah Bulan dan Stok

Cabai/KG terhadap Harga Cabai. Parameter feature selection seperti M5 Prime, Greedy, T-Test, dan Iterative T-Test dalam model algoritma Regresi Linear memberikan nilai Root Mean Squared Error (RMSE) yang serupa, yakni sebesar 14975.095.

4.6. Tahap Penyebaran (Deployment)

Setelah menyelesaikan tahap evaluasi, langkah terakhir dalam penelitian ini adalah melakukan *deployment* model yang telah dibangun. Penerapan ini akan dilakukan melalui pembuatan laporan yang merinci hasil penelitian sesuai dengan metode *Cross-Industry Standard Process for Data mining* (CRISP-DM). Laporan ini dimuat untuk membahas pencarian nilai *Root Mean Squared Error* (RMSE) terbaik, mengidentifikasi atribut yang memiliki pengaruh signifikan terhadap nilai RMSE, serta menganalisis peran *Parameter feature selection* dalam mempengaruhi nilai RMSE.

4.7. Hasil atribut yang mempengaruhi nilai *Root Mean Squared Error* (RMSE)

Berikut adalah hasil dari eksperimen tersebut, yang disajikan dalam tabel dengan label yang telah ditentukan, yaitu Harga Cabai.

Tabel 10. Hasil eksperimen atribut

Atribut	Label	Parameter Feature Selection	RMSE
Bulan	Stok Cabai	Harga Cabai	
		M5 Prime	14975.095
		Greedy	14975.095
		T-Test	14975.095
		Iterative T-Test	14975.095

5. KESIMPULAN DAN SARAN

Berikut adalah kesimpulan berdasarkan penelitian yang telah dilakukan pada dataset Harga Cabai Rawit di Pasar Indihiang menggunakan algoritma Regresi Linear atribut yang mempengaruhi nilai RMSE terbaik pada data harga cabai rawit di pasar Indihiang pada atribut Bulan dan Stok Cabai/KG, dengan nilai *Root Mean Squared Error* (RMSE) terbaik pada data harga cabai rawit di pasar Indihiang adalah 14975.095.

Dari hasil penelitian ini, beberapa saran strategis dapat diajukan untuk meningkatkan pemahaman dan pengembangan model Regresi Linear dalam menganalisis dinamika harga cabai rawit di Pasar Indihiang sebagai berikut : Rutin monitor dan perbarui dataset dengan informasi terkini untuk mempertahankan ketepatan model terhadap perubahan pasar. Pertimbangkan faktor-faktor eksternal seperti tren pasar global, perubahan kebijakan, atau faktor cuaca yang lebih rinci untuk meningkatkan ketepatan prediksi dan relevansi model dalam menghadapi dinamika pasar. Dukung pedagang dengan hasil analisis data mining dan Algoritma *Regresi Linear* untuk strategi penjualan. Membantu petani dalam mengoptimalkan musim tanam dengan memahami

faktor-faktor yang mempengaruhi harga cabai rawit. Membantu konsumen merencanakan anggaran rumah tangga secara lebih efektif saat harga cabai rawit naik.

DAFTAR PUSTAKA

- [1] N. A. Loanga, R. Indriani, and Y. Boekoesoe, "Analisis Faktor-Faktor Yang Mempengaruhi Fluktuasi Harga Cabai Rawit Di Kecamatan Sumalata Kabupaten Gorontalo Utara," *J. Sci. Soc. Res.*, vol. 4, no. 2, pp. 235–242, 2023.
- [2] M. Y. Wardhana, W. Widyawati, R. Hermawan, and T. M. Kesuma, "Analisis Faktor – Faktor Yang Mempengaruhi Harga Cabai Rawit (*Capsicum Frutescens* L.) Di Aceh," *Paradig. Agribisnis*, vol. 4, no. 2, p. 69, 2022.
- [3] A. Ramadhani, D. Susilowati, and S. Hindarti, "Studi Kasus Di Abang Sayur Organik Kota Malang," *J. Sos. Ekon. Pertan. Dan Agribisnis*, vol. 9, no. 2, pp. 1–10, 2021.
- [4] M. Royun Nuha, T. Andita Putri, and A. Dwi Utami, "Pendapatan Usahatani Cabai Merah Berdasarkan Musim di Provinsi Jawa Tengah," *J. Ilmu Pertan. Indones.*, vol. 28, no. 2, pp. 323–334, 2023.
- [5] N. R. Timisela, Y. E. Salampessy, and Y. M. T. N. Apituley, "Analisis Pembentukan Harga Komoditas Cabai Rawit dan Bawang Merah Pada Tingkat Eceran di Kota Ambon," *J. Budid. Pertan.*, vol. 16, no. 1, pp. 31–41, 2020.
- [6] R. A. Abdi, B. Siswadi, and N. Khoiriyah, "Deskripsi Harga Cabai Merah Di Jawa Timur," *SEAGRI*, vol. 10, no. 8, p. 9, 2022.
- [7] E. Terapan, "Model Regresi Linier," no. Eko 1708, 2021.
- [8] A. T. Z. Miftahuljannah, Aswan Supriyadi Sunge, "Analisis Prediksi Penjualan Dengan Metode Regresi Linear Di Pt. Eagle Industry Indonesia," vol. 5, no. July, pp. 1–23, 2020.
- [9] S. P. Dewi, N. Nurwati, and E. Rahayu, "Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor," *Build. Informatics, Technol. Sci.*, vol. 3, no. 4, pp. 639–648, 2022.
- [10] E. Triyanto, H. Sismoro, and A. D. Laksito, "Implementasi Algoritma Regresi Linear Berganda Untuk Memprediksi Produksi Padi Di Kabupaten Bantul," *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 4, no. 2, pp. 66–75, 2019.
- [11] D. S. O. Panggabean, E. Buulolo, and N. Silalahi, "Penerapan Data Mining Untuk Memprediksi Pemesanan Bibit Pohon Dengan Regresi Linear Berganda," *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 1, p. 56, 2020.
- [12] G. N. Ayuni and D. Fitrianah, "Penerapan metode Regresi Linear untuk prediksi penjualan properti pada PT XYZ," *J. Telemat.*, vol. 14, no. 2, pp. 79–86, 2019.
- [13] S. Adiguno, Y. Syahra, and M. Yetri, "Prediksi Peningkatan Omset Penjualan Menggunakan Metode Regresi Linier Berganda," *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 1, no. 4, p. 275, 2022.
- [14] A. N. Maharadja, I. Maulana, and B. A. Dermawan, "Penerapan Metode Regresi Linear Berganda untuk Prediksi Kerugian Negara Berdasarkan Kasus Tindak Pidana Korupsi," *J. Appl. Informatics Comput.*, vol. 5, no. 1, pp. 95–102, 2021.
- [15] V. Arinal, M. Azhari, and I. Pendahuluan, "Penerapan Regresi Linear Untuk Prediksi Harga Beras Di Indonesia," *J. Sains dan Teknol.*, vol. 5, no. 1, pp. 341–346, 2023.
- [16] 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. Tek. Inform.*, vol. 10, no. 1, p. 23, 2021.
- [17] A. Prawira, D. Arisandi, and T. Sutrisno, "Penerapan Algoritma Naive Bayes dan Multiple Linear Regression untuk Prediksi Status dan Plafon Kredit (Studi Kasus : Bank ABC)," *J. Appl. Informatics Comput.*, vol. 05, no. 01, pp. 1075–1087, 2022.
- [18] S. Isnanto and S. Widodo, "Penerapan Data Mining Pada Penerimaan Mahasiswa Baru Dengan Algoritma K-Means Clustering," *J. Tek. Inf. dan Komput.*, vol. 4, no. 2, p. 158, 2021.
- [19] Saputra Aldi Agusta, M. Munir, and Z. A. Daffa Rizki, "Peramalan Pendapatan dari Penjualan Bawang Merah Menggunakan Metode Regresi Linier Berganda," *Stain. (Seminar Nas. Teknol. Sains)*, vol. 2, no. 1, pp. 383–389, 2023.
- [20] K. Puteri and A. Silvanie, "Machine Learning untuk Model Prediksi Harga Sembako," *J. Nas. Inform.*, vol. 1, no. 2, pp. 82–94, 2020.