Techno.Com, Vol. 25 No. 2, May 2026: 367 – 382

Analisis Komparatif Model Random Forest dan XGBoost Berdasarkan Kinerja AUC Pada Fraud Detection

Comparative Analysis of Random Forest and XGBoost Models Based on AUC Performance in Fraud Detection

Muthiah As Saidah*[1], **Aggry Saputra**[2]
[1]*Program Studi Sistem Informasi, Sekolah Tinggi Teknologi Indonesia Tanjung Pinang, Tanjungpinang, Indonesia*
[2]*Program Studi Teknik Informatika, Sekolah Tinggi Teknologi Indonesia Tanjung Pinang, Tanjungpinang, Indonesia*
*E-mail : muthiahassaidah40@gmail.com *[1], aggrysaputra@gmail.com [2]*
**muthiahassaidah40@gmail.com*
Received 23 April 2026; Revised 12 May 2026; Accepted 15 May 2026

Abstrak - Fraud detection menjadi salah satu tantangan penting dalam sistem informasi modern, khususnya pada transaksi finansial dan digital. Berbagai penelitian telah menunjukkan performa tinggi model Random Forest dan XGBoost, namun sebagian besar evaluasi masih dilakukan pada dataset tertentu dan terbatas pada perbandingan deskriptif. Penelitian ini bertujuan melakukan analisis statistik komparatif lintas studi terhadap performa model Random Forest dan XGBoost berdasarkan nilai Area Under Curve (AUC) pada berbagai domain fraud detection. Penelitian menggunakan data sekunder dari 40 studi terdahulu yang terdiri dari 20 model Random Forest dan 20 model XGBoost. Analisis dilakukan menggunakan statistik deskriptif, uji asumsi, independent samples t-test, Mann-Whitney U test, dan effect size menggunakan Cohen's d. Hasil penelitian menunjukkan bahwa XGBoost memiliki rata-rata AUC yang sedikit lebih tinggi dibandingkan Random Forest. Namun, hasil uji statistik dan effect size menunjukkan bahwa perbedaan tersebut tidak signifikan secara statistik maupun praktis. Selain itu, Random Forest cenderung menunjukkan performa yang lebih stabil, sedangkan XGBoost lebih sensitif terhadap karakteristik dataset dan konfigurasi model. Penelitian ini menunjukkan bahwa evaluasi performa machine learning lintas studi tidak cukup hanya berdasarkan nilai rata-rata AUC, tetapi juga perlu mempertimbangkan signifikansi statistik, effect size, stabilitas performa, dan heterogenitas antar studi.
Kata Kunci – AUC; Fraud Detection; Machine Learning; Random Forest; XGBoost

Abstract - *Fraud detection has become a critical challenge in modern information systems, particularly in financial and digital transactions. Previous studies have reported high performance of Random Forest and XGBoost models; however, most evaluations have been conducted on specific datasets and limited to descriptive comparisons. This study aims to conduct a cross-study comparative statistical analysis of Random Forest and XGBoost performance based on Area Under Curve (AUC) values across various fraud detection domains. The study used secondary data from 40 previous studies consisting of 20 Random Forest models and 20 XGBoost models. The analysis was conducted using descriptive statistics, assumption testing, independent samples t-test, Mann-Whitney U test, and effect size analysis using Cohen's d. The results indicate that XGBoost achieved a slightly higher mean AUC than Random Forest. However, statistical testing and effect size analysis showed that the difference was not statistically or practically significant. In addition, Random Forest tended to demonstrate more stable performance, while XGBoost showed greater sensitivity to dataset characteristics and model configurations. This study highlights that cross-study evaluation of machine learning performance should not rely solely on average AUC values, but also consider statistical significance, effect size, performance stability, and heterogeneity across studies.*

Keywords - *AUC; Fraud Detection; Machine Learning; Random Forest; XGBoost*

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035

1. PENDAHULUAN

Fraud detection merupakan salah satu permasalahan krusial dalam sistem informasi modern, khususnya pada sektor keuangan, transaksi digital, dan layanan berbasis teknologi. Peningkatan volume transaksi serta kompleksitas pola kecurangan yang semakin berkembang menuntut adanya sistem deteksi yang akurat dan efisien. Salah satu tantangan utama dalam *fraud detection* adalah ketidakseimbangan data (imbalanced data), di mana jumlah data fraud sebagai kelas minoritas jauh lebih sedikit dibandingkan data non-fraud, sehingga menyebabkan model klasifikasi cenderung bias terhadap kelas mayoritas dan kesulitan dalam mendeteksi kasus fraud secara optimal.

Dalam beberapa tahun terakhir, pendekatan berbasis *machine learning* telah menjadi solusi utama dalam menangani permasalahan fraud detection. Di antara berbagai metode yang digunakan*, Random Forest* dan *XGBoost* merupakan dua *algoritma ensemble* yang paling banyak diterapkan karena kemampuannya dalam menangani data kompleks dan menghasilkan performa klasifikasi yang tinggi. Beberapa penelitian menunjukkan bahwa *Random Forest* mampu mencapai nilai *Area Under Curve* (AUC) yang tinggi, bahkan pada kisaran 0.90 hingga 0.99 pada kasus *fraud detection*, khususnya pada data transaksi kartu kredit [1]. Selain itu, penelitian lain juga menunjukkan bahwa optimasi dan tuning pada *Random Forest* dapat meningkatkan performa secara signifikan dalam mendeteksi transaksi fraud [2]. Temuan serupa juga menunjukkan bahwa *Random Forest* memiliki kemampuan yang baik dalam membedakan antara data fraud dan non-fraud [3].

Di sisi lain, *XGBoos*t dikenal sebagai algoritma berbasis *boosting* yang memiliki kemampuan optimasi yang lebih adaptif dalam memodelkan pola data yang kompleks. Beberapa penelitian melaporkan bahwa *XGBoost* mampu menghasilkan performa yang lebih unggul dibandingkan model *machine learning* lainnya, termasuk *Random Forest*, terutama dalam konteks fraud detection [4]. Selain itu, penelitian lain juga menunjukkan bahwa *XGBoost* mampu mencapai ROC-AUC sebesar 95.2% dan mengungguli model lain seperti *Random Forest* dalam kompetisi *fraud detection* serta memiliki keunggulan dalam hal *robustness* dan kemampuan generalisasi pada data yang kompleks [5].

Meskipun *Random Forest* dan *XGBoost* telah banyak digunakan dalam *fraud detection* dan sering menunjukkan performa yang tinggi, sebagian besar penelitian masih dilakukan pada dataset dan domain fraud tertentu, seperti fraud kartu kredit atau fraud transaksi finansial [6], [7]. Akibatnya, hasil performa antar penelitian menjadi sulit dibandingkan secara langsung karena adanya perbedaan karakteristik dataset, teknik *preprocessing*, penanganan data imbalance, serta strategi tuning model yang digunakan.

Selain itu, sebagian besar penelitian terdahulu hanya membandingkan nilai performa seperti AUC, akurasi, atau F1-score secara deskriptif tanpa melakukan evaluasi signifikansi statistik antar model maupun sintesis performa lintas studi. Padahal, beberapa penelitian metodologis menekankan pentingnya penggunaan pendekatan statistik yang lebih ketat dalam evaluasi model *machine learning* untuk menghindari interpretasi performa yang bias atau tidak konsisten [8].

Berdasarkan kondisi tersebut, masih terdapat kebutuhan terhadap suatu pendekatan *comparative statistical* yang mampu mengevaluasi performa *Random Forest* dan *XGBoost* secara lebih sistematis menggunakan data dari berbagai penelitian *fraud detection*. Oleh karena itu, penelitian ini tidak hanya membandingkan rata-rata nilai AUC kedua model, tetapi juga menganalisis signifikansi statistik dan stabilitas performa model lintas domain fraud.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035

2. METODE PENELITIAN

2.1. Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan desain komparatif yang bertujuan untuk membandingkan kinerja dua model *machine learning*, yaitu *Random Forest* dan *XGBoost*, dalam mendeteksi fraud berdasarkan nilai *Area Under Curve* (AUC). Pendekatan kuantitatif dipilih karena penelitian ini berfokus pada analisis data numerik dan pengujian hipotesis menggunakan metode statistik. Desain komparatif digunakan untuk mengidentifikasi perbedaan kinerja antara dua atau lebih kelompok berdasarkan variabel tertentu [9], [10], [11], [12]. Dalam konteks penelitian ini, perbandingan dilakukan terhadap nilai AUC yang dihasilkan oleh masing-masing model, sehingga dapat diketahui apakah terdapat perbedaan performa yang signifikan antara *Random Forest* dan *XGBoost*.

Penelitian ini juga menggunakan data sekunder yang diperoleh dari berbagai studi sebelumnya yang relevan dengan topik *fraud detection*. Dalam konteks umum riset, data sekunder adalah data yang sebelumnya telah dikumpulkan orang lain dan sudah melalui proses pengolahan statistik [13]. Selain itu, penelitian ini menggunakan analisis statistik inferensial untuk menguji hipotesis dan menggeneralisasikan temuan sampel ke populasi [14], [15].

Hipotesis dalam penelitian ini dirumuskan untuk menguji perbedaan kinerja antara model *Random Forest* dan *XGBoost* dalam fraud detection berdasarkan nilai *Area Under Curve* (AUC). Hipotesis yang digunakan adalah sebagai berikut:

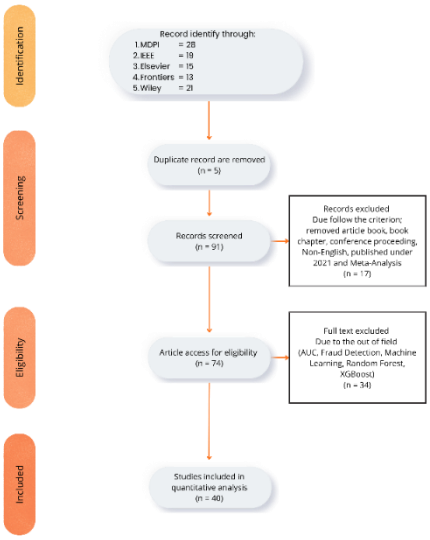
H_0 : Tidak terdapat perbedaan kinerja AUC antara *Random Forest* dan *XGBoost*

H_1 : Terdapat perbedaan kinerja AUC antara *Random Forest* dan *XGBoost*

Hipotesis tersebut digunakan sebagai dasar dalam pengujian statistik untuk menentukan apakah perbedaan performa kedua model signifikan atau tidak.

2.2. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder [13], yang diperoleh dari berbagai publikasi ilmiah yang relevan. Pengumpulan data dilakukan melalui studi literatur. Proses seleksi literatur dilakukan secara bertahap melalui identifikasi, *screening*, dan seleksi akhir studi yang memenuhi kriteria inklusi. Alur seleksi penelitian ditunjukkan melalui diagram PRISMA untuk meningkatkan transparansi proses pengumpulan data [16].



Gambar 1. Diagram Alir PRISMA

2.3. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri dari variabel dependen dan variabel independen yang digunakan untuk menganalisis dan membandingkan kinerja model dalam *fraud detection*. Variabel dependen dalam penelitian ini adalah nilai *Area Under Curve* (AUC), yang

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035

digunakan sebagai metrik evaluasi untuk mengukur performa model klasifikasi. AUC menggambarkan kemampuan model dalam membedakan antara kelas fraud dan non-fraud berdasarkan kurva *Receiver Operating Characteristic* (ROC). ROC curve merepresentasikan hubungan antara *True Positive Rate* (TPR) dan *False Positive Rate* (FPR) pada berbagai threshold klasifikasi.

Secara matematis, AUC dapat dinyatakan sebagai [17]:

$$AUC = P\,(f(x^{+}) > f(x^{-})) \tag{1}$$

Dimana, *f(x)* fungsi prediksi model, x^{+} adalah data fraud (kelas positif), dan x^{-} adalah data non-fraud (kelas negatif).

Persamaan tersebut menunjukkan bahwa AUC merupakan probabilitas bahwa model memberikan skor prediksi yang lebih tinggi pada data fraud dibandingkan data non-fraud. Nilai AUC berada pada rentang 0 hingga 1, di mana nilai yang lebih tinggi menunjukkan kemampuan model yang lebih baik dalam membedakan kedua kelas [18]. Nilai AUC mendekati 1 menunjukkan performa klasifikasi yang sangat baik, sedangkan nilai mendekati 0.5 menunjukkan bahwa model tidak lebih baik dari peluang acak dalam membedakan kelas (tidak punya kemampuan diskriminatif) sedangkan AUC bernilai 1 berarti *discriminasi* [19], [20], [21], [22].

Variabel independen dalam penelitian ini adalah jenis model *machine learning* yang digunakan, yaitu *Random Forest* dan *XGBoost*. *Random Forest* dan *XGBoost* adalah dua algoritma populer berbasis pohon keputusan (*tree-based*) yang banyak dipakai untuk klasifikasi dan regresi di berbagai bidang seperti kesehatan, geospasial, keuangan, pendidikan, hingga rekayasa teknik [23], [24], [25]. Kedua model tersebut menghasilkan fungsi prediksi $f(x)$yang berbeda berdasarkan pendekatan yang digunakan.

Pada *Random Forest*, fungsi prediksi diperoleh melalui agregasi sejumlah pohon keputusan, yang secara umum dapat dinyatakan sebagai[26]:

$$\hat{f}(x) = \frac{1}{T}\sum_{t=1}^{T} h_t(x) \tag{2}$$

di mana $h_t(x)$ merupakan fungsi prediksi dari pohon ke-t dan T adalah jumlah pohon dalam model. Sementara itu, pada *XGBoost*, fungsi prediksi dibangun secara iteratif melalui pendekatan boosting, yang dapat dinyatakan sebagai[27]:

$$\hat{y}(x) = \sum_{k=1}^{K} f_k(x) \tag{3}$$

di mana $f_k(x)$ merupakan fungsi model pada iterasi ke-k dan K adalah jumlah total pohon dalam *ensemble*. Perbedaan formulasi matematis tersebut menghasilkan karakteristik fungsi prediksi yang berbeda, yang pada akhirnya mempengaruhi nilai AUC yang dihasilkan oleh masing-masing model.

2.4. Prosedur Penelitian

Prosedur penelitian dalam studi ini dilakukan melalui beberapa tahapan yang sistematis yang dapat dilihat pada alur penelitian ditunjukkan pada gambar berikut :

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035

Gambar 2. Alur Penelitian

Berdasarkan alur penelitian diatas, tahap pertama adalah pengumpulan data dari berbagai sumber literatur yang relevan dengan topik *fraud detection*. Data yang dikumpulkan berupa nilai *Area Under Curve* (AUC) dari model *Random Forest* dan *XGBoost* yang dilaporkan dalam penelitian sebelumnya. Pada tahap ini, dilakukan pencarian literatur menggunakan kata kunci yang telah ditentukan untuk memperoleh data yang sesuai dengan tujuan penelitian.

Tahap kedua adalah seleksi data berdasarkan kriteria inklusi yang telah ditetapkan. Proses ini bertujuan untuk memastikan bahwa data yang digunakan memiliki kualitas yang baik dan relevan dengan penelitian. Selain itu, dilakukan identifikasi dan penghapusan data duplikat untuk menghindari bias dalam analisis.

Tahap ketiga adalah penyusunan dataset dalam bentuk yang terstruktur. Data yang telah dikumpulkan kemudian disusun ke dalam format dataset yang terdiri dari dua variabel utama, yaitu jenis model dan nilai AUC. Proses ini dilakukan untuk mempermudah analisis statistik yang akan dilakukan pada tahap berikutnya.

Tahap keempat adalah analisis statistik deskriptif. Pada tahap ini dilakukan perhitungan nilai rata-rata, median, standar deviasi, serta nilai minimum dan maksimum untuk memberikan gambaran awal mengenai distribusi data. Analisis ini digunakan untuk memahami karakteristik performa masing-masing model.

Tahap kelima adalah uji asumsi data yang meliputi uji normalitas dan uji homogenitas varians. Uji normalitas dilakukan untuk mengetahui apakah data berdistribusi normal, sedangkan uji homogenitas dilakukan untuk mengetahui apakah varians antar kelompok sama. Hasil dari uji asumsi ini digunakan untuk menentukan metode statistik yang tepat dalam pengujian hipotesis.

Tahap keenam adalah uji komparatif menggunakan metode statistik. Pada tahap awal, dilakukan uji independent samples t-test untuk membandingkan rata-rata nilai AUC antara kedua model. Namun, apabila asumsi normalitas tidak terpenuhi, maka digunakan uji nonparametrik *Mann-Whitney* U test sebagai metode utama dalam analisis [28], [29].

Tahap ketujuh adalah interpretasi hasil analisis. Hasil dari uji statistik kemudian dianalisis untuk menentukan apakah terdapat perbedaan yang signifikan antara kedua model. Selain itu, dilakukan pembahasan terhadap karakteristik performa masing-masing model berdasarkan hasil analisis deskriptif dan inferensial.

Secara keseluruhan, tahapan prosedur penelitian ini dirancang untuk memberikan alur analisis yang sistematis, mulai dari pengumpulan data hingga penarikan kesimpulan.

2.5 Teknik Analisis Data

Penelitian ini menggunakan pendekatan *comparative statistical* terhadap *nilai Area Under Curve* (AUC) yang diperoleh dari berbagai studi *fraud detection*. Pendekatan ini digunakan untuk mengevaluasi kecenderungan performa model *Random Forest* dan *XGBoost* secara lintas studi. Pemilihan pendekatan ini dipertimbangkan karena sebagian besar penelitian sumber hanya melaporkan metrik performa seperti AUC tanpa menyediakan informasi statistik lengkap seperti varians, confidence interval, atau ukuran sampel. Penggunaan AUC sebagai dasar perbandingan lintas studi didukung oleh penelitian sebelumnya yang menunjukkan bahwa AUC merupakan

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035

salah satu metrik evaluasi yang paling konsisten dalam membandingkan model klasifikasi pada berbagai dataset, termasuk pada kondisi data tidak seimbang (Li, 2024). Selain itu, penggunaan uji statistik pada evaluasi model *machine learning* juga direkomendasikan untuk menghindari interpretasi performa yang hanya didasarkan pada perbedaan nilai rata-rata semata (Rainio et al., 2024).

Berdasarkan pertimbangan tersebut, analisis statistik dalam penelitian ini digunakan untuk mengidentifikasi kecenderungan performa, signifikansi perbedaan, serta stabilitas performa model *Random Forest* dan *XGBoost* pada berbagai domain *fraud detection*. Selain uji signifikansi, penelitian ini juga menggunakan analisis *effect size* untuk mengukur besar perbedaan praktis antara kedua model. *Effect size* dihitung menggunakan Cohen's d pada *independent samples t-test*. Penggunaan *effect size* penting karena signifikansi statistik tidak selalu menunjukkan besarnya pengaruh praktis dari suatu perbedaan [30].

Seluruh proses analisis dilakukan menggunakan perangkat lunak Jamovi, yang merupakan aplikasi statistik berbasis *open-source* yang mendukung analisis parametrik dan nonparametrik. Tahapan analisis data dalam penelitian ini meliputi beberapa langkah sebagai berikut:

2.5.1 Statistik Deskriptif

Analisis statistik deskriptif dilakukan untuk memberikan gambaran awal mengenai karakteristik data yang digunakan. Parameter yang dihitung meliputi nilai rata-rata (mean), median, standar deviasi, serta nilai minimum dan maksimum. Statistik deskriptif berfungsi untuk mengidentifikasi pola umum dalam data, termasuk tingkat variasi dan distribusi nilai AUC pada masing-masing model. Analisis ini penting sebagai langkah awal sebelum dilakukan pengujian statistik inferensial [31].

2.5.2 Uji Normalitas

Uji normalitas dilakukan untuk mengetahui apakah data berdistribusi normal atau tidak. Dalam penelitian ini, uji normalitas dilakukan menggunakan metode *Shapiro-Wilk*, yang umum digunakan untuk ukuran sampel kecil hingga menengah [32], [33]. Apabila nilai signifikansi (p-value) kurang dari 0.05, maka data dianggap tidak berdistribusi normal [34].

2.5.3 Uji Homogenitas Varians

Uji homogenitas varians dilakukan untuk mengetahui apakah varians antara kelompok data adalah sama atau tidak [35]. Dalam penelitian ini, digunakan uji Levene untuk menguji kesamaan varians antara kelompok *Random Forest* dan *XGBoost*. Jika nilai p-value lebih besar dari 0.05, maka varians dianggap homogen [36], [37].

2.5.4 Uji Parametrik (Independent Samples T-Test)

Uji *independent samples t-test* digunakan sebagai analisis awal untuk membandingkan rata-rata nilai AUC antara dua kelompok independen, yaitu *Random Forest* dan *XGBoost*. Uji ini dilakukan dengan asumsi bahwa data berdistribusi normal dan memiliki varians yang homogen. Hipotesis yang digunakan pada *independent samples t-test* adalah [38], [39]:

H_0 : Tidak terdapat perbedaan kinerja AUC antara *Random Forest* dan *XGBoost*

H_1 : Terdapat perbedaan kinerja AUC antara *Random Forest* dan *XGBoost*

Namun, apabila asumsi normalitas tidak terpenuhi, maka hasil uji ini hanya digunakan sebagai indikasi awal dan dilanjutkan dengan uji non-parametrik [14].

2.5.5 Uji Non Parametrik (Mann-Whitney U Test)

Uji *Mann-Whitney U* adalah uji nonparametrik untuk membandingkan dua kelompok independen ketika data tidak memenuhi asumsi uji parametrik. Uji Mann-Whitney digunakan

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035

untuk membandingkan dua kelompok independen tanpa mengasumsikan distribusi normal. Uji ini didasarkan pada peringkat (*ranking*) data, bukan nilai absolut [28]. Jika nilai p-value lebih besar dari 0.05, maka tidak terdapat perbedaan yang signifikan antara kedua kelompok [40].

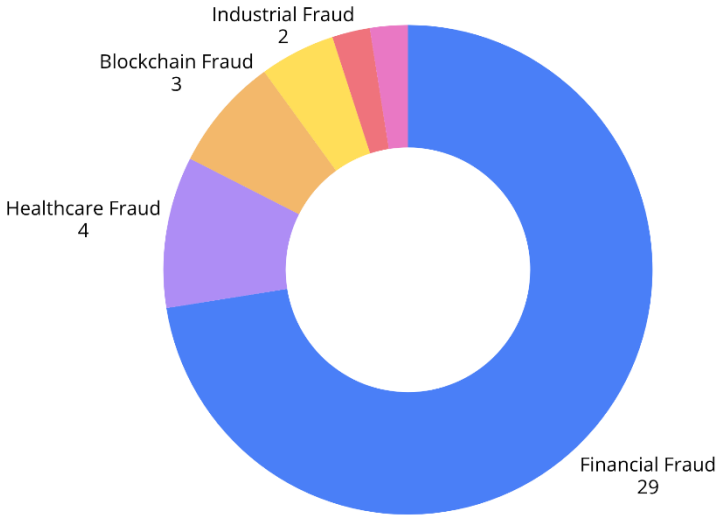
2.5.6 Interpretasi Hasil

Hasil dari seluruh pengujian statistik kemudian diinterpretasikan untuk menentukan apakah terdapat perbedaan kinerja yang signifikan antara model *Random Forest* dan *XGBoost*. Interpretasi dilakukan dengan mempertimbangkan hasil uji deskriptif, uji asumsi, serta uji komparatif yang telah dilakukan.

3. HASIL DAN PEMBAHASAN

3.1. Data Penelitian

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari berbagai penelitian terdahulu pada rentang tahun 2021-2026 terkait *fraud detection* menggunakan model *Random Forest* dan *XGBoost*. Data yang dikumpulkan berupa nilai *Area Under Curve* (AUC) dari masing-masing model.



Gambar 3. Klasifikasi Domain Fraud Data Penelitian

Jumlah data yang digunakan dalam penelitian ini sebanyak 40 data, yang terdiri dari 20 data untuk model *Random Forest* dan 20 data untuk model *XGBoost*. Data tersebut mencakup berbagai domain fraud, seperti transaksi kartu kredit, fraud laporan keuangan, transaksi *blockchain*, serta fraud pada sektor kesehatan. Berikut data untuk masing-masing model yang akan dianalisis pada penelitian ini :

Tabel 1. Data ROC-AUC Model *Random Forest* dan *XGBoost* Pada *Fraud Detection*

No	Penulis	Tahun	Domain Fraud	ROC-AUC	Model
1	Ashfaq dkk. [41]	2022	Transaksi Bitcoin/blockchain	0.92	Random Forest
2	Sagiraju dkk. [1]	2025	Credit card Eropa	0.96	Random Forest
3	Nabrawi & Alanazi [42]	2023	Klaim asuransi kesehatan Saudi Arabia	0.90	Random Forest
4	Albalawi & Dardouri [43]	2025	Credit card	0.9759	Random Forest
5	Kumar & Panwar [44]	2024	Credit card publik	0.959	Random Forest

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035

6	Hewiagh dkk [45]	2021	Fraud sistem pengelolaan sampah	0.9892	Random Forest
7	Alutaibi [46]	2025	Fraud transaksi blockchain/finansial	0.8	Random Forest
8	Ileberi dkk. [47]	2022	Credit card (Eropa)	1	Random Forest
9	Liu [48]	2021	Fraud laporan keuangan perusahaan	0.96	Random Forest
10	Korde dkk. [49]	2025	Transaksi finansial umum	0.987	Random Forest
11	Shahidullah dkk. [50]	2024	Transaksi kartu kredit (fraud finansial, U.S. finance)	0.98	Random Forest
12	Lee dkk. [51]	2025	Financial statement fraud (laporan keuangan perusahaan di Indonesia)	0.82	Random Forest
13	Balakishore dkk. [52]	2025	Fraud kartu kredit	0.94	Random Forest
14	Angelica dkk. [53]	2024	Fraud pembayaran e-commerce (online purchasing)	0.8869	Random Forest
15	Lin & Jiang [3]	2021	Fraud kartu kredit	0.962	Random Forest
16	Ahsan dkk. [54]	2022	Fraud kartu kredit	0.981	Random Forest
17	Afriyie dkk. [55]	2023	Transaksi kartu kredit (fraud)	0.989	Random Forest
18	Li [56]	2022	Fraud kartu kredit	0.97	Random Forest
19	Bounab dkk [57]	2024	Medicare fraud	0.80	Random Forest
20	Dake [58]	2023	Fraud lowongan kerja online (job recruitment fraud, Ghana)	0.963	Random Forest
21	Hájek dkk. [4]	2022	Mobile payment fraud	0.996	XGBoost
22	Tayebi dkk. [59]	2025	Kredit kartu	0.9088	XGBoost
23	Al-Hchaimi dkk. [60]	2026	Transaksi blockchain (Ethereum)	1	XGBoost
24	Zhao dkk. [61]	2022	Fraud laporan keuangan	0.725	XGBoost
25	Wei dkk. [62]	2025	Fraud laporan keuangan Tiongkok	0.9078	XGBoost
26	Xia dkk. [63]	2023	Kredit/Fraud keuangan	0.97	XGBoost
27	Hashemi dkk. [64]	2022	Fraud perbankan / transaksi kartu	0.95	XGBoost
28	Mehdary dkk. [65]	2024	Fraud smart grid / electricity theft	0.987	XGBoost
29	Abouelenein dkk. [66]	2025	Fraud klaim asuransi kesehatan	0.9187	XGBoost
30	Abdelghafour [67]	2024	Credit card fraud	0.9887	XGBoost
31	Jovanovic dkk. [68]	2022	Credit card fraud	1	XGBoost
32	Li dkk. [69]	2023	Financial statement fraud & pelanggaran emiten	0.90	XGBoost
33	Suarez dkk. [70]	2025	Bank transaction fraud	0.999506	XGBoost
34	Balayet dkk. [71]	2024	Healthcare provider / Medicare fraud	0.96	XGBoost
35	Dichev dkk. [72]	2025	Fraud risk perbankan (transaksi)	0.84	XGBoost
36	Han & Joe [73]	2025	Credit card / financial transaction fraud	0.9524	XGBoost
37	Salekshahrezaee dkk. [7]	2023	Credit card fraud	0.886	XGBoost
38	Damanik & Liu [74]	2024	Credit card fraud	0.965	XGBoost
39	Krishna dkk. [75]	2025	Credit card fraud dengan privasi & explainability	0.978	XGBoost
40	Nobel dkk. [76]	2024	Banking / mobile payment fraud	1	XGBoost

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035

3.2. Analisis Statistik Deskriptif

Analisis statistik deskriptif dilakukan untuk memberikan gambaran awal mengenai karakteristik data yang digunakan dalam penelitian ini.

Tabel 2. Statistik Deskriptif dari ROC-AUC Model *Random Forest* dan *XGBoost* Pada *Fraud Detection*

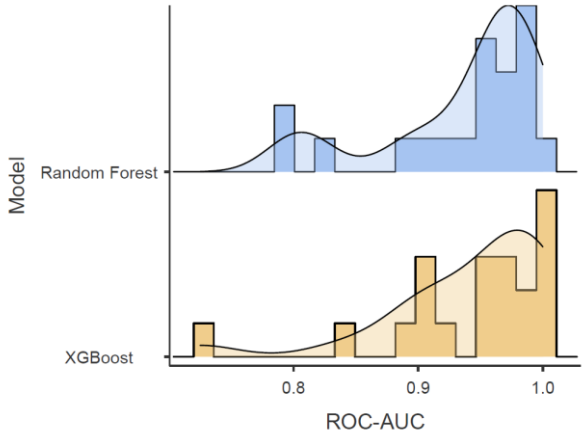
Statistik Deskriptif	Model	ROC-AUC
Mean	*Random Forest*	0.937
	XGBoost	0.942
Median	*Random Forest*	0.961
	XGBoost	0.962
Standard deviation	*Random Forest*	0.0635
	XGBoost	0.0684
Minimum	*Random Forest*	0.800
	XGBoost	0.725
Maximum	*Random Forest*	1.000
	XGBoost	1.000

Hasil analisis menunjukkan bahwa nilai rata-rata AUC model *Random Forest* sebesar 0.937, sedangkan model *XGBoost* memiliki nilai rata-rata sebesar 0.942. Perbedaan nilai rata-rata tersebut relatif kecil, yang menunjukkan bahwa kedua model memiliki performa yang hampir serupa. Nilai median kedua model juga menunjukkan nilai yang hampir identik, yaitu sebesar 0.961 untuk *Random Forest* dan 0.962 untuk *XGBoost*. Hal ini mengindikasikan bahwa distribusi pusat data dari kedua model tidak berbeda secara signifikan.

Dari sisi variabilitas, *Random Forest* memiliki standar deviasi sebesar 0.0635, lebih kecil dibandingkan *XGBoost* sebesar 0.0684. Hal ini menunjukkan bahwa *Random Forest* memiliki performa yang lebih stabil dibandingkan *XGBoost*. Sebaliknya, *XGBoost* menunjukkan variasi nilai yang lebih besar, dengan nilai minimum yang lebih rendah yaitu sebesar 0.725 dibandingkan *Random Forest* sebesar 0.800.

3.3. Analisis Visualisasi Data

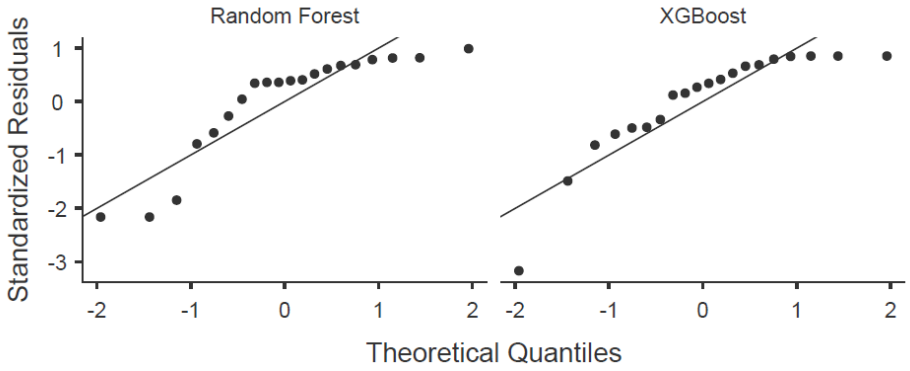
Analisis visualisasi dilakukan menggunakan histogram, Q-Q plot, dan boxplot untuk melihat distribusi data secara lebih mendalam.



Gambar 4. Histogram ROC-AUC Model *Random Forest* dan Model *XGBoost*

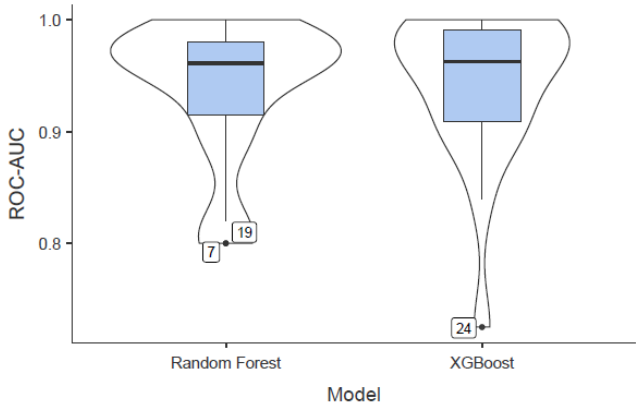
Hasil histogram menunjukkan bahwa sebagian besar nilai AUC terkonsentrasi pada rentang 0.9 hingga 1.0, yang mengindikasikan bahwa kedua model memiliki performa yang tinggi dalam mendeteksi fraud. Namun demikian, hasil Q-Q plot menunjukkan bahwa titik-titik data tidak mengikuti garis diagonal, yang mengindikasikan bahwa data tidak berdistribusi normal.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035



Gambar 5. Q-Q plot ROC-AUC Model *Random Forest* dan Model *XGBoost*

Selain itu, boxplot menunjukkan bahwa model *XGBoost* memiliki sebaran data yang lebih luas dibandingkan *Random Forest*, yang mengindikasikan bahwa *XGBoost* memiliki variasi performa yang lebih besar. Sementara itu, *Random Forest* menunjukkan distribusi yang lebih rapat dan konsisten.



Gambar 6. Boxplot ROC-AUC Model *Random Forest* dan Model *XGBoost*

3.4. Uji Normalitas dan Homogenitas Varians

Uji asumsi dilakukan untuk menentukan metode statistik yang tepat dalam analisis komparatif.

Tabel 3. Normality Test (Shapiro-Wilk)

	W	p
ROC-AUC	0.187	<.0001

Hasil uji normalitas menggunakan *Shapiro-Wilk* menunjukkan bahwa data tidak berdistribusi normal ($p < 0.05$). Hal ini menunjukkan bahwa asumsi normalitas tidak terpenuhi.

Tabel 4. Homogeneity of Variances Test (Levene's)

	F	df	df_2	p
ROC-AUC	0.00496	1	38	0.944

Sementara itu, hasil uji homogenitas varians menggunakan Levene menunjukkan bahwa varians antar kelompok homogen ($p > 0.05$). Dengan demikian, asumsi homogenitas terpenuhi.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035

Berdasarkan hasil tersebut, Asumsi data tidak memenuhi untuk dilakukan analisis parametrik, sehingga diperlukan pendekatan nonparametrik untuk analisis lebih lanjut.

3.5. Independent Samples T-Test dan Mann Whitney U Test

Analisis komparatif dilakukan untuk mengetahui apakah terdapat perbedaan kinerja antara model *Random Forest* dan *XGBoost* berdasarkan nilai AUC. Sebagai analisis awal, dilakukan uji *independent samples t-test*.

Tabel 5. Independent Samples T-Test

		Statistic	df	p		Effect Size
ROC-AUC	Student's t	-0.215	38	0.831	Cohen's d	-0,0681

Hasil independent samples t-test menunjukkan nilai signifikansi sebesar $p = 0.831$, yang menunjukkan bahwa tidak terdapat perbedaan yang signifikan antara nilai AUC Random Forest dan XGBoost. Selain itu, hasil effect size menggunakan Cohen's d menunjukkan nilai sebesar $d = -0.0681$, yang termasuk dalam kategori efek sangat kecil. Hasil ini menunjukkan bahwa perbedaan performa praktis antara kedua model relatif tidak signifikan.

Tabel 6. Mann-Whitney U Test

		Statistic	p
ROC-AUC	Mann- Whitney U	183	0.655

Hasil pengujian menunjukkan nilai signifikansi sebesar p = 0.655, yang lebih besar dari 0.05. Hal ini menunjukkan bahwa tidak terdapat perbedaan yang signifikan antara kinerja *Random Forest* dan *XGBoost* berdasarkan nilai AUC.

3.6. Pembahasan

Data penelitian mencakup berbagai domain fraud yang diklasifikasikan ke dalam beberapa kategori utama. Hasil klasifikasi menunjukkan bahwa sebagian besar data berasal dari domain financial fraud sebanyak 29 data, diikuti oleh *healthcare fraud* sebanyak 4 data, *blockchain fraud* sebanyak 3 data, serta industrial fraud sebanyak 2 data. Sementara itu, domain *e-commerce* dan *recruitment fraud* masing-masing hanya memiliki 1 data. Dominasi domain financial fraud menunjukkan bahwa penelitian terkait *fraud detection* masih berfokus pada sektor keuangan, khususnya transaksi kartu kredit dan perbankan.

Hasil penelitian menunjukkan bahwa *Random Forest* dan XGBoost memiliki performa yang relatif tinggi dalam *fraud detection*. Meskipun *XGBoos*t memiliki nilai rata-rata AUC yang sedikit lebih tinggi dibandingkan *Random Fore*st, hasil uji statistik menunjukkan bahwa perbedaan tersebut tidak signifikan secara statistik maupun praktis. Nilai *effect size* yang sangat kecil menunjukkan bahwa selisih performa kedua model hampir tidak memberikan dampak praktis yang berarti.

Selain itu, hasil penelitian menunjukkan bahwa *Random Forest* cenderung memiliki performa yang lebih stabil dibandingkan *XGBoost*. Hal ini terlihat dari nilai standar deviasi *Random Forest* yang lebih kecil dibandingkan *XGBoost*. Sebaliknya, *XGBoost* menunjukkan variasi performa yang lebih besar antar studi.

Heterogenitas antar studi juga menjadi faktor penting dalam interpretasi hasil penelitian ini. Setiap penelitian sumber menggunakan karakteristik dataset, teknik *preprocessing*, penanganan data imbalance, serta strategi tuning model yang berbeda. Perbedaan tersebut menyebabkan nilai AUC yang dihasilkan tidak sepenuhnya merepresentasikan kemampuan intrinsik model, tetapi juga dipengaruhi oleh desain eksperimen pada masing-masing studi. Oleh karena itu, hasil penelitian ini perlu dipahami sebagai kecenderungan performa umum lintas studi, bukan sebagai penentuan model terbaik secara absolut.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035

Secara metodologis, penelitian ini menunjukkan bahwa pendekatan *comparative statistical* dapat digunakan untuk mengevaluasi kecenderungan performa model *machine learning*. Pendekatan ini memberikan perspektif tambahan bahwa evaluasi model *machine learning* tidak hanya perlu mempertimbangkan nilai performa rata-rata, tetapi juga signifikansi statistik, effect size, dan stabilitas performa antar model.

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan. Data yang digunakan berasal dari berbagai studi dengan karakteristik dataset, preprocessing, dan strategi evaluasi yang berbeda sehingga berpotensi menimbulkan heterogenitas hasil. Selain itu, penelitian ini hanya menggunakan nilai AUC sebagai metrik evaluasi utama dan belum mempertimbangkan metrik lain seperti *precision*, recall, F1-score, maupun *Area Under Precision-Recall Curve* (AUPRC) yang juga penting dalam *fraud detection* pada data tidak seimbang.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa *Random Forest* dan *XGBoost* memiliki performa yang relatif sebanding dalam fraud detection berdasarkan nilai AUC pada berbagai studi. Meskipun *XGBoost* memiliki rata-rata AUC yang sedikit lebih tinggi, hasil uji statistik dan effect size menunjukkan bahwa perbedaan performa kedua model tidak signifikan secara statistik maupun praktis. Selain itu, *Random Forest* cenderung menunjukkan performa yang lebih stabil, sedangkan *XGBoost* lebih sensitif terhadap karakteristik data dan konfigurasi model.

Hasil penelitian ini menunjukkan bahwa evaluasi performa *machine learning* tidak cukup hanya berdasarkan nilai rata-rata AUC, tetapi juga perlu mempertimbangkan signifikansi statistik, effect size, serta heterogenitas dataset antar studi.

Penelitian ini memiliki keterbatasan karena menggunakan data dari berbagai studi dengan karakteristik dataset dan metode preprocessing, penanganan data imbalance, serta strategi tuning model yang berbeda. Selain itu, penelitian ini hanya menggunakan AUC sebagai metrik evaluasi utama. Penelitian selanjutnya dapat mengembangkan analisis dengan mempertimbangkan metrik evaluasi lain selain AUC, seperti precision, recall, dan F1-score, serta mengkaji performa model pada dataset yang lebih besar dan beragam.

DAFTAR PUSTAKA

[1] S. Sagiraju, J. Mohanty, and A. Naik, "Hyperparameter Tuning of Random Forest using Social Group Optimization Algorithm for Credit Card Fraud Detection in Banking Data," *Int. J. Comput. Exp. Sci. Eng.*, 2025, doi: 10.22399/ijcesen.777.

[2] S. Kumar, "Enhanced Fraud Detection in Financial Transactions Using Hyperparameter-Tuned Random Forests," 2024, pp. 1–7. doi: 10.1109/ICCCNT61001.2024.10725958.

[3] T.-H. Lin and J.-R. Jiang, "Credit Card Fraud Detection with Autoencoder and Probabilistic Random Forest," *Mathematics*, 2021, doi: 10.3390/math9212683.

[4] P. Hájek, M. Z. Abedin, and U. Sivarajah, "Fraud Detection in Mobile Payment Systems using an XGBoost-based Framework," *Inf. Syst. Front.*, pp. 1–19, 2022, doi: 10.1007/s10796-022-10346-6.

[5] S. Iqbal, K. M. Awan, and S. Kamal, "Interpretable Ensemble Learning Models for Credit Card Fraud Detection," pp. 1–30, 2025.

[6] A. Ali *et al.*, "Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review," *Appl. Sci.*, vol. 12, no. 19, 2022, doi: 10.3390/app12199637.

[7] Z. Salekshahrezaee, J. Leevy, and T. Khoshgoftaar, "The effect of feature extraction and data sampling on credit card fraud detection," *J. Big Data*, vol. 10, pp. 1–17, 2023, doi: 10.1186/s40537-023-00684-w.

[8] O. Rainio, J. Teuho, and R. Klén, "Evaluation metrics and statistical tests for machine learning," *Sci. Rep.*, vol. 14, no. 1, p. 6086, 2024, doi: 10.1038/s41598-024-56706-x.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035

[9] A. Aida, D. Hermina, and N. Norlaila, "Jenis Data Penelitian Kuantitatif (Korelasional, Komparatif, Dan Eksperimen)," *Al-Manba J. Ilm. Keislam. dan Kemasyarakatan*, 2025, doi: 10.69782/almanba.v10i1.48.

[10] B. Devi, "Application of Cross-National Comparative Research Design in Medical and Nursing Education," *J. Heal. Allied Sci. NU*, vol. 13, pp. 306–312, 2022, doi: 10.1055/s-0042-1757734.

[11] W.-F. Lai and M. Fong, "Use of comparative research in the study of chemistry education: A systematic analysis of the literature," *Heliyon*, vol. 10, 2023, doi: 10.1016/j.heliyon.2023.e22881.

[12] M. A. Saidah, A. Saputra, and H. Setiawan, "Analisis Komparasi Cybercrime Web Defacement dan Darknet Exposure di Indonesia (Studi Kasus : Lanskap Keamanan Siber di Indonesia Tahun 2022 dan Tahun 2023) | Prosiding Seminar Nasional Ilmu Sosial dan Teknologi (SNISTEK)," in *Prosiding Seminar Nasional Ilmu Sosial Dan Teknologi (SNISTEK)*, 2024, pp. 276–281.

[13] S. Mazhar, "Methods of Data Collection: A Fundamental Tool of Research," *J. Integr. Community Heal.*, 2021, doi: 10.24321/2319.9113.202101.

[14] E. S. Mukasa, W. Christospher, B. Ivan, and M. Kizito, "The Effects of Parametric, Non-Parametric Tests and Processes in Inferential Statistics for Business Decision Making — A Case of 7 Selected Small Business Enterprises in Uganda," *Open J. Bus. Manag.*, 2021, doi: 10.4236/ojbm.2021.93081.

[15] G. Kotronoulas *et al.*, "An Overview of the Fundamentals of Data Management, Analysis, and Interpretation in Quantitative Research.," *Semin. Oncol. Nurs.*, p. 151398, 2023, doi: 10.1016/j.soncn.2023.151398.

[16] M. Hiebl, "Sample Selection in Systematic Literature Reviews of Management Research," *Organ. Res. Methods*, vol. 26, pp. 229–261, 2021, doi: 10.1177/1094428120986851.

[17] E. C. Neto, V. Yadav, S. Sieberts, and L. Omberg, "A novel estimator for the two-way partial AUC," *BMC Med. Inform. Decis. Mak.*, vol. 24, 2024, doi: 10.1186/s12911-023-02382-2.

[18] D. Chicco and G. Jurman, "The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification," *BioData Min.*, vol. 16, 2023, doi: 10.1186/s13040-023-00322-4.

[19] M. Kahveci and L. Uğur, "Prediction and Stage Classification of Pressure Ulcers in Intensive Care Patients by Machine Learning," *Diagnostics*, vol. 15, 2025, doi: 10.3390/diagnostics15101239.

[20] A. Carrington *et al.*, "Deep ROC Analysis and AUC as Balanced Average Accuracy, for Improved Classifier Selection, Audit and Explanation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, pp. 329–341, 2021, doi: 10.1109/tpami.2022.3145392.

[21] Y. Lu, C. Yang, and Z. Meng, "Lithology Discrimination Using Sentinel-1 Dual-Pol Data and SRTM Data," *Remote. Sens.*, vol. 13, p. 1280, 2021, doi: 10.3390/rs13071280.

[22] Ş. Çorbacıoğlu and G. Aksel, "Receiver operating characteristic curve analysis in diagnostic accuracy studies: A guide to interpreting the area under the curve value," *Turkish J. Emerg. Med.*, vol. 23, pp. 195–198, 2023, doi: 10.4103/tjem.tjem_182_23.

[23] I. Ayulani, A. M. Yunawan, T. Prihutaminingsih, D. Sarwinda, G. Ardaneswari, and B. Handari, "Tree-Based Ensemble Methods and Their Applications for Predicting Students' Academic Performance," *Int. J. Adv. Sci. Eng. Inf. Technol.*, 2023, doi: 10.18517/ijaseit.13.3.16880.

[24] H. Baalousha, "Machine Learning-Driven Calibration of MODFLOW Models: Comparing Random Forest and XGBoost Approaches," *Geosciences*, 2025, doi: 10.3390/geosciences15080303.

[25] E. Ismanto and M. Novalias, "Komparasi Kinerja Algoritma C4 . 5 , Random Forest , dan Gradient Boosting untuk Klasifikasi Komoditas," *Techno.COM*, vol. 20, no. 3, pp. 400–

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035

410, 2021, doi: https://doi.org/10.33633/tc.v20i3.4576.
[26] Z. Farhadi, H. Bevrani, M. Feizi-Derakhshi, W. Kim, and M. F. Ijaz, "An Ensemble Framework to Improve the Accuracy of Prediction Using Clustered Random-Forest and Shrinkage Methods," *Appl. Sci.*, 2022, doi: 10.3390/app122010608.
[27] R. Natras, B. Soja, and M. Schmidt, "Ensemble Machine Learning of Random Forest, AdaBoost and XGBoost for Vertical Total Electron Content Forecasting," *Remote. Sens.*, vol. 14, p. 3547, 2022, doi: 10.3390/rs14153547.
[28] Y. Park, "Optimal two-stage group sequential designs based on Mann-Whitney-Wilcoxon test," *PLoS One*, vol. 20, 2025, doi: 10.1371/journal.pone.0318211.
[29] R. Tapio, "The Role of Data Assumptions in Selecting Between Parametric and Nonparametric Tests," *Asian J. Probab. Stat.*, 2025, doi: 10.9734/ajpas/2025/v27i11830.
[30] G. Sullivan and R. Feinn, "Using Effect Size-or Why the P Value Is Not Enough.," *J. Grad. Med. Educ.*, vol. 4 3, pp. 279–282, 2012, doi: 10.4300/jgme-d-12-00156.1.
[31] N. Bulanov *et al.*, "Basic principles of descriptive statistics in medical research," *Sechenov Med. J.*, 2021, doi: 10.47093/2218-7332.2021.12.3.4-16.
[32] N. Khatun, "Applications of Normality Test in Statistical Analysis," *Open J. Stat.*, 2021, doi: 10.4236/ojs.2021.111006.
[33] C. Avram and M. Mărușteri, "Normality assessment, few paradigms and use cases," *Rev. Rom. Med. Lab.*, vol. 30, pp. 251–260, 2022, doi: 10.2478/rrlm-2022-0030.
[34] A. Ritonga *et al.*, "Penerapan Distribusi Normal Dalam Pengukuran Tinggi Badan Mahasiswa FMIPA Universitas Negeri Medan 2024," *Bilangan J. Ilm. Mat. Kebumian dan Angkasa*, 2025, doi: 10.62383/bilangan.v3i2.465.
[35] A. Katsileros, N. Antonetsis, P. Mouzaidis, E. Tani, P. Bebeli, and A. Karagrigoriou, "A comparison of tests for homoscedasticity using simulation and empirical data," *Commun. Stat. Appl. Methods*, 2024, doi: 10.29220/csam.2024.31.1.001.
[36] G. Katz, A. Restori, and H. Lee, "A Monte Carlo Study Comparing the Levene Test to Other Homogeneity of Variance Tests," *N. Am. J. Psychol.*, vol. 11, p. 511, 2009.
[37] A. Yonar, H. Yonar, M. Demirsöz, and M. Tekindal, "A COMPARATIVE ANALYSIS FOR HOMOGENEITY OF VARIANCE TESTS," *J. Sci. Arts*, 2023, doi: 10.46939/j.sci.arts-24.2-a06.
[38] P. Sambaraju, "Two sample unpaired T-test power calculation using simulation," *Analecta Tech. Szeged.*, 2023, doi: 10.14232/analecta.2023.4.10-15.
[39] A. D. Putri, R. S. Hilmia, S. Almaliyah, S. Permana, and U. Pendidikan, "Pengaplikasian uji t dalam penelitian eksperimen," vol. 4, no. 3, pp. 1978–1987, 2023.
[40] D. Chicco, A. Sichenze, and G. Jurman, "A simple guide to the use of Student's t-test, Mann-Whitney U test, Chi-squared test, and Kruskal-Wallis test in biostatistics," *BioData Min.*, vol. 18, 2025, doi: 10.1186/s13040-025-00465-6.
[41] T. Ashfaq *et al.*, "A Machine Learning and Blockchain Based Efficient Fraud Detection Mechanism," *Sensors (Basel).*, vol. 22, 2022, doi: 10.3390/s22197162.
[42] E. Nabrawi and A. Alanazi, "Fraud Detection in Healthcare Insurance Claims Using Machine Learning," *Risks*, 2023, doi: 10.3390/risks11090160.
[43] T. Albalawi and S. Dardouri, "Enhancing credit card fraud detection using traditional and deep learning models with class imbalance mitigation," *Front. Artif. Intell.*, vol. 8, 2025, doi: 10.3389/frai.2025.1643292.
[44] A. Kumar and A. Panwar, "Voting Classifier as a Balanced Framework for Fraud Detection in Imbalanced Credit Card Transactions," *J. Inf. Syst. Eng. Manag.*, 2024, doi: 10.52783/jisem.v9i4s.12735.
[45] A. Hewiagh, K. Ramakrishnan, T. Yap, and C. S. Tan, "Waste Management System Fraud Detection Using Machine Learning Algorithms to Minimize Penalties Avoidance and Redemption Abuse," *Recycling*, 2021, doi: 10.3390/recycling6040065.
[46] A. Alutaibi, "Blockchain Analytics Based on Artificial Intelligence: Using Machine

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035

Learning for Improved Transaction Analysis," *IET Inf. Secur.*, vol. 2025, 2025, doi: 10.1049/ise2/5560771.
[47] E. Ileberi, Y. Sun, and Z. Wang, "A machine learning based credit card fraud detection using the GA algorithm for feature selection," *J. Big Data*, vol. 9, 2022, doi: 10.1186/s40537-022-00573-8.
[48] X. Liu, "Empirical Analysis of Financial Statement Fraud of Listed Companies Based on Logistic Regression and Random Forest Algorithm," *J. Math.*, 2021, doi: 10.1155/2021/9241338.
[49] M. Korde, S. Bhayal, R. Maheshwari, S. Pandya, and M. Raikwar, "Fraud Detection in Financial Systems Using Machine Learning Techniques," *J. Inf. Syst. Eng. Manag.*, 2025, doi: 10.52783/jisem.v10i33s.5737.
[50] M. Shahidullah *et al.*, "Intelligent Fraud Detection: Applying Advanced Analytics and Cybersecurity Insights in U.S. Finance," *J. Posthumanism*, 2024, doi: 10.63332/joph.v4i3.3593.
[51] C.-W. Lee, M.-W. Fu, C.-C. Wang, and M. I. Azis, "Evaluating Machine Learning Algorithms for Financial Fraud Detection: Insights from Indonesia," *Mathematics*, 2025, doi: 10.3390/math13040600.
[52] M. N. Balakishore, N. Supriya, G. HaleelBasha, Ds. Maheswari, and S. Vasu, "A Credit Card Fraud Detection Method Based on Hybrid Sampling and Random Forest Algorithm," *African J. Biomed. Res.*, 2025, doi: 10.53555/ajbr.v28i2s.7228.
[53] C. Angelica, C. Charleen, and A. Wibowo, "Elevating fraud detection: machine learning models with computational intelligence optimization," *IAES Int. J. Artif. Intell.*, 2024, doi: 10.11591/ijai.v13.i4.pp4273-4280.
[54] M. Ahsan, T. Y. Susanto, T. A. Virania, and A. I. Jaya, "Credit Card Fraud Detection Using Linear Discriminant Analysis (LDA), Random Forest, And Binary Logistic Regression," *BAREKENG J. Ilmu Mat. dan Terap.*, 2022, doi: 10.30598/barekengvol16iss4pp1337-1346.
[55] J. K. Afriyie *et al.*, "A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions," *Decis. Anal. J.*, 2023, doi: 10.1016/j.dajour.2023.100163.
[56] P. Li, "Credit Card Fraud Detection Based on Random Forest Model," *Acad. J. Comput. & Inf. Sci.*, 2022, doi: 10.25236/ajcis.2022.051309.
[57] R. Bounab, K. Zarour, B. Guelib, and N. Khlifa, "Enhancing Medicare Fraud Detection Through Machine Learning: Addressing Class Imbalance With SMOTE-ENN," *IEEE Access*, vol. 12, pp. 54382–54396, 2024, doi: 10.1109/access.2024.3385781.
[58] D. Dake, "Online Recruitment Fraud Detection: A Machine Learning-based Model for Ghanaian Job Websites," *Int. J. Comput. Appl.*, 2023, doi: 10.5120/ijca2023922639.
[59] M. Tayebi and E. Kafhali, "A Novel Approach based on XGBoost Classifier and Bayesian Optimization for Credit Card Fraud Detection," *Cyber Secur. Appl.*, 2025, doi: 10.1016/j.csa.2025.100093.
[60] A. A. J. Al-Hchaimi, M. Khalifa, and W. El-Shafai, "Explainable AI With Imbalanced Learning Strategies for Blockchain Transaction Fraud Detection," *Eng. Reports*, 2026, doi: 10.1002/eng2.70545.
[61] Z. Zhao and T. Bai, "Financial Fraud Detection and Prediction in Listed Companies Using SMOTE and Machine Learning Algorithms," *Entropy*, vol. 24, 2022, doi: 10.3390/e24081157.
[62] C. Wei and X. Qian, "Bridging the Semantic Gap: An Ensemble Learning Framework With Textual Topic-Raw Financial Feature Fusion to Enhance Fraud Detection in Chinese Markets," *J. Math.*, 2025, doi: 10.1155/jom/6643152.
[63] H. Xia, W. An, and Z. Zhang, "Credit Risk Models for Financial Fraud Detection: A New Outlier Feature Analysis Method of XGBoost With SMOTE," *J. Database Manag.*, vol.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035

34, pp. 1–20, 2023, doi: 10.4018/jdm.321739.
[64] S. K. Hashemi, S. L. Mirtaheri, and S. Greco, "Fraud Detection in Banking Data by Machine Learning Techniques," *IEEE Access*, vol. 11, pp. 3034–3043, 2023, doi: 10.1109/access.2022.3232287.
[65] A. Mehdary, A. Chehri, A. Jakimi, and R. Saadane, "Hyperparameter Optimization with Genetic Algorithms and XGBoost: A Step Forward in Smart Grid Fraud Detection," *Sensors (Basel).*, vol. 24, 2024, doi: 10.3390/s24041230.
[66] M. Abouelenein, H. Noaman, and G. S. S. A. Al Salmany, "Investigating the impact of feature engineering and machine learning model selection for real-world fraud detection systems in healthcare insurance claims," *Int. J. Innov. Res. Sci. Stud.*, 2025, doi: 10.53894/ijirss.v8i5.8858.
[67] E. B. Abdelghafour, C. Mohamed, A. Noura, and B. Abdelhamid, "Enhancing Credit Card Fraud Detection Using a Stacking Model Approach and Hyperparameter Optimization," *Int. J. Adv. Comput. Sci. Appl.*, 2024, doi: 10.14569/ijacsa.2024.01510110.
[68] D. Jovanovic, M. Antonijevic, M. Stanković, M. Zivkovic, M. Tanaskovic, and N. Bačanin, "Tuning Machine Learning Models Using a Group Search Firefly Algorithm for Credit Card Fraud Detection," *Mathematics*, 2022, doi: 10.3390/math10132272.
[69] W. Li and X. Xu, "Ensemble learning algorithm - research analysis on the management of financial fraud and violation in listed companies," *Decis. Mak. Appl. Manag. Eng.*, 2023, doi: 10.31181/dmame622023785.
[70] F. Becerra-Suarez, H. Alvarez-Vasquez, and M. Forero, "Improvement of Bank Fraud Detection Through Synthetic Data Generation with Gaussian Noise," *Technologies*, 2025, doi: 10.3390/technologies13040141.
[71] M. Balayet *et al.*, "Enhancing Medicare Fraud Detection With a CNN-Transformer-XGBoost Framework and Explainable AI," *IEEE Access*, vol. 13, pp. 79609–79622, 2025, doi: 10.1109/access.2025.3562577.
[72] A. Dichev, S. Zarkova, and P. Angelov, "Machine Learning as a Tool for Assessment and Management of Fraud Risk in Banking Transactions," *J. Risk Financ. Manag.*, 2025, doi: 10.3390/jrfm18030130.
[73] Y. Han and I. Joe, "Enhanced Predictive Modeling for Anomaly Detection in Financial Transactions Using Machine Learning," *IEEE Access*, vol. 13, pp. 154438–154449, 2025, doi: 10.1109/access.2025.3602236.
[74] N. Damanik and C.-M. Liu, "Advanced Fraud Detection: Leveraging K-SMOTEENN and Stacking Ensemble to Tackle Data Imbalance and Extract Insights," *IEEE Access*, vol. 13, pp. 10356–10370, 2025, doi: 10.1109/access.2025.3528079.
[75] B. Krishna, V. S. S. Nishwan, M. Sandireddy, and S. E., "A Hybrid Framework for Privacy-Preserving and Explainable Credit Card Fraud Detection Using XGBoost and Homomorphic Encryption," in *2025 Second International Conference on Intelligent Technologies for Sustainable Electric and Communications Systems (iTech SECOM)*, 2025, pp. 1–6. doi: 10.1109/itechsecom64750.2025.11307461.
[76] S. N. Nobel *et al.*, "Unmasking Banking Fraud: Unleashing the Power of Machine Learning and Explainable AI (XAI) on Imbalanced Data," *Inf.*, vol. 15, p. 298, 2024, doi: 10.3390/info15060298.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i2.16035