

ANALISIS PERFORMANSI ALGORITMA LINEAR REGRESSION DENGAN GENERALIZED LINEAR MODEL UNTUK PREDIKSI PENJUALAN PADA USAHA MIKRO, KECIL, DAN MENENGAH

Fitria Habibatul Hamdanah¹, Devi Fitriana²

^{1,2}Teknik Informatika, Universitas Mercu Buana
Jakarta, Indonesia

e-mail: fitriahabibatul4@gmail.com¹, devifitriana1@gmail.com²

Abstrak

Penjualan merupakan syarat mutlak kelangsungan suatu usaha, karena dengan penjualan maka akan didapatkan keuntungan. Metode *Linear Regression* dan *Generalized Linear Model* merupakan metode pendekatan yang didukung dengan perhitungan RSME. RMSE (*Root Mean Square Error*) berfungsi untuk mendapatkan besaran tingkat kesalahan dari hasil prediksi, dimana semakin kecil (mendekati 0) nilai RMSE maka semakin akurat nilai prediksinya. Pada setiap Usaha Mikro Kecil Menengah (UMKM) aktivitas transaksi dan pelayanan terhadap konsumen setiap harinya semakin lama semakin meningkat, sehingga tanpa disadari hal ini dapat menimbulkan tumpukan data yang semakin membesar. UMKM biasanya mengeluarkan beberapa item berbeda untuk ditawarkan ke pasar dengan harga yang berbeda, namun tidak semua barang banyak peminatnya. Keberhasilan penjualannya menentukan keberlanjutan untuk umkm itu sendiri. Pada penelitian ini akan dibandingkan penggunaan algoritma *Linear Regression* dengan *Generalized Linear Model* yang diimplementasikan pada data penjualan yang sudah diinputkan sebelumnya guna menghasilkan prediksi penjualan barang untuk tahun berikutnya. Hasil perhitungan menunjukkan bahwa algoritma *Linear Regression* dengan nilai RSME, MSE, MAPE sebesar 1,983; 3,933; dan 1,518 sedangkan hasil dari algoritma *Generalized Linear Model* dengan nilai RSME, MSE, MAPE sebesar 4,827; 23,295; dan 3,882. Berdasarkan perhitungan prediksi oleh algoritma *Linear Regression* dan *Generalized Linear Model* dapat disimpulkan bahwa nilai RSME pada algoritma *Linear Regression* menunjukkan perhitungan paling baik dikarenakan nilai RSME paling kecil.

Kata kunci: Linear Regression, GLM, RSME, UMKM

Abstract

Sales are an absolute requirement for the continuity of a business because with sales you will get profits. Linear Regression and Generalized Linear Model methods are approaches that are supported by RSME calculations. RMSE (Root Mean Square Error) functions to obtain the magnitude of the error rate from the prediction results, where the smaller (close to 0) the RMSE value, the more accurate the prediction value will be. In every Micro, Small, and Medium Enterprise (UMKM), transaction activities and services to consumers are increasing every day, so that without realizing it, this can lead to a growing pile of data. MSMEs usually issue several different items to be offered to the market at different prices, but not all goods are in high demand. The success of its sales determines the sustainability of the UMKM itself. The purpose of this study is to compare the use of the Linear Regression algorithm with the Generalized Linear Model which is implemented on previously inputted sales data to produce predictions of sales of goods for the following year. The calculation results show that the Linear Regression algorithm with RSME, MSE, MAPE values is 1.983, 3.933, 1.518 while the results of the Generalized Linear Model algorithm with RSME, MSE, MAPE values are 4.827, 23.295, 3.882. Based on the calculation of predictions by the Linear Regression and Generalized Linear Model algorithms, it can be concluded that the RSME value in the Linear Regression algorithm shows the best calculation because the RSME value is the smallest.

Keywords : Linear regression, GLM, RSME, UMKM

PENDAHULUAN

Penjualan adalah aktivitas atau bisnis yang menjual produk atau jasa. Dalam proses penjualan, penjual atau penyedia barang dan jasa memberikan kepemilikan suatu komoditas kepada pembeli dengan harga tertentu. Penjualan merupakan syarat mutlak kelangsungan suatu usaha, karena dengan penjualan maka akan didapatkan keuntungan[1]. Penjualan dapat dilakukan melalui berbagai cara, seperti penjualan langsung, dan melalui agen penjualan, salah satunya ada di PINTER USAHA ini. PINTER USAHA merupakan suatu aplikasi yang dikembangkan oleh PT. WELLCODE untuk mengakomodasi transaksi penjualan di UMKM. Sudah sekitar lebih dari 100 UMKM yang bergabung menggunakan aplikasi PINTER USAHA. Di setiap UMKM, aktivitas transaksi dan pelayanan kepada konsumen semakin meningkat setiap harinya, sehingga tanpa disadari dapat mengakibatkan semakin banyaknya tumpukan data. Data yang besar ini perlu pemrosesan *Data Mining* untuk dapat menghasilkan informasi yang penting bagi keberlanjutan bisnis UMKM. UMKM biasanya mengeluarkan beberapa item berbeda untuk ditawarkan ke pasar dengan harga yang berbeda, namun tidak semua barang banyak peminatnya. Keberhasilan penjualannya menentukan keberlanjutan untuk umkm itu sendiri.

Pada penelitian yang dilakukan oleh Tri Indarwati pada tahun 2018 dengan judul "PENGUNAAN METODE LINEAR REGRESSION UNTUK PREDIKSI PENJUALAN SMARTPHONE" dilakukan pengujian prediksi penjualan *smartphone* dengan menggunakan data riwayat penjualan selama periode tahun 2016, menunjukkan bahwa persentase kesalahan prediksi berdasarkan nilai MAPE (*Mean Absolute Percentage Error*) dan MSE (*Mean Squared Error*) kurang dari 10%. Hasil tersebut mengindikasikan bahwa prediksi penjualan *smartphone* dengan metode *Linear Regression* berada pada kategori sangat baik[5].

Penelitian serupa dilakukan oleh Reno Supardi pada tahun 2020 dengan judul "PENERAPAN METODE REGRESI LINEAR DALAM MEMPREDIKSI DATA PENJUALAN BARANG DI TOKO BANGUNAN VITA VIYA". Penelitian ini menghasilkan sebuah aplikasi yang dapat memprediksi data penjualan barang di Toko Bangunan Vita Viya yang dibuat menggunakan bahasa pemrograman Visual Basic .Net. Aplikasi tersebut menerapkan metode *Linear Regression*. Berdasarkan pengujian akurasi prediksi yang dilakukan

didapatkan hasil akurasi prediksi menggunakan MAPE adalah sebesar 0,1244575. Sedangkan menggunakan MSE adalah sebesar 14,25[6].

Penelitian yang dilakukan oleh Ghebyla Najla Ayuni dengan judul "Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti pada PT XYZ". Hasil dari penelitian ini adalah peramalan penjualan properti dengan menggunakan metode *Linear Regression* yang dapat dikatakan dalam kategori cukup baik. Hal ini berdasarkan hasil uji MSE, RMSE, dan MAPE semua tipe properti di PT. XYZ yang memenuhi standar. Hasil yang diperoleh dari penelitian ini hanya merupakan perkiraan berdasarkan data selama periode waktu tertentu [7].

Penelitian yang dilakukan oleh WU Jia dengan judul "*Personalized Collaborative Filtering Recommendation Algorithm based on Linear Regression*". Hasil dari penelitian ini adalah nilai MAE (*Mean Absolute Error*) ada di set validasi, rasio ini adalah yang terkecil hanya 70% dari hasil eksperimen yang menyebabkan kesalahan prediksi algoritma PCFLR lebih sedikit daripada dua algoritma lainnya, dan akurasi prediksi meningkat secara signifikan, sehingga memverifikasi keefektifan algoritma[8].

Penelitian yang dilakukan oleh Hamidreza Hakimdavoodi dengan judul "*Maximum Likelihood Estimation of Generalized Linear Models with Generalized Gaussian residuals*". Hasil dari penelitian ini adalah hasil GLM sangat dekat satu sama lain, dan hasil yang didapat hampir sama. Namun, dalam perilaku noise tail-berat (h^2), GLM standar jelas bukan pilihan yang baik, dan GG-GLM memiliki efek yang lebih baik. GLM berbasis GG dan distribusi normal memberikan hasil yang hampir sama pada distribusi derau Gaussian biasa. Selain itu, GG-GLM menyediakan alat yang ampuh untuk memperkirakan koefisien regresi model linier, terutama jika data memiliki perilaku kesalahan ekor yang parah dan SNR yang rendah. Hasil percobaan menunjukkan keefektifan metode ini[9].

Penelitian yang dilakukan oleh Nora Muda dengan judul "*The performance of M-based Generalized Linear Model (GLM) Procedures based on the coverage probability*". Penelitian ini mengusulkan suatu metode perancangan data eksperimen berbasis GLM berbasis M, GLM dan metode kuadrat terkecil, khususnya 2 2 memiliki perbedaan respon abnormal dan non konstan. Kami membandingkan kinerja ketiga perkiraan ini berdasarkan cakupan. Berdasarkan contoh numerik, panjang interval kepercayaan penduga berbasis M sedikit lebih baik daripada GLM, dan penaksir kuadrat terkecil dalam data berkinerja baik. Studi empiris menunjukkan bahwa ketika *outlier* muncul dalam data, estimasi yang

diberikan oleh *estimator* berbasis M. lebih baik daripada GLM dan estimasi kuadrat terkecil, sehingga beberapa stimulator sekunder sangat sensitif terhadap keberadaan outlier[10].

Penelitian lain dilakukan oleh Ehud Vinepinsky dengan judul “*A Generalized Linear Model of a Navigation Network*”. Hasil penelitiannya secara keseluruhan memberikan pendekatan baru untuk meneliti *Neuron MEC*. Dengan menggunakan pendekatan GLM, penelitian tersebut berhasil mengukur bagaimana setiap variabel mempengaruhi respon saraf. Menerapkan kerangka ini ke populasi neuron lengkap di MEC bisa menjelaskan interaksi jaringan dan mengarah pemahaman yang lebih baik tentang cara di mana neuron dengan MEC menyandikan informasi navigasi[11].

Penelitian yang dilakukan oleh Tejas Kulkarni dengan judul “*Differentially Private Bayesian Inference for Generalized Linear Models*” merumuskan model yang peka terhadap gangguan untuk GLMs untuk melakukan inferensi DP Bayesian dan mendemonstrasikan kemanjurannya untuk kumpulan data dengan koefisien regresi kecil. Hasil ini menunjukkan bahwa model yang dibuat akurat untuk set data dengan true kecil θ untuk sampel sedang hingga ukuran sampel yang besar bahkan ketika menggunakan persyaratan privasi ketat. Dalam eksperimen internal tambahan pada dataset sintesis, peneliti mempelajari akurasi sebagai fungsi dari norma L_2 dari true θ , memperbaiki parameter lain, dan menemukan bahwa keakuratan menurun tajam ketika $\|\theta\|_2 \geq 3$ [12].

Penelitian yang dilakukan oleh ACC Coolen dengan judul “*Replica analysis of overfitting in generalized linear regression models*” menunjukkan hasil Derivasi yang hanya bergantung pada bentuk linear tergeneralisasi dari GLM dan saat memilih prior L_2 . Karena itu replika perhitungannya tidak perlu diulangi untuk setiap contoh model GLM baru; seperti biasa metode replika berfungsi sebagai kendaraan yang relatif tidak menyakitkan dan elegan tetapi kuat untuk sampai pada kumpulan persamaan parameter orde tertutup, bersama dengan rumus mengungkapkan hubungan antara penduga parameter ML / MAP dan benar nilai-nilai parameter ini[13].

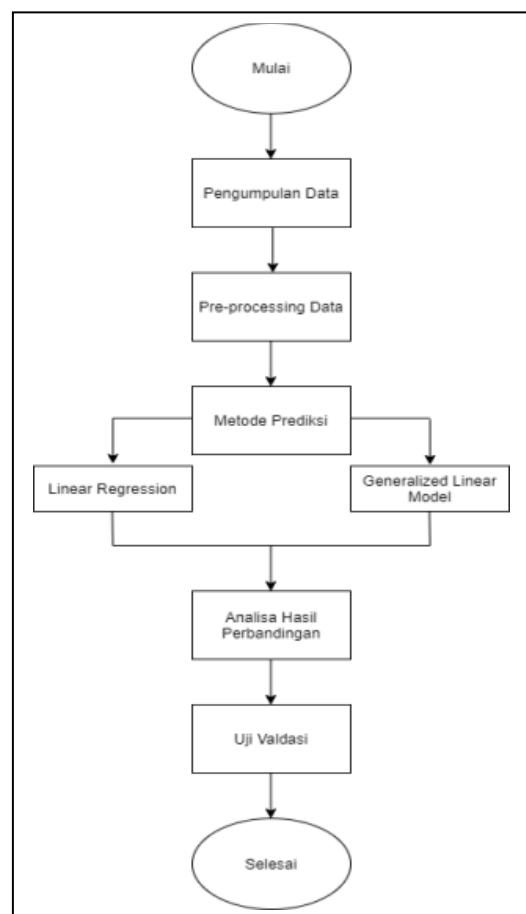
Tujuan dari penelitian ini adalah untuk mendapatkan prediksi penjualan barang untuk tahun berikutnya berdasarkan data penjualan yang telah diinputkan sebelumnya[2]. Data penjualan dimanfaatkan menjadi informasi dalam meningkatkan penjualan barang[3]. Penerapan untuk penelitian ini menggunakan sebuah metode prediksi dengan perbandingan

algoritma *Linear Regression* dan *Generalized Linear*.

Metode *Linear Regression* dan *Generalized Linear* yang didukung dengan perhitungan RSME[4]. RMSE (*Root Mean Square Error*) berfungsi untuk mendapatkan besaran tingkat kesalahan dari hasil prediksi, dimana semakin kecil (mendekati 0) nilai RMSE maka semakin akurat nilai prediksinya.

METODE

Dalam metodologi ini terdapat beberapa tahapan yang dilakukan selama penelitian. Pada Gambar 1, tahapan penelitian dijelaskan sebagai berikut:



Gambar 1. Tahapan Penelitian

A. Pengumpulan Data

Dalam penelitian ini data yang dibutuhkan adalah data penjualan barang pada UMKM di aplikasi PINTER USAHA dari bulan Januari hingga Desember Tahun 2019.

B. Preprocessing Data

Pada tahapan ini selanjutnya melakukan Preprocessing data yang dikelompokkan dari

data per-hari menjadi per-bulan. Kemudian menghilangkan atau mengeliminasi terhadap variabel dan informasi yang tidak dibutuhkan kembali.

C. Metode Prediksi

1. Algoritma *Linear Regression*

Metode *Linier Regression* adalah regresi yang melibatkan hubungan antara satu variabel dependen dengan satu variabel independen atau variabel dependen (Y) dan variabel independen (X)[14].

Hubungan variabel dependen dan variabel independen tergantung dalam beberapa bentuk persamaan, sebagai berikut hubungan *linear, eksponensial* dan yang terakhir berganda[15]. Tujuan penggunaan analisis regresi adalah untuk mengestimasi nilai variabel dependen yang didasarkan pada nilai variabel independen[16].

Metode *linier regression* didasarkan pada pola hubungan data terkait masa lalu. Secara umum variabel yang dapat diprediksi yang diwakili oleh variabel yang direpresentasikan oleh variabel (seperti persediaan) dipengaruhi oleh besar kecilnya variabel bebas. Hubungan yang terjadi antara variabel independen dengan variabel yang akan ditemukan adalah sebuah fungsi[17].

Regresi Linear memiliki model persamaan seperti berikut :

$$y = a + bx \quad (1)$$

Dimana:

Y = variabel Dependent

x = variabel Independent

a = konstanta

b = jumlah respon yang dihasilkan oleh prediktor[18].

Langkah-langkah metode yang diusulkan berdasarkan *Linier Regression* adalah sebagai berikut: (1) Pembuatan dataset yang terdiri dari data latih dan data uji. (2) Pembentukan model *Linier Regression* (model didasarkan pada data latih). Langkah-Langkah pembentukan model adalah sebagai berikut: (a) *Langkah 1*: Hitung X^2 , Y^2 , XY dan jumlah masing-masing. (b) *Langkah 2*: Hitung a yang menggunakan persamaan. (c) *Langkah 3*: Membuat model persamaan *Linier Regression*. (d) *Langkah 4*: Membuat prediksi atau ramalan dari variabel dependen dan variabel independen. (3) Pengujian kinerja didasarkan pada model prediksi yang telah dibuat dengan input data pengujian[19]. Pengujian prediksi meliputi mean square error (MSE), root mean square error

(RMSE), dan mean absolute persentase error (MAPE). [7].

2. Algoritma Generalized Linear Model

Generalized Linear Model

Merupakan perpanjangan dari model *Linier Regression* dengan asumsi bahwa prediktor memiliki efek linier tetapi tidak mengasumsikan distribusi spesifik dari variabel respon dan digunakan ketika variabel respon adalah anggota dari keluarga eksponensial.[20]. Variabel respon biasanya mengikuti anggota dari keluarga distribusi eksponensial, seperti normal, Poisson, binomial, gamma, dan invers Gaussian[21].

Metode GLM adalah Kuantifikasi Faktor independen dengan faktor (variabel respon) *Binding* (variabel prediktor), penggunaan Cara ini untuk menjelaskan perubahannya Jika ada, nilai faktor besar Perubahan faktor independen[22].

Generalized Linear Model adalah kelas model regresi yang banyak digunakan[23]. Pada model GLM, ada tiga bagian, salah satunya adalah fungsi link (fungsi *Contact person*) mampu menjelaskan ekspektasi Variabel tak bebas[24].

Model ini juga merupakan perpanjangan dari variabel dependen diperbolehkan memiliki distribusi yang berbeda[25]. Perspektif yang berbeda menyebabkan beberapa generalisasi pada koefisien penentuan[26]. Langkah-langkah metode yang diusulkan didasarkan pada *Generalized Linear Model* sebagai berikut: (1) *Mengumpulkan Data*. Salah satu komponen penting dalam proses penelitian. (2) *Menentukan Variabel dependent dan variable independent*. Variabel bebas (independent variable) merupakan jenis variabel yang menjelaskan atau mempengaruhi variabel lain. Variabel terikat (*dependent variable*) adalah jenis variabel yang dideskripsikan atau dipengaruhi oleh variabel bebas. (3) *Statistika Deskriptif*. Yang menggambarkan karakteristik data. (4) *Pengujian Linieritas*. Uji linieritas merupakan syarat sebelum dilakukan uji generalized linier models bertujuan untuk mengetahui apakah kedua variabel memiliki hubungan linier yang signifikan atau tidak. (5) *Identifikasi Distribusi Variabel dependent dan Pemilihan Fungsi Link*. Identifikasi variabel dependent perlu dilakukan dengan tepat distribusi dan mencocokkan data pada kurva

distribusi. (6) *Pemodelan Generalized Linier Model*. Dengan memasukkan semua variabel bebas dan variabel tak bebas. (7) *Pemilihan Model Terbaik*. Metode pemilihan model GLM terbaik ini digunakan untuk memilih model yang paling sesuai untuk menjelaskan hubungan antara variabel dependent (y) dan variabel independent (x)[20]. Model terbaik termasuk *mean square error* (MSE), *root mean square error* (RMSE), dan *mean absolute persentase error* (MAPE).

D. Analisa Hasil Perbandingan

Perbandingan hasil perhitungan nilai RSME, MAPE, MSE untuk metode prediksi dengan algoritma *Linear Regression* dan *Generalized Linear Model*. Diketahui dari dua algoritma tersebut yang dibandingkan, maka algoritma manakah yang lebih baik yang digunakan untuk penelitian ini.

E. Uji Validasi

Validasi prediksi adalah sebuah proses pengujian seberapa besar tingkat kesalahan yang terjadi antara data aktual dan data prediksi.

Terdapat berbagai cara untuk menghitung tingkat kesalahan tersebut yaitu MSE (*mean square error*) adalah rata-rata perbedaan antara nilai prediksi dan observasi, RMSE (*root mean square error*) adalah akar dari MSE, dan MAPE (*mean absolute persentase error*) adalah rata-rata dari diferensiasi absolut antara diprediksi dan aktual nilai-nilai[7].

MAPE menunjukkan seberapa besar kesalahan dalam prediksi dibandingkan dengan nilai sebenarnya[27]. Hasil nilai MSE dapat dikatakan sangat baik jika nilainya kurang dari 10%[5]. Sedangkan Hasil nilai RMSE dan MAPE dikatakan baik jika semakin kecil hasil nilai tersebut maka nilai prediksinya semakin akurat.

HASIL DAN PEMBAHASAN

Pada tahap ini hasil yang diperoleh pada tahap metodologi yang dijelaskan akan ditampilkan. Pertama dengan tahap pendataan yang diperoleh dari data penjualan barang pada UMKM pada aplikasi PINTER USAHA bulan Januari sampai dengan Desember 2019 dapat dilihat pada Tabel 1.

Tabel 1. Dataset

No. Pesanan	Nama Produk	Waktu Pengiriman Diatur	Waktu Pesanan Dibuat	Waktu Pembayaran Dilakukan	Harga Sebelum Diskon	Harga Setelah Diskon	Total Harga Produk	Total Diskon	Total Pembayaran
18081302111MDJ5	GRATIS ONGKIR!!! Ikea Pruta set isi 17 Tempat Makanan wadah toples makanan penyimpanan makanan	04/01/2019 09:21	03/01/2019 01:11	02/01/2019 01:15	Rp 77.870	Rp 62.500	Rp 62.500	Rp 15.370	Rp 63.500
180813679FADER	Packing Tambahan Bubble Wrap dan kardus, Packing Kado untuk Hadiah	04/01/2019 09:21	03/01/2019 01:11	02/01/2019 01:15	Rp 1.000	Rp 1.000	Rp 1.000	Rp 0	RP 3.000
18081307021Q154	GRATIS ONGKIR!!! Ikea Pruta set isi 17 Tempat Makanan wadah toples makanan penyimpanan makanan	04/01/2019 09:21	03/01/2019 01:11	02/01/2019 01:15	Rp 77.870	Rp 62.500	Rp 62.500	Rp 15.370	Rp 62.500
18081314121CCXF	IKEA POKAL Gelas Cangkir Mug Minuman 35 cl Kaca Bening	07/01/2019 09:21	06/01/2019 13:12	05/01/2019 13:14	Rp 12.870	Rp 12.870	Rp 38.610	Rp 0	Rp 38.610
18081317211PR2B	GRATIS ONGKIR!!! Ikea Pruta set isi 17 Tempat Makanan wadah toples makanan penyimpanan	07/01/2019 09:21	06/01/2019 13:12	05/01/2019 13:14	Rp 77.870	Rp 62.500	Rp 62.500	Rp 15.370	Rp 74.000

Tabel 2. Hasil *Preprocessing* Data

Nama Produk	Bulan											
	1	2	3	4	5	6	7	8	9	10	11	12
1	20	54	51	39	26	21	30	14	12	19	24	29
2	28	39	32	24	17	13	19	9	10	12	22	19
3	3	6	6	4	2	2	3	6	9	8	3	8
4	4	0	6	4	2	2	3	3	8	6	7	4
5	3	5	7	4	2	2	3	5	0	9	3	6
6	3	5	7	4	2	2	3	2	3	4	5	3
7	10	18	17	7	7	6	8	13	8	5	10	14
8	4	6	7	6	4	4	5	3	10	8	4	4

Tabel 3. Hasil prediksi Algoritma *Linear Regression*

Bulan	Prediction	Nama Produk	C	D	E	F	G	H	I	J	K	L	M
1	1.863	11	2	3	4	5	6	7	8	9	10	11	12
20	22.639	1	54	51	39	26	21	30	14	12	19	24	29
28	22.930	2	39	32	24	17	13	19	9	10	12	22	19
3	5.321	3	6	6	4	2	2	3	6	9	8	3	8

Kemudian tahap selanjutnya adalah tahap *preprocessing* data yang dikelompokkan dari data per-hari menjadi per-bulan. Kemudian menghilangkan atau mengeliminasi terhadap variabel dan informasi yang tidak dibutuhkan kembali sehingga menyisakan variabel x dan variabel y seperti pada Tabel 2.

Tahapan selanjutnya yaitu masuk kedalam model prediksi yaitu sebagai berikut:

A. Penerapan algoritma *Linear Regression*

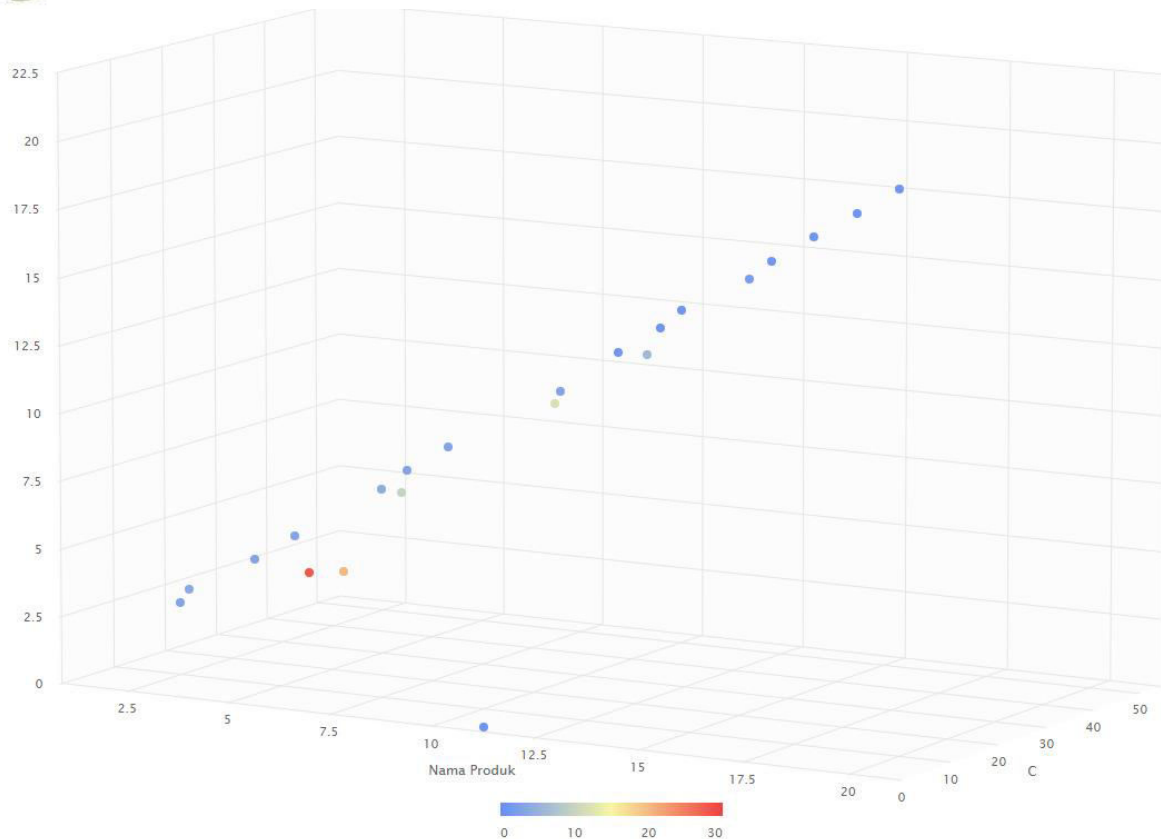
Pada Tabel 3 dapat dilihat hasil prediksi dari data penjualan yang telah melewati tahapan *preprocessing*. Dapat terlihat pada

bulan Januari dengan nilai prediksi sebesar 1.863.

Pada Gambar 2. Menunjukkan hasil prediksi yang menunjukkan bahwa hasil RSME, MSE, MAPE data yang diuji adalah sebesar 1,983; 3,933; 1,518.

B. Penerapan Algoritma *Generalized Linear Model*

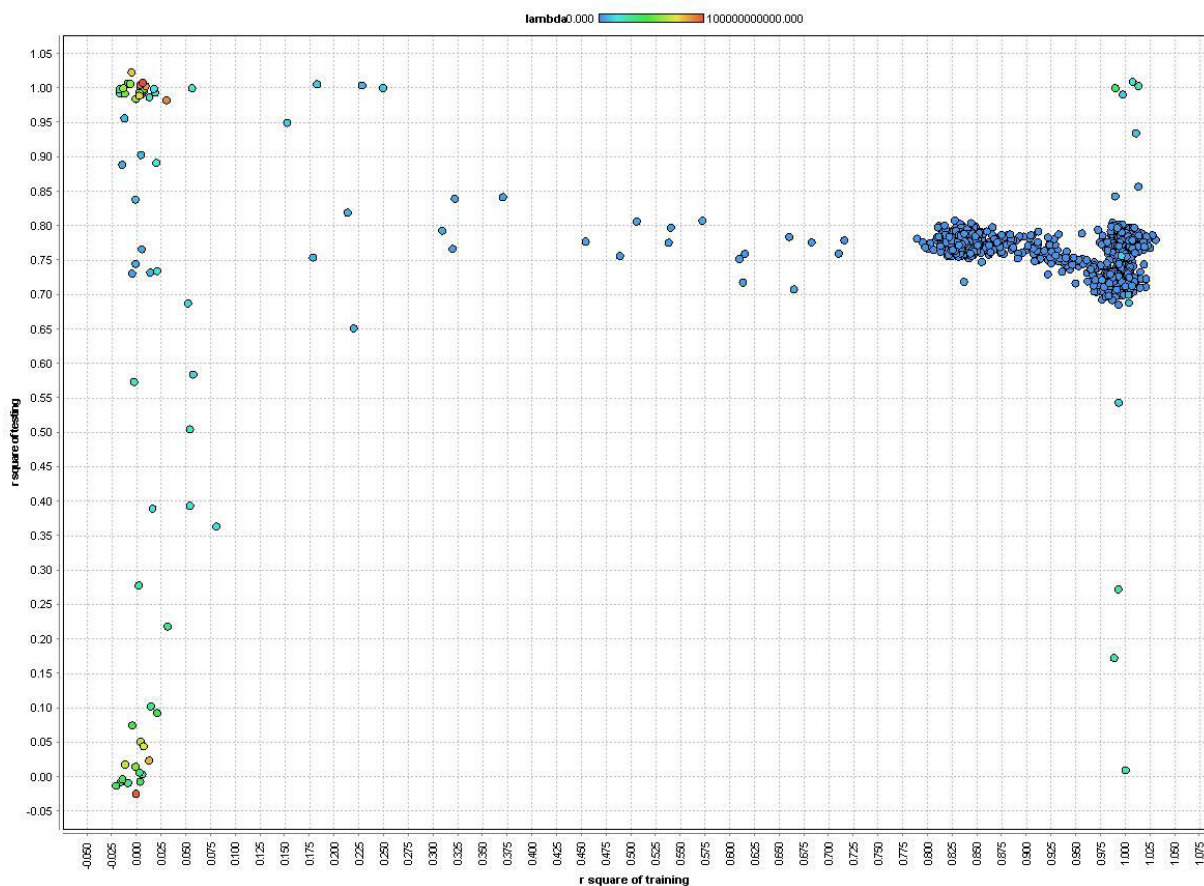
Pada Tabel 4 dapat dilihat hasil prediksi dari data penjualan yang juga telah melewati tahapan *preprocessing*. Dapat terlihat bahwa r squared ini menunjukkan seberapa dekat terhadap garis regresi.



Gambar 2. Grafik Hasil prediksi *Algoritma Linear Regression*

Tabel 4. Hasil prediksi *Algoritma Generalized Linear Model*

R square	R square	Lambda
0.999	0.778	0.000
0.999	0.778	0.000
0.999	0.778	0.000
0.999	0.778	0.000
0.999	0.778	0.000
0.999	0.778	0.000
0.999	0.778	0.000
0.999	0.778	0.000
0.999	0.778	0.000
0.999	0.778	0.000



Gambar 3. Grafik Hasil prediksi Algoritma *Generalized Linear Model*

Berdasarkan pada Gambar 3. Grafik hasil prediksi algoritma *Generalized Linear Model* dengan penjualan periode 2019 menunjukkan bahwa hasil nilai RSME, MSE, MAPE yang diuji adalah 4,827; 23,295; 3,882.

C. Membandingkan Metode Prediksi

Tahapan selanjutnya adalah Membandingkan antara metode *Linear Regression* dengan *Generalized Linear Model*. Kedua algoritma tersebut di implementasikan untuk mengetahui algoritma yang cocok digunakan dalam penelitian ini. Untuk melihat tingkat kesalahan hasil prediksi dengan menggunakan RSME (*root mean square error*) dan MAPE, dimana semakin kecil (mendekati 0) nilai RSME, semakin akurat nilai prediksi. Hasil nilai MSE dapat dikatakan sangat baik jika nilainya kurang dari 10%[5]. Hasil dari perbandingan tersebut dapat dilihat pada Gambar 8.

Berdasarkan pada Tabel 5 Algoritma *Linear Regression* mendapatkan nilai RSME yang kecil atau mendekati 0. Maka dapat disimpulkan dari perbandingan *Linear Regression* dan *Generalized Linear Model* bahwa algoritma *Linear Regression* lebih baik dibandingkan dengan *Generalized Linear Model*.

Tabel 5. Hasil Pengujian Keakurasian

Metode	Nilai RSME	Nilai MSE	Nilai MAPE
Linear Regression	1,983	3,933	1,518
Generalized Linear Model	4,827	23,295	3,882

Penelitian yang dilakukan oleh Andi Bode menggunakan metode mesin vektor dukungan pabrikan untuk mendapatkan kumpulan data pesanan penjualan dengan RMSE 0,127, menggunakan mesin vektor dukungan dan metode eliminasi mundur untuk mendapatkan RMSE 0,155, dan metode *Linier Regression* untuk mendapatkan RMSE 0,188, dan dengan menggunakan RMSE diperoleh *Linier Regression* dan metode eliminasi mundur sebesar 0,188[16].

KESIMPULAN

Dari hasil penelitian dapat disimpulkan bahwa metode *Linear Regression* dengan nilai RSME, MSE, MAPE sebesar 1,983; 3,933; 1,518. Dibandingkan dengan metode

Generalized Linear Model dengan hasil RSME, MSE, MAPE sebesar 4,827; 23,295; 3,882. Hal ini terjadi dikarenakan nilai RSME mendekati 0. Penelitian ini juga membuktikan bahwa dengan menggunakan algoritma *Linear Regression* dan *Generalized Linear Model* dapat memberikan prediksi terhadap penjualan pada UMKM di PINTER USAHA sehingga menghasilkan prediksi penjualan barang untuk tahun berikutnya. Hasil penelitian ini hanya estimasi berdasarkan data untuk periode tersebut dalam periode waktu tertentu, pengembang dapat menggunakan informasi tersebut sebagai pertimbangan untuk meminimalisir kelebihan barang pada barang yang penjualannya menurun, dan memaksimalkan barang yang penjualannya meningkat. Saran yang didapat pada penelitian ini adalah agar menambahkan variabel yang lebih banyak sehingga membuat hasilnya lebih akurat.

UCAPAN TERIMA KASIH

Dalam penelitian ini penulis juga dibantu oleh beberapa pihak yang mendukung kelancaran pelaksanaan penelitian ini. Oleh karena itu, penulis mengucapkan terima kasih kepada PT Wellcode yang telah memberikan izin kepada penulis untuk melakukan pengambilan data. Penulis juga mengucapkan terima kasih kepada Dr. Devi Fitriana selaku pembimbing.

REFERENSI

- [1] R. Gusrizaldi and E. Komalasari, "Analisis Faktor-Faktor Yang Mempengaruhi Tingkat Penjualan Di Indrako Swalayan Teluk Kuantan," *J. Valuta*, vol. 2, no. 2, pp. 286–303, 2016.
- [2] E. Kwok and W. Susanti, "Penerapan Metode Regresi Linier dalam Aplikasi Sistem Peramalan Jumlah Bahan Baku untuk Produksi Tahu," *Mhs. Apl. Teknol. Komput. dan Inf.*, vol. 1, no. 2, pp. 1–8, 2019.
- [3] A. F. Lestari and M. Hafiz, "Penerapan Algoritma Apriori Pada Data Penjualan Barbar Warehouse," *INOVTEK Polbeng - Seri Inform.*, vol. 5, no. 1, p. 96, 2020, doi: 10.35314/isi.v5i1.1317.
- [4] M. Merfin and R. S. Oetama, "Prediksi Harga Saham Perusahaan Perbankan Menggunakan Regresi Linear Studi Kasus Bank BCA Tahun 2015-2017," *Ultim. J. Tek. Inform.*, vol. XI, no. 1, pp. 2015–2019, 2019.
- [5] T. Indarwati, T. Irawati, and E. Rimawati, "Penggunaan Metode Linear Regression Untuk Prediksi Penjualan Smartphone," *J. Teknol. Inf. dan Komun.*, vol. 6, no. 2, pp. 2–7, 2019, doi: 10.30646/tikomsin.v6i2.369.
- [6] B. V. Viya, "Penerapan Metode Regresi Linear Dalam Memprediksi Data Penjualan Barang Di Toko," vol. 3, no. x, pp. 11–18, 2020.
- [7] G. N. Ayuni and D. Fitriana, "Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti pada PT XYZ," *J. Telemat.*, vol. 14, no. 2, pp. 79–86, 2019.
- [8] J. Wu, C. Liu, W. Cui, and Y. Zhang, "Personalized Collaborative Filtering Recommendation Algorithm based on Linear Regression," *2019 IEEE Int. Conf. Power Data Sci. ICPDS 2019*, no. 1, pp. 139–142, 2019, doi: 10.1109/ICPDS47662.2019.9017166.
- [9] H. Hakimdavoodi and M. Amirmazighani, "Maximum likelihood estimation of generalized linear models with generalized Gaussian residuals," *Proc. - 2016 2nd Int. Conf. Signal Process. Intell. Syst. ICSPIS 2016*, pp. 14–15, 2017, doi: 10.1109/ICSPIS.2016.7869893.
- [10] N. Muda, "The performance of M-based Generalized Linear Model (GLM) procedures based on the coverage probability," *2009 World Congr. Nat. Biol. Inspired Comput. NABIC 2009 - Proc.*, pp. 596–599, 2009, doi: 10.1109/NABIC.2009.5393419.
- [11] E. Vinepinsky, S. Perchik, and R. Segev, "A Generalized Linear Model of a Navigation Network," *Front. Neural Circuits*, vol. 14, no. September, pp. 1–17, 2020, doi: 10.3389/fncir.2020.00056.
- [12] T. Kulkarni, J. Jälkö, A. Koskela, S. Kaski, and A. Honkela, "Differentially Private Bayesian Inference for Generalized Linear Models," pp. 1–20, 2020.
- [13] A. C. C. Coolen, M. Sheikh, A. Mozeika, F. Aguirre-Lopez, and F. Antenucci, "Replica analysis of overfitting in generalized linear regression models," *J. Phys. A Math. Theor.*, vol. 53, no. 36, 2020, doi: 10.1088/1751-8121/aba028.
- [14] A. Kurniadi and Y. Novianto, "Penerapan Metode Regresi Linier untuk Memprediksi Kebiasaan Pelanggan Studi Kasus : PT . Mensa Binasukses," *J. Ilm. Mhs. Tek. Inform.*, vol. 2, no. 2, p. 107, 2020.
- [15] T. Rizgitta, V. C. M, and J. Hendryli, "Sistem Analisis Kinerja Sales Berdasarkan Transaksi Penjualan Dengan Regresi Linear Dan Algoritma Apriori," pp. 173–178.
- [16] A. Bode, "Perbandingan Metode Prediksi Support Vector Machine Dan Linear Regression Menggunakan Backward Elimination Pada Produksi Minyak Kelapa," *J. Sist. Inf. dan Tek. Komput.*, vol. 4, no. 2,

- pp. 104–107, 2019.
- [17] D. A. Trianggana, “Peramalan Jumlah Siswa-Siswi Melalui Pendekatan Metode Regresi Linear,” *J. Media Infotama*, vol. 16, no. 2, pp. 115–120, 2020.
- [18] R. Septyawan, “Analisis Peramalan Kebutuhan Energi Listrik PLN Area Batam Menggunakan Metode Regresi Linear,” 2018.
- [19] A. Bengnga and R. Ishak, “Prediksi Jumlah Mahasiswa Registrasi Per Semester Menggunakan Linier Regresi Pada Universitas Ichsan Gorontalo,” *Ilk. J. Ilm.*, vol. 10, no. 2, pp. 136–143, 2018, doi: 10.33096/ilkom.v10i2.274.136-143.
- [20] Jamilatuzzahro., R. Herliansyah, and R. E. Caraka, *Aplikasi generalized linear model pada R*, no. February. 2018.
- [21] Y. Wilandari, S. H. Kartiko, and A. R. Effendie, “Estimasi Cadangan Klaim Menggunakan Generalized Linear Model (Glm) Dan Copula,” *J. Gaussian*, vol. 9, no. 4, pp. 411–420, 2020, doi: 10.14710/j.gauss.v9i4.29260.
- [22] G. L. Arnenda, “INTERAKSI BEBERAPA FAKTOR YANG BERPENGARUH TERHADAP HASIL TANGKAPAN PER UNIT USAHA (CATCH PER UNIT EFFORT) ALBAKORA (Thunnus Alalunga Bonnaterre,1788) DI SAMUDRA HINDIA,” *JFMR-Journal Fish. Mar. Res.*, vol. 3, no. 3, pp. 382–389, 2019, doi: 10.21776/ub.jfmr.2019.003.03.14.
- [23] D. J. Eck and F. W. Crawford, “Conformal prediction for exponential families and generalized linear models,” *arXiv*, no. May, 2019.
- [24] J. Matematika, E. Saut, P. Situmorang, B. Susanto, and L. Ricky, “d ’ Cartesian Estimasi Parameter Copula Plackett Untuk Data Bivariat Melalui Metode Generalized Linear Model Pada Regresi Mediannya.”
- [25] M. Marzjarani, “A Comparison of a General Linear Model and the Ratio Estimator,” *Int. J. Stat. Probab.*, vol. 9, no. 3, p. 54, 2020, doi: 10.5539/ijsp.v9n3p54.
- [26] D. Zhang, “A Coefficient of Determination for Generalized Linear Models,” *Am. Stat.*, vol. 71, no. 4, pp. 310–316, 2017, doi: 10.1080/00031305.2016.1256839.
- [27] J. Adhiva, S. A. Putri, and S. G. Setyorini, “Prediksi Hasil Produksi Kelapa Sawit Menggunakan Model Regresi Pada PT . Perkebunan Nusantara V,” pp. 155–162, 2020.