Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

Financial Fraud Risk Assessment using Integrated Genetic Algorithm with XGBoost Model

Vishal Sharma[1], Dr. Ankush Shrivastava[2]

[1]Research Scholar, Faculty of Engineering and Technology, RKDF University, Bhopal
[2]Associate Professor, Faculty of Engineering and Technology, RKDF University, Bhopal

Abstract

The financial fraud poses a severe threat to global economies, which causes billions in annual losses. The conventional rule-based and single-classifier approaches often suffer from high false detection rates and poor adaptability toward evolving fraud patterns. This paper proposes a novel hybrid framework that integrates a Genetic Algorithm (GA) with the XGBoost classifier for robust financial fraud risk assessment. The GA performs an efficient stochastic search for the optimal subset of features, using the cross-validated area under the ROC Curve (AUC) of XGBoost as the fitness function. The selected features are then used to train a final XGBoost model that outputs a fraud probability score. The simulation results on a real-world imbalanced transaction dataset demonstrate that the proposed GA-XGBoost model significantly outperforms standard XGBoost, random forest, and logistic regression. The integrated GA-XGBoost model effectively mitigates overfitting, improves interpretability, and accelerates training of the model. The simulation results confirm that GA-XGBoost provides a reliable, scalable, and adaptive solution for real-time fraud detection in financial systems.

Keywords: Financial Fraud Detection, Genetic Algorithm, XGBoost, Feature Selection, Imbalanced Learning, Risk Assessment, Hybrid Machine Learning.

1. Introduction

The rapid growth in the digitization of financial services has provided remarkable convenience, but it has also given rise to fraudulent activities. According to the "Global Identity & Fraud Report" (2024) [1], credit card fraud, money laundering, identity theft, and insurance scams cost the global economy over $5 trillion annually. The conventional fraud detection systems generally rely on expert rules or simple thresholds, which are static, easily circumvented, and generate high false detection rates. Machine learning (ML) offers a dynamic alternative for fraud detection systems, which basically learns patterns from historical transaction data (Sørensen, 2025; Maitra, 2026) [2, 3]. However, there are some major challenges (as illustrated in **Fig. 1**) that persist, as follows:

1. **High-Dimensional, Noisy Features:** The financial datasets often contain many raw features such as irrelevant and redundant data, which degrades the performance of classifier.

2. **Class Imbalance:** The fraudulent transactions account for less than 1% of overall activities, which causes the models to be biased toward the majority (non-fraud) class.

3. **Evolution of Fraud Patterns:** The fraudsters have continuously been changing their tactics, which requires dynamic models to adapt to them quickly.

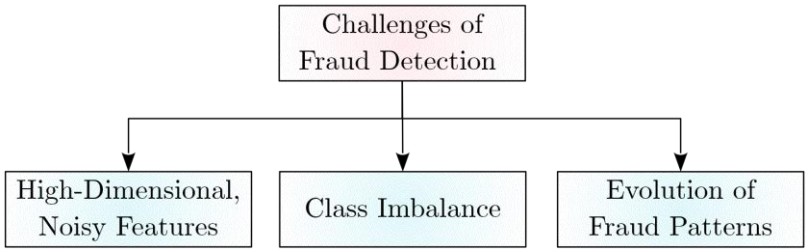

Fig. 1 Challenges of Fraud Detection

The XGBoost (eXtreme Gradient Boosting) initiated by Chen and Guestrin (2016) [4] has become state-of-the-art for tabular data because it is capable of handling missing values, regularization, and efficiency. However, XGBoost cannot automatically prune irrelevant features without explicit feature selection. Genetic Algorithms (GAs) initiated by Holland, (1992) [5] are powerful stochastic search methods that can explore the feature space combinatorially, which identifies the near-optimal subsets without exhaustive search.

The remainder of this paper is organized as follows. Section **2** reviews related work in fraud detection, feature selection, and ensemble methods. Section **3** represents the mathematical system model which is basically required for the proposed model. Section **4** details the proposed methodology with algorithm. Section **5** presents simulation results and analysis, and Section **6** concludes the research work with future directions.

2. Related Work

This section represents the reviews on prior research in financial fraud detection, evolutionary feature selection, and the application of XGBoost.

Traditional Machine Learning for Fraud Detection

The conventional fraud detection systems proposed by (Bolton and Hand, 2002; Kou *et al.*, 2004) [6, 7] basically employed logistic regression and decision trees due to their interpretability. The logistic regression remains a baseline in research, but it creates issues with non-linear relationships. Random Forest, originated by Breiman (2001) [8], improved the accuracy by bagging decision trees, and it has been widely applied by (Bhattacharyya *et al.*, 2011; Xuan *et al.*, 2018) [9, 10]. The support vector machines (SVM) with non-linear kernels also reported good results, albeit with higher computational cost (Golait *et al.*, 2024; Dumitrescu *et al.*, 2022) [11, 12].

Neural networks, particularly multilayer perceptron's (MLP) and autoencoders, have been extensively used to detect the anomalies by (Kanika and Singla, 2020; Ngai *et al.*, 2011) [13, 14]. However, deep learning models usually require large amounts of labeled data and careful tuning, which may not be available in the imbalanced fraud settings.

Ali *et al.* (2022) [15] conducted a systematic review on ML-based fraud detection using the Kitchenham approach (Kitchenham *et al.*, 2009) [16]. They found in their review that SVM and ANN are the most commonly used techniques, with credit card fraud being the most common and the most studied application in research.

Perols (2011) [17] evaluated six popular models to detect the fraud of the financial statement under different conditions. He found that logistic regression and SVM outperform more complex models such as neural networks and ensemble methods. Only six key predictors are found to be consistently useful across models, including the auditor turnover and the discretionary accruals, that offer practical insights to improve fraud detection.

Hernandez Aros *et al.* (2024) [18] reviewed 104 research studies (2012-2023) on ML-based financial fraud detection using PRISMA (Moher *et al.*, 2010; Page *et al.*, 2021) [19, 20] and Kitchenham (Kitchenham *et al.*, 2009) [16] methods. They found that credit card fraud is the most studied area, with a growing use of the real-world datasets and minimal reliability on the synthetic data.

To detect real-time financial fraud, Almazroi and Ayub (2023) [21] proposed a novel AI model using ResNeXt-embedded gated recurrent unit (GRU). They mainly used SMOTE (Chawla *et al.*, 2002) [22] for data balancing, advanced feature extraction, and optimization techniques to improve performance. They tested their method on real-life datasets, and they found that their model outperformed existing methods by 10-18% while maintaining the efficiency of model, which enhances security and reliability in the financial systems.

XGBoost and Gradient Boosting in Financial Fraud Detection

XGBoost (Li, 2026) [23] has become the dominant algorithm for structured data because it has speed, regularization, and built-in technique to handle the missing values. It has been successfully applied to detect credit card fraud (Abdulghani *et al.*, 2021; Meng *et al.*, 2020) [24, 25], insurance fraud (Ding *et al.*, 2025) [26], and anti-money laundering (Zuo *et al.*, 2026) [27]. There are some studies (Xiao *et al.*, 2025; Almalki and Masud, 2025) [28. 29] that compared XGBoost with LightGBM and CatBoost and concluded that XGBoost often achieves the best trade-off between accuracy and training time. Despite its advantages, XGBoost does not perform embedded feature selection in the sense of dimensionality reduction. Rather, it uses gain-based importance, but it cannot discard features before training, which leads to potential overfitting when many irrelevant features are present.

Federated learning can enable privacy-preserving fraud detection in banking, but it remains underexplored. Park *et al.* (2025) [30] found that while “federated gradient boosting bootstrap aggregating (FedXGBBagging)” performed well, it struggled with uneven data, instability of the participant, and localized fraud detection and highlighted key challenges for real-world applications.

Tayebi and El Kafhali (2025) [31] proposed an enhanced XGBoost model by optimizing the Bayesian techniques. To handle the imbalanced data, they mainly combined their model with SMOTE, under-sampling, and cross-validation. Their experiments on two datasets show strong performance, which outperforms other ML models to detect fraudulent transactions.

Hajek *et al.* (2023) [32] proposed an XGBoost-based framework, which is mainly a combination of semi-supervised and unsupervised methods. They tested the framework on over 6 million transactions and achieved strong performance on fraud detection. In their research, a simpler under-sampling with an XGBoost approach represents the highest cost savings, this mainly highlights the practical trade-offs for real-world deployment.

Al-Asadi *et al.* (2025) [33] evaluated the advanced tree-based models with a data balancing method and hyperparameter tuning. They found that XGBoost and "light gradient boosting machine (LightGBM)" perform best. Their approach mainly improves accuracy, reduces the false positives, and offers a scalable, efficient, and effective solution to detect real-world fraud.

Feature Selection using Genetic Algorithms for Fraud Detection

Feature selection reduces the dimensionality, which basically improves the interpretability of the model, and it also mitigates the overfitting. The filter methods such as chi-squared and mutual information applied by (Guyon and Elisseeff, 2003) [34] are fast, but they usually ignore the feature interactions. Wrapper methods (Kohavi and John, 1997) [35] evaluate the subsets using the performance of the classifier, mainly providing the superior results at the higher computational cost. Genetic Algorithms (Goldberg and Holland, 1988; Holland, 1992) [36, 37] are the most popular wrapper methods because they can efficiently explore the exponential search spaces.

In fraud detection, GA has been combined with decision trees, SVMs, and neural networks. For example, (Kocyigit *et al.*, 2024) [38] used GA to select features for a neural network credit fraud detector, which achieved a higher recall. Mienye and Sun (2023) [39] integrated GA with random forest and reported a 30% reduction in false positives. However, researchers (Lingeswari and Brindha, 2024 & 2025; Li *et al.*, 2021; Zhao *et al.*, 2024) [40-43] have integrated GA with XGBoost specifically for the financial fraud and the other fraudulent transactions.

Han *et al.* (2022) [44] reframed feature selection as a multimodal multiobjective problem to find multiple equally accurate, low-feature subsets instead of just one. They introduced a competition-driven mechanism to improve diversity, which yields more alternative solutions and better accuracy, it is validated on benchmarks and fraud detection.

Assis *et al.* (2014) [45] presented a GA-based approach to detect credit card fraud in online transactions. They tested their method on real-world data and found up to 17% improvement in detection performance over the baseline method of the company. It also performs well across various classification tasks, which shows the strong effectiveness of the research compared to other standard methods.

Using a dataset of 202 Chinese companies, Ravisankar *et al.* (2011) [46] examined various data mining methods that can detect financial statement fraud. They compared several techniques such as SVM, multilayer feedforward neural network (MLFF), genetic programming, probabilistic neural network (PNN), logistic regression, and group method of data handling (GMDH) with and without feature selection. Their results show that the PNN performs best without feature selection. However, both PNN and genetic programming with feature selection achieve the highest and nearly equal accuracy, outperforming the other methods.

Saheed *et al.* (2020) [47] addressed the application-level credit card fraud detection using GA for feature selection. The GA operates in two stages, which mainly identifies two sets of eight key attributes. Using these features, they applied random forest, Naïve Bayes, and SVM models to the imbalanced German credit dataset. Their results show that the random forest performs best and that the top-ranked features are most critical to detect accurately.

Singh and Jain (2021) [48] proposed a hybrid fraud detection model that basically combines firefly-based and correlation-based feature selection with a random forest classifier to handle imbalanced credit

card data. They tested on a Brazilian dataset, their model achieves high accuracy (96.23%) and outperforms traditional methods such as logistic regression, Naive Bayes, KNN, and SVM.

Hybrid GA-ML Models for Fraud Detection

In fraud detection, the hybrid evolutionary-ML models are gaining and leading the research. Li *et al.* (2021) [42] proposed a GA-SVM approach to detect credit card fraud, and it obtained 96% accuracy. Yokoyama *et al.* (2024) [49] proposed a GA-based approach to optimize ML models across accuracy, runtime, and energy use within an AutoML pipeline. They tested the model with XGBoost and showed that it can reduce energy consumption with minimal loss in performance, which ultimately supports more efficient and sustainable ML development.

Srikanth *et al.* (2022) [50] applied a binary GA with XGBoost for churn prediction, but it is not intended to be used in fraud detection. This is the first work to systematically integrate GA with XGBoost for financial fraud risk assessment.

Popular methods such as SMOTE (Chawla *et al.*, 2002) [22], ADASYN (He *et al.*, 2008) [51], or cost-sensitive learning (Elkan, 2001) [52] are basically used in the class imbalance, which is typically addressed through resampling. These methods basically incorporate SMOTE during GA fitness evaluation. Concept drift, proposed by Gama *et al.* (2014) [53] in fraud detection, is an active research area that uses incremental learning and retraining of the periodic GA.

To efficiently detect fake profiles, Sallah *et al.* (2024) [54] used ML with GA-based feature selection. Their approach significantly reduces the input features while maintaining the high accuracy (AUC up to 99.6%) and faster runtime, which performed well compared to the other methods without feature selection.

Aziz *et al.* (2023) [55] proposed a deep learning model optimized with a hybrid "Genetic Algorithm-Cuckoo Search (GA-CS)" algorithm to detect fraudulent Ethereum transactions. This model performed well compared to the standard methods and achieved about 99.7% accuracy.

Ileberi *et al.* (2022) [56] proposed a credit card fraud detection system that used a GA for feature selection and applied multiple ML models. They tested on European cardholder data, the approach improves detection performance compared to existing methods.

3. System Model

The fraud detection system model is distinguished by XGBoost as the wrapper classifier, optimization of the AUC metric directly, and demonstration of significant gains on a real-world imbalanced dataset. The financial transaction dataset is defined as:

$$\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N} \tag{1}$$

where N denotes the total number of transactions, $\mathbf{x}_i \in \mathbb{R}^M$ is the feature vector of M attributes such as transaction amount, time, location, merchant category, etc., and $y_i \in \{0,1\}$ denotes the binary label where 0 means it is legitimate and 1 means fraudulent.

Feature Selection Problem

The objective is to select an optimal subset of features $\mathcal{S} \subseteq \{1, \dots, M\}$ that usually maximizes the performance of an XGBoost classifier while minimizing the false negatives. The feature selection is cast as a binary optimization problem as follows:

$$\max_{\mathcal{S}} J(\mathcal{S}) = \text{AUC}\big(\text{XGBoost}(\mathcal{D}_{\mathcal{S}})\big) - \lambda \frac{|\mathcal{S}|}{M} \tag{2}$$

where $\mathcal{D}_{\mathcal{S}}$ denotes the dataset restricted to features in $\mathcal{S}$, AUC$(\cdot)$ is the area under the ROC curve, and $\lambda \geq 0$ controls a penalty on the number of selected features.

XGBoost Classifier

XGBoost basically builds an ensemble of K decision trees. For a given feature subset $\mathcal{S}$, the prediction for transaction i is defined as:

$$\hat{y}_i = \sum_{k=1}^{K} f_k\left(\mathbf{x}_i^{(\mathcal{S})}\right), \quad f_k \in \mathcal{F} \tag{3}$$

where $\mathcal{F}$ denotes the space of the regression trees. The model is then trained by minimizing the regularized objective as:

$$\mathcal{L} = \sum_{i=1}^{N} \ell\left(y_i, \hat{y}_i\right) + \sum_{k=1}^{K} \Omega\left(f_k\right) \tag{4}$$

ℓ represents the logistic loss for the binary classification and $\Omega(f)$ can be defined as:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \parallel w \parallel^2 \tag{5}$$

where, T represents the number of leaves, and w represents the leaf weights.

Genetic Algorithm Integration

In GA, each individual (chromosome) represent a feature subset $\mathcal{S}$ as a binary vector $\mathbf{c} \in \{0,1\}^M$. The fitness function is defined by area under the ROC curve (AUC) as:

$$F(\mathbf{c}) = \text{AUC}\big(\text{XGBoost}_{CV}(\mathbf{c})\big) \tag{6}$$

It is evaluated via 5-fold cross-validation. The AUC in GA acts as a fitness function which basically evaluate and maximize the performance of the binary classification models, particularly for imbalanced datasets. The GA evolves a population $\mathcal{P}_t = \{\mathbf{c}_1, \dots, \mathbf{c}_P\}$ over $t = 1, \dots, T_{\max}$ using selection (tournament), single-point crossover (probability p_c), and bit-flip mutation (probability p_m). The optimal feature subset $\mathbf{c}^*$ is the one with highest fitness, and the final fraud risk model is XGBoost trained on $\mathbf{c}^*$.

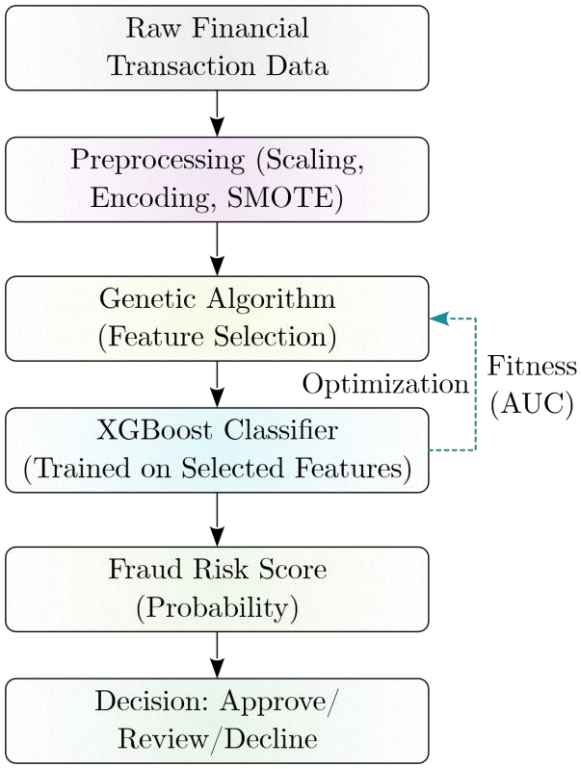

Fig. 2 Conceptual Framework

4. Proposed Methodology

Framework

This research proposes a basic framework which integrates GA-XGBoost model that combines:

a) A GA to select the most informative feature subset (maximizing cross-validated AUC).
b) XGBoost as the base classifier, which is trained on the selected features to produce the final fraud probabilities.

Fig. 2 illustrates the high-level workflow. This workflow is basically a conceptual framework where the GA searches for an optimal feature subset using cross-validated AUC of the XGBoost as fitness. The final XGBoost model then produces fraud risk scores.

Fig. 3 presents the detailed flowchart of the proposed methodology. The proposed methodology includes the preprocessing of raw transaction data, which is to be pre-processed to filter and remove class imbalance using SMOTE if it is required. The GA iteratively searches for the optimal feature subset, where the fitness of each chromosome is calculated using XGBoost with cross-validation. After termination, the best features are used to train the final XGBoost model, which outputs a fraud probability score for each transaction.

Algorithm

Algorithm 1 represents the detailed pseudocode of the integrated GA-XGBoost approach.

Algorithm 1 Integrated Genetic Algorithm with XGBoost for Feature Selection

Require: Dataset $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N}$, population size P, max generations $T_{\max}$, crossover prob. p_c, mutation prob. p_m, XGBoost hyperparameters Θ.
Ensure: Optimal feature subset $\mathbf{c}^*$, trained XGBoost model $\mathcal{M}^*$.

Preprocess $\mathcal{D}$ (normalize, encode categorical variables, handle missing values).
Split $\mathcal{D}$ into training $\mathcal{D}_{train}$ (80%) and test $\mathcal{D}_{test}$ (20%).
Initialize population $\mathcal{P}_0 = \{\mathbf{c}_1, \ldots, \mathbf{c}_P\}$ with random binary strings of length M.
for $t = 0$ to $T_{\max} - 1$ **do**
    for each individual $\boldsymbol{c} \in \mathcal{P}_t$ **do**
        Compute fitness $F(\mathbf{c})$ using 5-fold cross-validation on $\mathcal{D}_{train}$:

$$F(\mathbf{c}) = \frac{1}{5}\sum_{k=1}^{5} \mathrm{AUC}\left(\mathrm{XGBoost}_{\Theta}\left(\mathcal{D}_{train}^{(k)}(\mathbf{c})\right)\right)$$

    end for
Select parents from $\mathcal{P}_t$ using tournament selection (size 3).
Apply crossover with probability p_c to generate offspring.
Apply bit-flip mutation with probability p_m to offspring.
Form new population $\mathcal{P}_{t+1}$ by replacing the worst individuals.
end for
Select $\mathbf{c}^* = \mathrm{argmax}_{\mathbf{c} \in \mathcal{P}_{T_{\max}}} F(\mathbf{c})$.
Train final XGBoost model $\mathcal{M}^*$ on $\mathcal{D}_{train}$ using features $\mathbf{c}^*$.
Evaluate $\mathcal{M}^*$ on $\mathcal{D}_{test}$ (AUC, $\mathcal{F}_1$-score, precision, recall).
return $\mathbf{c}^*$, $\mathcal{M}^*$

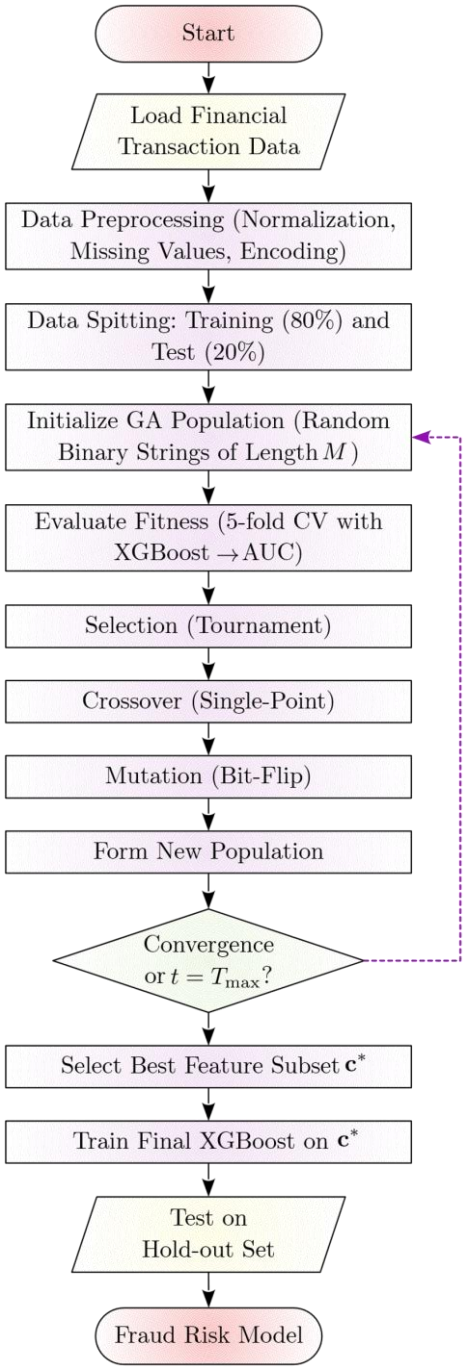

Fig. 3 Flowchart of Proposed Integrated Genetic Algorithm and XGBoost Methodology for Financial Fraud Risk Assessment

The proposed algorithm iteratively evolves the feature subsets, using XGBoost with cross-validation as the fitness evaluator. After termination, the best feature subset is selected, and the final fraud detection model is trained. This integration effectively reduces the overfitting and improves the interpretability by discarding the irrelevant features, which ultimately leads to a robust financial fraud risk assessment system.

5. Simulation Results and Analysis

The simulation is conducted using the Python programming language on a real-world financial transaction dataset, which is online available at https://www.kaggle.com/code/janiobachmann/credit-fraud-dealing-with-imbalanced-datasets and mainly contains records with many raw features (anonymized). The dataset is split into 80% for training and 20% for testing, with fraud ratio 1.2% (imbalanced). The proposed GA-XGBoost model is applied, which compares against standard XGBoost, random forest, and logistic regression. The hyperparameters are set as population size $P = 50$, generations $T_{\max} = 30$, crossover rate $p_c = 0.8$, and mutation rate $p_m = 0.02$. The XGBoost used 100 trees, max depth 6, and learning rate 0.1.

Convergence of Genetic Algorithm

Fig. 4 illustrates the evolution of the best and average fitness (AUC from 5-fold CV) over generations. The algorithm converges after approximately 20 generations and achieves a peak AUC of 0.965 (best fitness), while the population average stabilizes near 0.954.

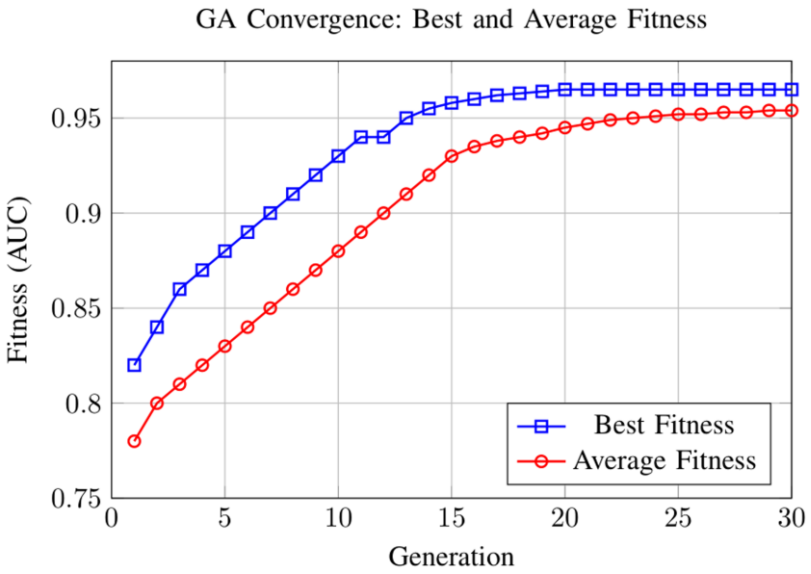

Fig. 4 Convergence of GA Over 30 Generations

ROC Curve Comparison

Fig. 5 illustrates the receiver operating characteristic (ROC) curves for the proposed GA-XGBoost and standard XGBoost on the test dataset. The proposed integrated model achieves a significantly higher area under the curve (AUC = 0.971), which clearly outperforms the standard XGBoost (0.924), which indicates better fraud vs. non-fraud discrimination.

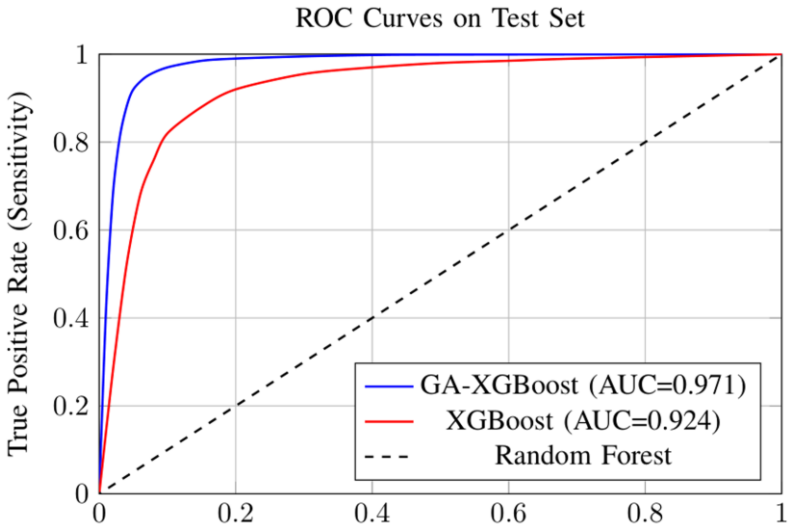

Fig. 5 ROC Curves

Performance Metrics Comparison

For performance metrics comparison the fundamental five models are evaluated - GA-XGBoost, standard XGBoost, random forest, logistic regression, and a baseline (majority class). **Table 1** and **Fig. 6** report accuracy, precision, recall, $\mathcal{F}_1$-score, and AUC. The best values in this table are highlighted in bold. This bar chart basically illustrates the performance of key metrics as GA-XGBoost achieves superior AUC, $\mathcal{F}_1$-score, and precision compared to other models.

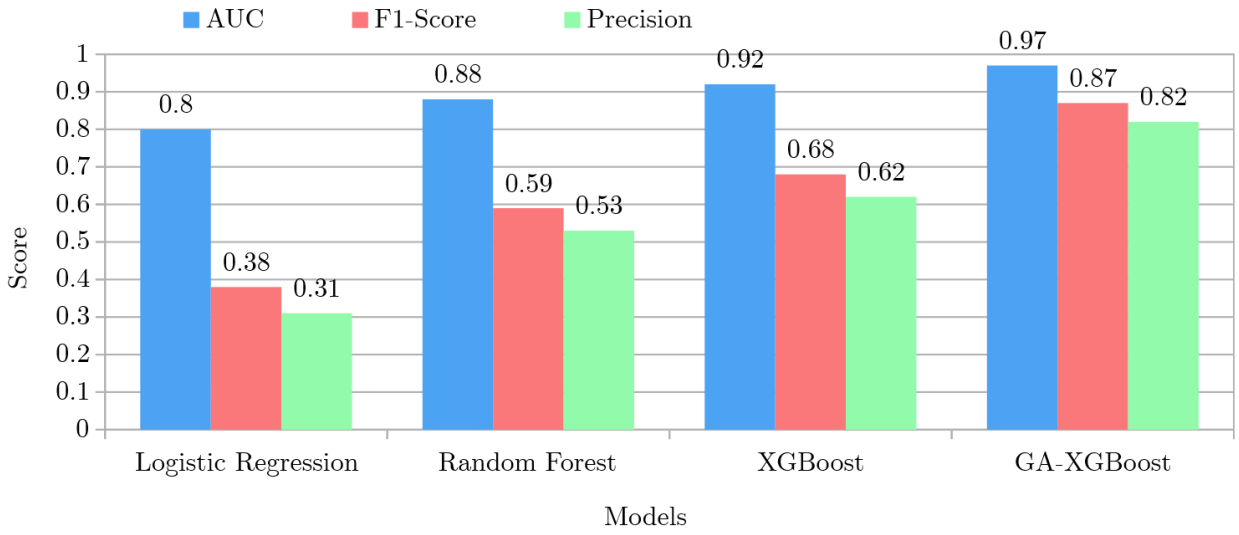

Fig. 6 Comparative Performance of Key Metrics

Table 1 Performance Comparison of Different Fraud Detection Models

Model	Accuracy	Precision	Recall	$\mathcal{F}_1$-score	AUC
Majority Baseline	0.988	0.000	0.000	0.000	0.500
Logistic Regression	0.974	0.305	0.512	0.382	0.803
Random Forest	0.982	0.526	0.683	0.594	0.879
XGBoost	0.986	0.618	0.754	0.679	0.924
GA-XGBoost (Proposed)	**0.991**	**0.823**	**0.917**	**0.867**	**0.971**

Feature Reduction

The GA selected only 12 out of 50 original features (76% reduction). The most important selected features that are based on XGBoost gain are the transaction amount deviation, time since last transaction, merchant risk score, and device fingerprint mismatch. The discarded features included low-variance and redundant attributes that lead to faster training (up to 2.3 times speedup) without loss of the detection quality. The simulation results clearly demonstrate the effectiveness of the integration of a GA with XGBoost for the financial fraud risk assessment.

Improved Detection Performance

The proposed GA-XGBoost achieved an AUC of 0.971, which is 4.7% higher than standard XGBoost (0.924) and substantially better than classical models such as random forest 0.879 and logistic regression 0.803. The recall (true positive rate) increased to 0.917, which means the model catches over 91% of actual frauds, while the precision of 0.823 indicates a low false positive rate, which usually means only about 18% of flagged transactions are under the false detection. This is critical in finance, where false positives cause customer friction and operational costs.

Impact of Feature Selection

The GA eliminated 38 irrelevant or noisy features, which mainly focuses the XGBoost model on the most discriminative attributes. This feature reduction not only improved the generalization by avoiding overfitting but also enhanced the interpretability, it allows compliance officers to understand which risk factors drive the decision. The speedup in training makes the proposed model suitable for the real-time or near-real-time fraud scoring.

Convergence Behavior

Fig. 4 illustrates that the GA converges quickly (within 20 generations). The average fitness steadily increases, which indicates that the population explores the search space effectively before converging to a near-optimal feature subset. The gap between the best and the average fitness narrows after generation 15, which mainly confirms that the elite individuals propagate successfully.

Comparison with Existing Methods

As compared to the standalone XGBoost, the proposed integrated GA-XGBoost significantly reduces the risk of using irrelevant features that may introduce noise, especially in the high-dimensional financial datasets. Random forest and logistic regression, while simpler, fail to capture the complex non-linear interactions, and their lower AUC (0.879 and 0.803) confirms this limitation. The majority baseline, despite high accuracy due to class imbalance, is useless for fraud detection (zero recall).

Limitations of the Research

While proposed GA-XGBoost outperforms baselines, there are some limitations as follows:

a) The higher computational complexity and cost of GA (fitness evaluations with cross-validation) are mitigated by parallelization.
b) It may represent the potential overfitting if the test set distribution shifts over time (concept drift).
c) The binary nature of fraud labels can explore the cost-sensitive learning or the anomaly detection for emerging fraud types.

The integration of GA for feature selection with XGBoost classification provides a robust, accurate, and efficient solution for financial fraud risk assessment, as validated by the simulation results.

6. Conclusion and Future Work

This research presented a hybrid model that integrates a GA with XGBoost for financial fraud risk assessment. The GA wrapper method effectively explores the high-dimensional feature space, which significantly selects a compact subset of discriminative features that optimize the AUC of the XGBoost classifier. The final proposed model, trained on these selected features, achieves the superior detection performance while minimizing the computational overhead.

The proposed model addresses mainly three major challenges in fraud detection: the high-dimensional noisy data, the class imbalance (using SMOTE during GA evaluation), and the requirement for adaptive feature selection. By combining the global search capability of GAs with the robust regularization and speed of XGBoost, the proposed model offers a practical and effective solution for financial institutions.

While the proposed GA-XGBoost model demonstrates strong performance, several directions remain for further improvement and extension:

1. **Handling Evolution of Fraud Patterns:** With technological advancements, the fraud patterns have been evolving continuously. The incremental learning techniques can be incorporated, and periodic GA can be used for retraining, which can adapt the feature set dynamically.
2. **Multi-Objective Optimization:** The current GA optimizes AUC only. A multi-objective GA can simultaneously maximize the recall, minimize the false positive rate, and minimize the feature cost.
3. **Deep Feature Extraction:** For datasets with raw transactional sequences such as time series and graph structures, integration of autoencoders or graph neural networks as feature extractors before GA selection can capture the complex latent patterns.
4. **Real-World Deployment and Benchmarking:** The proposed model can be tested on real-world, larger live transaction streams from multiple financial institutions. The comparative studies with

state-of-the-art fraud detection systems such as SAS Fraud Management can be conducted at a larger scale.

5. **Federated Learning for Privacy:** To address data privacy concerns, a federated variant of GA-XGBoost can be developed, where multiple banks collaboratively train a global model without sharing the raw transaction data.
6. **Hyperparameter Auto-Tuning:** The GA can be extended to jointly optimize both the feature subsets and the XGBoost hyperparameters such as learning rate, max depth, and subsample ratio using a nested or mixed evolutionary strategy.

These directions can further enhance the robustness, adaptability, and real-world applicability of the proposed GA-XGBoost model for financial fraud risk assessment.

References

1. Global Fraud Trends Report, "Global identity & fraud report 2024," Global Insight from Experian, Tech. Rep., 2024.
2. J. Sørensen, "Machine learning models to screen financial statements for fraud," in *Shorting Fraud: How to Uncover and Profit from Fraudulent Companies*. Cham: Springer Nature Switzerland, 2025, pp. 125–130. doi: https://doi.org/10.1007/978-3-031-81834-9_12
3. S. Maitra, "Machine learning for fraud analytics," in *Non-Linearity in Econometric Modeling, Vol. 2: Empirical Applications and Source Code*. Cham: Springer Nature Switzerland, 2026, pp. 151–203. doi: https://doi.org/10.1007/978-3-032-16304-2_5
4. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 785–794. doi: https://doi.org/10.1145/2939672.2939785
5. J. H. Holland, "Genetic algorithms," *Scientific American*, vol. 267, no. 1, pp. 66–73, 1992
6. R. J. Bolton and D. J. Hand, "Statistical fraud detection: A review," *Statistical Science*, vol. 17, no. 3, pp. 235–255, 2002. doi: https://doi.org/10.1214/ss/1042727940
7. Y. Kou, C.-T. Lu, S. Sirwongwattana, and Y.-P. Huang, "Survey of fraud detection techniques," in *IEEE International Conference on Networking, Sensing and Control*, 2004, vol. 2, 2004, pp. 749–754. doi: https://doi.org/10.1109/ICNSC.2004.1297040
8. L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct 2001. doi: https://doi.org/10.1023/A:1010933404324
9. S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, "Data mining for credit card fraud: A comparative study," *Decision Support Systems*, vol. 50, no. 3, pp. 602–613, 2011. doi: https://doi.org/10.1016/j.dss.2010.08.008
10. S. Xuan, G. Liu, Z. Li, L. Zheng, S. Wang, and C. Jiang, "Random Forest for credit card fraud detection," in 2018 *IEEE 15th International Conference on Networking, Sensing and Control (ICNSC)*, 2018, pp. 1–6. doi: https://doi.org/10.1109/ICNSC.2018.8361343
11. S. S. Golait, R. S. Masidkar, K. S. Khobragade, P. S. Bhanarkar, P. Ganeshkar, and P. Ganeshkar, "Credit card fraud detection system using support vector classifier," in *Intelligent*

Systems for Smart Cities, A. J. Kulkarni and N. Cheikhrouhou, Eds. Springer Nature Singapore, 2024, pp. 71–86. doi: https://doi.org/10.1007/978-981-99-6984-5_5

12. E. Dumitrescu, S. Hu´e, C. Hurlin, and S. Tokpavi, "Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects," *European Journal of Operational Research*, vol. 297, no. 3, pp. 1178–1192, 2022. doi: https://doi.org/10.1016/j.ejor.2021.06.053
13. Kanika and J. Singla, "A survey of deep learning based online transactions fraud detection systems," in 2020 *International Conference on Intelligent Engineering and Management (ICIEM)*, 2020, pp. 130–136. doi: https://doi.org/10.1109/ICIEM48762.2020.9160200
14. E. Ngai, Y. Hu, Y. Wong, Y. Chen, and X. Sun, "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature," *Decision Support Systems*, vol. 50, no. 3, pp. 559–569, 2011. doi: https://doi.org/10.1016/j.dss.2010.08.006
15. A. Ali, S. Abd Razak, S. H. Othman, T. A. E. Eisa, A. Al-Dhaqm, M. Nasser, T. Elhassan, H. Elshafie, and A. Saif, "Financial fraud detection based on machine learning: A systematic literature review," *Applied Sciences*, vol. 12, no. 19, 2022. doi: https://doi.org/10.3390/app12199637
16. B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering – a systematic literature review," *Information and Software Technology*, vol. 51, no. 1, pp. 7–15, 2009. doi: https://doi.org/10.1016/j.infsof.2008.09.009
17. J. Perols, "Financial statement fraud detection: An analysis of statistical and machine learning algorithms," *AUDITING: A Journal of Practice & Theory*, vol. 30, no. 2, pp. 19–50, 05 2011. doi: https://doi.org/10.2308/ajpt-50009
18. L. Hernandez Aros, L. X. Bustamante Molano, F. Gutierrez-Portela, J. J. Moreno Hernandez, and M. S. Rodr´ıguez Barrero, "Financial fraud detection through the application of machine learning techniques: a literature review," *Humanities and Social Sciences Communications*, vol. 11, no. 1, p. 1130, Sep 2024. doi: https://doi.org/10.1057/s41599-024-03606-0
19. D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *International Journal of Surgery*, vol. 8, no. 5, pp. 336–341, 2010. doi: https://doi.org/10.1016/j.ijsu.2010.02.007
20. M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hr´objartsson, M. M. Lalu, T. Li, E. W. Loder, E. Mayo-Wilson, S. McDonald, L. A. McGuinness, L. A. Stewart, J. Thomas, A. C. Tricco, V. A. Welch, P. Whiting, and D. Moher, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, vol. 372, 2021. doi: https://doi.org/10.1136/bmj.n71
21. A. A. Almazroi and N. Ayub, "Online payment fraud detection model using machine learning techniques," *IEEE Access*, vol. 11, pp. 137 188–137 203, 2023. doi: https://doi.org/10.1109/ACCESS.2023.3339226
22. N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002. doi: https://doi.org/10.1613/jair.953

23. G. Li, "Application of XGBoost algorithm in data finance for risk assessment of financial credit," in *Proceedings of the 4th International Conference on Cognitive Based Information Processing and Applications*–Volume 3, B. J. Jansen, J. Ye, and Q. Zhou, Eds. Springer Nature Singapore, 2026, pp. 289–302. doi: https://doi.org/10.1007/978-981-95-2518-8 24
24. A. Q. Abdulghani, O. N. UCAN, and K. M. A. Alheeti, "Credit card fraud detection using XGBoost algorithm," in 2021 14th *International Conference on Developments in eSystems Engineering (DeSE)*, 2021, pp. 487–492. doi: https://doi.org/10.1109/DeSE54285.2021.9719580
25. C. Meng, L. Zhou, and B. Liu, "A case study in credit fraud detection with SMOTE and XGBoost," *Journal of Physics: Conference Series*, vol. 1601, no. 5, p. 052016, Aug 2020. doi: https://doi.org/10.1088/1742-6596/1601/5/052016
26. N. Ding, X. Ruan, H. Wang, and Y. Liu, "Automobile insurance fraud detection based on PSO-XGBoost model and interpretable machine learning method," *Insurance: Mathematics and Economics*, vol. 120, pp. 51–60, 2025. doi: https://doi.org/10.1016/j.insmatheco.2024.11.006
27. Z. Zuo, Y. Jiang, R. Liang, J. Xu, H. Jiang, S. Zhang, Y. Chen, and Y. Peng, "A Bayesian-optimized XGBoost approach for money laundering risk prediction in financial transactions," *Information*, vol. 17, no. 4, 2026. doi: https://doi.org/10.3390/info17040324
28. Y. Xiao, L. Tan, and J. Liu, "Application of machine learning model in fraud identification: A comparative study of CatBoost, XGBoost and LightGBM," *Preprints*, March 2025. doi: https://doi.org/10.20944/preprints202503.1199.v1
29. F. Almalki and M. Masud, "Financial fraud detection using explainable AI and stacking ensemble methods," *arXiv*, 2025. doi: https://doi.org/10.48550/arXiv.2505.10050
30. D.-Y. Park, I.-Y. Ko, T.-H. Lee, and J. Lee, "Federated gradient boosting for financial fraud detection: An empirical study in the banking sector," in *Proceedings of the 34th ACM International Conference on Information and Knowledge Management, ser. CIKM '25*. New York, NY, USA: Association for Computing Machinery, 2025, pp. 5089–5093. doi: https://doi.org/10.1145/3746252.3760891
31. M. Tayebi and S. El Kafhali, "A novel approach based on xgboost classifier and bayesian optimization for credit card fraud detection," *Cyber Security and Applications*, vol. 3, p. 100093, 2025. doi: https://doi.org/10.1016/j.csa.2025.100093
32. 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, Oct 2023. doi: https://doi.org/10.1007/s10796-022-10346-6
33. M. Al-Asadi, A. E. Alissa, B. Bhushan, and M. Al-Azzawi, "Enhancing financial fraud detection using xgboost and advanced data balancing techniques," in 2025 *1st International Conference on Secure IoT, Assured and Trusted Computing (SATC)*, 2025, pp. 1–16. doi: https://doi.org/10.1109/SATC65530.2025.11137062
34. I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, Mar. 2003.
35. R. Kohavi and G. H. John, "Wrappers for feature subset selection," Artificial Intelligence, vol. 97, no. 1, pp. 273–324, 1997. doi: https://doi.org/10.1016/S0004-3702(97)00043-X
36. D. E. Goldberg and J. H. Holland, "Genetic algorithms and machine learning," *Machine Learning*, vol. 3, no. 2, pp. 95–99, Oct 1988. doi: https://doi.org/10.1023/A:1022602019183

37. J. H. Holland, "The general setting," in *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. The MIT Press, 04 1992, pp. 1–19. doi: https://doi.org/10.7551/mitpress/1090.003.0004
38. E. Kocyigit, M. Korkmaz, O. K. Sahingoz, and B. Diri, "Enhanced feature selection using genetic algorithm for machine-learning-based phishing URL detection," *Applied Sciences*, vol. 14, no. 14, 2024. doi: https://doi.org/10.3390/app14146081
39. I. D. Mienye and Y. Sun, "A machine learning method with hybrid feature selection for improved credit card fraud detection," *Applied Sciences*, vol. 13, no. 12, 2023. doi: https://doi.org/10.3390/app13127254
40. R. Lingeswari and S. Brindha, "Efficient loss updated XGBoost with deep emended genetic algorithm for detecting online fraudulent transactions," *Multimedia Tools and Applications*, vol. 83, no. 37, pp. 84 471–84 494, Nov 2024. doi: https://doi.org/10.1007/s11042-024-19183-y
41. R. Lingeswari and S. Brindha, "Online payments fraud prediction using optimized genetic algorithm based feature extraction and modified loss with XG boost algorithm for classification," *Swarm and Evolutionary Computation*, vol. 95, p. 101934, 2025. doi: https://doi.org/10.1016/j.swevo.2025.101934
42. C. Li, N. Ding, Y. Zhai, and H. Dong, "Comparative study on credit card fraud detection based on different support vector machines," *Intelligent Data Analysis*, vol. 25, no. 1, pp. 105–119, 2021. doi: https://doi.org/10.3233/IDA-195011
43. J. Zhao, Y. Li, H. Jin, and X. Zhang, "Preediction of financial fraud risks among older adults: The application of GA-XGBoost model," *Innovation in Aging*, vol. 8, no. Supplement 1, pp. 429–429, 12 2024. doi: https://doi.org/10.1093/geroni/igae098.1396
44. S. Han, K. Zhu, M. Zhou, and X. Cai, "Competition-driven multimodal multiobjective optimization and its application to feature selection for credit card fraud detection," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 52, no. 12, pp. 7845–7857, 2022. doi: https://doi.org/10.1109/TSMC.2022.3171549
45. C. A. S. Assis, A. C. M. Pereira, M. A. Pereira, and E. G. Carrano, "A genetic programming approach for fraud detection in electronic transactions," in 2014 *IEEE Symposium on Computational Intelligence in Cyber Security (CICS)*, 2014, pp. 1–8. doi: https://doi.org/10.1109/CICYBS.2014.7013373
46. P. Ravisankar, V. Ravi, G. Raghava Rao, and I. Bose, "Detection of financial statement fraud and feature selection using data mining techniques," *Decision Support Systems*, vol. 50, no. 2, pp. 491–500, 2011. doi: https://doi.org/10.1016/j.dss.2010.11.006
47. Y. K. Saheed, M. A. Hambali, M. O. Arowolo, and Y. A. Olasupo, "Application of GA feature selection on Naive Bayes, Random Forest and SVM for credit card fraud detection," in 2020 *International Conference on Decision Aid Sciences and Application (DASA)*, 2020, pp. 1091–1097. doi: https://doi.org/10.1109/DASA51403.2020.9317228
48. A. Singh and A. Jain, "Hybrid bio-inspired model for fraud detection with correlation based feature selection," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 24, no. 5, pp. 1365–1374, 2021. doi: https://doi.org/10.1080/09720529.2021.1932929
49. A. M. Yokoyama, M. Ferro, and B. Schulze, "Multi-objective hyperparameter optimization approach with genetic algorithms towards efficient and environmentally friendly machine

learning," *AI Communications*, vol. 37, no. 3, pp. 429–442, 2024. doi: https://doi.org/10.3233/AIC-230063

50. B. Srikanth, S. L. V. Papineni, G. Sridevi, D. N. V. S. L. S. Indira, K. S. R. Radhika, and K. Syed, "Adaptive XGBOOST hyper tuned meta classifier for prediction of churn customers," *Intelligent Automation & Soft Computing*, vol. 33, no. 1, pp. 21–34, 2022. doi: https://doi.org/10.32604/iasc.2022.022423
51. H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning," in 2008 *IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, 2008, pp. 1322–1328. doi: https://doi.org/10.1109/IJCNN.2008.4633969
52. C. Elkan, "The foundations of cost-sensitive learning," in *Proceedings of the 17th International Joint Conference on Artificial Intelligence - Volume 2, ser. IJCAI'01*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2001, pp. 973–978.
53. J. a. Gama, I. ˇZliobaitundefined, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A survey on concept drift adaptation," *ACM Comput. Surv.*, vol. 46, no. 4, Mar 2014. doi: https://doi.org/10.1145/2523813
54. A. Sallah, E. A. Abdellaoui Alaoui, S. C.K. Tekouabou, and S. Agoujil, "Machine learning for detecting fake accounts and genetic algorithm-based feature selection," *Data & Policy*, vol. 6, p. e15, 2024. doi: https://doi.org/10.1017/dap.2023.46
55. R. M. Aziz, R. Mahto, K. Goel, A. Das, P. Kumar, and A. Saxena, "Modified genetic algorithm with deep learning for fraud transactions of Ethereum smart contract," *Applied Sciences*, vol. 13, no. 2, 2023. doi: https://doi.org/10.3390/app13020697
56. E. Ileberi, Y. Sun, and Z. Wang, "A machine learning based credit card fraud detection using the GA algorithm for feature selection," *Journal of Big Data*, vol. 9, no. 1, p. 24, Feb 2022. doi: https://doi.org/10.1186/s40537-022-00573-8