

STAB-AD: Framework Deteksi Anomali Mobile Banking Berbasis Perilaku Dan Spatio-Temporal

Tri Wiyono*, Asrul Helmandi, Ibnu Afandi Manao, Andika Dwi Aryo, Ardiansyah

Pascasarjana, Magister Teknologi Informasi, Universitas Pembangunan Pancabudi, Medan, Indonesia

Email: ^{1,*}tri82.wiyono@gmail.com, ²andihelmandi@gmail.com, ³afandiibnu05@gmail.com, ⁴andhika.basisgb@gmail.com, ⁵andhika.basisgb@gmail.com

Email Penulis Korespondensi: tri82.wiyono@gmail.com*

Submitted: 09/04/2026; Accepted: 25/04/2026; Published: 30/06/2026

Abstrak– Perkembangan mobile banking meningkatkan risiko fraud, seperti *unauthorized transaction* dan *account takeover* (ATO), yang semakin kompleks dan sulit dideteksi. Penelitian ini mengusulkan metode *Spatio-Temporal and Behavioral Agreement-based Anomaly Detection* (STAB-AD) untuk mendeteksi anomali transaksi melalui integrasi fitur perilaku dan *spatio-temporal*. Model mengombinasikan *Isolation Forest*, *Local Outlier Factor*, dan *One-Class SVM* dengan mekanisme *agreement-based* untuk mengukur konsistensi antar model dan mengklasifikasikan tingkat risiko. Fitur yang digunakan mencakup frekuensi transaksi, deviasi nilai, *time gap*, perubahan perangkat, serta indikator *spatio-temporal* seperti jarak geografis, *velocity*, dan *impossible travel*. Evaluasi dilakukan menggunakan *precision*, *recall*, *F1-score*, dan *ROC-AUC* dengan pendekatan *pseudo-labeling* berbasis aturan sebagai proksi *ground truth* pada data tidak berlabel. Hasil menunjukkan bahwa STAB-AD mampu meningkatkan kinerja deteksi dibandingkan baseline serta efektif dalam *risk stratification*, dengan hanya 0,72% transaksi teridentifikasi sebagai risiko tinggi. Fitur *spatio-temporal* terbukti signifikan dalam mengidentifikasi pola perpindahan tidak wajar yang mengindikasikan potensi ATO. Namun, penelitian ini masih terbatas pada penggunaan *threshold* berbasis asumsi dan validasi tanpa label aktual, sehingga diperlukan pendekatan yang lebih *data-driven* serta evaluasi berbasis *ground truth* pada penelitian selanjutnya.

Kata Kunci: STAB-AD; Deteksi Anomali; Mobile Banking; Spatio-Temporal; Fraud

Abstract– The rapid growth of mobile banking has increased the risk of fraud, such as unauthorized transactions and account takeover (ATO), which are becoming more complex and difficult to detect. This study proposes a method called *Spatio-Temporal and Behavioral Agreement-based Anomaly Detection* (STAB-AD) to identify anomalous transactions by integrating behavioral and spatio-temporal features. The model combines *Isolation Forest*, *Local Outlier Factor*, and *One-Class SVM* with an agreement-based mechanism to measure inter-model consistency and classify risk levels. The features used include transaction frequency, value deviation, time gap, device change, as well as spatio-temporal indicators such as geographic distance, velocity, and impossible travel. The evaluation is conducted using precision, recall, F1-score, and ROC-AUC, employing a rule-based pseudo-labeling approach as a proxy for ground truth on unlabeled transaction data. The results show that STAB-AD improves detection performance compared to baseline methods and is effective in risk stratification, with only 0.72% of transactions identified as high risk. Spatio-temporal features are found to be significant in detecting abnormal movement patterns that may indicate potential ATO. However, this study is limited by the use of assumption-based thresholds and validation without actual labeled data. Therefore, future work should focus on more data-driven approaches and validation using ground truth labels.

Keywords: STAB-AD; Anomaly Detection; Mobile Banking; Spatio-Temporal; Fraud

1. PENDAHULUAN

Perkembangan *mobile banking*[1] meningkat pesat seiring transformasi digital di sektor perbankan[2]. Kemudahan dan kecepatan transaksi menjadikannya sebagai kanal utama aktivitas finansial[3], namun disisi lain turut meningkatkan risiko keamanan[4], khususnya fraud seperti *unauthorized transaction*[5] dan *account takeover* (ATO)[6] yang semakin kompleks dan sulit dideteksi. Deteksi anomali [7] menjadi pendekatan penting dalam mitigasi *fraud*[8]. Dibandingkan sistem berbasis aturan statis, pendekatan *machine learning* [9],[10],[11], khususnya *unsupervised learning* [12], lebih adaptif terhadap dinamika data dan tidak bergantung pada ketersediaan label. Berbagai metode seperti *Z-Score*, *Isolation Forest*[13], *Local Outlier Factor*[14], dan *One-Class SVM*[15][16] telah banyak digunakan dan terbukti efektif dalam mendeteksi *outlier*[17], namun pendekatan tersebut umumnya bekerja secara independen dan belum secara eksplisit mempertimbangkan konsistensi hasil antar model, sehingga rentan terhadap variasi sensitivitas dan *false positive*.

Sejumlah penelitian terkini mulai mengintegrasikan *behavioral profiling*[18] dan analisis *spatio-temporal* [19] untuk meningkatkan kemampuan deteksi anomali[20] khususnya dalam mengidentifikasi pola pergerakan tidak wajar seperti *impossible travel*. Selain itu, pendekatan *ensemble* juga digunakan untuk meningkatkan performa. Namun demikian, pendekatan yang ada umumnya masih terbatas pada integrasi parsial fitur atau agregasi hasil tanpa mempertimbangkan tingkat kepercayaan (*confidence*) dari masing-masing model. Akibatnya, hasil deteksi sering kali sulit diinterpretasikan dan kurang mendukung proses pengambilan keputusan.

Berdasarkan kajian tersebut, terdapat tiga kesenjangan utama penelitian (*research gap*): (1) belum adanya integrasi yang komprehensif antara fitur *behavioral* dan *spatio-temporal* dalam satu kerangka deteksi, (2) belum

adanya mekanisme untuk mengukur konsistensi dan reliabilitas hasil deteksi antar model, serta (3) keterbatasan dalam klasifikasi tingkat risiko anomali berbasis konsensus model. Kesenjangan ini berdampak pada rendahnya kepercayaan terhadap hasil deteksi dan keterbatasan dalam prioritas investigasi, terutama pada skenario ATO yang bersifat kompleks dan multi-dimensi.

Untuk mengatasi permasalahan tersebut, penelitian ini mengusulkan kerangka *Spatio-Temporal and Behavioral Agreement-based Anomaly Detection (STAB-AD)* yang mengintegrasikan fitur behavioral, spasial, dan temporal secara simultan dengan pendekatan *agreement-based anomaly detection*. Berbeda dengan pendekatan sebelumnya, STAB-AD tidak hanya menggabungkan beberapa model *unsupervised learning*, tetapi juga mengukur konsistensi antar model sebagai indikator kepercayaan (*confidence-aware detection*). Fitur yang digunakan meliputi frekuensi transaksi, deviasi nilai, perubahan perangkat, serta indikator spatio-temporal seperti jarak geografis, kecepatan perpindahan (*velocity*), dan interval waktu transaksi, [21].

Kontribusi utama penelitian ini adalah: (1) pengembangan *framework* STAB-AD yang mengintegrasikan fitur *behavioral* dan *spatio-temporal* secara komprehensif, (2) penerapan mekanisme *agreement-based* sebagai indikator reliabilitas deteksi, (3) pengembangan skema *risk stratification* berbasis konsensus multi-model, serta (4) analisis efektivitas fitur spatio-temporal dalam mengidentifikasi pola anomali transaksi dan potensi ATO.

2. METODOLOGI PENELITIAN

2.1 Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif [22] dengan metode *unsupervised machine learning*[23],[24] untuk mendeteksi anomali pada transaksi mobile banking[25]. Pendekatan ini dipilih karena data yang digunakan tidak memiliki label (*unlabeled*)[26],[27], sehingga model difokuskan pada pembelajaran pola transaksi normal dan identifikasi penyimpangan (*outlier*).

Penelitian ini mengusulkan kerangka *Spatio-Temporal and Behavioral Agreement-based Anomaly Detection (STAB-AD)* yang mengintegrasikan tiga komponen utama yaitu *behavioral features*, *spatio-temporal features*, dan *agreement-based detection*. Pendekatan ini bertujuan meningkatkan akurasi dan konsistensi deteksi, serta mendukung identifikasi potensi *Account Take Over (ATO)*. Alur penelitian ditunjukkan pada Gambar 1



Gambar 1. Diagram Model STAB-AD

2.2 Dataset dan Karakteristik Data

Data yang digunakan merupakan data transaksi *mobile banking* yang telah melalui proses anonimisasi untuk menjaga kerahasiaan nasabah. Struktur data disajikan pada Tabel 1.

Tabel 1. Atribut Transaksi

Nama Atribut	Keterangan
<i>syslogno</i>	Identifikasi unit transaksi
<i>device_code</i>	Identitas perangkat
<i>transaksi</i>	Jenis Transaksi
<i>srac</i>	Nomor rekening sumber
<i>bank_code, bank_name</i>	Informasi bank tujuan
<i>trx_value</i>	Nilai transaksi
<i>received_time</i>	Waktu transaksi
<i>last_state, last_rc</i>	Status transaksi
<i>longitude, latitude</i>	Lokasi transaksi

Data bersifat tidak berlabel (*unlabeled*) dan memiliki karakteristik tidak seimbang (*imbalanced*), dimana proporsi anomali relatif kecil dibandingkan Transaksi normal. *Atribut* tersebut merepresentasikan pola perilaku transaksi (*transaction behavioral pattern*) yang menjadi dasar dalam deteksi anomali dan potensi *fraud*.

2.3 Data Preprocessing

Tahap *preprocessing* dilakukan untuk meningkatkan kualitas data sebelum pemodelan:

- Data *cleaning*, menghapus duplikasi dan menangani inkonsistensi data.
- Penanganan *Missing Value*, Menggunakan metode imputasi sederhana (*mean/median*) atau penghapusan data jika proporsinya kecil.
- Normalisasi menggunakan *Min-Max Scaling*.

2.4 Feature Engineering

Feature engineering mencakup frekuensi transaksi, deviasi nilai, *time gap*, dan *device change* untuk merepresentasikan pola perilaku dan mendeteksi anomaly.

a. Frekuensi Transaksi

$$f = \frac{n}{t} \quad (1)$$

b. Deviasi Nilai Transaksi

$$d = |x - \mu| \quad (2)$$

c. *Time Gap*

$$\Delta t = t_i - t_{i-1} \quad (3)$$

d. *Device Change* Indikator

$$DC(x) = \begin{cases} 1, & \text{jika perangkat berubah} \\ 0, & \text{lainnya} \end{cases} \quad (4)$$

2.5 Spatio-Temporal Feature

Untuk mendeteksi anomaly berbasis lokasi dan waktu, digunakan fitur tambahan sebagai berikut :

a. Jarak Geografis (*Haversine Distance*)

$$d = 2r \cdot \arcsin \left(\sqrt{\sin^2 \left(\frac{\Delta \varphi}{2} \right) + \cos(\varphi_1) \cos(\varphi_2) \sin^2 \left(\frac{\Delta \lambda}{2} \right)} \right) \quad (5)$$

b. *Euclidean Distance*

$$d_E = \sqrt{(x_1 - x_2)^2 + (y_2 - y_1)^2} \quad (6)$$

c. Kecepatan Perpindahan Transaksi

$$v = \frac{d}{\Delta t} \quad (7)$$

d. *Impossible Travel* Indikator

$$IT(x) = \begin{cases} 1, & \text{jika } v > v_{max} \\ 0, & \text{lainnya} \end{cases} \quad (8)$$

2.6 Pemodelan Anomali

Didalam penelitian ini menggunakan tiga algoritma *unsupervised learning*, yaitu:

a. *Isolation Forest*

$$S(x, n) = 2^{-\frac{E(h(x))}{c(n)}} \quad (9)$$

Parameter :

1. `n_estimators` = 100
2. `contamination` = 0.05

b. *Local Outlier Factor*

$$LOF(x) = \frac{\sum_{o \in N_k(x)} \frac{lr_d(o)}{lr_d(x)}}{|N_k(x)|} \quad (10)$$

Parameter :

1. `n_neighbors` = 20

c. *One Class Support Vector Machine* (OCSVM)

$$f(x) = \text{sign}(w \cdot \phi(x) - \rho) \quad (11)$$

Parameter :

2. `Kernel` = RBF
3. `v = 0.05`, `v = scale`

Parameter dipilih berdasarkan praktik umum pada deteksi anomaly dan disesuaikan dengan karakteristik data yang tidak berlabel dan *imbalanced*.

2.7 Agreement-Based Anomaly Detection

Hasil dari masing-masing model digabungkan menggunakan *agreement score* untuk mengukur konsistensi deteksi antar model:

$$A(x) = \sum_{i=1}^k I_i(x) \quad (12)$$

$$I_i(x) = \begin{cases} 1, & \text{anomali} \\ 0, & \text{normal} \end{cases} \quad (13)$$

Berdasarkan nilai *agreement*, transaksi diklasifikasikan sebagai berikut :

- a. 0 = Normal
- b. 1 = *Low* Anomali
- c. 2 = *Medium* Anomali
- d. 3 = *High* Anomali

Pendekatan ini digunakan untuk mengurangi *false positive* dari model tunggal dan meningkatkan kepercayaan (*confidence*) terhadap hasil deteksi.

2.8 Account Take Over (ATO) Detection

Pendekatan deteksi ATO dilakukan secara *rule-based* dengan kombinasi indikator perilaku dan *spatio-temporal*.

$$ATO(x) = \begin{cases} 1, & \text{jika } DC(x) = 1 \wedge IT(x) = 1 \wedge trx_value > T \\ 0, & 0 \text{ lainnya} \end{cases} \quad (14)$$

Suatu transaksi dikategorikan sebagai potensi ATO jika memenuhi kombinasi indikator berikut: (1) perubahan *device_code*, (2) perubahan lokasi signifikan, (3) kecepatan perpindahan tidak wajar, dan (4) nilai transaksi tinggi. Pendekatan ini memungkinkan identifikasi anomali yang lebih kontekstual terhadap skenario *fraud*.

Rule ini dirancang berdasarkan pola *fraud* umum dalam sistem perbankan digital, khususnya kombinasi perubahan perangkat dan anomali lokasi dalam waktu singkat.

2.9 Penentuan Threshold Anomali

Karena data tidak berlabel, *threshold* ditentukan menggunakan pendekatan statistik berbasis distribusi skor (misalnya kuantil) untuk merepresentasikan nilai ekstrem (*outlier*) pada data

$$T = Q_{0.95}(\text{score}) \quad (15)$$

Transaksi dengan skor melewati *threshold* dikategorikan sebagai anomali.

Pemilihan kuantil 95% didasarkan pada asumsi bahwa anomali merupakan kejadian langka (*rare event*), sehingga hanya nilai ekstrem yang dikategorikan sebagai anomali. Pendekatan ini umum digunakan dalam *unsupervised anomaly detection* untuk menghindari bias akibat ketiadaan label.

2.10 Evaluasi Model

Evaluasi model dilakukan menggunakan metrik *precision*, *recall*, *F1-score*, dan *ROC-AUC* untuk mengukur kinerja deteksi anomali. Karena data tidak memiliki label, evaluasi dilakukan menggunakan

- a. *Pseudo-labeling* berbasis *threshold* dan rule ATO sebagai proksi *ground truth*.
- b. Analisis distribusi skor anomali.
- c. Perbandingan relatif antar model (*relative performance comparison*)

Pendekatan ini memungkinkan evaluasi kinerja model secara indikatif, meskipun belum menggantikan validasi berbasis label aktual (*ground truth*).

3. HASIL DAN PEMBAHASAN

3.1 Karakteristik Dataset

Dataset terdiri dari 27.230 transaksi mobile banking, 18.592 perangkat unik, dan 51 bank tujuan dalam periode satu hari (27 Maret 2026), yang merepresentasikan aktivitas transaksi harian. Rasio perangkat yang tinggi menunjukkan frekuensi transaksi per pengguna relatif rendah, sedangkan variasi bank tujuan mencerminkan heterogenitas pola transaksi. Kondisi ini menjadikan dataset relevan untuk analisis perilaku dan deteksi anomali.

Tabel 2. Statistik Deskriptif Fitur STAB-AD

Fitur	Mean	Std	Min	Max
<i>trx_value</i>	8.36×10^{10}	7.09×10^{12}	1.0	7.80×10^{14}
<i>time_gap</i> (Δt)	1.91×10^3	6.92×10^3	0.0	8.35×10^4
<i>freq</i> (f)	1.67	1.60	1.0	28.0
<i>deviasi</i> (d)	9.73×10^{10}	5.48×10^{12}	0.0	5.45×10^{14}
<i>velocity</i> (v)	7.09×10^2	4.36×10^4	0.0	6.60×10^6

Statistik pada Tabel 2 menunjukkan variabilitas tinggi pada *trx_value*, *deviasi*, dan *velocity*, yang mengindikasikan heterogenitas perilaku transaksi.

Tabel 3. Spatio-Temporal Features

Fitur	Jumlah (IT/DC=1)	Persentase (%)	Mean	Max
Impossible Travel (IT)	244	0.896	–	–
Device Change (DC)	18,592	68.28	–	–
distance_km	–	–	10.27	11,131.55

Berdasarkan Tabel 3, *impossible travel* (0,896%) bersifat jarang namun indikatif, sedangkan *device change* (68,28%) kurang diskriminatif. Nilai maksimum jarak yang tinggi menunjukkan perpindahan ekstrem, sehingga fitur spasial berperan penting dalam mendeteksi potensi ATO.

3.2 Hasil Deteksi Anomali per Model

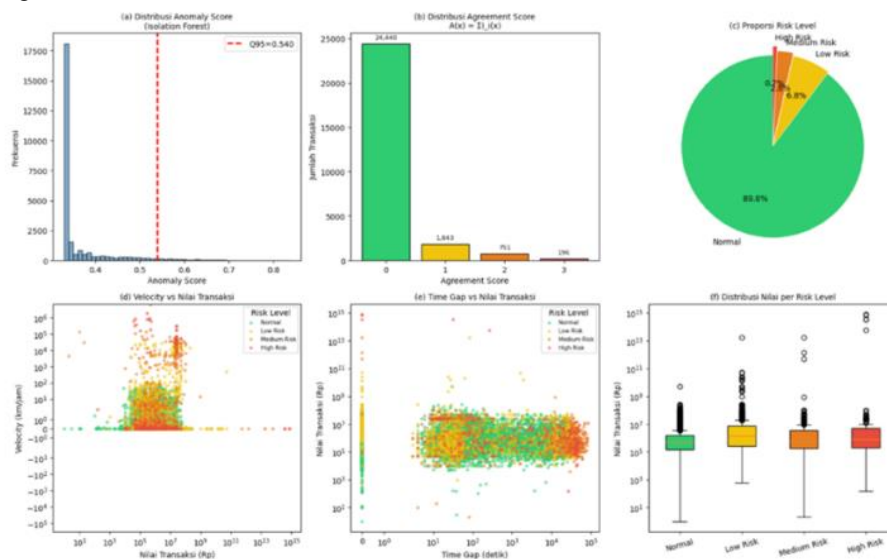
Pada tahap ini dilakukan evaluasi hasil deteksi anomali menggunakan tiga algoritma utama, yaitu *Isolation Forest*, *Local Outlier Factor* (LOF), dan *One-Class SVM* (OCSVM). Masing-masing model memiliki pendekatan yang berbeda dalam mengidentifikasi anomali, sehingga penting untuk membandingkan performanya baik dari sisi jumlah anomali yang terdeteksi maupun karakteristik transaksi yang dihasilkan.

Perbandingan ini bertujuan untuk mengidentifikasi pola deteksi yang dihasilkan oleh setiap model, termasuk tingkat sensitivitas terhadap data outlier serta kecenderungan dalam menangkap anomali berdasarkan nilai transaksi. Selain itu, analisis ini juga menjadi dasar dalam mengevaluasi kontribusi masing-masing model dalam skema *agreement-based* pada framework STAB-AD. Hasil perbandingan deteksi anomali dari ketiga model disajikan pada Tabel 4.

Tabel 4. Perbandingan Deteksi 3 Model

Algoritma	n Anomali	n Normal	% Anomali	Rata-rata Nilai Transaksi Anomali (Rp)
<i>Isolation Forest</i>	1.362	25.868	5,00	1.643.105.852.240
<i>Local Outlier Factor</i>	1.362	25.868	5,00	1.670.382.966.382
<i>One-Class SVM</i>	1.209	26.021	4,44	1.867.092.714.270

Tabel 4 menunjukkan bahwa *Isolation Forest* dan LOF menghasilkan proporsi anomali yang sama (5%) sesuai parameter *contamination*, sedangkan OCSVM lebih konservatif (4,44%). Rata-rata nilai transaksi pada kelompok anomali lebih tinggi dibandingkan transaksi normal, menunjukkan bahwa nilai transaksi merupakan indikator penting dalam karakterisasi anomali.



Gambar 2. Visualisasi STAB-AD

3.3 Agreement-Based Classification

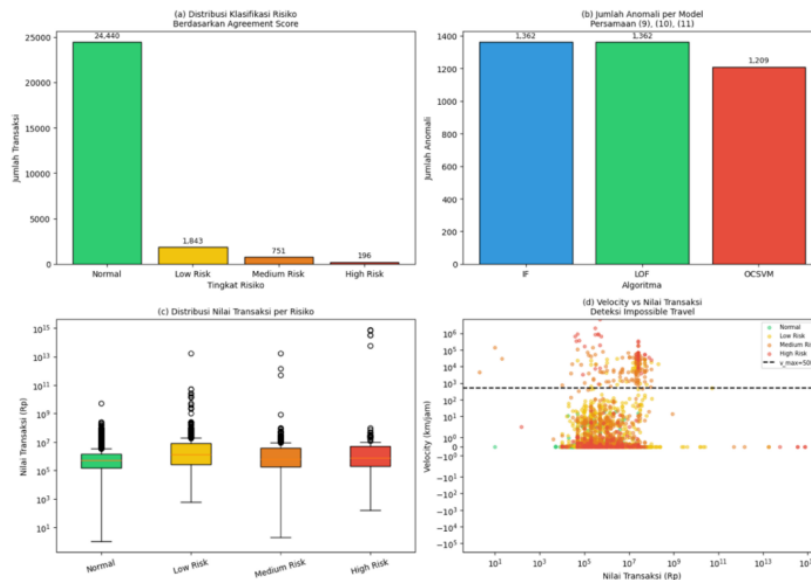
Setelah setiap model menghasilkan label anomali, hasilnya digabungkan melalui pendekatan *agreement-based classification*. Pendekatan ini melihat seberapa banyak model yang sepakat menandai suatu transaksi sebagai anomali (*agreement score*), sehingga keputusan yang dihasilkan lebih stabil dibandingkan satu model saja. Semakin tinggi tingkat kesepakatan, semakin besar indikasi bahwa transaksi tersebut benar-benar menyimpang.

Berdasarkan nilai tersebut, transaksi dikelompokkan ke dalam kategori *normal*, *low risk*, *medium risk*, dan *high risk*. Dengan cara ini, proses penyaringan anomali menjadi lebih terstruktur dan membantu penentuan prioritas dalam investigasi fraud. Distribusi hasilnya disajikan pada Tabel 5.

Tabel 5. Distribusi *Agreement Score*

Agreement (A)	Klasifikasi	Jumlah (n)	Persentase (%)
0	Normal	24.440	89,75
1	Low Risk	1.843	6,77
2	Medium Risk	751	2,76
3	High Risk	196	0,72

Mayoritas transaksi berada pada kategori normal (89,75%), yang konsisten dengan karakteristik fraud sebagai *rare event*. Pendekatan *agreement-based* memungkinkan proses *risk stratification*, di mana hanya 0,72% transaksi dikategorikan sebagai *high risk* dan dapat diprioritaskan untuk investigasi lebih lanjut.



Gambar 3. Distribusi Risiko

Gambar 3 menyajikan visualisasi hasil deteksi anomali menggunakan STAB-AD yang mencakup distribusi risiko, perbandingan model, nilai transaksi, dan hubungan dengan velocity.

Distribusi risiko (Gambar 3a) menunjukkan bahwa mayoritas transaksi berada pada kategori normal, sedangkan anomali relatif kecil, mencerminkan karakteristik fraud sebagai *rare event*. Perbandingan model (Gambar 3b) menunjukkan bahwa *Isolation Forest* dan LOF menghasilkan deteksi yang serupa, sementara OCSVM lebih konservatif, menegaskan perbedaan sensitivitas antar model dan pentingnya pendekatan multi-model.

Distribusi nilai transaksi (Gambar 3c) menunjukkan bahwa kelompok anomali cenderung memiliki nilai lebih tinggi, dengan *outlier* dominan pada kategori risiko tinggi, yang mengindikasikan keterkaitan antara nilai transaksi besar dan potensi *fraud*. Hubungan nilai transaksi dan *velocity* (Gambar 3d) menunjukkan bahwa transaksi dengan kecepatan tinggi lebih banyak muncul pada kategori risiko menengah hingga tinggi, mengindikasikan efektivitas fitur *spatio-temporal* dalam mendeteksi pola perpindahan tidak wajar.

Secara keseluruhan, STAB-AD mampu mengintegrasikan berbagai dimensi fitur untuk mendeteksi anomali secara lebih komprehensif serta meningkatkan kepercayaan deteksi melalui mekanisme *agreement-based*.

3.4 Evaluasi Performa Model

Pada tahap ini dilakukan evaluasi performa masing-masing model dalam mendeteksi anomali menggunakan metrik *silhouette score*, *accuracy*, *precision*, *recall*, dan *F1-score*. Karena dataset tidak memiliki label ground truth, evaluasi menggunakan pendekatan *pseudo ground truth* berbasis *agreement* ($A \geq 2$) sebagai referensi anomali. Hasil evaluasi masing-masing model disajikan pada Tabel 6.

Tabel 6. Evaluasi Model

Model	Silhouette	Accuracy*	Precision*	Recall*	F1-Score*
<i>Isolation Forest</i>	0,8007	0,9837	0,6850	0,9852	0,8081
<i>Local Outlier Factor</i>	0,4459	0,9402	0,2496	0,3590	0,2945
<i>One-Class SVM</i>	0,8077	0,9808	0,6758	0,8627	0,7579

Keterangan: Evaluasi dilakukan terhadap *pseudo ground truth* ($Agreement \geq 2$)

Berdasarkan Tabel 5, *Isolation Forest* memiliki *recall* tertinggi, sedangkan *One-Class SVM* menunjukkan keseimbangan terbaik antara *precision* dan *recall*. Sebaliknya, LOF memiliki performa terendah, mengindikasikan keterbatasan pendekatan berbasis kepadatan lokal pada dataset ini.

Evaluasi dilakukan menggunakan *pseudo ground truth* berbasis *agreement* ($A \geq 2$) karena data tidak berlabel. Pendekatan ini umum digunakan pada *unsupervised learning* sebagai estimasi awal performa, namun berpotensi menghasilkan bias optimistik (*self-referential bias*). Oleh karena itu, hasil evaluasi lebih merepresentasikan konsistensi antar model dibandingkan kondisi aktual, sehingga diperlukan validasi lanjutan menggunakan data berlabel atau expert judgment.

Tabel 7. *Inter-Rater Agreement (Cohen's Kappa)*

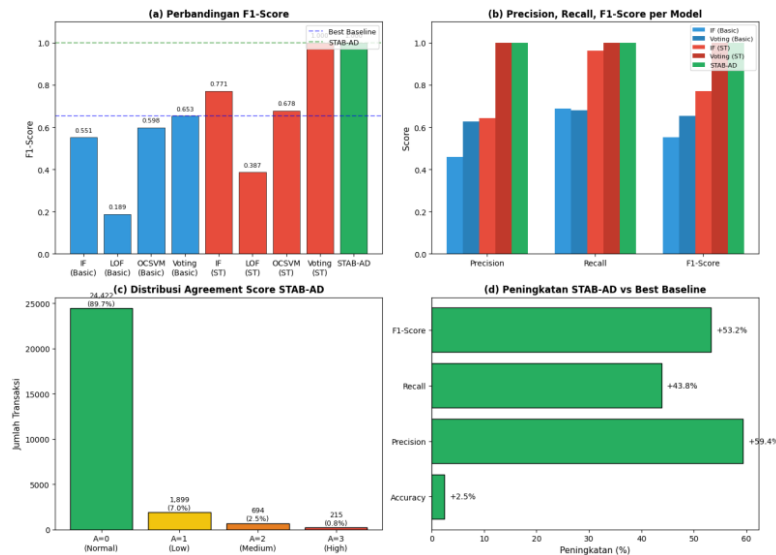
Model 1	Model 2	Cohen's Kappa	% Agreement
<i>Isolation Forest</i>	LOF	0,1993	92,39
<i>Isolation Forest</i>	<i>One-Class SVM</i>	0,6061	96,46
LOF	<i>One-Class SVM</i>	0,1221	92,10

Hasil *inter-rater agreement* pada Tabel 6 menunjukkan tingkat kesepakatan moderat hingga tinggi antar model, yang mendukung penggunaan pendekatan multi-model.

Sebagai pembandingan, dilakukan evaluasi terhadap beberapa *baseline*, meliputi model tunggal, model dengan penambahan fitur *spatio-temporal*, serta ensemble berbasis *voting*.

3.5 Baseline Comparison

Pada tahap ini dilakukan perbandingan performa antara model *baseline* dan framework STAB-AD untuk mengevaluasi kontribusi integrasi fitur *spatio-temporal* dan mekanisme *agreement-based*. Perbandingan dilakukan menggunakan metrik *precision*, *recall*, dan *F1-score* guna mengukur peningkatan kinerja deteksi anomali secara komprehensif. Visualisasi hasil perbandingan disajikan pada Gambar 4.



Gambar 4. Visualisasi Perbandingan *Baseline vs STAB-AD*

Gambar 4 membandingkan performa model *baseline* dan STAB-AD menggunakan *precision*, *recall*, dan *F1-score*. Model *baseline* tanpa fitur *spatio-temporal* menunjukkan performa terendah (*F1-score* 0,189–0,598), sedangkan penambahan fitur *spatio-temporal* meningkatkan performa hingga *F1-score* 0,771, menegaskan kontribusi fitur seperti *velocity* dan *impossible travel*.

STAB-AD menghasilkan performa terbaik pada seluruh metrik, dengan peningkatan *precision* 59,4%, *recall* 43,8%, dan *F1-score* 53,2% dibandingkan *baseline* terbaik. Peningkatan *accuracy* relatif kecil (2,5%) akibat ketidakseimbangan data. Hal ini menunjukkan peningkatan kemampuan deteksi sekaligus penurunan *false positive*.

Distribusi *agreement score* menunjukkan mayoritas transaksi berada pada kategori normal, dengan proporsi risiko tinggi sekitar 0,8%, mencerminkan karakteristik fraud sebagai *rare event*.

Secara keseluruhan, STAB-AD meningkatkan performa deteksi sekaligus menyediakan mekanisme *confidence scoring*, sehingga lebih andal dibandingkan model *baseline*.

3.6 Deteksi *Spatio-Temporal* & ATO

Bagian ini memfokuskan analisis pada pola anomali yang berkaitan dengan aspek spasial dan temporal transaksi, khususnya untuk mengidentifikasi indikasi *account takeover* (ATO). Beberapa indikator kunci seperti *impossible travel*(IT), *device change* (DC), serta transaksi bernilai tinggi dianalisis baik secara individual maupun kombinatif untuk melihat kontribusinya dalam mendeteksi risiko. Ringkasan hasil identifikasi tersebut disajikan pada Tabel 8.

Tabel 8. Hasil Deteksi *Impossible Travel* dan *Account Take Over* (ATO)

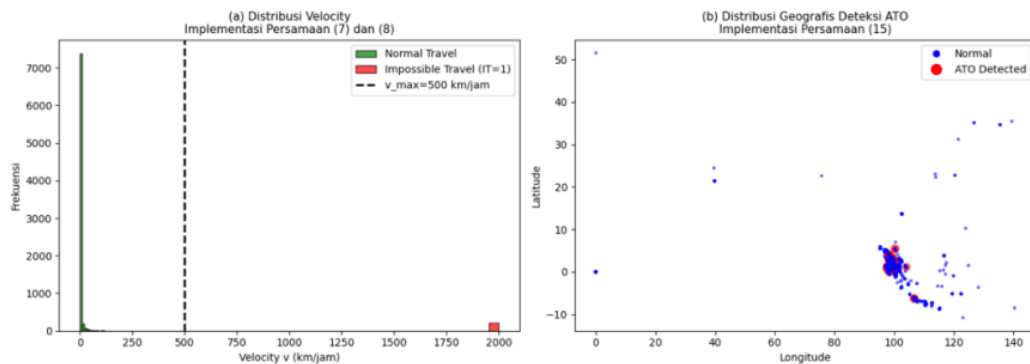
Indikator	Jumlah (n)	Persentase (%)
Transaksi dengan $IT = 1 (v > v_{max})$	244	0,896
Transaksi dengan $DC = 1$ (Device Change)	18.592	68,28
Transaksi dengan $trx_value > Q95$	1.266	4,65
ATO ($DC = 1 \wedge IT = 1 \wedge High Value$)	0	0,000
Alternatif ($IT = 1 \wedge High Value$)	144	0,5288

Hasil pada Tabel 8 menunjukkan bahwa *impossible travel* (IT) memiliki proporsi rendah (0,896%) namun bernilai diagnostik tinggi dalam mengidentifikasi anomali berbasis pergerakan geografis. Sebaliknya, *device change* (DC) memiliki frekuensi tinggi (68,28%) sehingga kurang diskriminatif jika digunakan secara tunggal. Transaksi bernilai tinggi ($>Q95$) sebesar 4,65% mengindikasikan adanya segmen berisiko.

Tidak ditemukan transaksi yang memenuhi seluruh kriteria ATO ($DC \wedge IT \wedge High Value$), menunjukkan bahwa kombinasi tersebut sangat ketat (*high specificity*) dan berpotensi menurunkan sensitivitas. Alternatif kombinasi ($IT \wedge High Value$) menghasilkan 144 transaksi (0,5288%) yang dapat dikategorikan sebagai kandidat anomali berisiko tinggi, sehingga lebih efektif dalam mempersempit ruang deteksi tanpa kehilangan kandidat relevan.

Temuan ini menegaskan bahwa deteksi ATO memerlukan pendekatan multi-indikator dengan keseimbangan sensitivitas dan spesifisitas. Oleh karena itu, hasil deteksi perlu dilengkapi dengan analisis lanjutan atau investigasi manual.

Threshold Q95 digunakan untuk merepresentasikan nilai ekstrem, sementara batas kecepatan maksimum didasarkan pada mobilitas realistis. Penggunaan *agreement score* ≥ 2 bertujuan meningkatkan reliabilitas dengan mengurangi bias model tunggal.



Gambar 5. ATO Detection

Gambar 5 (a) menunjukkan bahwa sebagian besar transaksi memiliki *velocity* di bawah ambang batas $v_{max} = 500$ km/jam, mencerminkan pola perpindahan normal. Namun, sejumlah kecil transaksi dengan *velocity* tinggi dikategorikan sebagai *impossible travel*, yang mengindikasikan potensi anomali berbasis lokasi. Gambar 5 (b) menunjukkan bahwa transaksi normal tersebar luas, sedangkan kandidat ATO membentuk kluster pada area tertentu, yang mengindikasikan konsentrasi aktivitas mencurigakan. Secara keseluruhan, fitur *spatio-temporal*, khususnya *velocity* dan lokasi, efektif dalam mengidentifikasi perpindahan tidak wajar dan memperkuat deteksi potensi ATO.

3.7 Uji Signifikansi Fitur

Untuk memastikan bahwa fitur yang digunakan mampu membedakan tingkat risiko transaksi, dilakukan pengujian statistik menggunakan uji Kruskal-Wallis. Pengujian ini bertujuan melihat apakah terdapat perbedaan distribusi fitur antar kelompok risiko yang terbentuk dari hasil klasifikasi sebelumnya. Ringkasan hasil pengujian disajikan pada Tabel 9.

Tabel 9. Uji *Kruskal-Wallis* Antar Tingkat Risiko

Fitur	H-statistic	p-value	Signifikan
<i>trx_value</i>	437,0252	$2,11 \times 10^{-94}$	Ya
<i>time_gap</i>	4179,5051	0,0000	Ya

<i>freq</i>	4251,4151	0,0000	Ya
<i>deviasi</i>	2406,0261	0,0000	Ya
<i>velocity</i>	3982,3582	0,0000	Ya

Uji *Kruskal–Wallis* digunakan untuk menguji perbedaan distribusi fitur antar kelompok risiko tanpa asumsi normalitas. Hasil pada Tabel 9 menunjukkan seluruh fitur signifikan ($p\text{-value} < 0,05$), sehingga terdapat perbedaan distribusi yang bermakna antar kelompok.

Nilai *H-statistic* yang tinggi pada fitur *freq*, *time_gap*, dan *velocity* menunjukkan kontribusi kuat dalam membedakan pola transaksi, khususnya pada aspek temporal dan perilaku. Sementara itu, *trx_value* dan deviasi juga signifikan, menegaskan pentingnya karakteristik nilai transaksi.

Secara keseluruhan, hasil ini mengonfirmasi bahwa fitur *behavioral* dan *spatio-temporal* memiliki daya diskriminatif tinggi dalam mendukung deteksi anomali.

3.8 Pembahasan Temuan

Bagian ini merangkum hasil utama yang diperoleh dari seluruh tahapan analisis, mulai dari deteksi anomali hingga klasifikasi risiko berbasis *agreement*. Ringkasan ini memberikan gambaran menyeluruh mengenai distribusi anomali, tingkat risiko, serta karakteristik transaksi yang teridentifikasi dalam framework STAB-AD. Detailnya disajikan pada Tabel 10.

Tabel 10. Ringkasan Temuan Utama Penelitian STAB-AD

Aspek	Nilai
<i>Total Transaksi Dianalisis</i>	27.230
<i>Anomali (≥ 1 model)</i>	2.790 (10,25%)
<i>High Risk (3 model)</i>	196 (0,72%)
<i>Impossible Travel (IT=1)</i>	244 (0,896%)
<i>Account Take Over (ATO)</i>	0 (0,000%)
<i>Rata-rata trx_value (Normal)</i>	Rp 2.522.536
<i>Rata-rata trx_value (High Risk)</i>	Rp 11.417.353.299.679

Hasil pada Tabel 10 menunjukkan bahwa dari 27.230 transaksi, sebesar 10,25% terdeteksi sebagai anomali oleh minimal satu model, sementara hanya 0,72% yang tergolong *high risk* berdasarkan konsensus tiga model. Temuan ini mengindikasikan bahwa pendekatan *agreement-based* pada STAB-AD efektif dalam melakukan *risk stratification*, dengan menyaring anomali menjadi subset yang lebih spesifik dan relevan untuk investigasi. Dibandingkan pendekatan model tunggal, mekanisme ini mampu mengurangi potensi *false positive* akibat sensitivitas berlebih dari masing-masing algoritma.

Secara empiris, hasil ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa pendekatan *ensemble* atau *multi-model* dapat meningkatkan stabilitas dan konsistensi deteksi anomali. Namun, berbeda dengan pendekatan *ensemble* konvensional yang umumnya berbasis *voting* atau agregasi skor, STAB-AD menambahkan dimensi *agreement* sebagai indikator kepercayaan (*confidence*), sehingga tidak hanya meningkatkan performa deteksi, tetapi juga memberikan interpretabilitas dalam klasifikasi risiko.

Fitur *spatio-temporal*, khususnya *velocity* dan *impossible travel*, terbukti memiliki peran signifikan dalam mengidentifikasi pola anomali berbasis pergerakan geografis. Hal ini konsisten dengan studi sebelumnya yang menekankan pentingnya analisis mobilitas dalam deteksi *fraud*, terutama pada skenario ATO. Nilai maksimum jarak dan kecepatan yang tinggi menunjukkan adanya pola perpindahan yang tidak realistis, yang sulit ditangkap oleh fitur berbasis nilai transaksi saja.

Namun demikian, tidak ditemukannya transaksi yang memenuhi seluruh kriteria ATO ($DC \wedge IT \wedge High Value$) menunjukkan bahwa pendekatan *rule-based* yang digunakan memiliki spesifisitas tinggi tetapi sensitivitas rendah. Hal ini mengindikasikan bahwa kombinasi indikator yang terlalu ketat berpotensi mengabaikan kasus *fraud* yang lebih subtil. Oleh karena itu, pendekatan alternatif seperti $(IT \wedge High Value)$ yang menghasilkan 0,5288% kandidat anomali dapat menjadi kompromi yang lebih seimbang antara sensitivitas dan spesifisitas.

Dari sisi evaluasi, penggunaan *pseudo ground truth* berbasis *agreement* ($A \geq 2$) memberikan indikasi awal performa model, namun memiliki keterbatasan karena bersifat *self-referential*. Nilai metrik seperti *precision*, *recall*, dan *F1-score* dalam konteks ini lebih merepresentasikan konsistensi antar model dibandingkan akurasi terhadap kondisi nyata. Hal ini berpotensi menghasilkan bias optimistik, sehingga interpretasi hasil perlu dilakukan secara hati-hati.

Selain itu, penggunaan *threshold* berbasis kuantil (Q95) dan batas kecepatan tetap (v_{max}) masih bergantung pada asumsi statistik dan domain *knowledge*, yang mungkin tidak sepenuhnya merepresentasikan variasi perilaku pengguna secara dinamis. Oleh karena itu, pendekatan yang lebih *data-driven*, seperti *adaptive thresholding* atau pembelajaran berbasis distribusi dinamis, diperlukan untuk meningkatkan generalisasi model.

Secara keseluruhan, STAB-AD menunjukkan kemampuan yang baik dalam mengintegrasikan fitur *behavioral* dan *spatio-temporal* serta meningkatkan reliabilitas deteksi melalui mekanisme *agreement-based*. Namun, validasi lebih lanjut menggunakan data berlabel atau *expert judgment* tetap diperlukan untuk memastikan

bahwa peningkatan performa yang diperoleh benar-benar mencerminkan kemampuan deteksi fraud pada kondisi nyata.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa metode STAB-AD mampu mendeteksi anomali transaksi mobile banking melalui integrasi multi-model (*Isolation Forest*, *Local Outlier Factor*, dan *One-Class SVM*), sehingga dapat menangkap berbagai karakteristik penyimpangan data secara lebih komprehensif dibandingkan pendekatan model tunggal. Penerapan mekanisme *agreement-based* memungkinkan klasifikasi risiko secara terstruktur (Normal, Low, Medium, High), yang tidak hanya meningkatkan reliabilitas deteksi tetapi juga memberikan tingkat kepercayaan (*confidence*) terhadap hasil yang diperoleh. Secara ilmiah, kontribusi utama penelitian ini terletak pada pengembangan kerangka deteksi anomali yang mengintegrasikan fitur behavioral dan spatio-temporal dengan mekanisme *agreement-based* sebagai indikator konsistensi antar model. Pendekatan ini memperluas metode *unsupervised anomaly detection* konvensional yang umumnya hanya berfokus pada skor anomali tanpa mempertimbangkan reliabilitas hasil deteksi. Dari sisi praktis, hasil penelitian menunjukkan bahwa fitur spatio-temporal, khususnya *velocity* dan *impossible travel*, efektif dalam mengidentifikasi pola perpindahan tidak wajar, sehingga dapat dimanfaatkan oleh institusi perbankan untuk meningkatkan sistem monitoring fraud secara real-time. Selain itu, skema *risk stratification* yang dihasilkan memungkinkan prioritas investigasi terhadap transaksi berisiko tinggi, sehingga meningkatkan efisiensi operasional dalam penanganan fraud. Namun demikian, penelitian ini memiliki beberapa keterbatasan. Evaluasi model masih menggunakan *pseudo ground truth* berbasis *agreement*, sehingga berpotensi menghasilkan bias optimistik. Selain itu, penggunaan threshold berbasis kuantil dan aturan tetap pada deteksi ATO masih bergantung pada asumsi statistik dan domain knowledge, yang mungkin belum sepenuhnya merepresentasikan kondisi nyata. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan pendekatan yang lebih *data-driven*, seperti adaptive thresholding atau semi-supervised learning, serta melakukan validasi menggunakan data berlabel atau *expert judgment* guna meningkatkan akurasi dan generalisasi model dalam skenario fraud yang lebih kompleks.

REFERENCES

- [1] S. D. Safira, D. Ernawati, and M. Iqbal, "Penerapan Technology Acceptance Model dalam Meningkatkan Minat Menggunakan Kembali M-Banking Livin by Mandiri," *JURNAL AKUNTANSI, EKONOMI dan MANAJEMEN BISNIS*, vol. 11, no. 1, 2023, doi: 10.30871/jaemb.v11i1.3961.
- [2] J. Hartono, K. Brian, E. Ferdian, J. Valentino, V. Felix, and N. Linawati, "Dampak Pandemi Covid-19 terhadap Percepatan Transformasi Digital di Sektor Perbankan Indonesia," *Jurnal Ekonomi, Manajemen, Akuntansi dan Keuangan*, vol. 6, no. 3, 2025, doi: 10.53697/emak.v6i3.2459.
- [3] N. Sari and A. Zaerofi, "Pengaruh Kemudahan Penggunaan, Kecepatan Transaksi, dan Keamanan Transaksi Terhadap Kepuasan Konsumen dalam Melakukan Pembayaran Menggunakan QRIS," *Empiricism Journal*, vol. 6, no. 2, 2025.
- [4] R. Setiawan and Rahmadsyah, "Digitalisasi Perbankan dan Ancaman Keamanan Siber: Tantangan dan Strategi Mitigasi Risiko Operasional," *ASEFBA: Advanced Studies in Economic, Finance and Banking*, vol. 1, 2025.
- [5] R Venkata Krishna, Nakka Maisaiah, Shaik Afshan Rehaan, Sara Ismath Alvi, Sofiya Ayesha Siddiqua, and Umamah shoukatullah, "Solutions for Suspicious Transactions and Unauthorized Debits in Credit Card Disputes," *International Research Journal on Advanced Engineering Hub (IRJAEH)*, vol. 3, no. 03, 2025, doi: 10.47392/irjaeh.2025.0153.
- [6] Chandrasekhar Anuganti, "AI-Driven Real-Time Fraud Detection in Digital Banking Using Explainable Graph Neural Networks and Behavioral Biometrics," *Int. J. Sci. Res. Sci. Eng. Technol.*, vol. 11, no. 6, 2024, doi: 10.32628/ijsrset25121181.
- [7] S. Japit, Y. Risyani, T. Selamat, C. Bombongan, and Y. Yuliana, "Deteksi Anomali Transaksi E-Commerce Menggunakan Support Vector Machine Berbasis Data Mining," *Jurnal Minfo Polgan*, vol. 13, no. 2, 2024, doi: 10.33395/jmp.v13i2.14325.
- [8] H. Mardiansyah, S. Suwilo, E. B. Nababan, and S. Efendi, "The role of Louvain-coloring clustering in the detection of fraud transactions," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 1, p. 608, Feb. 2024, doi: 10.11591/ijece.v14i1.pp608-616.
- [9] M. S. Novelan and S. Aryza, "OPTIMIZATION CVRP WITH MACHINE LEARNING FOR IMPROVED CLASSIFICATION OF IMBALANCED DATA FOOD DISTRIBUTION," *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, vol. 10, no. 4, pp. 917–925, Jun. 2025, doi: 10.33480/jitk.v10i4.6467.
- [10] Zulham Sitorus, Eko Hariyanto, and Fahmi Kurniawan, "Analysis of Artificial Intelligence Machine Learning Technology for Mapping and Predicting Flood Locations in Pahlawan Batu Bara Village," *International Journal Of Computer Sciences and Mathematics Engineering*, vol. 2, no. 2, 2023, doi: 10.61306/ijecom.v2i2.54.
- [11] M. Rasyid, Z. Sitorus, R. Farta Wijaya, and M. Iqbal, "MACHINE LEARNING ANALYSIS IN IMPROVING THE EFFICIENCY OF THE STUDENT ADMISSION DECISION MAKING PROCESS NEW AT PANCA BUDI MEDAN DEVELOPMENT UNIVERSITY," *Bulletin of Engineering Science, Technology and Industry*, vol. 2, no. 3, 2024.

- [12] E. Karnavou, G. Cascavilla, G. Marcelino, and Z. Geradts, "I know you're a fraud: Uncovering illicit activity in a Greek bank transactions with unsupervised learning," *Expert Syst. Appl.*, vol. 288, 2025, doi: 10.1016/j.eswa.2025.128148.
- [13] H. Abbassi, S. E L Mendili, and Y. Gahi, "Digital banking fortification: a real-time isolation forest architecture for detecting online transaction fraud," *Engineering Research Express*, vol. 6, no. 2, 2024, doi: 10.1088/2631-8695/ad4958.
- [14] F. Chettiar, "Integrating Autoencoders with Local Outlier Factor and Isolation Forest for Effective Fraud Detection in Imbalanced Datasets," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 12, no. 10, 2024, doi: 10.22214/ijraset.2024.64502.
- [15] Z. Hamid, F. Khalique, S. Mahmood, A. Daud, A. Bukhari, and B. Alshemaimri, "Healthcare insurance fraud detection using data mining," *BMC Med. Inform. Decis. Mak.*, vol. 24, no. 1, 2024, doi: 10.1186/s12911-024-02512-4.
- [16] N. Mardiah, L. Marlina, K. Khairul, Z. Sitorus, and M. Iqbal, "Analysis Of Indonesian People's Sentiment Towards 2024 Presidential Candidates On Social Media Using Naive Bayes Classifier and Support Vector Machine," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 2, 2024, doi: 10.47065/bits.v6i2.5766.
- [17] M. I. Burhan, A. N. Ali, A. I. Auliyah, and M. Hading, "Comparative Analysis of Algorithms for Sensitive Outlier Protection in Privacy Preserving Data Mining," *Jurnal Tekno Kompak*, vol. 19, no. 2, 2025, doi: 10.33365/jtk.v19i2.52.
- [18] Gowtham Chilakapati, "AI-Driven Fraud Detection: Enhancing Financial Security," *International Journal of Advanced Research in Science, Communication and Technology*, 2025, doi: 10.48175/ijarsct-24603.
- [19] J. Mao, M. Zhu, Y. Sun, L. Li, and H. Zhu, "Transaction Spatio-Temporal Distribution for Permissioned Blockchain Performance Profiling," *Concurr. Comput.*, vol. 37, no. 27–28, 2025, doi: 10.1002/cpe.70316.
- [20] S. Japit, Y. Risyani, T. Selamat, C. Bombongan, and Y. Yuliana, "Deteksi Anomali Transaksi E-Commerce Menggunakan Support Vector Machine Berbasis Data Mining," *Jurnal Minfo Polgan*, vol. 13, no. 2, pp. 1976–1980, Dec. 2024, doi: 10.33395/jmp.v13i2.14325.
- [21] Andrew Harper and Miriam D. Lee, "Scalable Blockchain Fraud Detection Using Spatial-Temporal Graph Neural Networks," *Frontiers in Applied Physics and Mathematics*, vol. 2, no. 1, 2025, doi: 10.71465/fapm141.
- [22] Haryanti, Syamsuddin, and Sahrullah, "Pengaruh Good Corporate Governance dan Corporate Social Responsibility Terhadap Kinerja Keuangan Pada Perusahaan Perbankan," *Jurnal Teknologi dan Manajemen Industri Terapan*, vol. 4, no. 2, 2025, doi: 10.55826/jtmit.v4i2.597.
- [23] C. M. Parlett-Pelleriti, E. Stevens, D. Dixon, and E. J. Linstead, "Applications of Unsupervised Machine Learning in Autism Spectrum Disorder Research: a Review," 2023. doi: 10.1007/s40489-021-00299-y.
- [24] M. S. Novelan, S. Efendi, P. Sihombing, and H. Mawengkang, "VEHICLE ROUTING PROBLEM OPTIMIZATION WITH MACHINE LEARNING IN IMBALANCED CLASSIFICATION VEHICLE ROUTE DATA," *Eastern-European Journal of Enterprise Technologies*, vol. 5, no. 3(125), 2023, doi: 10.15587/1729-4061.2023.288280.
- [25] C. A. R. Dinda, "Deteksi Transaksi Penipuan pada Sektor Perbankan Menggunakan Ruled-Based Model dan Pembelajaran Mesin," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 14, no. 2, 2025.
- [26] S. Lian, Y. Xu, and C. Zhang, "Family profile mining in retailing," *Decis. Support Syst.*, vol. 118, 2019, doi: 10.1016/j.dss.2019.01.007.
- [27] J. Huang *et al.*, "GAPLG: Graph Augmented With Pseudolabels Generation for Blockchain Anomaly Transaction Detection," *IEEE Trans. Comput. Soc. Syst.*, vol. 12, no. 6, 2025, doi: 10.1109/TCSS.2025.3555658.