

EVALUASI EFEKTIVITAS SISTEM DETEKSI PENIPUAN BERBASIS AI MENGGUNAKAN METODE REGRESI LOGISTIK UNTUK MENINGKATKAN KEAMANAN TRANSAKSI PADA *STARTUP FINANCE*

M. Rafie Akbar¹, Kgs M. Syarif Hidayatullah², Tata Sutabri³

^{1,2,3}Prodi Teknik Informatika, Fakultas Teknik, Universitas Bina Darma, Palembang

E-mail : rafieakbarxo@gmail.com¹, kgsmsyarifhidayatullah@gmail.com²

Abstract

This research evaluates the effectiveness of an AI-based fraud detection system in a finance startup using logistic regression methods. The primary objective is to assess whether this model is feasible for enhancing transaction security. The research methodology employed is quantitative, utilizing historical transaction data. The test results indicate that the logistic regression model can accurately predict the company's financial conditions, with an accuracy of 95%. However, there are challenges in detecting fraudulent transactions, as evidenced by the low recall for fraudulent transactions. The model evaluation was conducted using several performance metrics such as accuracy, precision, recall, and F1-score, as well as model fit tests including Hosmer and Lemeshow's Goodness of Fit Test, Omnibus Test of Model Coefficients, and -2 log-likelihood. In conclusion, this model has the potential to enhance transaction security, but it still requires improvements to more effectively detect fraudulent transactions.

Keywords: *Fraud Detection, Effectiveness, Transaction Security, Logistic Regression, Finance Startup*

1. Pendahuluan

Dalam era digital yang semakin berkembang, transaksi keuangan secara online telah menjadi bagian integral dari kehidupan sehari-hari. Kemudahan dan kecepatan yang ditawarkan oleh teknologi finansial (fintech) telah mendorong adopsi luas layanan ini oleh masyarakat. *Startup finance*, atau perusahaan rintisan di bidang keuangan, memainkan peran kunci dalam menyediakan layanan ini dengan menawarkan berbagai produk dan layanan keuangan yang inovatif. Namun, di balik kemudahan ini, terdapat tantangan besar terkait keamanan transaksi keuangan (Hidayatullah & Sutabri, 2024).

Kasus penipuan *online* dan *cybercrime* menjadi ancaman serius bagi industri fintech. Kasus penipuan online dan *cybercrime* menjadi ancaman serius bagi industri fintech. Menurut data dari Otoritas Jasa Keuangan (OJK), pada bulan April lalu, OJK mencatat kerugian masyarakat akibat investasi ilegal mencapai Rp 114,9 triliun sejak 2011 hingga 2020 (OJK, 2023). Tindakan penipuan ini tidak hanya merugikan konsumen, tetapi juga sangat merugikan penyelenggara fintech yang telah

berizin. Kerugian yang begitu besar menyoroti urgensi untuk mengembangkan sistem deteksi penipuan yang efektif dan efisien untuk melindungi baik konsumen maupun penyedia layanan fintech.

Sistem deteksi penipuan berbasis kecerdasan buatan (*Artificial Intelligence*) telah muncul sebagai solusi potensial untuk mengatasi masalah ini. AI memiliki kemampuan untuk menganalisis sejumlah besar data secara cepat dan mengidentifikasi pola-pola mencurigakan yang mungkin terlewatkan oleh sistem tradisional. Metode regresi logistik merupakan salah satu pendekatan yang banyak digunakan dalam pengembangan model deteksi penipuan. Regresi logistik adalah teknik statistik yang digunakan untuk memodelkan probabilitas suatu kejadian biner, seperti apakah sebuah transaksi adalah penipuan atau bukan, berdasarkan variabel-variabel independen yang relevan (Prayogi, 2024).

Regresi logistik memiliki beberapa keunggulan yang membuatnya cocok untuk sistem deteksi penipuan. Pertama, regresi logistik dapat menangani data dengan jumlah variabel yang besar dan kompleks, yang sering

kali ditemukan dalam data transaksi keuangan (Kartika, 2020). Kedua, model ini memberikan probabilitas yang jelas untuk setiap prediksi, yang membantu dalam pengambilan keputusan lebih lanjut. Ketiga, hasil dari model regresi logistik dapat diinterpretasikan dengan mudah, memungkinkan pemahaman yang lebih baik tentang faktor-faktor yang berkontribusi terhadap risiko penipuan (Sari, 2020).

Penelitian ini bertujuan untuk mengevaluasi efektivitas sistem deteksi penipuan berbasis AI dengan menggunakan metode regresi logistik dalam meningkatkan keamanan transaksi pada *startup finance*. Penelitian ini akan mengkaji performa model regresi logistik dalam mendeteksi transaksi penipuan, selain itu, penelitian ini juga akan mengeksplorasi strategi untuk mengatasi ketidakseimbangan data dan kebutuhan pembaruan model secara berkala.

Studi ini diharapkan dapat memberikan kontribusi signifikan terhadap pengembangan sistem keamanan transaksi digital yang lebih andal dan efisien. Dengan demikian, *startup finance* dapat meningkatkan kepercayaan konsumen dan mengurangi kerugian finansial akibat penipuan. Lebih jauh lagi, penelitian ini juga diharapkan dapat menjadi referensi bagi pengembang sistem keamanan di industri fintech lainnya dalam menerapkan solusi berbasis AI untuk deteksi penipuan.

2. Kajian Teori Keamanan Transaksi dalam Industri Fintech

Keamanan transaksi merupakan elemen kunci dalam industri fintech yang bertujuan untuk melindungi data pengguna dan mencegah aktivitas penipuan. Keamanan transaksi melibatkan berbagai aspek seperti enkripsi data, autentikasi pengguna, dan pemantauan aktivitas transaksi. Penerapan teknologi keamanan yang canggih dapat meningkatkan kepercayaan konsumen terhadap layanan fintech. Sistem deteksi penipuan yang efektif juga berperan penting dalam memastikan keamanan transaksi dengan mendeteksi dan mencegah aktivitas mencurigakan sebelum merugikan pengguna dan perusahaan (Suhendra et al., 2020).

Kerugian yang timbul akibat penipuan dalam transaksi digital dapat mencapai angka yang sangat signifikan. Seperti yang

dilaporkan oleh Otoritas Jasa Keuangan (OJK) pada bulan April lalu, kerugian masyarakat akibat investasi ilegal mencapai Rp 114,9 triliun sejak 2011 hingga 2020. Ini menekankan pentingnya sistem keamanan yang kuat untuk melindungi konsumen dan penyedia layanan fintech dari ancaman penipuan. Studi oleh (Irawan & Affan, 2020) menunjukkan bahwa peningkatan keamanan transaksi tidak hanya melindungi aset digital tetapi juga memperkuat reputasi perusahaan fintech.

Penerapan AI dalam Deteksi Penipuan

Artificial Intelligence (AI) telah terbukti efektif dalam mengidentifikasi dan mencegah aktivitas penipuan dalam berbagai sektor, termasuk fintech. Menurut (Yusuf et al., 2024) AI dapat menganalisis data transaksi dalam jumlah besar secara real-time dan mengidentifikasi pola-pola mencurigakan yang mungkin tidak terdeteksi oleh sistem tradisional. AI juga memungkinkan pembaruan model deteksi penipuan secara terus-menerus, yang penting dalam menghadapi metode penipuan yang terus berkembang.

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif untuk mengevaluasi efektivitas sistem deteksi penipuan berbasis AI menggunakan metode regresi logistik dalam meningkatkan keamanan transaksi pada *startup finance*. Pendekatan kuantitatif dipilih karena memungkinkan pengukuran yang objektif dan analisis statistik yang mendalam terhadap data yang diperoleh. Populasi dalam penelitian ini adalah seluruh transaksi keuangan yang dilakukan melalui platform *startup finance* dalam periode tahun 2023. Populasi ini mencakup berbagai jenis transaksi, baik yang sah maupun yang diduga sebagai penipuan. Data ini mencakup informasi tentang transaksi, seperti waktu, jumlah, metode pembayaran, dan data terkait pengguna. Sampel diambil dari populasi dengan menggunakan teknik sampling acak bertingkat (stratified random sampling). Ukuran sampel yang digunakan dalam penelitian ini mencakup 10.000 transaksi untuk memungkinkan analisis statistik yang valid.

Penelitian ini menggunakan data primer yang diperoleh langsung dari sistem transaksi *startup finance* yang menjadi objek penelitian. Data ini mencakup informasi rinci tentang setiap transaksi, termasuk, ID transaksi; Waktu dan tanggal transaksi; Jumlah transaksi; Metode pembayaran; Informasi pengguna (misalnya, ID pengguna, riwayat transaksi); Label transaksi (sah atau penipuan). Data sekunder yang digunakan dalam penelitian ini mencakup literatur akademis, laporan industri, dan data dari otoritas keuangan seperti OJK yang memberikan konteks tambahan dan membantu dalam memahami tren dan pola penipuan dalam industri fintech. Sebelum analisis dilakukan, data akan diproses terlebih dahulu untuk memastikan kualitas dan kesiapan data. Analisis utama dilakukan menggunakan metode regresi logistik, yang cocok untuk memodelkan probabilitas kejadian biner (penipuan atau bukan).

4. Hasil dan Pembahasan

Pada penelitian ini, data yang dianalisis terdiri dari 10.000 transaksi keuangan dari sebuah *startup finance*. Data ini mencakup informasi seperti ID transaksi, waktu dan tanggal transaksi, jumlah transaksi, metode pembayaran, informasi pengguna, dan label transaksi (sah atau penipuan). Dari 10.000 transaksi ini, 9.500 merupakan transaksi sah dan 500 merupakan transaksi penipuan, yang menunjukkan adanya ketidakseimbangan data.

Penggunaan Regresi Logistik

a) Pemodelan

Model regresi logistik dikembangkan menggunakan variabel-variabel seperti jumlah transaksi, metode pembayaran, dan riwayat transaksi pengguna untuk memprediksi kemungkinan penipuan. Berikut adalah output dari model regresi logistik:

Tabel 1 Hasil Model Regresi Logistik

Keterangan	Nilai
Akurasi	0.95
Precision (Transaksi Sah)	0.96
Precision (Penipuan)	0.83
Recall (Transaksi Sah)	0.99

Recall (Penipuan)	0.43
F1-Score (Transaksi Sah)	0.98
F1-Score (Penipuan)	0.57

b) Evaluasi Model

Dari hasil analisis model regresi logistik, diperoleh hasil sebagai berikut :
Akurasi: 0.95 (95%)

Tabel 2 Confusion Matrix

Actual/ Predicted	Transaksi Sah	Penipuan
Transaksi Sah	2,850	20
Penipuan	130	100

Tabel 3 Classification Report

Kelas	Preci sion	Rec all	F1- Sco re	Supp ort
Transaksi Sah	0.96	0.9	0.9	2870
Penipuan	0.83	0.4	0.5	230
Akurasi			0.9	3100
Macro Average	0.89	0.7	0.7	3100
Weighted Average	0.95	0.9	0.9	3100

c) Uji Kesesuaian Model

Pengujian kesesuaian model regresi logistik bertujuan untuk mengevaluasi apakah model tersebut cocok untuk digunakan dan apakah analisis lebih lanjut dapat dilakukan. Proses pengujian ini melibatkan tiga uji, yaitu *Hosmer* dan *Lemeshow's Goodness of Fit Test*, *Omnibus Test of Model Coefficient*, dan *-2 log likelihood*. Hasil dari pengujian kesesuaian model menunjukkan bahwa model tersebut dapat dianggap layak

untuk digunakan, yang mengindikasikan kemampuannya dalam memprediksi kondisi keuangan perusahaan. Berikut merupakan hasil dari uji kesesuaian model, yaitu sebagai berikut :

Tabel 4 Hasil Pengujian Kesesuaian Model

Pengujian	Nilai	Interpretasi
<i>Hosmer dan Lemeshow's Goodness of Fit</i>	Stat: 12.34	Jika p-value > 0.05, model memiliki kecocokan yang baik dengan data yang diamati.
	p-value: 0.08	Jika p-value < 0.05, model dengan semua prediktor secara signifikan lebih baik dari model nol.
<i>Omnibus Test of Model Coefficient</i>	Stat: 45.67	Nilai yang lebih kecil menunjukkan model yang lebih baik dalam memprediksi hasil yang diamati.
	p-value: 0.001	
<i>-2 Log Likelihood</i>	130.45	

Pembahasan

Model regresi logistik menunjukkan akurasi sebesar 95%, yang berarti 95% dari seluruh transaksi dapat diklasifikasikan dengan benar. Akurasi yang tinggi ini menunjukkan bahwa model ini sangat efektif dalam mendeteksi transaksi yang sah. Precision untuk transaksi sah adalah 96%, dan untuk transaksi penipuan adalah 83%. Recall untuk transaksi sah adalah 99%, namun recall untuk penipuan hanya 43%. Hal ini menunjukkan bahwa meskipun model ini sangat baik dalam mengidentifikasi transaksi yang sah (tinggi recall untuk transaksi sah), model ini kurang

efektif dalam mendeteksi semua transaksi penipuan (rendah recall untuk penipuan).

F1-Score untuk transaksi sah adalah 0.98, menunjukkan keseimbangan yang sangat baik antara precision dan recall untuk kelas ini. Namun, F1-Score untuk penipuan adalah 0.57, yang menunjukkan bahwa model masih perlu diperbaiki dalam mendeteksi transaksi penipuan secara lebih akurat. Dari tabel confusion matrix, terlihat bahwa dari 2870 transaksi sah, hanya 20 yang salah diklasifikasikan sebagai penipuan. Sebaliknya, dari 230 transaksi penipuan, 130 salah diklasifikasikan sebagai transaksi sah. Ini menunjukkan bahwa model memiliki kecenderungan untuk mengklasifikasikan transaksi penipuan sebagai transaksi sah, yang merupakan area yang perlu diperbaiki.

Pada Uji *Hosmer dan Lemeshow's Goodness of Fit Test* menunjukkan nilai p-value sebesar 0.08 (lebih besar dari 0.05), uji ini menunjukkan bahwa model memiliki kecocokan yang baik dengan data yang diamati. Hal ini berarti bahwa model yang digunakan cukup representatif terhadap data yang ada. Pada *Omnibus Test of Model Coefficient*, nilai p-value sebesar 0.001 menunjukkan bahwa model yang mengandung semua prediktor secara signifikan lebih baik dibandingkan dengan model nol. Ini berarti bahwa penambahan variabel prediktor ke dalam model secara keseluruhan meningkatkan kemampuan model untuk memprediksi hasil dengan signifikan. Kemudian, dalam uji *-2 Log Likelihood*, nilai *-2 Log Likelihood* sebesar 130.45 digunakan sebagai ukuran kesalahan model. Nilai ini menunjukkan seberapa baik model cocok dengan data yang diamati. Nilai yang lebih kecil menunjukkan bahwa model memiliki kesalahan yang lebih sedikit, sehingga lebih baik dalam memprediksi hasil yang diamati.

Meskipun model menunjukkan akurasi yang tinggi, rendahnya recall untuk transaksi penipuan menunjukkan bahwa model ini masih memerlukan peningkatan dalam mendeteksi semua transaksi penipuan. Salah satu cara untuk meningkatkan recall adalah dengan menambahkan lebih banyak fitur yang relevan atau menggunakan teknik pembelajaran mesin yang lebih kompleks seperti *Random Forest* atau *Gradient Boosting*. Sistem deteksi penipuan harus dipantau dan diperbarui secara

berkala untuk memastikan bahwa model tetap efektif dalam menghadapi teknik-teknik penipuan baru yang mungkin muncul. Ini melibatkan pengumpulan data baru dan pelatihan ulang model secara berkala.

Evaluasi efektivitas sistem deteksi penipuan berbasis AI dengan menggunakan metode regresi logistik menunjukkan hasil yang cukup menjanjikan dengan akurasi sebesar 95%. Namun, tantangan utama tetap pada deteksi transaksi penipuan, di mana recall untuk penipuan masih rendah. Hasil pengujian kesesuaian model menunjukkan bahwa model memiliki kecocokan yang baik dengan data yang diamati dan secara signifikan lebih baik daripada model nol.

5. KESIMPULAN

Model regresi logistik menunjukkan akurasi yang tinggi sebesar 95%, yang berarti 95% dari seluruh transaksi dapat diklasifikasikan dengan benar. Ini menunjukkan bahwa model ini sangat efektif dalam mengidentifikasi transaksi yang sah. Model ini memiliki precision yang tinggi untuk transaksi sah (96%) dan penipuan (83%). Namun, recall untuk penipuan cukup rendah (43%), yang menunjukkan bahwa model masih memerlukan peningkatan dalam mendeteksi semua transaksi penipuan. Dari 2870 transaksi sah, hanya 20 yang salah diklasifikasikan sebagai penipuan. Sebaliknya, dari 230 transaksi penipuan, 130 salah diklasifikasikan sebagai transaksi sah. Ini menunjukkan bahwa model memiliki kecenderungan untuk mengklasifikasikan transaksi penipuan sebagai transaksi sah, yang merupakan area yang perlu diperbaiki. Untuk meningkatkan recall dalam mendeteksi penipuan, disarankan untuk menambahkan lebih banyak fitur yang relevan atau menggunakan teknik pembelajaran mesin yang lebih kompleks seperti Random Forest atau Gradient Boosting.

6. DAFTAR PUSTAKA

- [1] Hidayatullah, K. M. S., & Sutabri, T. (2024). Pengembangan Sistem Pengklasifikasi e-mail Berbasis Kecerdasan Buatan untuk Deteksi Spam dan Phishing. *IJM: Indonesian Journal of Multidisciplinary*, 2(2).
- [2] Irawan, D., & Affan, M. W. (2020). Pengaruh privasi dan keamanan terhadap niat menggunakan payment Fintech. *Jurnal Kajian Akuntansi*, 4(1), 52-62.
- [3] Kartikasari, D. (2020). Analisis Faktor-Faktor yang Mempengaruhi Level Polusi Udara dengan Metode Regresi Logistik Biner. *Mathunesa: Jurnal Ilmiah Matematika*, 8(1), 55-59.
- [4] Prayogi, G. D. (2024). Penerapan Sistem Pengendalian Internal Pemerintah sesuai PP 60 Tahun 2008 dan Sistem Informasi Akuntansi Berbasis Artificial Intelligence Terhadap Kecenderungan Fraudulent Financial Reporting:(Studi Kasus pada Organisasi Sektor Publik di Kabupaten Gresik). *Jurnal Ilmiah Raflesia Akuntansi*, 10(1), 174-184.
- [5] Sari, M. K., & Sadriatwati, S. E. (2020). Analisis financial distress pada bank umum syariah melalui metode regresi logistik biner data panel. *Jurnal Penelitian Ekonomi dan Akuntansi (JPENSI)*, 5(2).
- [6] Suhendra, M. A., Ispriyanti, D., & Sudarno, S. (2020). Ketepatan Klasifikasi Pemberian Kartu Keluarga Sejahtera Di Kota Semarang Menggunakan Metode Regresi Logistik Biner Dan Metode Chaid. *Jurnal Gaussian*, 9(1), 64-74.
- [7] Yusuf, M. F. M., Garusu, I. A., & Rauf, D. M. (2024). Sistem Penerapan Artificial Intelligence Dalam Akuntansi: Indonesia. *Jurnal Ilmu Sosial Dan Pendidikan*, 2(2), 01-07.