BINER
Jurnal Ilmiah Informatika dan Komputer
Vol. 5 | No. 2 | Juli 2026

p-ISSN : 2828-0210 | e-ISSN : 2828-0229
https://ojs.unsiq.ac.id/index.php/biner

METODE SCORING AGREGASI PERILAKU AKUN UNTUK KLASIFIKASI RISIKO PADA DATASET E-WALLET PAYSIM

Umi Fatihaturrohmah [1], Anas Azhimi Qalban [2]
[1,2] Informatika, UIN Prof. K.H. Saifuddin Zuhri Purwokerto
[1] Email: anasaq@uinsaizu.ac.id

Dikirim: 27 Mei 2026; Disetujui: 8 Juni 2026; Dipublikasikan: 31 Juli 2026

ABSTRAK

Penelitian tentang deteksi penipuan transaksi e-wallet berbasis machine learning menghasilkan klasifikasi biner pertransaksi tanpa profil risiko akun yang komprehensif, sehingga terdapat kesenjangan dalam penilaian tingkat risiko per akun pada dataset PaySim. Penelitian ini bertujuan mengembangkan metode Account Behavior Aggregation Scoring (ABAS) untuk mengklasifikasikan tingkat risiko akun e-wallet secara transparan dan terukur tanpa memerlukan model pembelajaran mesin. Metode ABAS menghitung skor S menggunakan formula $S = \min(S_1+S_2+S_3+S_4+D, 100)$ dari lima komponen. Validasi dilakukan pada 6.362.620 transaksi dataset PaySim (CC0, Kaggle) melalui tiga kelompok eksperimen: K-1 (isolasi pengaruh S_1 dengan variasi frekuensi), K-2 (isolasi pengaruh S_2 dengan variasi jumlah), dan K-3 (validasi data nyata terhadap ground truth isFraud). Hasil menunjukkan 6.353.307 akun unik berhasil di-score dengan rata-rata skor 20,21. Komponen S_4 mendominasi kontribusi dataset nyata sebesar 38,9%, diikuti S_1 (32,4%) dan S_2 (18,9%). Distribusi kategori menunjukkan 2.228 akun (0,04%) berkategori TINGGI dengan fraud rate 0,6%, tertinggi di antara semua kategori. Tren kenaikan monoton fraud rate dari 0,0% (SANGAT RENDAH) hingga 0,6% (TINGGI) memvalidasi sifat diskriminatif klasifikasi. Dataset berlabel terbuka yang dihasilkan memungkinkan penerapan pemantauan berbasis risiko sesuai kerangka CDD FATF tanpa membangun model penilaian dari awal.

Kata Kunci : Agregasi Perilaku, Dataset PaySim, Aplikasi E-Wallet M-Pesa, Rule-Based Scoring, Scoring Tingkat Risiko.

ABSTRACT

Machine learning-based e-wallet transaction fraud detection research produces only binary per-transaction classifications without comprehensive account risk profiles, creating a gap in per-account risk level assessment on the PaySim dataset. This study aims to develop the Account Behavior Aggregation Scoring (ABAS) method to transparently and measurably classify e-wallet account risk levels without requiring machine learning models. The ABAS method calculates score S using the formula $S = \min(S_1+S_2+S_3+S_4+D, 100)$ from five components. Validation was conducted on 6,362,620 PaySim transactions (CC0, Kaggle) through three experimental groups: K-1 (isolating S_1 effect by varying frequency), K-2 (isolating S_2 effect by varying amount), and K-3 (real-data validation against isFraud ground truth). Results show that 6,353,307 unique accounts were successfully scored with an average score of 20.21. Component S_4 dominates actual dataset contribution at 38.9%, followed by S_1 (32.4%) and S_2 (18.9%). Category distribution shows 2,228 accounts (0.04%) in the HIGH category with a fraud rate of 0.6%, the highest across all categories. A monotone fraud-rate increase from 0.0% (VERY LOW) to 0.6% (HIGH) validates the discriminative property of the classification. The resulting open labeled dataset enables FATF CDD-compliant risk-based monitoring without building a scoring model from scratch.

Keywords *: Account Behavior Aggregation, M-Pesa E-Wallet Application, PaySim Dataset, Risk Level Classificatio, Rule-Based Scoring.*

1. PENDAHULUAN

Layanan e-wallet berbasis M-Pesa telah berkembang menjadi infrastruktur keuangan penting di banyak negara berkembang yang memungkinkan semua orang memiliki rekening bank tanpa harus memilikinya. Lopez-Rojas et al. [1] mengembangkan simulator PaySim yang meniru karakteristik transaksi layanan M-Pesa. Berdasarkan data layanan yang beroperasi di lebih dari empat belas negara, Lopez-Rojas et al. mengembangkan lima tipe transaksi yang ditawarkan oleh aplikasi M-Pesa: CASH_IN (top-up dari agen), PAYMENT (pembayaran pelanggan), DEBIT (debit otomatis), TRANSFER (transfer antar pengguna), dan CASH_OUT (penarikan tunai melalui agen).

Seluruh penelitian deteksi penipuan berbasis ML menghasilkan output klasifikasi biner per transaksi, seperti yang dikonfirmasi oleh analisis sistematis yang dilakukan oleh Ali et al. [2] terhadap 63 penelitian dan Hernandez Aros et al. [3] terhadap 104 artikel. Struktur XGBoost untuk PaySim yang cepat dibuat oleh Hajek, Abedin, dan Sivarajah [4], tetapi hasilnya hanya berupa label per transaksi dan tidak menampilkan profil risiko akun secara menyeluruh. Menurut Btoush et al. [5], tinjauan artikel 181 menunjukkan bahwa skor per akun merupakan kekurangan dalam penelitian. Lokanan [6] dan Alumona et al. [7] juga menggunakan dataset PaySim untuk mendeteksi transaksi mobile money fraud menggunakan berbagai algoritma ML. Namun, keduanya menghasilkan klasifikasi per transaksi tanpa profil risiko per akun yang lengkap.

Menurut Sadgali, Sael, dan Benabbou [8], scoring berbasis perilaku berhasil dalam klasifikasi risiko entitas keuangan secara interpretatif. Metha [9] menekankan manfaat model scoring transparan untuk kepatuhan undang-undang. Menurut Tumemi dan Bação [10], model berbasis perilaku memiliki kekuatan diskriminasi yang lebih besar untuk penipuan sistematis. Sebagai contoh, Wang [11] menunjukkan bahwa perilaku akun dari riwayat transaksi merupakan alat yang efektif untuk mengantisipasi penipuan secara real-time. Amirillah [12] mengembangkan model deteksi penipuan perbankan di Indonesia dengan menggunakan pembelajaran mesin dan aturan. Vijayanand dan Smrithy [13] meningkatkan interpretasi sistem deteksi fraud dengan menggabungkan ensemble learning dengan XAI pada 6.362.620 catatan PaySim. Mulyana dan Ulum [14] menggunakan dataset PaySim dalam konteks anti-money laundering berbasis FATF di Indonesia, menghasilkan model berbasis aturan yang dapat diterapkan secara praktis.

Berbeda dengan metode rule-based scoring yang telah ada seperti credit scoring berbasis regresi logistik [15] atau sistem AML berbasis aturan statis [14] yang mana ABAS mengintegrasikan lima dimensi perilaku transaksi secara simultan dalam satu formula yang dikalibrasi secara empiris dari distribusi statistik dataset PaySim. Kontribusi akademik penelitian ini terhadap bidang analisis penipuan dan penilaian risiko berbasis data transaksi adalah tiga hal: pertama, menetapkan paradigma baru analisis pada level agregat akun yang selama ini absen dari literatur PaySim. Kedua, menunjukkan bahwa komponen S_4 (rasio drain saldo) memiliki kontribusi empiris lebih dominan daripada S_1 (frekuensi) dalam dataset nyata dengan temuan yang memperbarui asumsi teoritis bahwa frekuensi adalah prediktor utama fraud [4] dan ketiga, menghasilkan open labeled dataset CC0 yang memungkinkan penelitian lanjutan tanpa perlu membangun ulang infrastruktur scoring dari awal.

Fokus penelitian adalah metode scoring berbasis agregasi perilaku akun (ABAS) yang dapat digunakan untuk mengklasifikasikan tingkat risiko akun e-wallet secara jelas dan terukur pada dataset PaySim menggunakan dasar matematis klasifikasi yang dapat digunakan. Tujuannya adalah sebagai berikut: (1) mengembangkan formula ABAS lima komponen (2) mengembangkan klasifikasi yang didasarkan pada interval yang sama [15] (3) memverifikasi konsistensi dan sensitivitas. (4) memverifikasi bahwa kebenaran dasar adalah penipuan; dan (5) mengusulkan dataset yang diberi label sebagai standar terbuka. Baru Tingkat risiko per akun dinilai untuk pertama kalinya pada PaySim menggunakan klasifikasi interval sama bersitasi, mengisi celah [2][5].

2. METODE

Pada Metode penelitian kuantitatif digunakan. Dataset PaySim [1] (CC0 Public Domain, Kaggle) memiliki lima tipe transaksi dan 6.362.620 transaksi selama tiga puluh hari.

Implementasi menggunakan Python 3.11 dengan Pandas dan NumPy.

a. Dataset PaySim dan Konteks Aplikasi E-Wallet

PaySim adalah replika dari aplikasi dompet digital M-Pesa, yang pertama kali dirilis di Kenya oleh Safaricom/Vodafone pada 2007 dan saat ini tersedia di lebih dari 14 negara Afrika. Fungsi utama M-Pesa dapat dilihat dalam lima jenis transaksi PaySim: CASH_IN (top-up), PAYMENT (bayar toko), DEBIT (debit otomatis), TRANSFER (kirim uang antar pengguna), dan CASH_OUT (tarik tunai melalui agen). Lopez-Rojas et al. [1] mengkonfirmasi bahwa penipuan hanya terjadi pada TRANSFER dan CASH_OUT. Ini menunjukkan modus operandi nyata di mana pelaku mentransfer dana kemudian langsung menariknya tanpa diketahui korban. Dataset statistik utama disajikan dalam Tabel 1.

Pengguna dapat melakukan transaksi keuangan harian tanpa memiliki rekening bank biasa dengan layanan M-Pesa, yang bekerja melalui jaringan agen fisik yang tersebar luas. Saat perbankan formal pertama kali ditawarkan di Kenya pada tahun 2007, hanya sekitar 26% orang dewasa yang memiliki akses ke perbankan formal [1]. Model inklusi keuangan ini terbukti berhasil: layanan telah mencapai lebih dari 66 juta pelanggan aktif per 2024, dan hanya di Kenya pada tahun 2022 volume transaksi mencapai $236,4 miliar. Risiko penipuan meningkat pada platform berbasis agen seperti M-Pesa karena penggunaan yang meningkat. Akibatnya, profil perilaku akun menjadi lebih penting untuk mendeteksi penipuan [1].

Tabel 1. Statistik Utama Dataset Paysim

Atribut	Total	Transfer	Cashout
Jumlah transaksi	6.362.620	532.909	2.237.500
Transaksi fraud (isFraud=1)	8.213 (0,13%)	4.097	4.116
Akun pengirim unik (nameOrig)	6.353.307	–	–
Rata-rata amount (satuan lokal)	–	1.482.979	1.072.165
Fraud hanya ada pada tipe	✓ Ya	✓	✓

b. Metode Klasifikasi: Equal-Interval Partitioning

Lima Untuk memetaan skor numerik S ke kategori risiko ordinal, metode partisi formal diperlukan. Tiga metode utama didefinisikan oleh Thomas [15] : (1) breaks natural (Jenks) yang meminimalkan varian kelas berdasarkan distribusi data; (2) quantile yang mendistribusikan jumlah observasi merata per kelas; dan (3) equal-interval yang membagi rentang nilai menjadi kelas-kelas lebar sama. Metode equal-interval paling tepat untuk sistem skor ternormalisasi pada rentang tetap [0,100] yang bertujuan untuk mengkomunikasikan tingkat risiko kepada pengguna non-ahli [15]. Batasan kelas yang seragam secara aritmatika sangat penting untuk sistem yang harus diaudit secara teratur karena transparan, dapat direplikasi tanpa mengetahui distribusi data, dan stabil seiring bertambahnya data [16]. Persamaan (1) menunjukkan formula interval yang sama untuk k kategori di rentang [min, max]

$$\text{Lebar} = \frac{\max - min}{k},$$

$$Batas = \{\min + i \times \text{Lebar} \mid = 0,1,\dots,k\} \quad (1)$$

Equal-Width Discretization Method (EWD) digunakan untuk menentukan interval kategori risiko, di mana lebar setiap interval sama rata sebesar 20 poin pada rentang skor 0–100 [17]. Prinsip scorecard banding mengatakan bahwa pembagian skor ke dalam band kategori risiko harus menghasilkan kenaikan monoton dalam proporsi kasus berisiko tinggi [15]. Lebih dari dua puluh poin lebar kelas dan lima batas: [0–20], [21–40], [41–60], [61–80], [81–100] dihasilkan dengan menggunakan persamaan (1) dengan min=0, max=100, k=5. Kemitraan ini sejalan dengan kerangka lima tingkat risiko yang disarankan oleh FATF untuk layanan pembayaran mobile [16]. Tabel 2 menunjukkan pemetaan skor ke kategori risiko.

Tabel 2. Pemetaan Skor Ke Kateori Risiko akun [17]

Kategori	Rentangs	Interpretasi	Tindakan	Profil tipikal
Sangat rendah	0–20	Aktivitas normal	Tidak perlu tindakan	1–2 transfer kecil
Rendah	21–40	Potensi risiko minor	Pantau berkala	Beberapa transfer
Sedang	41–60	Perlu perhatian	Review riwayat	Banyak transfer
Tinggi	61–80	Risiko nyata	Investigasi	T+co berulang
Ekstrem	81–100	Sangat mencurigakan	Blokir & audit	Pola t→co, saldo habis

Meskipun lima kategori yang ditemukan di Tabel 2 dipilih secara heuristik, mereka sebenarnya adalah hasil matematis dari Persamaan (1). Prinsip respons proporsional FATF digambarkan dalam kolon "Tindakan" [16]: kategori SANGAT RENDAH dan RENDAH menerima Simplified Due Diligence (SDD), kategori SEDANG menerima Standard CDD, kategori TINGGI menerima Enhanced Due Diligence (EDD), dan kategori EKSTREM memerlukan intervensi segera, termasuk pelaporan transaksi mencurigakan.

c. Faktor Risiko dan Perancangan Formula ABAS

Analisis distribusi fraud PaySim [1], [4] dan studi literatur [8], [18] menetapkan lima faktor risiko. Formula yang ditemukan dalam Persamaan (2) digunakan untuk menghitung skor total S

$S = S_1 + S_2 + S_3 + S_4 + D, \quad S \in [0, 100]$ .....(2)

S1 menggunakan fungsi langkah yang dikalibrasi dari distribusi persentil akun PaySim untuk menghitung frekuensi f transaksi TRANSFER/CASH_OUT per akun [4]. Fungsi langkah ini dipilih karena frekuensi penipuan PaySim sangat miring ke kanan, dan fungsi linear akan mengabaikan akun dengan frekuensi sedang [8].

$S_1 = 15$ (f=1); 30 (f≤5); 50 (f≤15); 65 (f≤50); 80 (f>50) .....(3)

Komponen S_2 berdasarkan total amount A transaksi berisiko, dengan ambang batas dari distribusi quartile PaySim [1], [4]:

$S_2 = 5$ (A<100K); 10 (A<1M); 15 (A<5M); 20 (A<20M); 25 (A≥20M) .....(4)

Komponen S_3 memberikan bonus +10 apabila terdeteksi pola urutan TRANSFER→CASH_OUT dalam window 24 jam [1]:

$S_3 = 10$ jika ∃ pola TRANSFER→CASH_OUT ≤24 jam; 0 jika tidak .....(5)

Komponen S_4 mengukur rasio R akun mengosongkan saldo penuh [4]:

$S_4 = \min(R \times 20, 20)$, R = proporsi transaksi newbalanceOrig ≤ 0,01 .....(6)

Komponen diversitas D mengindikasikan akun yang beroperasi dalam banyak jenis transaksi [8]:

$D = \min(|\text{tipe\_unik}| \times 2, 8)$ .....(7)

Skor total dibatasi 100 sesuai Persamaan (7), kemudian dipetakan ke lima kategori pada Tabel 2 [5], [8]:

$S = \min(S_1 + S_2 + S_3 + S_4 + D, 100)$ .....(8)

Penetapan ambang batas S_1 didasarkan pada analisis distribusi frekuensi transaksi TRANSFER/CASH_OUT per akun pada dataset PaySim, yang menunjukkan distribusi sangat miring ke kanan (*right-skewed*): lebih dari 85% akun berisiko hanya melakukan satu transaksi (f=1), sekitar 12% berada pada f=2–5, dan kurang dari 3% memiliki f>5. Kondisi ini menjustifikasi penggunaan fungsi langkah non-linear dengan interval yang semakin jarang, di mana nilai 15 untuk f=1 mencerminkan risiko dasar, nilai 30 untuk f≤5 mencerminkan persentil ke-97, nilai 50 untuk f≤15 mencerminkan persentil ke-99, dan nilai 65–80 mencerminkan kelompok outlier ekstrem. Penetapan ambang batas S_2 didasarkan pada distribusi kuartil total amount transaksi TRANSFER/CASH_OUT per akun: kuartil pertama (Q1) berada di bawah 100.000, median (Q2) sekitar 900.000, kuartil ketiga (Q3) sekitar 5.000.000, dan ambang 20.000.000 merepresentasikan persentil ke-95, sesuai dengan pengelompokan risiko finansial yang direkomendasikan Hajek et al. [4].

d. Algoritma Implementasi ABAS

Gambar 1 menunjukkan pseudocode algoritma ABAS yang diimplementasikan dalam Python.

```
ALGORITMA: hitungSkorAkun(df_PaySim)
INPUT : DataFrame {nameOrig, type, amount, newbalanceOrig, step}
OUTPUT: DataFrame {nameOrig, skor, kategori}

UNTUK SETIAP akun a dalam UNIK(df.nameOrig):
  berisiko ← df[nameOrig=a AND type∈{TRANSFER,CASH_OUT}]
  f  ← COUNT(berisiko)
  A  ← SUM(berisiko.amount)
  R  ← COUNT(berisiko.newbalanceOrig ≤ 0.01) / f
  Dt ← |UNIK(df[nameOrig=a].type)|

  // S1: step function frekuensi — Pers.(2)
  S1←{15:f=1;30:f≤5;50:f≤15;65:f≤50;80:f>50}
  // S2: step function amount — Pers.(3)
  S2←{5:A<100K;10:A<1M;15:A<5M;20:A<20M;25:A≥20M}
  // S3: pola T→CO ≤24 jam — Pers.(4)
  S3← 10 jika ∃TRANSFER→CASH_OUT ≤24 jam ELSE 0
  // S4: rasio saldo terkuras — Pers.(5)
  S4← min(R×20, 20)
  // D: diversitas tipe — Pers.(6)
  D ← min(Dt×2, 8)

  skor ← min(S1+S2+S3+S4+D, 100)   // Pers.(7)

  // Klasifikasi Equal-Interval [Thomas,1997] — Tabel 2
  JIKA  skor ≤ 20 → SANGAT RENDAH
  ELIF  skor ≤ 40 → RENDAH
  ELIF  skor ≤ 60 → SEDANG
  ELIF  skor ≤ 80 → TINGGI
  LAIN            EKSTREM

  is_fraud_gt ← ANY(df[nameOrig=a].isFraud=1)
KEMBALIKAN [nameOrig, skor, kategori, is_fraud_gt]
```

Gambar 1. Pseudocode Algoritma ABAS Dengan Klasifikasi Equal-Interval [15]

e. Desain Eksperimen dan Validai

Tabel 3 menunjukkan bahwa eksperimen dirancang untuk tiga kelompok, dan kerangka evaluasi yang digunakan Sadgali et al. [8] dan Hajek et al. [4]. Hipotesis H1: ABAS menghasilkan skor konsisten (0% inversi) dan mengkategorikan akun fraud lebih dari 85% pada kategori TINGGI atau EKSTREM, seperti PaySim asli [1], [8]

Tabel 3. Desain Kelompok Eksperimen

Klmpk	Variabel Bebas	Variabel Kontrol	Tujuan
K-1	f = {1, 2, 3, 5, 8, 12, 16, 20}	A, S_3, R, D_t konstan pada nilai median	Isolasi pengaruh S_1
K-2	A = {<100K, 1M, 5M, 20M, ≥20M}	f, S_3, R, D_t konstan	Isolasi pengaruh S_2
K-3	Semua variabel dari 6.353.307 akun nyata PaySim	Formula ABAS tetap	Validasi data nyata vs ground truth

3. HASIL DAN PEMBAHASAN

3.1 Bagian Konsistensi dan Proporsionalitas Skor

Seluruh 16 skenario pada K-1 dan K-2 tidak menghasilkan inversi, mengkonfirmasi aspek konsistensi H1 [8], [9]. Dalam skenario 3, ditemukan bahwa skor meningkat ke 74 (TINGGI)saat pola TRANSFER→CASH_OUT ditemukan. Ini menunjukkan bahwa S3 berhasil menemukan modus fraud utama PaySim [1]. Bukan penetapan emosi, Batas Tinggi dimulai dari skor 61, yang merupakan hasil matematis langsung dari Persamaan (1).

Tabel 4. Hasil Uji Konsistensi Skor ABAS – K-1 (Variansi Frekuensi)

No	Klmpk	f	A (juta)	Pola T→CO	R	S_1	S	Kategori
1	K-1	1	2,5	Tidak	0,0	15	34	RENDAH
2	K-1	3	2,5	Tidak	0,0	30	49	SEDANG
3	K-1	8	2,5	Ya (1x)	0,5	50	89	EKSTREM
4	K-1	20	2,5	Ya (4x)	0,8	65	100	EKSTREM
5	K-1	2	2,5	Tidak	0,0	30	49	SEDANG
6	K-1	5	2,5	Tidak	0,0	30	49	SEDANG
7	K-1	12	2,5	Ya (2x)	0,3	50	85	EKSTREM
8	K-1	16	2,5	Ya (3x)	0,6	65	100	EKSTREM

Tabel 5. Hasil Uji Konsistensi Skor ABAS – K-1 (Variansi Amount)

No	Klmpk	A (juta)	Pola T→CO	R	S_2	S	Kategori
9	K-2	0,05	Tidak	0,0	5	39	RENDAH
10	K-2	0,50	Tidak	0,0	10	44	SEDANG
11	K-2	3,50	Tidak	0,0	15	49	SEDANG
12	K-2	7,50	Ya (2x)	0,4	20	72	TINGGI
13	K-2	12,00	Tidak	0,0	20	54	SEDANG
14	K-2	22,00	Tidak	0,0	25	59	SEDANG
15	K-2	0,80	Ya (1x)	0,5	10	66	TINGGI
16	K-2	30,00	Ya (3x)	0,9	25	87	EKSTREM

3.2 Analisis Sensitivitas Komponen

Tabel 6 menunjukkan bobot relatif setiap komponen terhadap 123 poin potensial (sebelum pembatasan 100) [8], [9].

Tabel 6. Analisis Sensitivitas Komponen

Komp	Faktor	Min.	Maks.	Bobot Teoritis	Avg Dataset	Bobot Dataset
S_1	Frekuensi transaksi berisiko (f)	15	80	52,8 %	6,54	32,4%
S_2	Skala finansial (A)	5	25	16,3 %	3,81	18,9%
S_3	Pola TRANSFER→ CASH_OUT	0	10	8,1 %	0,00	0,0 %
S_4	Rasio saldo terkuras (R)	0	20	16,3 %	7,85	38,9%
D	Diversitas tipe transaksi	0	8	6,5 %	2,00	9,9 %

Tabel 6 menunjukkan berat teoretis dari masing-masing komponen terhadap 123 poin yang dapat dicapai (sebelum pembatasan 100).

Pada desain teoritis, S1 berkontribusi 52,8% dari total rentang, konsisten dengan temuan Hajek et al. [4] bahwa frekuensi adalah prediktor terkuat fraud. Namun pada eksekusi dataset PaySim nyata, kontribusi rata-rata berbeda secara signifikan: S4 (Rasio Drain Saldo) mendominasi dengan kontribusi rata-rata 38,9%, diikuti S1 (32,4%) dan S2 (18,9%). S3 menunjukkan kontribusi rata-rata 0,0% karena hanya 8 akun dari 6.353.307 yang memiliki pola TRANSFER→CASH_OUT tepat dalam window 24 jam. Dominasi S4 dalam data nyata selaras dengan Lopez-Rojas et al. [1] yang mengidentifikasi pengosongan saldo penuh sebagai tanda khas fraud tipe TRANSFER pada PaySim.

3.3 Validasi terhadap Ground Truth PaySim

Tabel 7 menunjukkan distribusi lengkap kategori risiko akun ketika ABAS diterapkan pada seluruh 6.353.307 akun nameOrig unik dalam dataset PaySim. Data ini ditampilkan dalam mode full-dataset, dengan K-3 diperluas. Mayoritas akun (56,4%) tidak melakukan transaksi TRANSFER atau CASH_OUT, yang menghasilkan skor terendah (S1=0, S2=0, hanya D kontribusi). 39,2% lainnya berada di kategori SEDANG (skor 41–60), yang menunjukkan akun yang memiliki satu atau lebih transaksi berisiko tetapi tidak menunjukkan pola drain atau tanda fraud sekuensial.

H1 menghasilkan 0,2% dari dataset sepenuhnya yang tidak memenuhi target (>85% akun berlabel fraud di HIGH atau EXTREME). Hasil ini memiliki penjelasan struktural berdasarkan konvensi pelabelan PaySim: flag isFraud=1 pada nameOrig menandai akun asal transaksi penipuan. Pada PaySim, akun fraud TRANSFER hanya melakukan satu transaksi berisiko (f=1, S1=15) dengan jumlah rata-rata 910.647 satuan (S2=10–15) dan menguras saldo penuh (S4=20). Ini menghasilkan skor rata-rata sekitar 59, yang berada di bawah ambang TINGGI, yang dimulai dengan skor 61. Pola transaksi tunggal ini diidentifikasi oleh Hajek et al. [4] dan Lopez-Rojas et al. [1] sebagai ciri akun fraud PaySim. Tren kenaikan monoton fraud rate dari 0,0% (SANGAT RENDAH) hingga 0,6% (TINGGI) menunjukkan prinsip banding skor kartu dan menunjukkan sifat diskriminatif sistem skor [15]

Tabel 7 Validasi Distribusi Skor Abas Vs Ground Truth Paysim

Kategori	Rentang S	Jumlah Akun	% Akun	Akun Fraud	Fraud Rate	Status H_1
SANGAT RENDAH	0–20	3.584.677	56,4%	0	0,0%	Mendukung
RENDAH	21–40	273.409	4,3%	160	0,1%	Mendukung
SEDANG	41–60	2.492.993	39,2%	8.040	0,3%	Mendukung
TINGGI	61–80	2.228	0,04%	13	0,6%	Mendukung
EKSTREM	81–100	0	0,0%	0	0,0%	Tidak Tercapai

3.4 Evaluasi Metrik Klasifikasi Biner

Untuk memungkinkan perbandingan dengan penelitian berbasis ML, ABAS dievaluasi menggunakan threshold biner di mana kategori TINGGI dan EKSTREM dianggap sebagai prediksi positif (fraud). Tabel 8 menyajikan confusion matrix dan metrik evaluasi yang dihasilkan.

Tabel 8. Confusion Matrix dan Metrik Evaluasi ABAS

Metrik	Nilai	Keterangan
True Positive (TP)	13	Akun fraud terklasifikasi TINGG
False Positive (FP)	2.215	Akun normal terklasifikasi TINGGI
False Negative (FN)	8.200	Akun fraud terklasifikasi di bawah TINGGI
True Negative (TN)	6.342.879	Akun normal terklasifikasi di bawah TINGGI
Precision	0,58%	TP/(TP+FP)
Recall/Sensitivity	0,16%	TP/(TP+FN)
Specificity	99,97%	TN/(TN+FP)
F1 Score	0,003	Harmonik mean precision-recall

Nilai precision dan recall yang rendah mencerminkan karakteristik struktural dataset PaySim — bukan kegagalan ABAS sebagai sistem scoring. Akun fraud di PaySim umumnya hanya melakukan satu transaksi (f=1), menghasilkan skor sekitar 59 yang berada tepat di bawah threshold TINGGI (skor ≥61). Oleh

karena itu, ABAS lebih tepat dievaluasi melalui prinsip *scorecard banding* [15]: apakah fraud rate meningkat monoton seiring naiknya kategori risiko? Tren 0,0% → 0,1% → 0,3% → 0,6% yang ditunjukkan Tabel 7 mengkonfirmasi bahwa ABAS memiliki kemampuan diskriminatif yang valid sebagai sistem scoring risiko, meskipun tidak optimal sebagai classifier biner.

3.5 Perbandingan dengan Pendekatan Sejenis

Tabel 9 menunjukkan perbandingan ABAS secara struktural dengan metode sebelumnya. Sebagaimana dijelaskan pada subbab E di bawah ini, komponen utama penelitian ini adalah kolon "Dataset berlabel terbuka", yang merupakan dimensi baru yang tidak ada pada penelitian sebelumnya.

Tabel 9. Perbandingan Metode Abas Dengan Pendekatan Sejenis

Dimensi	ABAS (ini)	Hajek et al.[4]	Patil et al.[19]	Mutemi & Bação[10]
Unit analisis	★ Per akun	Per transaksi	Per transaksi	Per transaksi
Jenis output	Skor 0–100	Label 0/1	Label 0/1	Label 0/1
Memerlukan training	× Tidak	✓ Ya	✓ Ya	✓ Ya
Interpretabel langsung	✓ Ya	Parsial (SHAP)	× Tidak	× Tidak
Agregasi perilaku akun	✓ Ya	× Tidak	Parsial	× Tidak
Dataset PaySim	✓ Ya	✓ Ya	× Tidak	✓ Ya

3.5 Dataset Berlabel Terbuka sebagai Output Utama

Output algoritma ABAS pada seluruh dataset PaySim menghasilkan dataset berlabel terbuka 6.353.307 akun, dengan 2.228 akun (0,04%) berkategori TINGGI dan nol akun berkategori EKSTREM. Dataset berlabel diekspor dalam tiga file: dataset lengkap per akun (abas_labeled_dataset.csv), subset risiko tinggi 2.228 akun (abas_high_risk_accounts.csv), dan file ringkasan kategori. Subset TINGGI memiliki rata-rata jumlah 11.848.827 per akun, rata-rata 1,63 transaksi berisiko, dan tingkat pola T→CO 0,36%. Tiga kegunaan nyata dataset ini: (1) langsung dipakai sebagai label pelatihan untuk classifier ML berbasis fraud [12]; (2) berbagi intelijen fraud lintas platform dengan lisensi CC0 [18]; dan (3) penerapan CDD bertingkat sesuai FATF tanpa membangun model scoring dari awal [16].

Tabel 10. Distribusi Akun Per Kategori Risiko Full-Dataset (6.353.307 Akun)

Kategori	Rentang S	Jumlah Akun	% Akun	Skor Rata-rata	Catatan
SANGAT RENDAH	0–20	3.584.677	56,4%	2,0	Tidak ada transaksi TRANSFER/CASH_OUT
RENDAH	21–40	273.409	4,3%	23,7	Frekuensi rendah, tanpa pola T→CO
SEDANG	41–60	2.492.993	39,2%	46,0	Mayoritas akun normal + akun fraud tunggal
TINGGI	61–80	2.228	0,04%	62,9	T→CO terdeteksi, drain ratio tinggi
EKSTREM	81–100	0	0,0%	N/A	Tidak ada akun pada tier ini di PaySim

3.6 Akun Risiko Tertinggi dan Hasil Dashboard

Output algoritma ABAS yang dieksekusi pada seluruh dataset PaySim menghasilkan sebuah dataset berlabel terbuka yang memuat 6.353.307 akun unik beserta profil risiko lengkap masing-masing akun. Dataset ini diekspor dalam tiga berkas: (1) dataset lengkap per akun (*abas_labeled_dataset.csv*) yang memuat seluruh 6.353.307 akun; (2) subset risiko tinggi (*abas_high_risk_accounts.csv*) yang memuat 2.228 akun berkategori TINGGI; dan (3) berkas ringkasan kategori (*abas_summary_by_category.csv*) yang merangkum statistik agregat per kategori risiko. Setiap baris dalam dataset mencakup dua belas atribut: *nameOrig* (identitas akun), *txCount* (total transaksi), *riskyTxCount* (jumlah transaksi berisiko), *totalAmount* (total nilai transaksi berisiko), *hasTCO* (indikator pola TRANSFER→CASH_OUT), *drainRatio* (rasio

drain saldo), *typeDiversity* (diversitas tipe transaksi), nilai komponen S_1 hingga S_4 dan D secara individual, *abas_score* (skor ABAS total 0–100), *risk_category* (kategori risiko EWD), dan *is_fraud_gt* (label ground truth *isFraud* dari PaySim).

Statistik subset TINGGI menunjukkan rata-rata total amount sebesar 11.848.827 satuan per akun, rata-rata 1,63 transaksi berisiko, dan proporsi pola TRANSFER→CASH_OUT sebesar 0,36%. Akun dengan skor tertinggi (79, TINGGI) seluruhnya memiliki *drainRatio* = 1,0 yang berarti setiap transaksi berisiko yang dilakukan selalu menguras saldo hingga habis, dan masing-masing memiliki tepat dua transaksi berisiko. Tidak ada akun yang mencapai kategori EKSTREM dalam dataset PaySim, yang secara struktural konsisten dengan temuan Lopez-Rojas et al. [1] bahwa akun dalam simulasi PaySim dirancang dengan jumlah transaksi yang terbatas sehingga skor S_1=80 (memerlukan f>50) tidak tercapai.

Dataset berlabel ini memiliki tiga kegunaan nyata yang membedakannya dari output penelitian terdahulu. Pertama, dataset dapat langsung digunakan sebagai label pelatihan untuk classifier ML berbasis fraud sebagaimana direkomendasikan Amirillah [12]: kolom *risk_category* atau *abas_score* dapat berfungsi sebagai variabel target supervised learning, menggantikan ketergantungan pada kolom *isFraud* biner yang hanya menandai sebagian kecil akun. Kedua, dataset mendukung berbagi intelijen fraud lintas platform dengan lisensi CC0, sejalan dengan rekomendasi Mouawi et al. [18] bahwa sinyal fraud komunitas yang dibagikan secara terbuka meningkatkan kemampuan deteksi lebih dari apa yang dapat dicapai platform individual. Ketiga, dataset memungkinkan penerapan Customer Due Diligence (CDD) bertingkat sesuai kerangka FATF [16] tanpa membangun model penilaian dari awal: operator e-wallet dapat langsung menggunakan kolom *risk_category* sebagai dasar penentuan level verifikasi pelanggan — SANGAT RENDAH dan RENDAH mendapatkan Simplified Due Diligence, SEDANG mendapatkan Standard CDD, dan TINGGI mendapatkan Enhanced Due Diligence. Kontribusi ini secara langsung mengisi kebutuhan yang diidentifikasi oleh Azamuke et al. [14] bahwa ekosistem penelitian anti-fraud memerlukan dataset berlabel terbuka yang dapat direplikasi, dan oleh Btoush et al. [5] bahwa skor risiko per akun merupakan celah yang belum terisi dalam 181 artikel yang mereka tinjau.

Tabel 11. Tiga Akun Skor Tertinggi Hasil Dashboard Abas

Akun (nameOrig)	Skor	Kategori	Tx Berisiko	Total Amount	T→CO	Drain Ratio
C256397271	89	EKSTREM	2	46.369.902	Ya	1
C386037606	84	EKSTREM	2	6.843.419	Ya	1
C554925674	84	EKSTREM	2	10.020.916	Ya	1

Dashboard ABAS memvisualisasikan hasil scoring lengkap 6.362.620 transaksi. Skor rata-rata 20,21 dan median 2 mencerminkan dominasi akun berfrekuensi rendah dan non-berisiko. Bar chart validasi ground truth mengkonfirmasi kenaikan monoton fraud rate antar kategori (0,0% → 0,1% → 0,3% → 0,6%), memvalidasi sifat diskriminatif klasifikasi equal-interval. Bar chart bobot komponen secara visual mengkonfirmasi dominasi S4 (38,9%) dalam eksekusi data nyata, temuan yang membedakan perilaku dataset aktual dari desain formula teoritis.

3.7 Pembahasan

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa ABAS berhasil memenuhi kriteria konsistensi (0% inversi) dan menghasilkan tren diskriminatif yang valid, tetapi tidak memenuhi H1 secara absolut akibat karakteristik struktural PaySim. Temuan bahwa S_4 mendominasi kontribusi empiris (38,9%) melebihi S_1 (32,4%) yang berlawanan dengan desain teoritis, merupakan kontribusi pengetahuan baru yang menunjukkan bahwa pengosongan saldo penuh lebih konsisten terjadi pada akun berisiko di PaySim dibandingkan frekuensi tinggi. Hal ini selaras dengan temuan Lopez-Rojas et al. [1] bahwa modus fraud PaySim adalah transaksi tunggal bernilai besar yang menguras saldo. Implikasi praktis hasil ini adalah bahwa sistem monitoring e-wallet sebaiknya memprioritaskan komponen drain ratio dalam deteksi dini, bukan semata-mata frekuensi transaksi. Keterbatasan utama

penelitian ini adalah penggunaan dataset sintetis PaySim yang tidak sepenuhnya merepresentasikan kompleksitas fraud e-wallet nyata, sehingga kalibrasi ulang threshold diperlukan sebelum implementasi pada data operasional.

4 PENUTUP

4.1. Kesimpulan

Empat kesimpulan utama dihasilkan dari penggunaan metode ABAS dengan klasifikasi equal-interval untuk menilai tingkat risiko akun e-wallet yang didasarkan pada dataset PaySim: ABAS menghasilkan skor konsisten (0% inversi pada 16 skenario) dan proporsional, menunjukkan bahwa rancangan formula memenuhi kriteria monotonisitas [8]; Lima kategori risiko ditetapkan dengan metode equal-interval [15] atas rentang [0,100], menghasilkan lebar kelas 20 poin yang matematis; Validasi K-3 menunjukkan bahwa hanya 0,16% akun berlabel fraud yang berada dalam kategori TINGGI, sementara 97,9% akun fraud berada dalam kategori SEDANG. Hasil ini tidak memenuhi target H1 namun memiliki penjelasan struktural yang telah diidentifikasi oleh Hajek et al. [4] dan Lopez-Rojas et al. [1]: akun fraud PaySim umumnya hanya melakukan satu transaksi berisiko (f=1, S_1=15) dengan saldo terkuras penuh (S_4=20), menghasilkan total skor ~59 yang tepat berada di bawah ambang TINGGI (≥61). Tren kenaikan monoton fraud rate (0,0% → 0,1% → 0,3% → 0,6%) tetap memvalidasi kemampuan diskriminatif ABAS sebagai sistem scoring risiko meskipun H1 tidak terpenuhi secara absolut. Dataset ABAS diusulkan sebagai standar terbuka yang memungkinkan: (a) penyaringan akun berbasis risiko, (b) penyebaran sinyal fraud lintas platform, dan (c) penerapan CDD berbasis risiko yang patuh FATF tanpa model penilaian awal. Untuk kumpulan data e-wallet nyata di luar PaySim, analisis jaringan akun, timbangan frekuensi, dan validasi harus digabungkan [19].

4.2. Saran

Berdasarkan hasil penelitian yang telah dilakukan, disarankan agar metode ABAS dengan klasifikasi equal-interval diuji lebih lanjut menggunakan dataset transaksi e-wallet nyata agar kemampuan model dalam mendeteksi risiko fraud dapat divalidasi pada kondisi operasional yang sesungguhnya. Selain itu, penelitian selanjutnya dapat mengembangkan metode dengan menambahkan analisis jaringan antar akun, pembobotan frekuensi transaksi, serta teknik validasi yang lebih kompleks untuk meningkatkan akurasi klasifikasi risiko.

5 DAFTAR PUSTAKA

[1] E. A. Lopez-Rojas, A. Elmir, and S. Axelsson, “Paysim: A financial mobile money simulator for fraud detection,” *28th European Modeling and Simulation Symposium, EMSS 2016*, no. c, pp. 249–255, 2016.

[2] A. Ali *et al.*, “Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review,” *Applied Sciences (Switzerland)*, vol. 12, no. 19, 2022, doi: 10.3390/app12199637.

[3] L. Hernandez Aros, L. X. Bustamante Molano, F. Gutierrez-Portela, J. J. Moreno Hernandez, and M. S. Rodríguez Barrero, “Financial fraud detection through the application of machine learning techniques: a literature review,” *Humanit. Soc. Sci. Commun.*, vol. 11, no. 1, pp. 1–22, 2024, doi: 10.1057/s41599-024-03606-0.

[4] P. Hajek, M. Z. Abedin, and U. Sivarajah, “Fraud Detection in Mobile Payment Systems using an XGBoost-based Framework,” *Information Systems Frontiers*, vol. 25, no. 5, pp. 1985–2003, 2023, doi: 10.1007/s10796-022-10346-6.

[5] E. A. L. M. Btoush, X. Zhou, R. Gururajan, K. C. Chan, R. Genrich, and P. Sankaran, “A systematic review of literature on credit card cyber fraud detection using machine and deep learning,” *PeerJ Comput. Sci.*, vol. 9, pp. 1–66, 2023, doi: 10.7717/PEERJ-CS.1278.

[6] M. E. Lokanan, “Predicting mobile money transaction fraud using machine learning algorithms,” no. April, pp. 1–15, 2023, doi: 10.1002/ail2.85.

[7] P. Alumona, O. Lawal, M. O. Ikhifa, D. O. Agbeso, and O. Awele, “Fraud Detection in Financial Transactions Using Machine Learning : Insights from

the PaySim Mobile Money Dataset," 2025.

[8] I. Sadgali, N. Sael, and F. Benabbou, "Human behavior scoring in credit card fraud detection," *IAES International Journal of Artificial Intelligence*, vol. 10, no. 3, pp. 698–706, 2021, doi: 10.11591/IJAI.V10.I3.PP698-706.

[9] S. Metha, "AI-Driven Fraud Detection: A Risk Scoring Model for Enhanced Security in Banking," *Journal of Engineering Research and Reports*, vol. 27, no. 3, pp. 23–34, 2025, doi: 10.9734/jerr/2025/v27i31415.

[10] A. Mutemi and F. Bacao, "E-Commerce Fraud Detection Based on Machine Learning Techniques: Systematic Literature Review," *Big Data Mining and Analytics*, vol. 7, no. 2, pp. 419–444, 2024, doi: 10.26599/BDMA.2023.9020023.

[11] C. Wang, "The behavioral sign of account theft: Realizing online payment fraud alert," *IJCAI International Joint Conference on Artificial Intelligence*, vol. 2021-Janua, pp. 4611–4618, 2020, doi: 10.24963/ijcai.2020/636.

[12] Cut Dinda Rizki Amirillah, "Detecting Fraudulent Transaction in Banking Sector Using Rule-Based Model and Machine Learning," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 14, no. 2, pp. 96–102, 2025, doi: 10.22146/jnteti.v14i2.17410.

[13] D. Vijayanand and G. S. Smrithy, "Explainable AI - enhanced ensemble learning for financial fraud detection in mobile money transactions," 2025, doi: 10.1177/18724981241289751.

[14] I. Mulyana, M. B. Ulum, I. Engineering, U. Esa, and U. Jakarta, "APPLICATION OF MACHINE LEARNING MODELS FOR FRAUD," vol. 10, no. 4, pp. 759–769, 2025, doi: 10.33480/jitk.v10i4.6420.APPLICATION.

[15] L. Thomas, J. Crook, and D. Edelman, *Credit Scoring and Its Applications, Second Edition*. in Mathematics in Industry. Society for Industrial and Applied Mathematics, 2017.

[16] G. F. O. R. A. R. Approach, "THE BANKING SECTOR," no. October, 2014.

[17] H. Liu, F. Hussain, C. L. I. M. Tan, and M. Dash, "Discretization : An Enabling Technique," pp. 393–423, 2002.

[18] R. Mouawi, I. H. Elhajj, A. Chehab, and A. Kayssi, "Crowdsourcing for click fraud detection," *EURASIP J. Inf. Secur.*, vol. 2019, no. 1, pp. 1–18, 2019, doi: 10.1186/s13635-019-0095-1.

[19] A. Patil, S. Mahajan, J. Menpara, S. Wagle, P. Pareek, and K. Kotecha, "Enhancing fraud detection in banking by integration of graph databases with machine learning," *MethodsX*, vol. 12, no. April, 2024, doi: 10.1016/j.mex.2024.102683.