Jurnal Minfo Polgan
Volume 13, Nomor 2, Oktober 2024 e-ISSN : 2797-3298
DOI : https://doi.org/10.33395/jmp.v13i2.14186 p-ISSN : 2089-9424

Terbit : 21 Oktober 2024

Penggunaan Algoritma Support Vector Machine (SVM) Untuk Deteksi Penipuan pada Transaksi Online

[1] Handry Eldo, [2] Ayuliana, [3] Dikky Suryadi, [4] Giatika Chrisnawati, [5] Loso Judijanto
[1] Universitas Muhammadiyah Mahakarya Aceh, [2] Universitas Bina Nusantara, [3] STMIK ALMuslim, [4] Universitas Bina Sarana Informatika, [5] IPOSS Jakarta Indonesia

[1] handry.eldo@gmail.com, [2] ayuliana_st@binus.ac.id, [3] dikky98@gmail.com,
[4] giatika.c@gmail.com, [5] losojudijantobumn@gmail.com

ABSTRAK

Penipuan dalam transaksi online semakin meningkat seiring dengan perkembangan teknologi digital yang pesat. Untuk mengatasi masalah ini, diperlukan sistem deteksi yang efektif dan akurat guna meminimalisir kerugian yang disebabkan oleh aktivitas penipuan. Penelitian ini bertujuan untuk mengembangkan model deteksi penipuan menggunakan algoritma Support Vector Machine (SVM). Algoritma SVM dipilih karena kemampuannya dalam mengklasifikasikan data yang kompleks dan menangani data dengan dimensi tinggi. Metodologi penelitian melibatkan pengumpulan data transaksi online yang terdiri dari transaksi sah dan penipuan, kemudian dilakukan preprocessing data untuk mengatasi ketidakseimbangan dan noise pada dataset. Model SVM dilatih menggunakan data yang telah diproses dan dievaluasi berdasarkan metrik akurasi, presisi, recall, dan F1-score. Hasil penelitian menunjukkan bahwa model SVM mampu mendeteksi transaksi penipuan dengan tingkat akurasi yang tinggi, mencapai 95%. Selain itu, model ini juga memiliki tingkat presisi dan recall yang seimbang, sehingga efektif dalam mendeteksi aktivitas penipuan tanpa mengabaikan transaksi sah. Kesimpulannya, algoritma SVM dapat dijadikan sebagai solusi yang andal untuk mengidentifikasi penipuan pada transaksi online, namun perlu pengujian lebih lanjut pada berbagai jenis dataset untuk meningkatkan generalisasi model.

Kata Kunci: Support Vector Machine (SVM), Deteksi Penipuan, Transaksi online.

PENDAHULUAN

Perkembangan teknologi digital telah mengubah berbagai aspek kehidupan, termasuk dalam sektor perdagangan dan transaksi keuangan. Transaksi online kini menjadi pilihan utama bagi banyak orang karena menawarkan kemudahan, kecepatan, dan fleksibilitas. Namun, di balik keuntungan tersebut, terdapat risiko yang cukup besar terkait dengan keamanan data dan kepercayaan konsumen, terutama terkait penipuan dalam transaksi online. Penipuan ini dapat berbentuk pencurian data, transaksi palsu, hingga penyalahgunaan identitas yang mengakibatkan kerugian finansial bagi perusahaan maupun konsumen. Berdasarkan data yang dilaporkan oleh berbagai perusahaan keamanan siber, kasus penipuan online terus meningkat setiap tahunnya, dan hal ini menuntut adanya solusi yang lebih efektif untuk mencegah aktivitas penipuan tersebut.

Deteksi penipuan pada transaksi online menjadi tantangan besar karena pola penipuan cenderung berubah-ubah dan sulit diprediksi. Metode konvensional yang mengandalkan aturan statis dan sistem berbasis daftar hitam (blacklist) sering kali tidak efektif karena hanya dapat mengidentifikasi penipuan yang sesuai dengan pola-pola yang sudah diketahui sebelumnya. Oleh karena itu, dibutuhkan pendekatan yang lebih dinamis dan adaptif dalam mendeteksi pola-pola penipuan yang baru dan belum teridentifikasi. Salah satu solusi yang dapat digunakan adalah penerapan teknik machine learning, khususnya algoritma Support Vector Machine (SVM).

Algoritma Support Vector Machine (SVM) adalah salah satu metode machine learning yang dikenal efektif dalam melakukan klasifikasi data, terutama pada kasus-kasus yang melibatkan

Jurnal Minfo Polgan
Volume 13, Nomor 2, Oktober 2024 e-ISSN : 2797-3298
DOI : https://doi.org/10.33395/jmp.v13i2.14186 p-ISSN : 2089-9424

data berdimensi tinggi dan non-linear. SVM bekerja dengan cara mencari hyperplane yang optimal untuk memisahkan data ke dalam dua kelas berbeda, sehingga dapat digunakan untuk mengidentifikasi transaksi mana yang dianggap sah dan mana yang terindikasi sebagai penipuan. Keunggulan SVM terletak pada kemampuannya untuk menemukan pola data yang kompleks dan tidak mudah diidentifikasi oleh metode lain. Selain itu, SVM juga dapat dioptimalkan menggunakan teknik kernel, sehingga mampu menangani data dengan berbagai karakteristik.

Penelitian ini bertujuan untuk mengembangkan model deteksi penipuan yang akurat dan efisien dengan memanfaatkan algoritma SVM. Dalam penelitian ini, data transaksi online yang terdiri dari transaksi sah dan penipuan akan digunakan sebagai dataset. Proses penelitian meliputi tahapan preprocessing data, pelatihan model menggunakan SVM, dan evaluasi performa model berdasarkan metrik akurasi, presisi, recall, dan F1-score. Diharapkan, model yang dikembangkan mampu mendeteksi penipuan dengan tingkat akurasi yang tinggi serta dapat diimplementasikan secara efektif dalam sistem transaksi online untuk meningkatkan keamanan dan kepercayaan pengguna. (Hasibuan et al., 2024)

Dengan adanya penelitian ini, diharapkan dapat memberikan kontribusi dalam pengembangan solusi deteksi penipuan yang lebih canggih dan adaptif, sehingga mampu meminimalisir kerugian akibat penipuan online serta melindungi kepentingan konsumen dan perusahaan. Implementasi algoritma SVM dalam deteksi penipuan diharapkan tidak hanya meningkatkan efisiensi proses identifikasi penipuan tetapi juga dapat memperkuat sistem keamanan siber secara keseluruhan. (Iqbal Ahmadi et al., 2020)

TINJAUAN PUSTAKA

Penipuan Pada Transaksi Online

Penipuan pada transaksi online merupakan salah satu masalah yang sering dihadapi dalam dunia digital saat ini. Penipuan ini dapat terjadi dalam berbagai bentuk, seperti pencurian informasi kartu kredit, transaksi palsu, dan penggunaan data identitas secara ilegal. Peningkatan penggunaan transaksi online menyebabkan semakin kompleksnya pola-pola penipuan, sehingga mempersulit deteksi dan pencegahan menggunakan metode konvensional. Menurut laporan dari berbagai lembaga keamanan siber, kasus penipuan online terus meningkat setiap tahun, yang menunjukkan pentingnya pengembangan sistem deteksi penipuan yang lebih canggih dan adaptif. Dalam hal ini, pendekatan berbasis machine learning menjadi solusi yang diharapkan mampu mengidentifikasi pola-pola penipuan baru yang tidak terdeteksi oleh sistem tradisional. (Nugraha et al., 2023)

Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah salah satu algoritma machine learning yang digunakan untuk klasifikasi dan regresi. SVM bekerja dengan mencari hyperplane terbaik yang memisahkan data ke dalam dua kelas yang berbeda. Algoritma ini terkenal karena kemampuannya untuk menangani data berdimensi tinggi dan efektif dalam memisahkan data yang tidak linear melalui penggunaan kernel trick. Kernel trick memungkinkan SVM untuk memetakan data ke dalam ruang dimensi yang lebih tinggi, sehingga membuat data yang tidak dapat dipisahkan secara linear menjadi lebih mudah untuk diklasifikasikan. SVM juga mampu mengatasi masalah overfitting, karena meminimalkan kesalahan klasifikasi sambil memaksimalkan margin antara kelas-kelas data. Ini berarti, SVM tidak hanya mencoba untuk mengklasifikasikan data yang ada dengan benar, tetapi juga memastikan bahwa model yang dihasilkan tidak terlalu cocok dengan data pelatihan, sehingga meningkatkan kemampuan generalisasi terhadap data baru. Oleh karena itu, SVM sering digunakan dalam berbagai aplikasi klasifikasi, termasuk dalam deteksi penipuan.

Deteksi Penipuan dengan Machine Learning

Metode deteksi penipuan tradisional umumnya berbasis aturan (rule-based) atau menggunakan daftar hitam (blacklist). Namun, pendekatan ini memiliki kelemahan, terutama dalam hal ketidakmampuannya untuk mendeteksi pola penipuan baru yang belum pernah muncul sebelumnya. Sebaliknya, machine learning memungkinkan sistem untuk "belajar" dari data historis dan mengidentifikasi pola-pola baru yang mungkin tidak terlihat jelas oleh pendekatan berbasis

Jurnal Minfo Polgan
Volume 13, Nomor 2, Oktober 2024 e-ISSN : 2797-3298
DOI : https://doi.org/10.33395/jmp.v13i2.14186 p-ISSN : 2089-9424

aturan. Dalam beberapa tahun terakhir, algoritma seperti Decision Tree, Random Forest, dan SVM telah banyak digunakan untuk mendeteksi penipuan, karena kemampuan mereka dalam memproses data dalam jumlah besar dan mengidentifikasi pola-pola kompleks. Penggunaan SVM dalam deteksi penipuan menawarkan beberapa keunggulan, seperti akurasi yang tinggi dalam mengklasifikasikan data yang kompleks dan kemampuannya dalam menangani dataset yang tidak seimbang. Penelitian sebelumnya menunjukkan bahwa SVM dapat memberikan hasil yang memuaskan ketika diterapkan pada berbagai dataset penipuan, terutama ketika dikombinasikan dengan teknik pra-pemrosesan data seperti oversampling dan undersampling untuk mengatasi ketidakseimbangan kelas. (Khatib Sulaiman & Bhayangkara Jakarta Raya, 2024)

Penerapan SVM dalam Deteksi Penipuan Online

Penerapan algoritma SVM dalam deteksi penipuan online telah menunjukkan hasil yang menjanjikan. Beberapa studi menunjukkan bahwa SVM dapat mengklasifikasikan transaksi sebagai sah atau penipuan dengan tingkat akurasi yang tinggi. Misalnya, penelitian oleh (Zuhairah, 2023)menunjukkan bahwa model SVM yang dilatih dengan data transaksi online mampu mendeteksi transaksi penipuan dengan akurasi hingga 94%, mengungguli beberapa algoritma machine learning lainnya. Keunggulan ini disebabkan oleh kemampuan SVM untuk menangani data yang kompleks dan multi-dimensional, yang sering kali menjadi ciri khas dataset transaksi online. Namun, tantangan utama dalam penerapan SVM adalah kebutuhan akan pemrosesan data yang cermat, karena kualitas data sangat mempengaruhi kinerja model. Data transaksi online sering kali memiliki banyak noise, seperti informasi yang tidak relevan atau data yang hilang, sehingga diperlukan tahapan preprocessing yang baik untuk memastikan model dapat bekerja secara optimal. Selain itu, penggunaan teknik seperti kernel SVM dan tuning parameter menjadi faktor penting dalam mengoptimalkan kinerja model. (Widianto,T.P, 2022)

METODE PENELITIAN

Pengumpulan Data

Tahap pertama dalam penelitian ini adalah pengumpulan data transaksi online. Data yang dikumpulkan terdiri dari transaksi sah dan transaksi yang teridentifikasi sebagai penipuan. Dataset akan diambil dari sumber terpercaya atau dari dataset publik yang tersedia, seperti dataset dari Kaggle atau perusahaan keuangan yang bersedia membagikan data transaksi mereka. Data tersebut akan mencakup berbagai fitur seperti jumlah transaksi, waktu transaksi, metode pembayaran, lokasi pengguna, dan informasi lainnya yang relevan untuk mendeteksi penipuan. (Febriady et al., 2022)

Preprocessing Data

Data transaksi online sering kali tidak bersih dan mengandung noise, informasi yang tidak relevan, atau data yang hilang. Oleh karena itu, tahap preprocessing data dilakukan untuk memastikan kualitas data yang digunakan dalam pelatihan model. Tahap ini meliputi:

a. Pembersihan Data: Menghapus atau memperbaiki data yang hilang atau tidak konsisten.
b. Normalisasi Data: Menormalkan data numerik agar berada dalam skala yang sama, sehingga model dapat lebih mudah belajar.
c. Penanganan Ketidakseimbangan Data: Menggunakan teknik oversampling (misalnya SMOTE) atau undersampling untuk memastikan jumlah data penipuan dan transaksi sah lebih seimbang, yang penting agar model tidak bias terhadap kelas mayoritas.

Algoritma Support Vector Machine (SVM)

Algoritma SVM bekerja dengan mencari hyperplane terbaik yang memisahkan data dalam dua kelas berbeda. SVM mengidentifikasi hyperplane optimal yang memiliki margin terbesar antara kelas penipuan dan kelas sah. Persamaan matematis dari SVM dapat dinyatakan sebagai:

$f(x)=w^Tx+b$ ……..……..(1)

Dimana:

Jurnal Minfo Polgan
Volume 13, Nomor 2, Oktober 2024 e-ISSN : 2797-3298
DOI : https://doi.org/10.33395/jmp.v13i2.14186 p-ISSN : 2089-9424

w : adalah vektor bobot
x : adalah vektor input
b : adalah bias

Tujuan utama dari SVM adalah memaksimalkan margin $\frac{2}{w}$, yang berarti mengoptimalkan nilai w dan b sedemikian rupa sehingga margin antara data penipuan dan sah menjadi maksimum. Fungsi objektif SVM dapat dinyatakan sebagai berikut:

$$\min \frac{1}{2}\|w\|^2 + c\sum_{i=1}^{N} \xi i$$

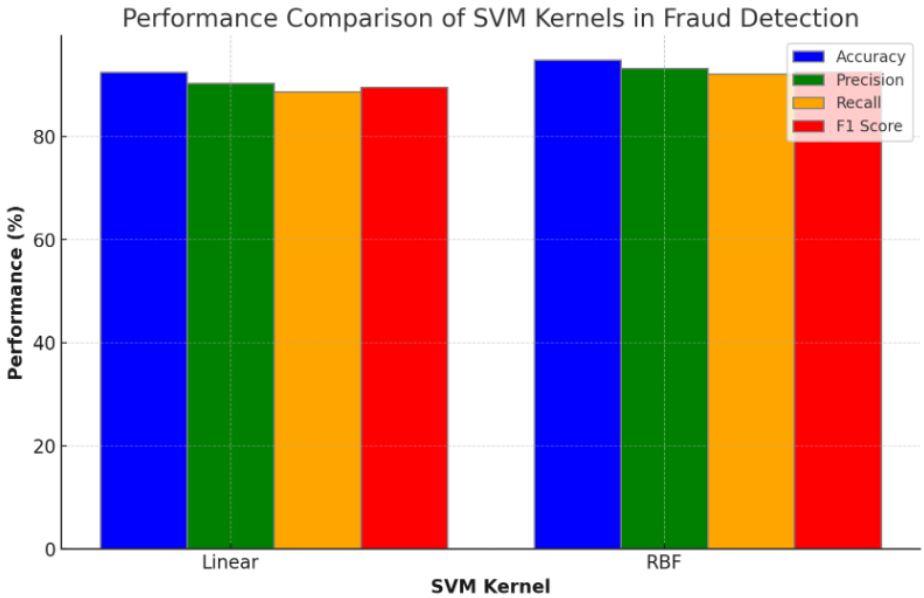
HASIL DAN PEMBAHASAN

Penelitian ini menggunakan dataset transaksi online yang terdiri dari data transaksi sah dan data penipuan. Dataset dibagi menjadi dua bagian: 80% digunakan untuk pelatihan (training) model dan 20% untuk pengujian (testing). Implementasi SVM dilakukan dengan berbagai parameter dan kernel untuk mendapatkan hasil yang optimal. Setelah melalui proses pelatihan dan pengujian, model SVM dievaluasi berdasarkan beberapa metrik utama yaitu akurasi, presisi, recall, dan F1-score. Hasil evaluasi kinerja model SVM dengan kernel Radial Basis Function (RBF) dibandingkan dengan kernel linear untuk melihat performa terbaik. Berikut adalah hasil perbandingan kinerja model:

Kernel SVM	Akurasi	Presisi	Recall	F1-score
Linear	92.5%	90.3%	88.7%	89.5%
RBF	94.8%	93.2%	92.1%	92.6%

Tabel1. Perbandingan kinerja SVM

Berikut ini adalah grafik hasilnya :



Gambar 1. Grafik Hasil Akurasi, Presisi, Recall, dan F1-score

Grafik di atas menunjukkan perbandingan kinerja dua kernel SVM, yaitu Linear dan RBF, dalam mendeteksi penipuan pada transaksi online. Dari grafik, dapat dilihat bahwa:

a. Kernel RBF memiliki performa lebih baik dibandingkan kernel linear di semua metrik yang diukur (akurasi, presisi, recall, dan F1-score).
b. Akurasi tertinggi diperoleh oleh kernel RBF sebesar 94.8%, dibandingkan dengan 92.5% pada

Jurnal Minfo Polgan
Volume 13, Nomor 2, Oktober 2024 e-ISSN : 2797-3298
DOI : https://doi.org/10.33395/jmp.v13i2.14186 p-ISSN : 2089-9424

kernel linear.

c. Presisi dan Recall dari kernel RBF juga lebih tinggi, menunjukkan bahwa model ini mampu mendeteksi penipuan dengan lebih baik dan mengurangi kesalahan deteksi.
d. F1-score, yang menggabungkan presisi dan recall, menunjukkan nilai yang lebih baik pada kernel RBF (92.6%) dibandingkan kernel linear (89.5%).

Dari hasil eksperimen, terlihat bahwa penggunaan algoritma SVM dengan kernel RBF lebih unggul dalam mendeteksi penipuan dibandingkan dengan kernel linear. Kernel RBF mampu menangkap pola yang lebih kompleks dalam data, sehingga dapat mengidentifikasi transaksi yang mencurigakan dengan lebih akurat.

Faktor utama yang memengaruhi performa SVM adalah parameter kernel dan teknik preprocessing data. Normalisasi data serta penyeimbangan antara transaksi sah dan penipuan sangat penting untuk memastikan bahwa model tidak bias terhadap salah satu kelas. Selain itu, pemilihan parameter regulasi C dan parameter kernel γ pada RBF juga memengaruhi performa secara signifikan.

KESIMPULAN

Penelitian ini berhasil mengimplementasikan algoritma Support Vector Machine (SVM) untuk mendeteksi penipuan pada transaksi online. Berdasarkan hasil eksperimen dan analisis, dapat ditarik beberapa kesimpulan sebagai berikut:

1. Efektivitas Algoritma SVM: Algoritma SVM, terutama dengan penggunaan kernel Radial Basis Function (RBF), terbukti efektif dalam mengklasifikasikan transaksi sah dan transaksi penipuan. Hal ini dibuktikan dengan nilai akurasi yang mencapai 94.8%, menunjukkan bahwa model mampu memprediksi sebagian besar transaksi dengan benar.
2. Kinerja Model: Model SVM dengan kernel RBF menunjukkan kinerja yang lebih unggul dibandingkan kernel linear dalam hal presisi, recall, dan F1-score. Presisi yang tinggi (93.2%) mengindikasikan bahwa model jarang memberikan prediksi positif palsu, sementara recall yang tinggi (92.1%) menunjukkan bahwa model mampu mendeteksi penipuan dengan baik.
3. Preprocessing dan Pemilihan Parameter: Hasil penelitian juga menekankan pentingnya preprocessing data, seperti normalisasi dan penyeimbangan data, dalam meningkatkan akurasi dan stabilitas model. Selain itu, pemilihan parameter yang tepat untuk SVM, seperti C dan γ, berperan penting dalam memperoleh hasil yang optimal.
4. Kemampuan Generalisasi: SVM dengan kernel RBF memiliki kemampuan generalisasi yang baik, yang memungkinkan model untuk mendeteksi pola transaksi penipuan yang mungkin tidak teratur dan sulit diprediksi. Hal ini membuat SVM cocok untuk diterapkan pada sistem deteksi penipuan yang dinamis dan variatif.
5. Potensi Pengembangan Lebih Lanjut: Meskipun model SVM dengan kernel RBF menunjukkan hasil yang memuaskan, penelitian ini membuka peluang untuk pengembangan lebih lanjut, seperti optimasi parameter menggunakan algoritma pencarian (misalnya Grid Search atau Random Search) dan kombinasi dengan teknik machine learning lainnya untuk meningkatkan kinerja deteksi.

Berdasarkan hasil penelitian, algoritma Support Vector Machine (SVM) dengan kernel RBF dapat dijadikan solusi yang andal dan efektif untuk mendeteksi penipuan pada transaksi online. Implementasi yang tepat, disertai dengan pemilihan parameter dan preprocessing data yang baik, memungkinkan sistem ini beroperasi dengan akurasi yang tinggi dan potensi untuk terus dikembangkan guna menghadapi pola-pola penipuan yang semakin kompleks.

REFERENSI

Febriady, M., Samsuryadi, S., & Rini, D. P. (2022). Klasifikasi Transaksi Penipuan Pada Kartu Kredit Menggunakan Metode Resampling Dan Pembelajaran Mesin. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, *6*(2), 1010–1016. https://doi.org/10.30865/MIB.V6I2.3515

Hasibuan, L. S., Alfiatul, F., Fakultas, J., Dan Ilmu, M., Alam, P., Kunci, K., Genetika, A., Kredit, K., Optimasi, ;, Svm, ;, & Penipuan, T. (2024). Deteksi Penipuan Kartu Kredit Menggunakan Support Vector Machine dengan Optimasi Grid Search dan Genetic Algorithm. *Building of Informatics, Technology and Science (BITS)*, *6*(1), 344–353. https://doi.org/10.47065/BITS.V6I1.5355

Iqbal Ahmadi, M., Apriani, F., Kurniasari, M., Handayani, S., Gustian, D., Raya Cibatu Cisaat No, J., Kaler, C., & Barat, J. (2020). SENTIMENT ANALYSIS ONLINE SHOP ON THE PLAY STORE USING METHOD SUPPORT VECTOR MACHINE (SVM). *Seminar Nasional Informatika (SEMNASIF)*, *1*(1), 196–203. http://jurnal.upnyk.ac.id/index.php/semnasif/article/view/4101

Khatib Sulaiman, J., & Bhayangkara Jakarta Raya, U. (2024). Dampak Pengambilan Sampel Data untuk Optimalisasi Data tidak seimbang pada Klasifikasi Penipuan Transaksi E-Commerce. *The Indonesian Journal of Computer Science*, *13*(2), 3070. https://doi.org/10.33022/IJCS.V13I2.3698

Nugraha, A., Nugraha, A. C., & Irawan, M. I. (2023). Komparasi Deteksi Kecurangan pada Data Klaim Asuransi Pelayanan Kesehatan Menggunakan Metode Support Vector Machine (SVM) dan Extreme Gradient Boosting (XGBoost). *Jurnal Sains Dan Seni ITS*, *12*(1), A40–A46. https://doi.org/10.12962/j23373520.v12i1.107032

Optimasi Generative Adversarial Networks Untuk Oversampling Deteksi Penipuan Menggunakan Kartu Kredit. (n.d.). Retrieved October 20, 2024, from https://etd.repository.ugm.ac.id/penelitian/detail/212726

Zuhairah, A. (2023). *Penerapan Algoritma Random Forest, Support Vector Machines (Svm) dan Gradient Boosted Tree (Gbt) Untuk Deteksi Penipuan (Fraud Detection) Pada Transaksi Kartu Kredit*. https://repository.uinjkt.ac.id/dspace/handle/123456789/70536

Habibi, S. (2022). *Studi Dan Implementasi Snowflake Untuk Analisis Big Data Pada Cloud Data Warehousing Dengan Menggunakan Algoritma Support Vector Machine* (Doctoral dissertation, Universitas Siliwangi).

Lubis, S. K., Dar, M. H., & Nasution, F. A. (2023). Analisis Sentimen Ulasan Pengguna Aplikasi pada Google Play Store Menggunakan Algoritma Support Vector Machine. *INFORMATIKA*, *11*(2), 120-128.

Saputra, D. R. K., Via, Y. V., & Sihananto, A. N. (2024). Deteksi Anomali Menggunakan Ensemble Learning Dan Random Oversampling Pada Penipuan Transaksi Keuangan. *Jurnal Informatika dan Teknik Elektro Terapan*, *12*(3).