

Prediksi Penjualan Live Streaming dengan Regresi Linear Menggunakan Stochastic Gradient Descent

Ginda M. A. Siregar¹, Muhammad Farid Fitriansyah^{2*}, Rizalul Akram³, Irwanda Syahputra⁴, Ade Zulkarnain Hasibuan⁵

^{1,2,3,4,5}Program Studi Informatika, Universitas Samudra

E-mail: , ¹ginda.mas@unsam.ac.id, ²1mhdFarid1811@gmail.com, ³rizalulakram@unsam.ac.id, ⁴irwandasyahputra@unsam.ac.id, ⁵adezulhsb@unsam.ac.id

Abstrak

Live streaming commerce menjadi strategi pemasaran digital yang memungkinkan interaksi langsung antara penjual dan konsumen secara *real-time*. Peningkatan durasi siaran dan jumlah penonton aktif tidak selalu diikuti peningkatan jumlah produk terjual. Penelitian ini bertujuan membangun model prediksi jumlah produk terjual berdasarkan durasi *live* (X_1) dan jumlah penonton aktif (X_2) menggunakan regresi linear yang dioptimasi dengan algoritma *Stochastic Gradient Descent* (SGD). *Dataset* historis aktivitas *live streaming* digunakan sebagai data pelatihan dan evaluasi model. Kinerja model diukur menggunakan metrik R^2 , *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE). Hasil pengujian menunjukkan model memperoleh nilai R^2 sebesar 0.4995, MAE sebesar 10.0113, RMSE sebesar 14.0243, dan MAPE sebesar 66.94%. Model kemudian diimplementasikan dalam *dashboard* berbasis *Streamlit* untuk menyediakan simulasi prediksi penjualan secara *real-time* guna mendukung perencanaan strategi *live streaming* yang lebih efektif dan berbasis data.

Kata kunci— *Dashboard, Live Streaming, Prediksi Penjualan, Regresi Linear, Stochastic Gradient Descent*

Abstract

Live streaming commerce has become an effective digital marketing strategy that enables *real-time* interaction between sellers and consumers. However, longer broadcast duration and higher numbers of active viewers do not always lead to increased product sales. This study aims to develop a prediction model for the number of products sold based on live duration (X_1) and the number of active viewers (X_2) using linear regression optimized with the *Stochastic Gradient Descent* (SGD) algorithm. Historical live streaming data were used as training and evaluation datasets. Model performance was evaluated using R^2 , *Mean Absolute Error* (MAE), *Root Mean Square*

Error (RMSE), and Mean Absolute Percentage Error (MAPE). The results show that the model achieved an R^2 of 0.4995, MAE of 10.0113, RMSE of 14.0243, and MAPE of 66.94%. The model was implemented in a Streamlit-based dashboard to provide real-time sales prediction simulations for data-driven live streaming strategies.

Keywords— *Dashboard , Live streaming, Sales Prediction, Linear Regression, Stochastic Gradient Descent*

1. PENDAHULUAN

Perkembangan *e-commerce* dalam beberapa tahun terakhir telah mendorong perubahan signifikan dalam strategi pemasaran digital. Salah satu inovasi yang berkembang pesat adalah *live streaming commerce*, yaitu metode pemasaran yang memungkinkan penjual mempromosikan produk secara langsung melalui siaran daring serta berinteraksi secara *real-time* dengan konsumen. Pendekatan ini dinilai efektif dalam meningkatkan keterlibatan audiens dan mempercepat proses pengambilan keputusan pembelian [1], [2]. Platform seperti *ShopeeLive* menjadi salah satu media yang banyak dimanfaatkan pelaku usaha untuk meningkatkan performa penjualan melalui komunikasi dua arah dan demonstrasi produk secara langsung [3].

Meskipun aktivitas *live streaming* semakin populer, peningkatan durasi siaran dan jumlah penonton aktif tidak selalu berbanding lurus dengan peningkatan jumlah produk terjual. Dalam praktiknya, terdapat sesi siaran dengan jumlah penonton tinggi namun menghasilkan konversi penjualan yang rendah. Kondisi ini menunjukkan bahwa hubungan antara durasi *live*, jumlah penonton aktif, dan jumlah produk terjual tidak selalu bersifat linier sederhana, sehingga diperlukan pendekatan kuantitatif berbasis data untuk memahami pola dan pengaruh antar variabel tersebut [4], [5]. Analisis berbasis data menjadi penting untuk menghasilkan model prediksi yang dapat membantu pengambilan keputusan secara lebih objektif dan terukur [6], [7].

Berbagai penelitian sebelumnya telah memanfaatkan metode data mining dan machine learning untuk menganalisis serta memprediksi penjualan. Pendekatan yang umum digunakan meliputi regresi linear untuk estimasi numerik, algoritma *Long Short-Term Memory* (LSTM) untuk prediksi berbasis deret waktu, serta teknik *clustering* dan asosiasi untuk menemukan pola transaksi [8], [9]. Meskipun metode tersebut menunjukkan performa yang baik, sebagian besar penelitian lebih berfokus pada klasifikasi atau analisis pola pembelian tanpa mengkaji secara mendalam proses optimasi parameter dalam model regresi. Metode berbasis *deep learning* seperti LSTM cenderung membutuhkan kompleksitas komputasi yang lebih tinggi serta jumlah data yang besar [10], [11].

Salah satu pendekatan optimasi yang banyak digunakan dalam pembelajaran mesin adalah *Stochastic Gradient Descent* (SGD). Algoritma ini bekerja dengan memperbarui parameter model secara iteratif berdasarkan gradien *error* pada setiap data atau mini-batch, sehingga memungkinkan proses pelatihan yang lebih efisien dibandingkan metode *batch gradient descent* [12], [13]. Sejumlah penelitian

menunjukkan bahwa performa SGD sangat dipengaruhi oleh pengaturan *hyperparameter* seperti *learning rate* dan jumlah iterasi, serta karakteristik *dataset* yang digunakan [14], [15]. Penerapan dan pengaturan optimasi SGD secara tepat berpotensi meningkatkan akurasi model regresi dalam konteks prediksi numerik.

Penelitian ini bertujuan untuk menganalisis pengaruh durasi *live* dan jumlah penonton aktif terhadap jumlah produk terjual menggunakan model regresi yang dioptimasi dengan algoritma *Stochastic Gradient Descent* (SGD). Perbandingan antara model regresi linear, polinomial, dan logaritmik juga dilakukan untuk menentukan model terbaik berdasarkan metrik evaluasi R^2 , *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE). Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan model prediksi penjualan berbasis *live streaming* yang lebih efisien, terukur, serta aplikatif dalam mendukung strategi pemasaran digital.

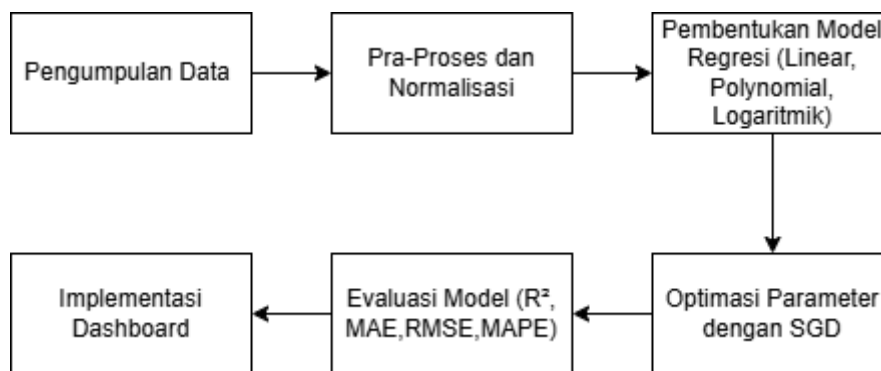
2. METODE

2.1 Perancangan Model

Penelitian ini mengusulkan model prediksi penjualan berbasis regresi multivariabel yang dioptimasi menggunakan algoritma *Stochastic Gradient Descent* (SGD). Model dirancang untuk menganalisis pengaruh durasi *live streaming* dan jumlah penonton aktif terhadap jumlah produk terjual. Pendekatan ini dipilih karena mampu memberikan estimasi numerik yang terukur serta memungkinkan proses optimasi parameter secara iteratif melalui minimisasi fungsi kerugian.

Arsitektur model terdiri dari beberapa tahapan utama, yaitu pengumpulan data, pra-proses dan normalisasi, pembentukan model regresi, optimasi parameter menggunakan SGD, evaluasi performa model, serta implementasi model dalam bentuk *dashboard* prediksi berbasis web. Tahapan tersebut disusun secara sistematis untuk memastikan proses analisis berjalan terstruktur dan menghasilkan model dengan performa optimal.

Diagram alur metode penelitian ditunjukkan pada Gambar 1, yang menggambarkan keterkaitan antar tahapan mulai dari input data hingga penentuan model terbaik.



Gambar 1 Diagram Alur Penelitian

Diagram tersebut menunjukkan bahwa proses pemodelan dilakukan secara sistematis dan terstruktur untuk memastikan hasil prediksi yang optimal.

2.2 Analisis Data

Tahap analisis data dalam penelitian ini bertujuan untuk membangun dan menguji model prediksi penjualan berdasarkan *dataset* aktivitas *live streaming* pada platform *ShopeeLive*. *Dataset* yang digunakan terdiri dari tiga variabel utama, yaitu durasi *live* (X_1) dalam satuan jam, jumlah penonton aktif (X_2) dalam satuan orang, dan jumlah produk terjual (Y) dalam satuan unit produk. Variabel X_1 dan X_2 berperan sebagai variabel independen, sedangkan variabel Y merupakan variabel dependen yang diprediksi menggunakan pendekatan regresi multivariabel. Proses analisis dilakukan secara sistematis melalui tahapan *preprocessing* data, pembentukan model regresi, optimasi parameter, dan evaluasi performa model.

2.2.1 Pra-Proses Data

Sebelum dilakukan pemodelan, *dataset* melalui tahap pra-proses untuk memastikan kualitas dan konsistensi data. Tahap ini mencakup seleksi variabel yang relevan sesuai tujuan penelitian, pemeriksaan kelengkapan data untuk mengidentifikasi kemungkinan nilai kosong (*missing values*) atau anomali, serta normalisasi data menggunakan metode *StandardScaler*. Normalisasi dilakukan untuk menyamakan skala antar variabel karena algoritma berbasis gradien seperti *Stochastic Gradient Descent* (SGD) sensitif terhadap perbedaan rentang nilai. Penyamaan skala membantu mempercepat proses konvergensi dan meningkatkan stabilitas pembaruan parameter selama pelatihan model.

Sebelum dilakukan proses normalisasi, data penelitian masih memiliki rentang nilai yang berbeda antar variabel. Perbedaan skala data tersebut berpotensi mempengaruhi proses optimasi pada algoritma *Stochastic Gradient Descent* (SGD), terutama pada kecepatan konvergensi model. Contoh data sebelum proses normalisasi ditunjukkan pada Tabel 1.

Tabel 1 Hasil Normalisasi Sampel Data

Durasi <i>Live</i>	Penonton Aktif	Setelah Normalisasi X_1	Setelah Normalisasi X_2
7.9128	19	-0.41	-0.52
11.7703	20	0.12	-0.49
12.3764	50	0.20	0.66

Sampel hasil normalisasi data menggunakan metode *StandardScaler* ditunjukkan pada Tabel 1. Proses normalisasi dilakukan untuk menyamakan rentang nilai antar variabel independen sehingga proses optimasi menggunakan algoritma *Stochastic Gradient Descent* (SGD) dapat berlangsung lebih stabil dan efisien.

2.2.2 Pembentukan Model Regresi

Untuk mengidentifikasi bentuk hubungan terbaik antara variabel independen dan variabel dependen, penelitian ini membangun tiga model regresi sebagai pembandingan, yaitu regresi linear, regresi polinomial, dan regresi logaritmik. Model regresi linear multivariabel dinyatakan dalam persamaan:

$$Y = a + b_1X_1 + b_2X_2 + c \quad (1)$$

di mana β_1 dan β_2 merupakan koefisien regresi yang menunjukkan kontribusi masing-masing variabel independen terhadap variabel dependen, dan b adalah *intercept*. Model ini mengasumsikan bahwa hubungan antara durasi *live*, jumlah penonton aktif, dan jumlah produk terjual bersifat linier. Untuk menangkap kemungkinan hubungan *nonlinier*, digunakan regresi polinomial orde dua yang dinyatakan sebagai:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_1^2 + b_4X_1X_2 + b_5X_2^2 \quad (2)$$

Model ini memungkinkan representasi hubungan yang lebih kompleks dengan mempertimbangkan efek kuadrat dari masing-masing variabel serta interaksi antara durasi *live* dan jumlah penonton aktif. Model regresi logaritmik juga di gunakan untuk mengakomodasi pola pertumbuhan yang tidak linier melalui transformasi logaritma pada variabel independen. Persamaan regresi logaritmik dinyatakan sebagai:

$$Y = a + b_1\ln(X_1) + b_2\ln(X_2) \quad (3)$$

Transformasi logaritma digunakan untuk menangani kondisi di mana perubahan pada variabel independen menghasilkan efek yang tidak proporsional terhadap variabel dependen.

2.2.3 Optimasi Parameter Menggunakan SGD

Parameter model regresi dioptimasi menggunakan algoritma *Stochastic Gradient Descent* (SGD) dengan tujuan meminimalkan fungsi kerugian yang dinyatakan sebagai Mean Squared Error (MSE). Proses pembaruan parameter dilakukan secara iteratif

berdasarkan gradien fungsi kerugian terhadap parameter model. Persamaan umum pembaruan parameter dinyatakan sebagai:

$$\theta_t = \theta_{t-1} - \eta \nabla J(\theta) \quad (4)$$

di mana θ merupakan parameter model, η adalah *learning rate*, dan $\nabla J(\theta)$ adalah gradien dari fungsi kerugian. Dalam penelitian ini digunakan *learning rate* sebesar 0.01, maksimum iterasi (*epoch*) sebanyak 1000, serta nilai tolerance sebesar 0.001 sebagai batas konvergensi. Pemilihan parameter tersebut dilakukan melalui pengujian awal untuk memperoleh proses optimasi yang stabil tanpa fluktuasi signifikan selama pelatihan model.

2.2.4 Evaluasi Model

Kinerja model dievaluasi menggunakan empat metrik kuantitatif, yaitu Koefisien Determinasi (R^2), *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE). Nilai R^2 digunakan untuk mengukur proporsi variasi data yang dapat dijelaskan oleh model, MAE dan RMSE mengukur besar kesalahan prediksi dalam satuan yang sama dengan variabel dependen. MAPE digunakan untuk mengukur tingkat kesalahan dalam bentuk persentase sehingga memudahkan interpretasi tingkat akurasi model. Model terbaik ditentukan berdasarkan nilai R^2 tertinggi serta nilai MAE, RMSE, dan MAPE terendah.

2.2.5 Implementasi Model

Model dengan performa terbaik selanjutnya diimplementasikan dalam bentuk *dashboard* berbasis web menggunakan Streamlit. Sistem ini memungkinkan pengguna memasukkan nilai durasi *live* dan jumlah penonton aktif untuk memperoleh estimasi jumlah produk terjual secara *real-time*. Implementasi ini menunjukkan bahwa model yang dibangun tidak hanya bersifat analitis, tetapi juga aplikatif dalam mendukung pengambilan keputusan pemasaran digital berbasis data.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pelatihan Model Regresi

Proses pelatihan model dilakukan menggunakan algoritma *Stochastic Gradient Descent* (SGD) untuk mengoptimasi parameter regresi linear, polinomial, dan logaritmik. *Dataset* yang telah melalui proses normalisasi digunakan sebagai data pelatihan dengan fungsi kerugian berupa *Mean Squared Error* (MSE). Berdasarkan hasil pelatihan pada model regresi linear, diperoleh parameter koefisien sebagaimana ditunjukkan pada Tabel 2.

Tabel 2 Parameter Regresi Linear dan Hasil Pelatihan Model

Parameter	Nilai
Jumlah Data	217
Rentang Durasi <i>Live</i>	0.2 – 22.9 jam
Rentang Penonton Aktif	0 – 270 orang
Rentang Produk Terjual	0 – 98 produk
Iterasi Konvergensi	4
Koefisien X_1 (Durasi)	0.043428
Koefisien X_2 (Penonton)	0.010585
<i>Intercept</i>	5.890686
R^2	0.4995
MAE	10.0113
RMSE	14.0243

Berdasarkan Tabel 2, proses pelatihan model regresi linear menggunakan algoritma *Stochastic Gradient Descent* (SGD) berhasil mencapai kondisi konvergen pada iterasi ke-4. Hasil tersebut menunjukkan bahwa proses optimasi parameter berlangsung secara stabil dengan nilai toleransi *error* sebesar 0.001. Koefisien variabel durasi *live* (X_1) dan jumlah penonton aktif (X_2) bernilai positif, yang menunjukkan bahwa peningkatan durasi siaran dan jumlah penonton aktif cenderung meningkatkan jumlah produk terjual. Model memperoleh nilai R^2 sebesar 0.4995 dengan nilai MAE sebesar 10.0113 dan RMSE sebesar 14.0243, sehingga model dinilai mampu merepresentasikan hubungan antar variabel meskipun masih terdapat tingkat kesalahan prediksi yang cukup tinggi.

Berdasarkan hasil optimasi parameter yang diperoleh, model regresi linear dapat direpresentasikan ke dalam bentuk persamaan matematis untuk menggambarkan hubungan antara durasi *live*, jumlah penonton aktif, dan jumlah produk terjual. Persamaan regresi yang dihasilkan ditunjukkan pada Persamaan (5).

$$Y = 0.043428.X_1 + 0.010585.X_2 + 5.890686 \quad (5)$$

Persamaan regresi tersebut menunjukkan bahwa setiap peningkatan durasi *live* dan jumlah penonton aktif memberikan kontribusi positif terhadap peningkatan jumlah produk terjual. Nilai *intercept* sebesar 5.890686 menunjukkan estimasi jumlah produk terjual ketika seluruh variabel independen bernilai nol. Untuk melihat kemampuan model dalam melakukan prediksi, dilakukan pengujian menggunakan beberapa sampel data pada *dataset* penelitian. Hasil perbandingan antara nilai aktual dan nilai prediksi model ditunjukkan pada Tabel 3.

Tabel 3 Sampel Hasil Prediksi Model

Durasi (Jam)	Penonton Aktif	Aktual	Prediksi	Error
7.9128	19	9	16.6899	-7.6899
11.7703	20	9	20.2207	-11.2207
12.3764	50	34	27.0376	6.9624
12.0003	21	16	20.6286	-4.6286

Berdasarkan Tabel 3, model mampu menghasilkan nilai prediksi yang relatif mendekati data aktual meskipun masih terdapat beberapa selisih pada beberapa data. Perbedaan antara nilai aktual dan prediksi menunjukkan bahwa pola penjualan pada *live streaming commerce* dipengaruhi oleh faktor lain di luar variabel penelitian. Model masih mampu menangkap pola umum hubungan antara durasi *live*, jumlah penonton aktif, dan jumlah produk terjual.

3.2 Perbandingan Kinerja Model

Evaluasi performa dilakukan untuk membandingkan kemampuan model regresi linear, polinomial, dan logaritmik dalam memprediksi jumlah produk terjual. Pengukuran performa dilakukan menggunakan metrik R^2 , *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), dan *Mean Absolute Percentage Error* (MAPE). Hasil evaluasi masing-masing model ditunjukkan pada Tabel 4.

Tabel 4 Perbandingan Kinerja Model

Model	R^2	MAE	RMSE	MAPE (%)
Linear	0.4995	10.0113	14.0243	66.94
Polynomial	0.4854	10.1924	14.2217	66.91
Logaritmik	0.4556	10.5492	14.6269	67.89

Berdasarkan Tabel 4, model regresi linear menunjukkan performa terbaik dibandingkan model regresi polinomial dan logaritmik. Nilai R^2 sebesar 0.4995 menunjukkan bahwa model mampu menjelaskan sekitar 49.95% variasi data penjualan. Model regresi linear juga menghasilkan nilai MAE dan RMSE yang lebih rendah dibandingkan model lainnya, sehingga menunjukkan tingkat kesalahan prediksi yang relatif lebih kecil. Kondisi ini mengindikasikan bahwa hubungan antara durasi *live*, jumlah penonton aktif, dan jumlah produk terjual pada *dataset* penelitian cenderung mengikuti pola linear sederhana.

Meskipun model regresi linear menunjukkan performa terbaik, nilai MAPE sebesar 66.94% mengindikasikan bahwa tingkat kesalahan prediksi masih tergolong tinggi. Kondisi ini disebabkan oleh karakteristik data *live streaming commerce* yang dinamis dan dipengaruhi oleh berbagai faktor eksternal yang belum dimasukkan ke dalam model penelitian. Faktor-faktor seperti jenis produk, strategi promosi, pemberian *voucher*, kualitas komunikasi *host*, waktu pelaksanaan *live streaming*, serta tingkat interaksi audiens dapat mempengaruhi jumlah produk terjual secara signifikan. Penelitian ini hanya menggunakan variabel durasi *live* dan jumlah penonton aktif sehingga model belum mampu merepresentasikan seluruh faktor yang mempengaruhi penjualan.

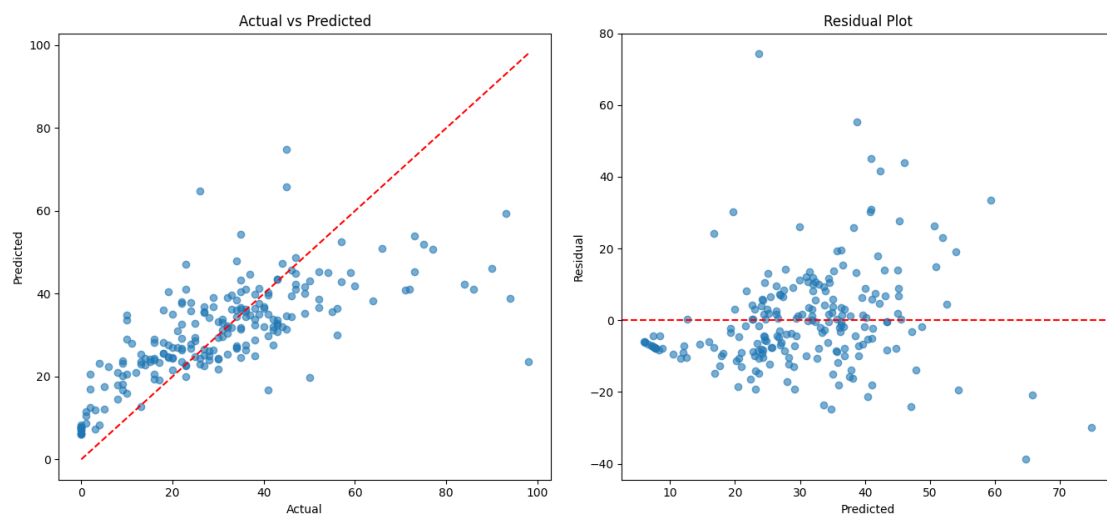
Hasil penelitian ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa model regresi linear efektif digunakan pada prediksi data numerik dengan hubungan variabel yang relatif sederhana. Penggunaan algoritma *Stochastic Gradient Descent* juga membantu proses optimasi parameter menjadi lebih stabil dan efisien pada proses pelatihan model. Hasil penelitian juga menunjukkan bahwa peningkatan jumlah

penonton aktif tidak selalu menghasilkan peningkatan penjualan secara proporsional. Kondisi tersebut menunjukkan bahwa tingkat konversi penjualan pada *live streaming commerce* tidak hanya dipengaruhi oleh jumlah audiens, tetapi juga kualitas strategi pemasaran yang diterapkan selama siaran berlangsung. Implementasi *dashboard* prediksi berpotensi membantu pelaku usaha dalam mengevaluasi efektivitas sesi *live streaming* secara lebih terukur dan berbasis data.

3.3 Analisis Visualisasi Prediksi

3.3.1 Grafik *Actual vs Predicted*

Visualisasi hasil prediksi dilakukan untuk melihat kedekatan antara nilai aktual dan nilai prediksi yang dihasilkan model regresi linear. Grafik *Actual vs Predicted* digunakan untuk menunjukkan kemampuan model dalam mengikuti pola data penjualan pada *dataset* penelitian. Hasil visualisasi prediksi ditunjukkan pada Gambar 2.



Gambar 2 Perbandingan Nilai Aktual dan Nilai Prediksi pada Regresi Linear

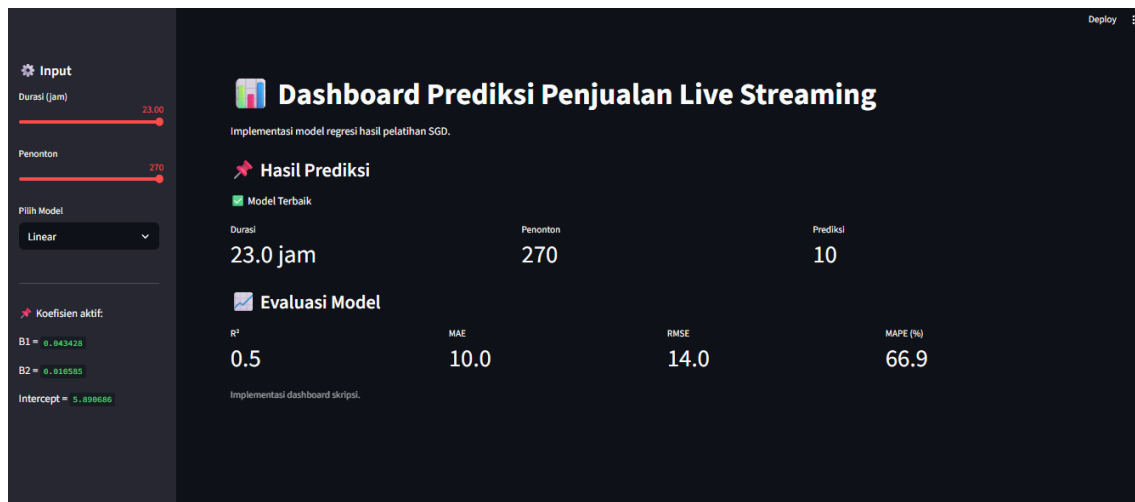
Berdasarkan Gambar 2, sebagian besar titik data berada di sekitar garis diagonal yang menunjukkan bahwa hasil prediksi model relatif mendekati nilai aktual. Kondisi tersebut menunjukkan bahwa model regresi linear mampu menangkap pola umum hubungan antara durasi *live*, jumlah penonton aktif, dan jumlah produk terjual. Meskipun demikian, masih terdapat beberapa titik yang berada cukup jauh dari garis diagonal sehingga menunjukkan adanya deviasi antara nilai aktual dan nilai prediksi. Deviasi tersebut mengindikasikan bahwa terdapat faktor lain di luar variabel penelitian yang turut mempengaruhi tingkat penjualan pada sesi *live streaming commerce*.

Penyebaran residual yang cenderung berada di sekitar garis nol menunjukkan bahwa model tidak mengalami bias prediksi yang ekstrem. Pola residual yang tidak membentuk pola tertentu juga menunjukkan bahwa asumsi linearitas pada model masih dapat diterima. Dengan demikian, visualisasi hasil prediksi mendukung hasil evaluasi

model sebelumnya bahwa regresi linear dengan optimasi *Stochastic Gradient Descent* mampu memberikan performa yang cukup stabil pada *dataset* penelitian.

3.4 Implementasi *Dashboard*

Model regresi linear terbaik selanjutnya diimplementasikan ke dalam bentuk *dashboard* berbasis *Streamlit* untuk mempermudah proses simulasi prediksi penjualan secara *real-time*. Implementasi ini bertujuan untuk membantu pengguna melakukan estimasi jumlah produk terjual berdasarkan input durasi *live* dan jumlah penonton aktif. Tampilan antarmuka aplikasi prediksi ditunjukkan pada Gambar 3.



Gambar 3 *Dashboard* Prediksi Penjualan

Berdasarkan Gambar 3, pengguna dapat memasukkan nilai durasi *live* dan jumlah penonton aktif melalui antarmuka aplikasi untuk memperoleh hasil prediksi jumlah produk terjual secara otomatis. Sistem akan memproses input menggunakan model regresi linear yang telah dioptimasi menggunakan algoritma *Stochastic Gradient Descent*. Implementasi *dashboard* ini menunjukkan bahwa model yang dibangun tidak hanya bersifat teoritis, tetapi juga dapat diterapkan sebagai alat bantu pengambilan keputusan berbasis data pada aktivitas *live streaming commerce*.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa penerapan algoritma *Stochastic Gradient Descent* (SGD) pada model regresi linear efektif digunakan untuk memprediksi jumlah produk terjual berdasarkan durasi *live* dan jumlah penonton aktif. Model yang dihasilkan memperoleh nilai R^2 sebesar 0.4995, MAE sebesar 10.0113, RMSE sebesar 14.0243, dan MAPE sebesar 66.94%, yang menunjukkan bahwa model mampu merepresentasikan hubungan antar variabel penelitian meskipun tingkat kesalahan prediksi masih tergolong cukup tinggi akibat karakteristik data *live streaming commerce* yang dinamis. Dibandingkan dengan regresi polinomial dan logaritmik, regresi linear

menunjukkan performa yang lebih stabil dan konsisten sehingga dinyatakan sebagai model terbaik dalam penelitian ini. Implementasi model ke dalam *dashboard* berbasis *Streamlit* juga menunjukkan bahwa sistem prediksi yang dibangun tidak hanya bersifat konseptual, tetapi juga aplikatif dan berpotensi membantu pelaku usaha dalam mendukung pengambilan keputusan berbasis data pada strategi *live streaming commerce*.

5. ACKNOWLEDGEMENTS

Penulis menyampaikan penghargaan dan terima kasih yang sebesar-besarnya kepada Universitas Samudra atas dukungan pendanaan melalui Program Hibah Skripsi yang telah mendukung pelaksanaan penelitian ini. Ucapan terima kasih juga disampaikan kepada para dosen pembimbing yang telah memberikan bimbingan, arahan ilmiah, serta masukan yang konstruktif selama proses penelitian dan penyusunan artikel ini. Penulis juga mengapresiasi pihak KawanNgampus yang telah memberikan akses terhadap data *live streaming* sebagai sumber utama penelitian, sehingga penelitian ini dapat diselesaikan dengan baik.

REFERENSI

- [1] I. Zuhdi, L. M. Furkan, And H. Hilmianti, 'Vol 2 No.3 Sept Pengaruh *Live streaming* Shopping Dan Motivasi Belanja Hedonis Terhadap Impulsif Buying Pada Pengguna Aplikasi Tiktok Shop Di Kota Mataram: Pengaruh *Live streaming* Shopping Dan Motivasi Belanja Hedonis Terhadap Impulsif Buying Pada Pengguna Aplikasi Tiktok Shop Di Kota Mataram', *J. Ris. Pemasar.*, Vol. 2, No. 3, Pp. 66–72, Sep. 2023, Doi: 10.29303/Jrp.V2i3.3286.
- [2] M. B. Rosa, 'Pengaruh *Live streaming* Shopping Tiktok Terhadap Minat Pembelian Pada Konsumen Akun Produk Atap Upvc @Dr.Shield', Vol. 9, 2025.
- [3] M. P. Kinasih And N. E. R. Wuryandari, 'Pengaruh Price Discount, Product Knowledge, Dan Time Pressure Terhadap Impulse Buying Pada Fitur *ShopeeLive*', *J. Ilm. Multidisiplin*, Vol. 2, No. 05, Pp. 88–101, Sep. 2023, Doi: 10.56127/Jukim.V2i05.940.
- [4] N. Nurliana, N. Suarna, And W. Prihartono, 'Analisis Keputusan Pembelian Produk Terhadap Promosi Dan Viral Marketing Di Tiktok Shop Menggunakan Algoritma C4.5: Studi Kasus (Kuesioner Mahasiswa Stmik Ikmi Cirebon)', *Jati J. Mhs. Tek. Inform.*, Vol. 8, No. 3, Pp. 2870–2876, May 2024, Doi: 10.36040/Jati.V8i3.9583.
- [5] T. Yu, A. P. Teoh, Q. Bian, J. Liao, And C. Wang, 'Can Virtual Influencers Affect Purchase Intentions In Tourism And Hospitality E-Commerce *Live streaming*? An Empirical Study In China', *Int. J. Contemp. Hosp. Manag.*, Vol. 37, No. 1, Pp. 216–238, Jan. 2025, Doi: 10.1108/Ijchm-03-2024-0358.
- [6] K. M. Arrazaq, D. R. S. Saputro, And R. Setiyowati, 'Penerapan *Stochastic Gradient Descent* Support Vector Regression Pada Data Laju Pertumbuhan Produk Domestik Bruto Di Indonesia'.
- [7] N. C. Florensa Nainggolan, A. F. Boy, And E. Elfitriani, 'Penerapan Data Mining Untuk Prediksi Export Penjualan Produk Kerajinan Rotan Menggunakan Metode Regresi Linear Berganda', *J. Sist. Inf. Triguna Dharma Jursi Tgd*, Vol. 2, No. 5, P. 743, Sep. 2023, Doi: 10.53513/Jursi.V2i5.6779.

- [8] R. Hermawan, N. Suarna, I. Ali, And D. Rohman, ‘Optimasi Prediksi Omset Penjualan Pada Pabrik Olahan Tahu Menggunakan Algoritma Regresi Linear’, *J. Inform. Dan Tek. Elektro Terap.*, Vol. 13, No. 1, Jan. 2025, Doi: 10.23960/Jitet.V13i1.5888.
- [9] D. Susilo, Ahmad Chusyairi, And M. I. Saputra, ‘Prediksi Penjualan Untuk Optimasi Stock Produk Menggunakan Algoritma Long Short Term Memory’, *Sist. J. Ilm. Sist. Inf.*, Vol. 1, No. 2, Pp. 119–134, Apr. 2025, Doi: 10.69533/56gyat30.
- [10] Q. Li And M. Yu, ‘Achieving Sales Forecasting With Higher Accuracy And Efficiency: A New Model Based On Modified Transformer’, *J. Theor. Appl. Electron. Commer. Res.*, Vol. 18, No. 4, Pp. 1990–2006, Nov. 2023, Doi: 10.3390/Jtaer18040100.
- [11] A. F. Ningrum, ‘Evaluation Of Adam, Nadam, Rmsprop And Sgd Optimization Algorithms On Lstm Model For Jisdor Exchange Rate Prediction’, Vol. 1, No. 1, 2024.
- [12] Y. Tian, Y. Zhang, And H. Zhang, ‘Recent Advances In *Stochastic Gradient Descent* In Deep Learning’, *Mathematics*, Vol. 11, No. 3, P. 682, Jan. 2023, Doi: 10.3390/Math11030682.
- [13] F. T. Admojo And Y. I. Sulistya, ‘Analisis Performa Algoritma *Stochastic Gradient Descent* (Sgd) Dalam Mengklasifikasi Tahu Berformalin’, *Indones. J. Data Sci.*, Vol. 3, No. 1, Pp. 1–8, Mar. 2022, Doi: 10.56705/Ijodas.V3i1.42.
- [14] Sutarman, M. A. I. Hutagalung, O. Darnius, And M. Y. P. El Sya’ban, ‘*Stochastic Gradient Descents* Optimizer And Its Variants: Performance Of The Optimizers For Multinomial Logistic Models On Large Data Sets By Simulation’, *Math. Model. Eng. Probl.*, Vol. 11, No. 10, Pp. 2823–2832, Oct. 2024, Doi: 10.18280/Mmep.11i1025.
- [15] A. T. N. Hartono And H. D. Purnomo, ‘Pengembangan *Stochastic Gradient Descent* Dengan Penambahan Variabel Tetap’, *J. Jtik J. Teknol. Inf. Dan Komun.*, Vol. 7, No. 3, Pp. 359–367, Jul. 2023, Doi: 10.35870/Jtik.V7i3.840.
- [16] A. Irianti, N. Nur, and A. Heri, ‘Prediksi Hasil Panen Kakao Di Desa Minanga Menggunakan Algoritma Random Forest Regression’.