

ESTIMASI PENJUALAN PADA TOKO DAMHIL CAKE AND BAKERY MENGUNAKAN ALGORITMA REGRESI LINIER SEDERHANA DALAM MENENTUKAN STRATEGI PENJUALAN KUE KERING

Kaslani, Regi Putri Siti Nur Hajjah, Sandy Eka Permana, Fatihanursari

Teknik Informatika, STMIK IKMI Cirebon

Jl. Perjuangan No.10 B Kec. Kesambi Kota Cirebon Tlp. (02311) 490480 - 490481

kaslani@gmail.com

ABSTRAK

Damhil cake and Bakery merupakan perusahaan kecil dan menengah yang bergerak di bidang industri roti dan kue di tengah persaingan yang semakin ketat. Penelitian ini bertujuan untuk mengembangkan strategi penjualan kue kering di Toko Damhil cake and Bakery melalui estimasi hasil penjualan produk menggunakan algoritma regresi linier sederhana. Penelitian ini melibatkan pengumpulan data penjualan produk kue kering di Toko Damhil Cake and Bakery selama periode 1 tahun yaitu tahun 2022. Data ini mencakup informasi tentang harga produk, dan dampak musiman terhadap penjualan. Dalam analisis ini, algoritma regresi linier sederhana dengan pendekatan Knowledge Discovery in Databases (KDD) digunakan untuk memodelkan hubungan antara variabel-variabel tersebut dengan hasil penjualan. Penelitian ini melibatkan analisis data historis penjualan kue kering, penerapan algoritma regresi linier sederhana untuk membangun model prediktif, dan penggunaan metode KDD untuk mengidentifikasi pola atau hubungan yang dapat meningkatkan strategi penjualan. Hasil estimasi penjualan diharapkan dapat membantu Toko Damhil cake and Bakery dalam mengambil keputusan yang lebih tepat terkait produksi, promosi, dan penetapan harga produk. Dengan menerapkan pendekatan ini, diharapkan toko dapat meningkatkan efisiensi operasionalnya dan merumuskan strategi penjualan yang lebih cerdas. Hasil penelitian ini dapat memberikan jawaban terhadap permasalahan yang dirumuskan dalam latar belakang, yaitu kesulitan untuk menentukan strategi penjualan yang efektif, adapun relevansi dari penelitian ini yaitu bertujuan untuk membantu Toko Damhil Cake and Bakery untuk membuat keputusan yang lebih akurat dalam merencanakan strategi penjualan kue kering dan meningkatkan efisiensi bisnis. Hasil yang diperoleh dari prediksi penjualan kue kering untuk periode waktu tahun 2022 sebanyak 10.458 pcs.

Kata kunci : Algoritma regresi linier, Strategi penjualan kue kering, dan Estimasi penjualan produk.

1. PENDAHULUAN

Dalam dunia bisnis yang kompetitif saat ini, masyarakat selalu perlu mengembangkan usahanya agar dapat bertahan dalam persaingan, apalagi persaingan penjualan menuntut para wirausahawan untuk memiliki model yang dapat meningkatkan penjualan dan pemasaran dalam perusahaan, salah satunya adalah pemanfaatan penjualan [1]. Penelitian yang sudah dilakukan sebelumnya oleh [2] Mencakup tentang cara melakukan prediksi penjualan yang diuji terhadap pemodelan data yang digunakan untuk mempermudah penelitian dan berjalan sesuai harapan maka dibuat alur atau tahapan penelitian. Objek yang digunakan adalah data penjualan perusahaan, yaitu data kuantitatif yang telah diolah untuk menjawab suatu hipotesis atau pertanyaan penelitian. Data ini mencakup atribut seperti produk, volume penjualan aktual, rencana penjualan bulanan, dan tahun.

Dalam penelitian ini, metode regresi linier digunakan untuk perhitungan dan perancangan sistem dan diimplementasikan dalam sistem. Tujuan dari penelitian ini adalah untuk memprediksi penjualan produk dan membuat perkiraan penjualan dengan menggunakan data penjualan bulan sebelumnya yang disediakan [1].

Pengolahan data menggunakan metode Data Mining dengan menerapkan Algoritma regresi linier

untuk meramalkan angka penjualan pada sistem ini memudahkan dalam meramalkan penjualan dan memungkinkan alokasi sumber daya yang lebih efisien, yang pada akhirnya meningkatkan pendapatan [3].

Masalah dalam penelitian ini yaitu bagaimana cara untuk mengimplementasikan algoritma regresi linier sederhana sebagai strategi pemasaran di toko damhil cake and bakery, yaitu dengan cara mengumpulkan data historis penjualan kue kering dan variabel pemasaran terkait seperti promosi, harga, dan faktor musiman. Selanjutnya, dengan menggunakan regresi linier sederhana untuk menganalisis hubungan antara variabel pemasaran tersebut dan penjualan kue kering. Jumlah data yang diperoleh pada penelitian ini adalah 3.009 data penjualan kue kering selama 1 tahun terakhir yaitu 2022 dari 6 varian kue kering yaitu diantaranya Kastangel, Putri Salju, Mini Pie Keju, Nastar Keju, Kue Semprit, dan Nastar Strawberry.

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data Mining adalah ekstraksi informasi atau pola yang penting dari data yang di database yang besar. Dalam jurnal ilmiah, data mining juga dikenal dengan nama Knowledge Discovery in Database (KDD).[4]

2.2. Regresi Linear

Regresi linear sederhana adalah sebuah metode pendekatan untuk pemodelan hubungan antara satu variabel dependen dan satu variabel independen, analisis regresi linear sederhana berfungsi untuk menguji hubungan sebab akibat antara variabel faktor penyebab terhadap variabel akibatnya. Regresi Linear Sederhana adalah salah satu metode least square yang berfungsi untuk menguji hubungan sebab akibat antara Variabel Penyebab (X) terhadap Variabel Akibat (Y). Variabel Penyebab pada umumnya dilambangkan dengan X sedangkan Variabel Akibat dilambangkan dengan Y [5]

Model regresi linier yang digunakan memiliki bentuk umum sebagai berikut:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

- Y Adalah hasil penjualan produk kue kering.
- β_0 Adalah intercept (konstanta).
- β_1 , β_2 , dan β_3 Adalah koefisien regresi untuk masing-masing variabel independen (X_1 , X_2 , X_3).
- ε Adalah kesalahan acak.

2.3. Perhitungan Root Mean Squared Error (RMSE)

Root mean squared error (RSME) adalah metric evaluasi yang mengukur sejauh mana perbedaan antara nilai prediksi dan nilai sebenarnya. Root Mean Squared Error (RMSE) merupakan salah satu cara untuk mengevaluasi model regresi linear dengan mengukur tingkat akurasi hasil perkiraan suatu model.[6]

RMSE dapat dihitung dengan mengambil akar kuadrat dari rata-rata kuadrat dengan selisih prediksi ($\hat{y}$) dan nilai sebenarnya (y) untuk setiap observasi, Rumus RMSE yang akan digunakan adalah sebagai berikut:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Dimana:

- n adalah jumlah observasi.
- y_i adalah nilai sebenarnya dari observasi ke- i .
- $\hat{y}_i$ adalah nilai prediksi dari observasi ke- i .

2.4. Contoh Gambar



Gambar 1. Hasil Root Mean Square Error (RMSE)

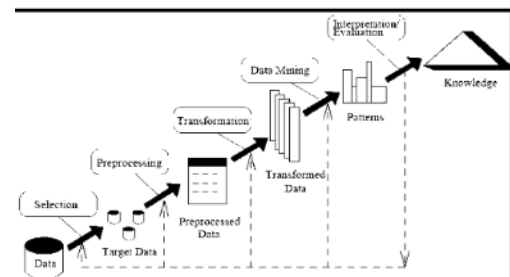
Jadi, hasil RMSE sebesar 2.675 +/- 0.000 menunjukkan bahwa RMSE yang rendah (2.675) menandakan bahwa model regresi linier sederhana cenderung memberikan prediksi yang akurat terhadap penjualan produk pada Toko Damhil Cake and Bakery.

2.5. Contoh Tabel

Tabel 1. Data Penjualan

Tanggal	Nama Produk	Qty	Harga	Total
18-01-2022	Kastangel	8	45.000	360.000
31-12-2022	Putri salju	6	40.000	240.000
	Mini pie keju	8	45.000	360.000
	Nastar Keju	7	50.000	350.000
	Sempurit	2	40.000	80.000
	Nastar Stw	5	40.000	200.000

3. METODE PENELITIAN

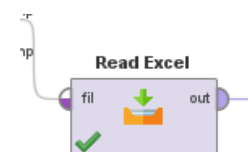


Gambar 2. Alur Model Penelitian

Metodologi penelitian merupakan kerangka kerja yang dilakukan peneliti. Kerangka kerja ini merupakan tahapan proses-proses pengelolaan data yang dijalankan. Ada beberapa tahapan yang digambarkan, yaitu mulai dari tahap studi literatur sampai tahap kesimpulan. Tahapan proses dilakukan secara terstruktur dan sistematis [7]. Tahapan Knowledge Discovery In Database (KDD) Knowledge Discovery In Databases (KDD) adalah keseluruhan proses nontrivial untuk mencari dan mengidentifikasi pola (pattern) dalam data, dimana pola yang ditemukan bersifat sah, baru, dapat bermanfaat dan dapat dimengerti. KDD berhubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola sejumlah kumpulan data [8]. Berikut adalah tahapan dari *Knowledge Discovery in Database* (KDD) :

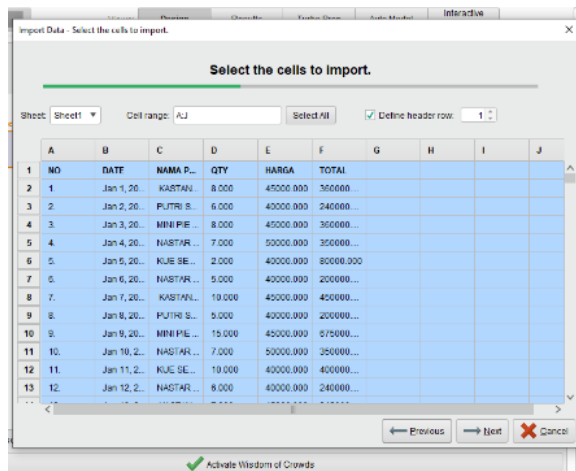
3.1. Data selection

Pengumpulan data dilakukan dengan mencatat penjualan harian kue kering selama periode tertentu, termasuk variabel-variabel seperti jumlah produk yang terjual, dan faktor-faktor lain yang mempengaruhi penjualan.



Gambar 3. Operator Read Excel

Proses Read excel digunakan untuk mengimpor dan membaca file excel sebagai input untuk membangun model regresi linier sederhana.



Gambar 4. Import Data

Setelah melakukan Read excel maka akan muncul tampilan select the cells to import untuk memastikan bahwa file data yang di upload sudah benar atau belum, jika dirasa data yang sudah di masukan ke dalam parameter Read excel maka selanjutnya tekan tombol next untuk melihat hasilnya seperti pada gambar 4

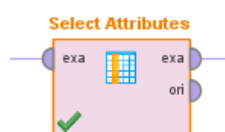
Row No.	QTY	NO	DATE	NAMA PRODUK	HARGA	TOTAL	G	H	I	J
1	8	1	4/8/22	KASTAN...	45000	360000	?	?	?	?
2	6	2	4/8/22	PUTRI S...	40000	240000	?	?	?	?
3	8	3	4/8/22	MINIPIE...	45000	360000	?	?	?	?
4	7	4	4/8/22	NASTAR...	50000	350000	?	?	?	?
5	2	5	4/8/22	KUE SE...	40000	80000	?	?	?	?
6	5	6	4/8/22	NASTAR...	40000	200000	?	?	?	?
7	10	7	4/8/22	KASTAN...	45000	450000	?	?	?	?
8	5	8	4/8/22	PUTRI S...	40000	200000	?	?	?	?
9	15	9	4/8/22	MINIPIE...	45000	675000	?	?	?	?
10	7	10	4/8/22	NASTAR...	50000	350000	?	?	?	?
11	10	11	4/8/22	KUE SE...	40000	400000	?	?	?	?
12	6	12	4/8/22	NASTAR...	40000	240000	?	?	?	?
13	7	13	4/8/22	KASTAN...	45000	315000	?	?	?	?
14	8	14	4/8/22	PUTRI S...	40000	320000	?	?	?	?
15	10	15	4/8/22	MINIPIE...	45000	450000	?	?	?	?

Gambar 5. Hasil Read Excel

Hasil read excel dalam rapidminer pada algoritma regresi linier sederhana mencakup dataset yang akan digunakan dalam proses pembuatan model regresi linier sederhana untuk melakukan prediksi atau analisis terhadap variabel independen yang terpilih yaitu Qty (Quantity).

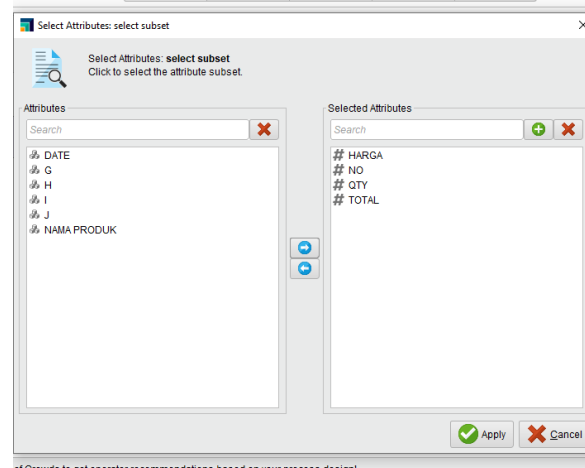
3.2. Data Processing

Mengidentifikasi dan menangani data yang hilang atau tidak valid, menghapus duplikat, dan melakukan normalisasi jika diperlukan untuk memastikan konsistensi dan kualitas data.



Gambar 6. Operator Select Atribut

Pada select attribute terdapat beberapa attribute yang dipilih yaitu Selected Attribute dan Attribute. Kemudian pada parameter Attribute filter opsi yang dipilih yaitu a subset. Selanjutnya setelah melakukan parameter a subset opsi yang dipilih bisa dilihat pada gambar 6



Gambar 7. Proses Select Atribut

Pada gambar diatas menampilkan atribut yang akan digunakan untuk dianalisis terdapat pada kolom Select Attribute yaitu Total, Harga, No, dan Qty (keterangan jumlah barang yang terjual) serta total. Sedangkan atribut yang tidak digunakan dalam analisis terdapat pada kolom attributes diantaranya Date, G, H, I, J dan Nama Produk.

Row No.	QTY	NO	HARGA	TOTAL
1	8	1	45000	360000
2	6	2	40000	240000
3	8	3	45000	360000
4	7	4	50000	350000
5	2	5	40000	80000
6	5	6	40000	200000
7	10	7	45000	450000
8	5	8	40000	200000
9	15	9	45000	675000
10	7	10	50000	350000
11	10	11	40000	400000
12	6	12	40000	240000
13	7	13	45000	315000
14	8	14	40000	320000
15	10	15	45000	450000

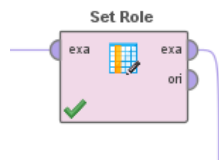
Gambar 8. Hasil Select Atribut

Hasil dari select attribute yang telah dilakukan menunjukan bahwa attribute yang akan digunakan pada penelitian ini yaitu Quantity (jumlah kue kering yang terjual).

3.3. Data Transformation

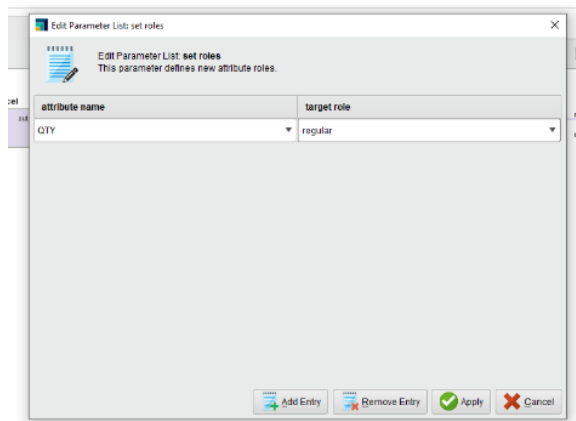
Mengonversi data ke dalam format yang sesuai untuk regresi linier, seperti mengubah variabel-variabel kategorikal menjadi variabel dummy, dan

mempersiapkan data agar sesuai dengan persyaratan algoritma regresi linier.



Gambar 9. Operator SetRole

Pada operator SetRole, terdapat parameter Editroles dengan 2 pilihan opsi yaitu attribute name dan target role, maka pada penelitian ini yang akan dijadikan sebagai target role nya adalah regular dan attribute name nya adalah Qty (Quantity).



Gambar 10. Parameter SetRole

Row No.	QTY	NO	HARGA	TOTAL
1	8	1	45000	360000
2	6	2	40000	240000
3	8	3	45000	360000
4	7	4	50000	350000
5	2	5	40000	80000
6	5	6	40000	200000
7	10	7	45000	450000
8	5	8	40000	200000
9	15	9	45000	675000
10	7	10	50000	350000
11	10	11	40000	400000
12	6	12	40000	240000
13	7	13	45000	315000
14	8	14	40000	320000
15	10	15	45000	450000

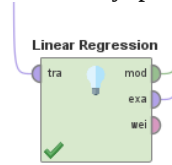
ExampleSet (288 examples, 1 special attribute, 3 regular attributes)

Gambar 11. Hasil SetRole

Hasil set role digunakan untuk mengidentifikasi atribut sebagai variabel independen (input) yang mencakup atribut-atribut yang dianggap sebagai variabel yang digunakan untuk memprediksi variabel dependen, Sedangkan variabel depeden (Output) mencakup atribut yang dianggap sebagai variabel yang akan di prediksi.

3.4. Data Mining

Data Mining adalah suatu proses ekstraksi atau penggalian data yang belum diketahui sebelumnya, namun dapat dipahami dan berguna dari database yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting. Data Mining biasa disebut dengan “Data atau Knowledge Discovery” atau menemukan pola tersembunyi pada data [9] .



Gambar 12. Operator Linear Regression

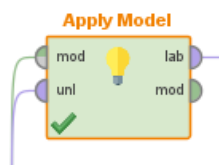
Operator Regresi linear digunakan untuk melakukan analisis regresi pada suatu variabel independen dan variabel dependen untuk menghasilkan persamaan garis regresi yang menggambarkan pola hubungan tersebut.

Row No.	QTY	NO	HARGA	TOTAL
1	8	1	45000	360000
2	6	2	40000	240000
3	8	3	45000	360000
4	7	4	50000	350000
5	2	5	40000	80000
6	5	6	40000	200000
7	10	7	45000	450000
8	5	8	40000	200000
9	15	9	45000	675000
10	7	10	50000	350000
11	10	11	40000	400000
12	6	12	40000	240000
13	7	13	45000	315000
14	8	14	40000	320000
15	10	15	45000	450000

ExampleSet (288 examples, 1 special attribute, 3 regular attributes)

Gambar 13. Hasil Linear Reggression

Hasil Regression linear dapat memberikan koefisien regresi dan konstanta yang digunakan dalam persamaan linier, dengan menggunakan regresi linier sederhana juga dapat digunakan untuk memahami seberapa kuat dan arah hubungan antara dua variabel.



Gambar 14. Operator Apply Model

Apply Model adalah tahapan terakhir yang mana pada proses ini akan menerapkan model data linear regression berdasarkan proses-proses sebelumnya mulai dari Select attribute, SetRole, dan Linear Regression. Setelah semua proses dilakukan, maka akan dihasilkan model prediksi dari data Exampleset sebesar 20%.

ExampleSet (Apply Model)					
Open in Turbo Prep Auto Model Interactive Analysis					
Row No.	QTY	prediction(Q...	NO	HARGA	TOTAL
1	8	8.569	1	45000	360000
2	6	6.772	2	40000	240000
3	8	8.560	3	45000	360000
4	7	7.520	4	50000	350000
5	2	3.277	5	40000	80000
6	5	5.883	6	40000	200000
7	10	10.498	7	45000	450000
8	5	5.873	8	40000	200000
9	15	15.382	9	45000	675000
10	7	7.491	10	50000	350000
11	10	10.208	11	40000	400000
12	6	6.724	12	40000	240000
13	7	7.533	13	45000	315000
14	8	8.454	14	40000	320000
15	10	10.459	15	45000	450000

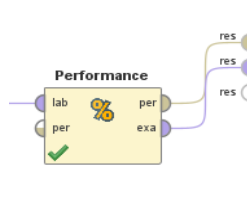
ExampleSet (288 examples, 2 special attributes, 3 regular attributes)

Gambar 15. Hasil Apply Model

Hasil dari Apply model diatas menunjukan bahwa Qty dari data asli yang berbentuk excel dengan Qty yang sudah diolah dengan tools Rapidminer memiliki persamaan dari hasilnya

3.5. Evaluation

Dalam tahap ini, merupakan hasil dari teknik data mining berupa pola pola yang khas maupun model dievaluasi untuk menilai apakah ada memang tercapai. Memeriksa akurasi dan kehandalan model regresi linier dengan menggunakan metric evaluasi seperti *Root mean squared error* (RMSE), dan melakukan validasi silang jika diperlukan. Jika model memberikan hasil yang memadai, dapat digunakan untuk membuat estimasi penjualan dan menentukan strategi penjualan kue kering di toko Damhil Cake and Bakery.



Gambar 16. Operator Performance

Selanjutnya, Operator performance akan menghasilkan hasil attribute, statistic performance, dan visualizations yang dapat dilihat seperti pada gambar

ExampleSet (Apply Model)					
Open in Turbo Prep Auto Model Interactive Analysis					
Row No.	QTY	prediction(Q...	NO	HARGA	TOTAL
1	8	8.569	1	45000	360000
2	6	6.772	2	40000	240000
3	8	8.560	3	45000	360000
4	7	7.520	4	50000	350000
5	2	3.277	5	40000	80000
6	5	5.883	6	40000	200000
7	10	10.498	7	45000	450000
8	5	5.873	8	40000	200000
9	15	15.382	9	45000	675000
10	7	7.491	10	50000	350000
11	10	10.208	11	40000	400000
12	6	6.724	12	40000	240000
13	7	7.533	13	45000	315000
14	8	8.454	14	40000	320000
15	10	10.459	15	45000	450000

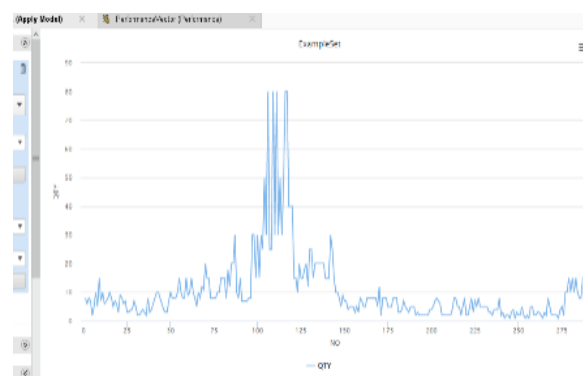
ExampleSet (288 examples, 2 special attributes, 3 regular attributes)

Gambar 17. Hasil Attribut Performance

Name	Type	Missing	Statistics	Filter (5 of 5 attributes)
Label QTY	Integer	0	Min: 1, Max: 80, Average: 10.458	
Prediction prediction(QTY)	Integer	0	Min: -0.220, Max: 86.413, Average: 10.458	
NO	Real	0	Min: 1, Max: 288, Average: 144.500	
HARGA	Integer	0	Min: 40000, Max: 50000, Average: 43333.333	
TOTAL	Integer	0	Min: 400000, Max: 4000000, Average: 466153.154	

Gambar 18. Hasil Statistic Performance

Statistic pada performance mencakup tentang hasil dari pengolahan data dengan menyatakan keterangan Qty berbentuk integer, Prediction(Qty) berbentuk integer, No berbentuk Real, Harga berbentuk integer, Total berbentuk integer, dan menunjukan bahwa prediksi penjualan kue kering pada toko damhil cake and bakery dengan menggunakan tools Rapidminer menunjukan hasil 0 untuk missing value.



Gambar 19. Hasil Visualizations

Hasil visualizations performance pada gambar diatas menunjukkan bahwa data excel dengan pengolahan data pada Rappidminer memiliki prediksi yang sama hasilnya.

4. HASIL DAN PEMBAHASAN

4.1. Hasil

Memuat hasil, Pengujian dan pembahasan tentang skripsi yang telah dilakukan Berdasarkan penelitian mengenai “Estimasi Penjualan Pada Toko Damhil Cake And Bakery Menggunakan Algoritma Regresi Linier Sederhana Dalam Menentukan Strategi Penjualan Kue Kering”.

- Hasil yang diperoleh dari prediksi penjualan kue kering untuk periode waktu tahun 2022 sebanyak 10.458pcs.
- Dari proses prediksi penjualan kue kering berdasarkan jenis nama produk kue kering tertinggi yaitu adalah kue kering mini pie keju, yang memiliki angka prediksi penjualan sekitar 5.546 pcs dengan total penjualan terendah untuk jenis kue kering nastar strawberry dengan angka penjualan 4.404 pcs.
- Pengukuran tingkat *error*, yang diperoleh *root mean squared error* (RMSE) dengan nilai 2.675+/- 0.000.

4.2. Pembahasan Pengujian

Root mean squared error (RSME) adalah metric evaluasi yang mengukur sejauh mana perbedaan antara nilai prediksi dan nilai sebenarnya. Hasil RMSE yang relative rendah menunjukkan model memiliki tingkat akurasi yang baik dalam memprediksi penjualan, dengan penggunaan algoritma regresi linier sederhana dinilai efektif dalam konteks estimasi penjualan kue kering serta model ini dapat digunakan untuk penyesuaian stok dan promosi yang dapat dilakukan lebih efisien dengan mengoptimalkan pendapatan dan tingkat kepuasan pelanggan. Dengan nilai RMSE yang rendah dan interval keyakinan yang sempit, model regresi linier sederhana dapat dianggap berhasil dalam mengestimasi penjualan produk pada Toko Damhil Cake and Bakery.

Langkah pengujian Root mean squared error untuk hasil *root_mean_squared_error* : 2.675 +/- 0.000 adalah sebagai berikut:

- Hitung Selisih Kuadrat:

$$\text{Squared Differences} = (2.675 - 2.675)^2 = 0$$
- Hitung Mean dari Selisih Kuadrat:

$$\text{Mean Squared Differences} = \frac{1}{1} \times 0 = 0$$
- Hitung RMSE:

$$\text{RMSE} = \sqrt{\text{Mean Squared Differences}} = \sqrt{0} = 0$$

Hasil Pengujian:

- Nilai Rata-rata RMSE: 2.6752.675
- Deviasi standar (ketidakpastian): 0.0000.000

Kesimpulan:

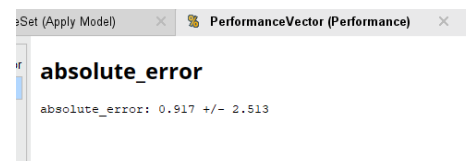
- Rata-rata RMSE dari model regresi adalah 2.6752.675, dengan ketidakpastian sekitar 0.0000.000.
- Mean dari selisih kuadrat yang dihitung adalah 00.
- RMSE yang dihitung berdasarkan rumus adalah 00, namun ini mungkin tidak merepresentasikan situasi yang umum dan dapat terjadi karena tidak adanya variasi dalam data (semua nilai sebenarnya dan prediksi sama).

root_mean_squared_error

```
root_mean_squared_error: 2.675 +/- 0.000
```

Gambar 20. hasil Root Mean Squared Error

Nilai 2.675 adalah nilai rata-rata dari sejumlah hasil RMSE, sedangkan 0.000 mengindikasikan deviasi standar atau kesalahan baku dari nilai-nilai tersebut. Hasil RMSE ini menunjukkan bahwa hasil RMSE dapat bervariasi dari nilai rata-rata, memberikan gambaran tentang konsistensi kinerja model, bahwa semakin kecil nilai RMSE, maka semakin baik model yang dapat memprediksi data dengan akurasi.



Gambar 21. Hasil Absolute Error

Diketahui :

- Rata-rata RMSE yang dilaporkan: 0.9170.917
- Deviasi standar (ketidakpastian): 2.5132.513

Maka langkah pengujian hasil absolute error adalah sebagai berikut:

1. Hitung Selisih Kuadrat

$$\text{Squared Differences} = (0.917 - 0)^2 = 0.841889$$

2. Hitung Mean dari Selisih Kuadrat:

$$\text{Mean Squared Differences} = \frac{1}{1} \times 0.841889 = 0.841889$$

3. Hitung RMSE: $\text{RMSE} =$

$$\sqrt{\text{Mean Squared Differences}} = \sqrt{0.841889} \approx 0.918$$

Hasil Pengujian:

- Nilai Rata-rata RMSE: 0.9170.917
- Deviasi standar (ketidakpastian): 2.5132.51

Kesimpulan:

- Rata-rata RMSE dari model regresi adalah 0.9170.917, dengan ketidakpastian sekitar 2.5132.513.
- Mean dari selisih kuadrat yang dihitung adalah 0.8418890.841889.

- RMSE yang dihitung berdasarkan rumus adalah sekitar 0.9180.918.

Maka dapat diketahui bahwa asil absolute error : 0.917 +/- 2.513 mencakup nilai rata-rata dari sejumlah hasil kesalahan mutlak (absolute error), dimana 0.917 adalah nilai rata-rata dan 2.513 adalah deviasi standar atau kesalahan baku dari nilai-nilai tersebut

squared_error

squared_error: 7.157 +/- 56.363

Gambar 22. Hasil Squared Error

Diketahui:

- Rata-rata dari squared error yang dilaporkan: 7.1577.157
- Deviasi standar (ketidakpastian): 56.36356.363

Langkah-langkah pengujian squared_error adalah sebagai berikut:

1. Hitung Mean Squared Differences:
Mean Squared Differences = 7.157
2. Hitung RMSE:

$$RMSE = \sqrt{\text{Mean Squared Differences}} = \sqrt{7.157} \approx 2.678$$

Hasil Pengujian:

- Nilai Rata-rata dari squared error: 7.1577.157
- Deviasi standar (ketidakpastian): 56.36356.363

Kesimpulan:

- Rata-rata dari squared error dari model adalah 7.1577.157, dengan ketidakpastian sekitar 56.36356.363.
- RMSE yang dihitung berdasarkan rumus adalah sekitar 2.6782.678.

Maka diketahui nilai squared_error 7.157 adalah nilai rata-rata dari sejumlah hasil kesalahan kuadrat (squared error), sementara 56.363 adalah deviasi standar atau kesalahan baku dari nilai-nilai tersebut. Kesalahan kuadrat mengukur seberapa jauh prediksi model dari nilai sebenarnya, dan deviasi standar memberikan gambaran tentang seberapa konsisten kesalahan dari satu pengukuran ke pengukuran lainnya.

Sedangkan, Model regresi linier sederhana digunakan untuk memodelkan hubungan antara variabel independen dan hasil penjualan, dengan mengestimasi koefisien regresi untuk setiap variabel independen. Regresi linier adalah metode statistik yang memungkinkan peneliti untuk memahami bagaimana perubahan dalam satu atau lebih variabel independen (harga produk, promosi dan waktu) dapat mempengaruhi variabel dependen (hasil penjualan).

LinearRegression

```
154.604 * NO
+ 43790.218 * QTY
+ 7.777 * HARGA
- 351135.576
```

Diketahui:

- Persamaan regresi linier sederhana: 154.645 x NO + 43790.218 x QTY + 7.777 x HARGA - 351135.576
- Dataset yang berisi nilai sebenarnya dan hasil prediksi untuk variabel NO, QTY, dan HARGA.

Langak-langkah pngujian Linear regression adalah sebagai berikut:

1. Hitung Prediksi Menggunakan Persamaan Regresi:

$$\text{Prediksi} = 154.645 \times \text{NO} + 43790.218 \times \text{QTY} + 7.777 \times \text{HARGA} - 351135.576.$$

2. Hitung Selisih Antara Data Aktual dan Prediksi:

$$\text{Selisih} = \text{Aktual} - \text{Prediksi}$$

3. Hitung selisih kuadrat:
Squared Differences = (Selisih)²

4. Hitung Mean dari Selisih Kuadrat:
Mean Squared Differences = $\frac{1}{n} \sum_i^n$

5. Hitung RMSE:

$$RMSE = \frac{\text{Mean Squared Differences}}{\text{Mean Squared Differences}} = \sqrt{\text{Mean Squared Differences}}$$

Hasil Pengujian:

- Nilai RMSE yang dihitung berdasarkan data aktual dan prediksi.

Maka dapat dilihat bahwa hasil regresi linier dihitung dengan mengalikan masing-masing koefisien (154.604, 43790.218, 7.777) dengan nilai yang sesuai dari variabel independen, dan kemudian menambahkan konstanta (-351135.576). Contohnya, jika nilai NO = a, QTY = b, dan HARGA = c.

5. KESIMPULAN DAN SARAN

Hasil penelitian ini dapat memberikan jawaban terhadap permasalahan yang dirumuskan dalam latar belakang, yaitu kesulitan untuk menentukan strategi penjualan yang efektif, adapun relevansi dari penelitian ini yaitu bertujuan untuk membantu Toko Damhil Cake and Bakery untuk membuat keputusan yang lebih akurat dalam merencanakan strategi penjualan kue kering dan meningkatkan efisiensi bisnis. Dapat disimpulkan bahwa penggunaan algoritma regresi linier sederhana dapat memberikan kontribusi yang positif dalam memprediksi penjualan kue kering menggunakan algoritma regresi linier sederhana dengan proses pengolahan data menggunakan tahapan dari Knowledge Discovery in Database (KDD), serta penggunaan tools Rapidminer

sebagai alat bantu dalam proses prediksi penjualan kue kering menghasilkan:

- a. Hasil yang diperoleh dari prediksi penjualan kue kering untuk periode waktu tahun 2022 sebanyak 10.458 pcs.
- b. Dari proses prediksi penjualan kue kering berdasarkan jenis nama produk kue kering tertinggi yaitu adalah kue kering mini pie keju, yang memiliki angka prediksi penjualan sekitar 5.546 pcs dengan total penjualan terendah untuk jenis kue kering nastar strawberry dengan angka penjualan 4.404 pcs.
- c. Pengukuran tingkat error, yang diperoleh root mean squared error (RMSE) dengan nilai 2.675+/- 0.000.

Sehubungan dengan fokus penelitian yang terkait dengan estimasi penjualan produk pada toko damhil cake and bakery menggunakan regresi linier sederhana, Penulis ingin memberikan beberapa saran yang dapat menjadi panduan dalam mengembangkan hasil penelitian.

Pertama, untuk meningkatkan keakuratan estimasi, disarankan untuk mengumpulkan data penjualan secara berkala dan memperbarui model regresi linier sesuai dengan perkembangan bisnis. Selanjutnya, untuk meningkatkan strategi penjualan kue kering, penulis menyarankan agar hasil estimasi digunakan untuk menyusun rencana promosi yang lebih terarah. Melakukan identifikasi periode dengan potensi penjualan yang tinggi dan memanfaatkan hasil estimasi untuk merencanakan promosi produk kue kering secara optimal, dan yang terakhir yaitu pentingnya untuk terus mengevaluasi strategi kinerja dengan melakukan penyesuaian yang diperlukan untuk tetap bersaing dan meningkatkan efektivitas penjualan. Dengan mempertimbangkan saran-saran ini, diharapkan penelitian ini dapat memberikan kontribusi yang signifikan dalam upaya meningkatkan strategi penjualan kue kering di Toko Damhil cake and Bakery.

DAFTAR PUSTAKA

- [1] T. Adianto, "Implementasi Data Mining untuk Prediksi Penjualan Produk Mayora Menggunakan Metode FP-Growth," vol. 3, no. 1, pp. 170–177, 2023, [Online]. Available: http://repository.uinsu.ac.id/id/eprint/20362%0Ahttp://repository.uinsu.ac.id/20362/3/BAB_II.pdf
- [2] Miftahuljannah, Aswan Supriyadi Sunge, and Ahmad Turmudi Zy, "Analisis Prediksi Penjualan Dengan Metode Regresi Linear Di Pt. Eagle Industry Indonesia," *J. Inform. Teknol. dan Sains*, vol. 5, no. 3, pp. 398–403, 2023, doi: 10.51401/jinteks.v5i3.3325.
- [3] A. D. A. Nasharudin and U. Ependi, "Analisis peramalan penjualan produk pada pt . enseval putera megatrading TBK menggunakan metode regresi linear sederhana," *JUPITER J. Penelit. ilmu dan Tek. komputar*, vol. 15, no. 1, pp. 317–326, 2023, [Online]. Available: <https://doi.org/10.5281/5318/15.jupiter.2023.04>
- [4] S. Supardi *et al.*, "Peran Data Mining dalam Memprediksi Tingkat Penjualan Sepatu Adidas Menggunakan Metode Algoritma Regresi Linear Sederhana," *J. Ekon. Manajemen Sist. Inf.*, vol. 4, no. 5, pp. 883–890, 2023, [Online]. Available: <https://creativecommons.org/licenses/by/4.0/>
- [5] C. E. Simbolon, "Penerapan Algoritma Regresi Linier Sederhana Dalam Memprediksi Keuntungan dan Kerugian Kelapa Sawit Pt . Sri Ulina," *J. Inf. Syst. Res.*, vol. 2, no. 2, pp. 169–172, 2021.
- [6] D. Setiawan, N. Surojudin, and W. Hadikristanto, "Prediksi Penjualan Obat Dengan Algoritma Regresi Linear," *Pros. Sains dan Teknol.*, vol. 1, no. 1, pp. 237–246, 2022.
- [7] M. Walid and F. Halimiyah, "Klasifikasi Kemandirian Siswa SMA/MA Double Track Menggunakan Metode Naive Bayes," *J. ICT Inf. Commun. Technol.*, vol. 22, pp. 190–197, 2022.
- [8] R. K. Purnomo, H. A. Saufi, and R. Hs, "Experimental Student Experiences," *Exp. Student Exp.*, vol. 1, no. 1, pp. 1–7, 2023.
- [9] O. Nurdian, A. Irma Purnamasari, and I. Ali, "Analisa Penjualan Mobil Dengan Menggunakan Algoritma K-Means Di PT. Mulya Putra Kencana," *J. Data Sci. dan Inform.*, vol. 1, no. 2, pp. 32–35, 2021.