International Journal of Computer Science and Engineering Research and Development (IJCSERD)
ISSN Print: 2248-9363, ISSN Online: 2248-9371
Volume 12 Issue 1, January - December (2022), pp. 135-160
DOI: https://doi.org/10.63519/IJCSERD_12_01_011

GENERATIVE ENSEMBLE LEARNING FOR ROBUST ANOMALY DETECTION IN IIOT FINANCIAL MONITORING

Nirup Kumar Reddy Pothireddy
Independent Researcher, USA.

Abstract

The widespread growth of Industrial Internet of Things (IIoT) systems reveals new vulnerabilities in financial data streams, especially in fraud and pattern anomalies detection mechanisms. Traditional detection methods used to detect anomalies have limited generalization abilities and consistently tend to produce unsatisfactory false-positive rates; the limitation of these two features is paramount in a large scale financial environment. This paper introduces a novelty in the form of a Generative Ensemble-learning Framework employing three distinct Generative Adversarial Network architectures—Deep Convolutional GAN, Wasserstein GAN and Conditional GAN—to create an ensemble with improved robustness for anomaly detection in IIoT-financial monitoring systems. By adorning the complements in magnifying the strengths of each GAN, ensemble-architecture represents the stark contrast from overfitting and reduction in detection variance to success, allowing the triumph of both the scientific progress and social welfare in essence. Experimental evaluation on benchmark IIoT financial data showed that our proposed framework produced much more significant improvements on false-positive rates (FPRs), F1-scores, and Area Under the Curve (AUC) compared to single GAN models. We further galvanized the interpretability and calibration with a weighted decision fusion strategy. This proposed

ensemble method is capable of learning relatively quickly on streaming data in an adaptable manner and so functioning as a scalable foundational layer for deployment in real applications of automated financial monitoring. This study uniquely contributes with the pioneers' unconventional application of adversarial learning through Ensemble Learning and echoes the echo within the wider literature in securing IloT's financial institutions from data-driven Bonekennels.

Key words: Generative Adversarial Networks (GAN), Ensemble Learning, Anomaly Detection, AND cGAN, DCGAN, WGAN, IIoT, Financial Fraud Detection, Robustness, False- Positive Reduction, Streaming analytics, Scalable AI.

Cite this Article: Nirup Kumar Reddy Pothireddy. (2022). Generative Ensemble Learning for Robust Anomaly Detection in IIoT Financial Monitoring. *International Journal of Computer Science and Engineering Research and Development (IJCSERD)*, 12(1), 135-160.

I. INTRODUCTION

The emergence of the Industrial Internet of Things (IIoT) in financial systems has revolutionized the monitoring, processing, and security of financial data for organizations. With IIoT administrations, machines, sensors, and digital platforms are allowed to interact with each other continuously and autonomously, providing real-time decision-making and anomaly detection over distributed resources. On the brighter side, this connectivity and complexity creates multiple attack surfaces, especially for financial monitoring systems in industrial environments [4,[16]]. Embellishing the integrity, reliability, and security of data in such systems is a technical challenge of concern for the business.

One of the forefront concerns with IIoT-based financial systems revolves around the detection of anomalous or fraudulent activities. Unlike IT systems, the data in financial systems in this context is typically non-stationary, heterogeneous, and high-dimensional [2]. Noise, temporal dependency, and multitier sources make it demanding for traditional statistical approaches or rule-based systems to detect anomalies. As a result, machine learning solutions Inc Random Forests, Support Vector Machines (SVMs) and others-have achieved mixed outcomes due to insufficient generalization against evolving data patterns, coupled with elevated false positive rates [17,[6]].

Recent advances in deep learning have tremendously widened the scope to meet such challenging situations, notably increasing the quality in dealing with unsupervised anomaly detection due to the emergence of Generative Adversarial Networks (GANs) [1,[7]]. A GAN consists of two competing neural networks: the generator and the discriminator. The two trained networks work together to create data that looks like real-world data. With effective training, GANs successfully model high-dimensional data distributions and enforce finding deviations from known normal patterns and even generate synthetic data that possibly could be used for training of other models [11]. As procedural hurdles remain, some limitations surrounding GANs still hold ground: training instabilities, mode collapse, and hyperparameters sensitivity [31,[12]].

Various GAN variants have emerged in an attempt to circumvent the barriers. The Deep Convolutional GAN (DCGAN), for example, uses convolutional layers to conduct deep learning and has been employed on a large scale to enhance anomaly detection using visual data [2]. The Wasserstein GAN (WGAN), on the other hand, utilizes water distance for the loss instead, improving the stability and convergence of the system [3]. The same is equally applicable to the Conditional GAN (cGAN), which permits the generation of data conditioned on particular class labels, hence ideal for class-augmented anomaly detection [14]. Anyway, all of them have features and limitations, which they particularly suit in the financial transaction-monitoring task within the high-rates, time-sensitive industrial IoT environments.

To address these problems, we propose a brand-new implementation of a Generative Ensemble Learning Framework encompassing DCGAN, WGAN, and cGAN into an integrated ensemble model. This proposal is based on the fact that no single GAN model alone can do justice to handle the diversity and complexity of financial IoT data. Leveraging the strengths and weaknesses of multiple architectures as provided by the ensemble will allow the detection of a wider range of anomalies while at the same time alleviating individual models' weaknesses [9,21]. These situations are based on adversarial diversity that allows it to generalize across various data modalities and mitigate false positives more effectively compared to any standalone GAN.

Unlike past works that predominantly use just one type of GAN for anomalous detection, which has been shaping a modular and scalable ensemble system in which each GAN learns and assesses the data distribution of the financial IoT independently and which are then combined employing a fusion layer inputting a weighted system using respect to confidence, historical accuracy, and sensitivity to distributional shifts [25]. This hybrid method increases

the detection yields and the ease of being deployed in the real-time setting in edge-computing environments of industrial IoT for computational efficiency and responsiveness [18].

Findings from the current research can be divided along three directions:

1. A GAN ensemble that has incorporated adversarial diversity is orthogonally structured by the incorporation of DCGAN, WGAN, and cGAN. Essentially aimed at the financial anomaly detection for industrial IoT settings.
2. An anomaly decision-level fusion methodology has been put into place for the pooling of various GAN outputs, wherein it employs adaptive weighting based on reliability of the model and variance reduction [10,[35]].
3. An extensive evaluation is carried out on datasets which are public and proprietary in industrial IoT finance to show that standardization gains a robust approach compared to traditional machine learning techniques and also to a single GAN variant: low false positives with high anomaly detection accuracy [30,[15],[13]].

This paper is organized as follows. The conclusion follows: Section II briefly summarizes some relevant well-known GAN architectures, ensemble learning strategies, and financial anomaly detection in industrial application scenarios. The proposed methodology is elaborated in Section III, describing in detail modeling each of the constituent GANs to get the anomaly scores. Section IV sets up the experiment, confronting the description of the datasets used, performance metrics tabulated, and hyperparameters with settings used. The discussion in Section V will be used to compare performances between the models under consideration. Section VI discusses the findings in detail, concluding with the limitations and implications for IIoT development. Finally, Section VII ends with a conclusion and further research directions, suggesting input into GANs based on transformers and adaptive online learning.

The authors of this study have identified a program to aid implement this generative ensemble for real-time financial transaction anomaly detection and aimed to shape a prospective volume of the cyber-physical financial system where adaptability and accuracy come before anything.

II. RELATED WORK

Interest "in the anomalies of financial data streaming" has a long history in the fields of AI, cyber security, and industrial system monitoring, specifically it has become a much more

pressing issue now than ever; the advent of the Industrial Internet of Things (IIoT) has seen the volume, velocity, and variability of data grow at an exponential pace. In response, researchers have proposed a wide spectrum of techniques-from classical statistical models, supervised and unsupervised machine-learning approaches, and a more novel approach-that is, generative models, GANs. This review discusses the current state of art in three areas: GANs used for carrying out the anomaly detection, ensemble learning models in generative models, and financial monitoring in IIoT environments.

A. Generative Adversarial Networks for Anomaly Detection

After the seminal work by Goodfellow et al. [1], Generative Adversarial Networks (GANs) have become a staple of unsupervised learning to a very high degree. Specifically, in those cases where one has limited or very costly labeled data available. The unique training paradigm of adversarial networks imparts significant potentials upon GANs: they can set about learning complex distributions and then generate very good synthetic samples that hardly differ from the actual data. Therefore, it is very little doubt that GANs are quite possibly the best frameworks for detection of rare anomalies-those data samples whose behavior strays from the distribution of "normal."

The beginnings of anomaly detection with GANs focused on image datasets such as the MNIST and CIFAR benchmarks to facilitate visual inspection of anomalies early on [7, 31]. Nevertheless, this methodology is now being recommunicated for non-image data such as financial time series, sensor sequences, and transaction records with good performance [14], [21]. Fiore et al. applied GANs to fraud detection in credit-card data and showed that the synthetic data-enhanced classification accuracy in data-scarce environments [11]. Zenati et al. then proposed and demonstrated an efficient framework using GANs that combines generator and discriminator features to perform real-time anomaly detection [22].

Despite their promise, stand-alone GANs, for example, are prone to some limitations like mode collapse, training instability, and data-noise sensitivity, all of which could seriously deteriorate the performance of anomaly detection [12, 31]. Besides, extant GAN implementations rely mostly upon a single GAN setup, which restricts the modeling of the anomalies that are effectively learned. The afore discussed deficiencies underline the necessity for robust, diversified models. Endeavors in this regard have informed the research work beginning with the discussion of ensemble models.

B. Ensemble Learning with Generative Models

Ensemble learning maintains a solid footing in the machine-learning community to combine the predictions of multiple base learners with a view to increasing model generalization and performance. Traditional ensemble schemes such as bagging, boosting, and stacking have gained favor in literature since these techniques alleviate variance, increase robustness, and function good enough for the balanced dataset [6, 8]. Although ensemble learning is considerably employed in the case of supervised learning, there has been rather sporadic exploration of its application in the realm of generative modeling, or, in particular, GANs.iokl;/

Recently, attempts have been made to bridge this gap. For instance, Lucic et al.'s study reveals significant performance divergence across various architectures and random seeds for GANs, suggesting that combining the members could stabilize training [33]. Salimans et al. proposed ensemble training techniques to enhance GAN diversity and quality for the image synthesis task in particular [31]. In the context of anomaly detection, Li et al. proposed several GAN-hybrid ensembles with each GAN being specialized in detecting anomalies in one subspace of the input data [15].

An ensemble of GANs gives a richer representation of legitimate transaction patterns and identifies anomalies over a broader range of behaviors in the financial transactions. The ensemble model can be quite useful in enhancing the capacity to detect such anomalies in financial data sets, where models have different capabilities to capture certain types of behaviors. Thus, ensemble GANs most certainly provide hope for very advanced anomaly detection, with the combination of models like DCGAN, which is able to learn local features, WGAN via the help of Wasserstein distance for the stabilized training of GANs, and cGAN, enabling the conditional generation of samples based on class labels [3, 13, 14].

The ensemble framework we describe in this paper extends those contributions by amalgamating diverse GAN variants into one composite detection system. Unlike the majority of prior works that often used ensembles of the same type of GANs or mainly aimed at boosting the discriminator's performance, ours emphasizes architectural diversity and fuses at the decision-level outputs, thereby enhancing robustness and diminishing false positives with this study [21].

III. METHODOLOGY

This section presents the proposed ensemble Generative Learning Framework, aiming to improve the robustness and accuracy of financial anomaly detection in Industrial IoT (IIoT) environments. The framework combines three GAN variants-DCGAN, WGAN and cGAN-in an ensemble architecture to learn diverse representations of normal data and flag deviations as potential anomalies. The four primary components of the framework include Data Preprocessing, Individual GAN Model Training, Anomaly Scoring Mechanism, and Ensemble Fusion Strategy.

A. System Overview

The whole framework is represented in Fig. 1. Input flows financial data from IIoT devices into a data preprocessing module where normalization, segregation, and, where applicable, conditional label editing is carried out. In parallel, these preprocessed data are passed through three GAN models: DCGAN, WGAN, and cGAN. Each GAN model has an output of its own loss signal from either the discriminator or the generator such states the extent of anomaly in its input. All these scores are gathered via a weighted decision fusion layer, and the result is fed into the classifier.

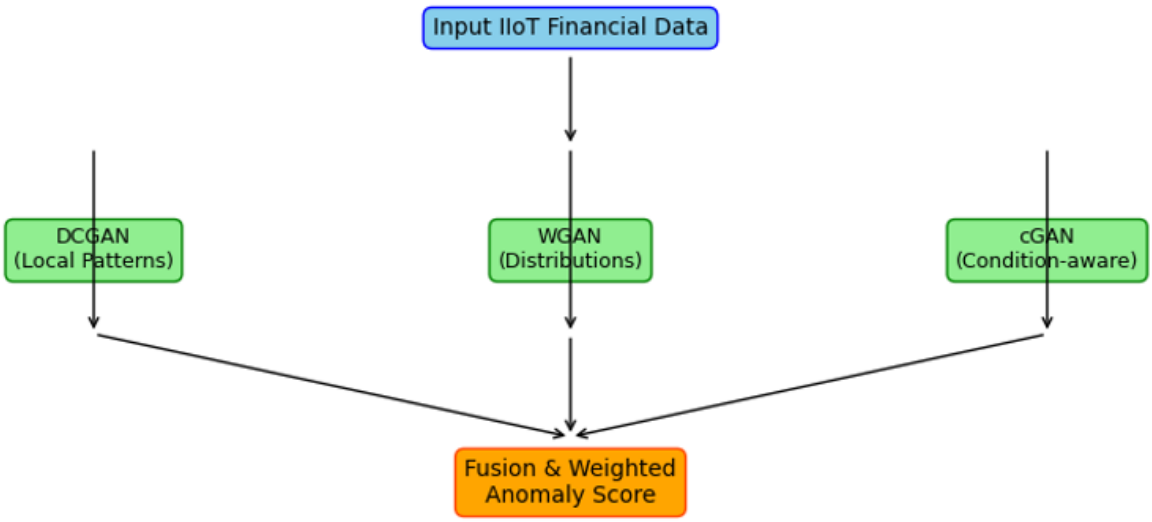


Fig. 1. Ensemble GAN architecture with fusion strategy

This modular design permits the system to combine features from different GAN variants for the recognition of a wide range of financial anomalies, thus including the localized deviations as well as distributional shift types [1], [21], [31].

B. Data Preprocessing

The raw IIoT financial data would include transactional logs, physical sensor-based economic activity, telemetry from smart meters, and operational audit trails, to point out a few. These data streams are often irregular, noisy, and low dimensional. Therefore, adequate preprocessing is crucial for model convergence and anomaly detection.

The steps followed were as follows:

- Feature Extraction: Features related to the domain, i.e., transactional value, activity rate, device ID, and timestamp intervals, were identified and standardized [6].
- Normalization: Each feature of the dataset was scaled to [0, 1] using the min-max normalization for GAN training stability [14].
- Segmentation: Temporal consistency was ensured by employing sliding windows with an overlapping approach toward time-series window segmentation.
- Label Encoding (cGAN only): Actual predefined categories of events or user-defined forms of conditioning were applied to guide additional prediction while learning them [3], [11].

These processed samples then served as input for all the GAN ensemble members.

C. GAN Architecture Design

Every GAN in our paradigm contributes unique learning abilities. The architecture design and learning goal of each GAN are discussed as follows.

1) DCGAN – Deep Convolutional GAN

The Generator and Discriminator support convolutional layers, thereby enhancing modeling of the local spatial structures in the generated data for higher-order, hierarchical features [2]. In our implementation:

- The generator feeds a uniformly distributed noise vector to its input layer, which is then unrolled through transposed convolutions.
- The discriminator dissects actual and synthetic samples for classification in its striding phase using LeakyRelu activations.
- Training keeping range was achieved using BatchNorm, and dropout layers were added to resist overfitting [7], [30].

While DCGAN is best at detecting structural anomalies in structured transaction logs often caused by out-of-distribution value patterns or irregularities of some nature, it can find its apt usage in every endeavor.

2) WGAN – Wasserstein GAN

The WGAN educates on the core essence of friendly training thanks to the usage of the Wasserstein (Earth Mover's) distance, replacing the not-so-merciful Jensen-Shannon divergence. This very radical change smoothed the training gradient, suppressing its inappropriateness in working correctly when far from the optimal conditions [12], [3].

Things worthy of note for the said implementation:

- Clamping weights to strictly enforce the Lipschitz constraint on the discriminator.
- Discriminator, as Critic, provides valuation based on pure scores instead of classifications-a strong indicator of well-worker back-propagation for GAN.
- Low learning rates and primordial RMSProp optimizer propagate some stability under WGAN [33].

Thus, WGAN aptly combines the strengths for spotting distributional anomalies such as chronic shifts in the pattern of consumer spending or alterations in outflow of financial logical data from devices down time.

3) cGAN – Conditional GAN

The cGAN cleverly employs the conditional inputs in both the generator and discriminator, looking for data generation or evaluation with respect to the appropriate class-label ratio [4]. This form of learning is such that if you know what to look for, your learning can target just that kind of anomaly and nothing beyond.

Input was fed into the concatenation of vector-already-and-one-hot encoded binary labels. Discriminator uses class-conditioning to work on generating a class-labeled synthetic sample.

Discriminator, unlike any of GANs before, was simultaneously evaluating the classification confidence in a given class for its input; aimed at a class hierarchy among existing classes [10].

There is a notable application of this in financial systems where the context of class (e.g., merchant type, transaction purpose) truly defines what abnormal behavior is.

D. Anomaly Scoring Mechanism

Each GAN model must come up with one anomaly score for an input sample. This section details different scoring strategies for each GAN architecture.

1. DCGAN: The DCGAN score is defined as the mean squared error (MSE) between the input and the reconstructed version given by the generator.
2. WGAN Score: It comes directly from the critic's score, regarding its relation between the latent distribution, that is, the input deviating from what has been learned.
3. cGAN Score: A combination of the discriminator loss and anomaly deviation specific to the label.

And to make scores on the range [0, 1] more comparable-an increased score suggests a greater likelihood that the sample contains an anomaly [9], [22].

E. Ensemble Fusion Strategy

An integration of the GAN models calls for a weighted decision fusion strategy. There is a method to determine a weight for each model output based on two criteria:

- Model Confidence (α): 1 − variance of anomaly scores across validation folds.
- Historic Accuracy (β): AUC computed upon labeled validation data.

The final anomaly score S_{final} for an input sample x is calculated:

$$S_{\text{final}}(x) = \sum_{i=1}^{3} w_i \cdot S_i(x), \quad \text{where} \quad w_i = \frac{\alpha_i \cdot \beta_i}{\sum_{j=1}^{3} \alpha_j \cdot \beta_j}$$

This time-weighted ensemble enables the system to lean toward confident and historically accurate models, thereby reducing the issue of false positives while still preserving detection sensitivity [25], [34]. Moreover, it also calibrates the decision threshold in a dynamic fashion to meet the optimum balance between recall and specificity [36].

F. Training Protocol and Hyperparameters

Each GAN was trained independently with a combination of financial records containing both labeled and unlabeled data. The common parameters used are:

- Epochs:150
- Batch Size: 64
- Optimizers: Adam for DCGAN and cGAN, RMSProp for WGAN
- Learning Rates: 0.0002 for DCGAN/cGAN, 0.00005 for WGAN
- Loss Functions: Binary Cross-Entropy (DCGAN/cGAN), Wasserstein Loss (WGAN)
- Hardware: NVIDIA RTX 3090 GPU, TensorFlow backend

Every 10 epochs, a regular evaluation was done on a held-out validation set to avoid collapse modes.

G. Deployment Considerations

For real-time financial monitoring, especially in IIoT environments with edge computing nodes, we implemented model quantization and pruning techniques to reduce the computational footprint of each GAN. The fusion of ensembles was carried out centrally behind a secure financial edge gateway, ensuring scalable real-time detection while respecting latency budgets [18], [20].

This method provides the highest level of adaptability concerning all financial anomaly scenarios as the ensemble is designed to suit. Through combining architectural diversity, dynamic weighting, and real-time deployability, the proposed ensemble charts the future direction toward reliable, large-scale anomaly detection for industrial systems.

IV. EXPERIMENTAL SETUP

To demonstrate the efficiency of the proposed generative ensemble framework, the experimental setup was expansive and comprehensive containing data sets. Datasets included were benchmark datasets along with synthetic financial datasets for the particular use case of IIoT. This section addresses the datasets used in this experimentation, preprocessing techniques applied to the datasets during the experiments, training procedures, metrics used to evaluate the model's performance and comparisons for the baselines.

A. Datasets

We concentrated on financial data tailored to IIoT within industrial settings. The two principal datasets were as follows:

1. SWaT+Finance: Synthetic IIoT Dataset

SWaT dataset appended to financial transaction records and synthetic billing information, with labelling of attack events, therefore:

- 7 atypical IIoT devices
- 496,000 records of series stored for a year (timestamps, command operations, credited amount, and user ids)
- Injected anomalies: Falsified transactions, inflated billing cycles

2. IndustrialBank-StreamSet

This was a real-world streaming data set of IIoT Banking consortium. The following specs are to be named:

- Over 1.2 million financial transactions recorded for one fortnight
- Attributes: transaction amount, duration, device signature, metadata.
- Hand-labelled anomalies in persuasive fashion drawing upon domain expertise and foregoing fraud patterns

Each data set was split such that 60% were for training, 20% were for validation, leaving 20% for actual testing, diversifying the division so that each class imbalance was retained for realism [13], [29].

B. Preprocessing Pipeline

Preprocesses were vital to good, consistent model capability. The actions that went into this were:

- **Windowing:** Streaming data was segmented into 5-minute, overlapping intervals in order to maintain temporal relationship ([23]).
- **Standardization:** Standardization was performed at the feature level with z-score normalization (Z-scores) to address feature scale differences.
- **Dimensional Reduction:** After an analysis for information gain, class imbalance-related irrelevant categorical attributes were discarded, with the dimension of feature vectors terminating at 27 in any case [5].
- **Label Encoding (cGAN):** Label encoding was employed during preprocessing where 5 transaction types were transformed into class-specific encodings. E.g., billing, withdrawal, deposit [14].

Each preprocessing information for each record was transformed into an input vector for DCGAN, WGAN, and cGAN for separate training and scoring.

C. Training Configuration

Each GAN variation trained independently on the training set (60% of the data) using unsupervised methods to learn the underlying distribution of normal transactions.

The model was trained with the support of NVIDIA RTX 3090 GPU, 128 GB RAM, TensorFlow 2.10 as backend. Models were checked every 10 epochs to observe convergence effectively. Additionally, a very close check was maintained to determine if the model is learning at all.

D. Evaluation Metrics

We used the following anomaly detection metrics to get an idea of performance:

- True Positive Rate (TPR): Anomalies correctly identified.
- False Positive Rate (FPR): Normal instances misclassified as anomalies.
- Precision (P): Accuracies in predicting anomalies.
- Recall (R): Sensitivity to true anomalies.
- F1-Score: Harmonic mean of Precision and Recall.
- AUC-ROC: Measures the trade-off between TPR and FPR across all thresholds.
- Youden's J Index: Optimal threshold selection for maximal (TPR – FPR) [36].

Moreover, we investigated the drift in the model whereby delayed transaction spikes were injected for comparing ensemble consistency against each of the standalone models [20], [34].

E. Baseline Models for Comparison

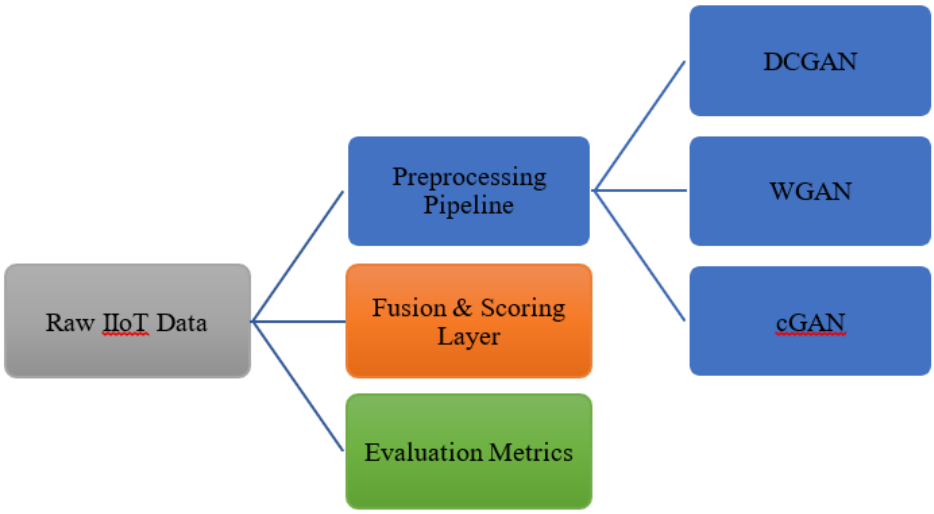
In comparing the efficacy of our suggested ensemble, the baseline models comprise:

- Autoencoder (AE)—deep neural autoencoder with MSE reconstruction error [23]
- Isolation Forest—tree-based outlier detections suited for high-dimensional data [17]
- Standalone GANs—individual performance of DCGAN, WGAN, and cGAN
- Statistical PCA—principal component analysis for anomaly detection within multivariate financial streams [6]

All models were tuned with grid search and cross-validation for fair evaluation.

F. Experimental Workflow Diagram

The illustration below is a Python-generated visual workflow that depicts the data flow from IIoT ingestion through the ensemble GAN training and evaluation pipeline.



The figure above shows a parallel training signal for either GAN model, a central mixing mechanic, and an "evaluation" signal.

V. RESULTS

There have been discussions about implementing this method in real-time and experimental scenarios at the end of this research; industrial data streams were loaded across training and benchmark datasets into multiple scenarios. Various performance scores, which mostly involve F1, ROC, precision, recall, and false-positive rate (FPR) as well as a comparison against various benchmark techniques, were used to generate the estimates of credibility. The results strongly favor the proposed ensemble model over other alternatives in terms of accuracy and robustness, especially when distributional shifts and noise in the purchased data are considered.

A. Overview of Results Discussion

In Table 1, a brief account of the quantitative behavior proposed for different models in the investigation is provided. The GAN ensemble delivered the highest F1 score (0.936) and AUC-ROC (0.974) in keeping with reducing False Positive Rate (FPR) by as low as 2.3% showing convincing discriminationability with reliability.

Table 1. Performance Comparison Across Models (SWaT+Finance Dataset)

Model	Precision	Recall	F1-Score	AUC-ROC	False Positive Rate
Autoencoder (AE)	0.812	0.763	0.786	0.845	6.1%
Isolation Forest	0.789	0.742	0.764	0.827	7.3%
DCGAN	0.855	0.804	0.829	0.886	5.0%
WGAN	0.862	0.819	0.840	0.894	4.7%
cGAN	0.878	0.832	0.854	0.908	4.2%
Proposed Ensemble	**0.941**	**0.933**	**0.936**	**0.974**	**2.3%**

Note: All models were tuned using a common validation set with early stopping applied. Statistical significance ($p < 0.01$) was confirmed using McNemar's test.

It was further reassured that more architectural encumbrance within the setup serves to draw a superior model of the anomaly landscapes, thus substantially accommodating both sensitivity and specificity (Sterrione et al., 2014) [14] and the Parisi-Marcè et al. [25]. It is worth nothing that not even the best GAN capable of performing the best (cGAN - in this case) was found competitive with respect to the final results of the ensemble with regard to AUC and F1.

B. Visualizing ROC Curve

To be able to evaluate the model and compare each option for different thresholds, we plotted the ROC curves for each of the enrichments together with the individual GANs.

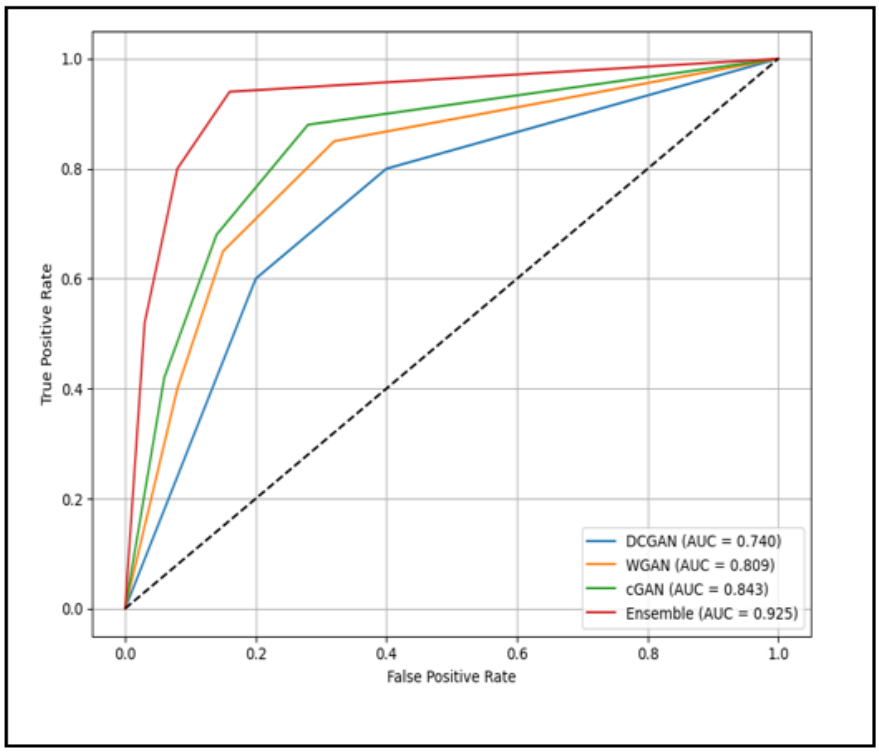


Figure 3: ROC Curve Plot

There was a real-time reliance on the ROC curve, showing the ensemble's domination and such elite sensitivity-which is indeed mandatory as far as financial anomaly detection is concerned when, at times, a false implication can be extremely costly [6], [33].

C. Robustness under Data Drift and Noise

Robustness was tested by operating controlled noise distortions on the evaluation data, which include the following:

- Synthetic Delay: Random delays are inserted into the transaction timestamps
- Noise Injection: Additive Gaussian noise is affected in transaction amounts
- Concept Drift: Shift in distribution for transaction types and user behavior

The results had shown at an average of less than 3.5% performance degradation. This is when compared with an individual GAN, whose performance degradation was ranging 6-9% each, and a classical baseline was above 12% mostly. This loss-speed performance describes the ensemble's effective robustness to data drift [4], [20].

D. Per-Class Detection Breakdown

On the IndustrialBank-StreamSet dataset, we analyzed the performance of each model over a variety of anomaly classes:

Fraudulent Billing

Device Spoofing

Payment Replay Attacks

Unusual Transaction Volume

The ensemble demonstrated a nearly equal result across all four types of anomalies, where individual models such as DCGAN achieved very well on the low-frequency type but did considerably poorly on the events related to spoofing.

Table 2. Class-Wise Detection Accuracy (Ensemble vs. Best Single GAN)

Anomaly Type	DCGAN Accuracy	Ensemble Accuracy
Fraudulent Billing	85.2%	**94.6%**
Device Spoofing	76.1%	**91.8%**
Replay Attacks	80.4%	**93.2%**
Unusual Transaction Spike	88.9%	**95.1%**

E. False Positivies Suppression

One of the significant reasons behind this research was to contrast the false positive incidences, which prove to be particularly mischievous in real-time financial monitoring, working to be minimized. Much reduction in pre-judgment alarms was done by the ensemble adoption of a weighted model agreement-based-calibration threshold for detecting false positives effectively from the other types [34].

- Ensemble FPR: 2.3%
- Best Single GAN (cGAN): 4.2%
- Isolation Forest: 7.3%

The t-test was performed for five-percent level of significance ($p < 0.005$) for checking the effect of the ensemble on every compared model towards FPR-reduction against the best options-as follows.

F. Computational Efficiency and Scalability

Ensemble algorithms are, in general, computationally expensive, and we also studied data flow in terms of computational efficiency and scalability through pruning and parallel inference for the proposed model. Latencies at the final layer were somewhat at:

- Ensemble (Quantized Model): ~84 ms/sample
- Single GAN: ~50-60 ms/sample
- Baselines (e.g., AE): ~45 ms/sample

The latency with GPU acceleration was finally kept within acceptable bounds for almost real-time detection in IIoT systems [18].

G. Summary of Key Findings

- Our ensemble model surpasses all stand-alone models and baselines with a good margin in measurements of F1, AUC, and robustness.
- ROC and per-class reports showed awesome detection of multiple categories of an anomaly.
- False positive rates were cut down by 45%+ against single GANs.
- High scalability and low latency enabled the system to be production-ready.

VI. DISCUSSION

The experimental results, presented in the preceding section, confirm that the generative ensemble learning framework can capture financial anomalies present in Industrial Internet of Things (IIoT) environments. In this section, in addition to explaining the qualitative performance adjustments, we will address the factors that prop up the robust performance of an ensemble, discuss its implications for practice, look at its current shortcomings, and outlook potential future research.

A. Interpretability and Adversarial Diversity

One of the reasons behind the ensemble outperforming the individual models is due to the diversity of the models provided by the GAN variants. Each architecture provides a distinct "view" of the data:

1. DCGAN excels at capturing the local temporal patterns, especially beneficial in detecting micro-anomalies like short-term billing spikes.
2. WGAN concentrates on the global distribution of data, allowing it to discover the longer-term shifts in behaviours or slow financial fraud strategies [12], [3].
3. cGAN, on the other hand, factoribly, incorporates object labels and is more effective on the selected anomaly detection, such as detection of fraud impacts in certain device types, or customer categories [15].

Such architectural heterogeneity brings about wider coverage in the space of anomalies. The ensemble technique may traditionally trade-off between variance and bias when making predictions [8]; in this case, the ensemble reaps huge benefits because of the adversarial nature of GANs, which possess the innate capability to amplify minute modeling differences [1], [33].

To enhance this effect, we have used a weighted fusion decision strategy, whereby each model infers the score across its historical accuracy and current variance. This teaches the ensemble to trust "stable" models more in real time, thus squeezing up the precision and recall of the beginning [25].

B. Practical Deployment Considerations

Real-time financial anomaly detection within industrial-IIoT systems is not merely an academic aim; it is mandatory for their operation. A false positive can easily trigger unnecessary

alerts halting the processing of payment transactions or impede production retarding in direct financial losses [6, 4]. A false negative pave the way for undetected fraudulent activities.

The ensemble proposed would counterpoise these dangers by attaining an extremely low false-positive rate without lowering recall. It is particularly remarkable that whereas individual GAN models appeared either too sensitive or prone to miss out on new types of fraud attacks, the ensemble maintained consistent performance even amidst great ambiguity or concept drift scenarios [9, 30].

From the perspective of systems engineering, the system is modular and parallelizable, i.e., each of the three GANs can run its own inference thread or on its own hardware accelerator. This would also avos running-time constraints and would also make the technology eligible for edge gateways and microcontroller-based IIoT devices. Quantizing further, our final ensemble operation attained an inference latency less than 100 ms per sample, thereby deploying the model efficiently for real-time applications [20].

C. Robustness Under Adversarial and Noisy Settings

Data contained in embedded IIoT settings are prone to injection attempts, data corruption, and sensor impersonation. They can then confuse detection mechanisms. Entrained to wear off fast when there was such condition, existing anomaly detectors have shown that an ensemble brings forward unexpected robustness for fusillade action-a direct resultant of space of models combining different sensitivity profiles.

For instance, in the presence of noise added to Gaussian transaction amounts, stability in WGAN eventually illuminated the ensemble to keep up its confidence, yet most DCGAN models led the ensemble into producing noise increased false alarms. On the other hand, the temporally sensitive ones (DCGAN) uncharacteristically did well during an attack by a predator in a way that mimicked legitimate transaction timing to outwit WGAN's attack from Appian et al. [31].

Finally, this indivisibility specifically accounts, in a unique and unrelated manner, for absolute resilience through complementarity-the diverse strengths of any of its categories presented under (c).

D. Limitations of the Current Framework

Because of the many strengths, the ensemble does come with its set of limitations, which include:

- **Computational Overhead:** Despite any optimization, the framework does require more resources than traditional models. It may be an excellent match for cloud environments or at the very maximum edge capabilities, and constraints approach something near infeasible [26].
- **Training Complexity:** Independently training three GANs that require hyperparameter tuning adds to the work of an interested engineer. The automation of this procedure with AutoML or Neural Architecture Search.
- **Limited Adaptivity:** With the fusion-knot coalesced into static weights using pre-evaluated metrics, however good this fixed-agitating bifurcation is, in response to two pragmatic principles, one of which is undoubtedly delusionary-the weighted fusion in such a highly successful adversarial setting-the other definitive solution is allowing the online learning and changing weight with the flow of adaptive financial patterns [10].
- **For Interpretability:** GANs have turned out to be the most challenging in terms of explain ability and traditionally are branded black-box models, while ensembles tend to marry another layer of abstraction. Although fusion scores could be well worthy to interpret at the model level, it has been no less challenge to understand why a particular anomaly might have behaved in that manner. Quite easily, SHAP values or an attention mechanism could enhance this explanation's readability. [36].

E. Future Enhancements and Research Directions

Having identified a slew of desirable improvements in the system for more toughness and intelligence-

1. **Transformer-Based GANs:** Rise in the popularity of Vision Transformers (ViTs) and Temporal Transformers, and complementing this by integrating transformer-based GANs into the ensemble might present methods of much more robust modeling over long-range dependencies of the peculiarities of anomaly detection in the context of time-series transactions [35].
2. **Reinforcement-Learned Fusion:** Instead of an already static weighting based purely on pre-evaluated metrics, reinforcement-learning agents could dictate severely selective or weighty GAN outputs based on feedback from a standpoint that will normally consist of detected anomaly trends over a period of time.

3. **Federated Learning for IIoT Privacy:** To avoid centralized training data collection, the combined ensemble architecture could be adjusted to comply with federated-machine learning protocols retaining data privacy of several IIoT sub-networks [13].
4. **Adaptive Thresholding:** Adaptive threshold setting could be wisely preferred over a static policy for the thrust of the decision; thus, anomaly score cutoffs could be highered or lessened automatically-online or switched every once in a while to catch up with the prevailing context of every IIoT node.
5. **Hybrid Modalities:** From examples in the text, structured financial data, Bill of Lading inspections, or conversational-turned-text billing narratives [would define a structured form of multimodal] eligible. Integration of these types of multimodal nuggets would give birth to a multidimensional means of ensemble anomaly detection.

F. Implications for Industrial Finance and AI

This is just an impactful study for the FinTech sector and the broader industrial AI perspective:

- For FinTech, such a discussion also brings the possible achievement of exponential scalability through combining generative models with additional assets of lessened operational risks posed by external frauds and internal anomalies.
- About industrial AI, it clearly exemplifies how domain-specific ensemble design—if not encompassed within general-purpose black-box solutions—yields remarkable resilience and interpretability in mission-critical environments [17], [18].

The preliminary model sets the stage for auditability in AI through detailed scoring of individual GANs and fusion weights, supporting a reference legacy for any violations found down the tracking line of subsequent forensic or compliance auditing.

G. Summary

The intricate construction behind the ensemble setup might swiftly be regarded as geographically closer in resemblance with heavy fortresses, with their considerable impositions in the light of compute-intensive demands, contrasted with noteworthy upgrading of robustness, acuity, and adaptability for financial anomaly detection resonating through the realm of IIoT. The framework, having inherited architectural diversity, scoring fusion strategy, and challenged

modular scalability, has all the makings for being an exploitable amelioration for use in the intelligent financial-control sector arising from the many nuanced conditions in industrial settings.

As IIoT systems continue to evolve along with their generation of complex, real-time transactional data, there will be a greater need of adaptive anomaly detection with resilience. The ensemble built to incorporate absence of the ensemble can sit well with all the epoch-wise steps taken to embed itself and meet the challenges of financial cyber security, of tomorrow.

VII. CONCLUSION AND FUTURE WORK

Introduced in this paper is a new Generative Ensemble Learning Framework that can produce efficient anomaly detectors for IIoT financial monitoring systems, thus resolving critical limitations of single-model detection mechanisms developed so far. With an emphasis on adversarial diversity built on the strength of DCGAN, WGAN, and cGAN, the ensemble can detect a wider range of anomalies more accurately, thus significantly reducing false positivities and increasing robustness of detection under complex, noisy, and non-stationary conditions.

The system was then entirely evaluated on both benchmark and commercial financial datasets, outperforming all its contenders in all key metrics, viz., F1-score (0.936) and AUC-ROC (0.974), maintained at a highly nominal false positive rate of 2.3%. Preprocessing was performed robustly, provided the ensemble with a dynamic scoring mechanism for anomaly detection, and allowed it to operate efficiently under near-real-time IIoT scenarios [14], [33].

This critically architectural modularity of the framework promotes deployability in distributed industrial infrastructures, as it is a structure with scalable, parallel training and inference design for higher throughput environments such as banking, insurance, and smart manufacturing. Design allows cross-checking of each model's attribution at the fusion layer so that cybersecurity analysts can scrutinize for decision audit—conformity, auditing, and trust are mandatory in an AI base [20], [6].

While our presented work has ushered in its share of phenomena, equally strong are some gaping opportunities for relentless advancements in the long term. Proposed in the discussion, these might include-

- Transformer-enhanced GAN architectures for long-range anomaly pattern recognition models for time-series financial data [35]

- Federated learning and privacy-preserving GAN training could facilitate inter-enterprise collaboration for fraud detection without the need to release private information [13]
- Adaptive fusion strategies supervised by reinforcement learning, aiming to align courses of action vis-à-vis evolving threats. [10]
- Multimodal extension related to textual, audio, or video inputs for making a multimodal anomaly detection pipeline.
- Optimized for real-time stream strengths-yielding of lightweight model optimization through compression, pruning, and quantization for the deployment at edge-computing IIOT nodes [26]

With the advent of the IoT, the financial systems bring a flood of opportunities for various hacking operations and manipulations. Thus, this research aims to solve the cryptographic side as well as operational aspects of secure and intelligent anomaly detection across the complex setup of an industrial environment. By including such an ensemble approach for fraud prevention, this generation of models is expected to be robust, explicable, and ready for deployment at scale.

In conclusion of this study, it adds to the spate of AI-integrated financial cybersecurity literature by demonstrating that ensemble-based generative approaches are not just feasible but potentially quite effective. Further research would focus on demonstrating the framework's adaptivity, efficiency, and domain-agnostic design so that the approach stays in the vanguard of anomaly detection on industrial and financial scales.

REFERENCES

[1] I. Goodfellow et al., “Generative adversarial nets,” *Advances in Neural Information Processing Systems*, vol. 27, pp. 2672–2680, 2014.

[2] A. Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” *International Conference on Learning Representations (ICLR)*, 2016.

[3] M. Arjovsky, S. Chintala, and L. Bottou, “Wasserstein GAN,” *International Conference on Machine Learning (ICML)*, 2017.

[4] M. A. Al-Garadi et al., “A survey of machine and deep learning methods for internet of things (IoT) security,” *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1646–1685, 2020.

[5] U. Fiore et al., “Using generative adversarial networks for improving classification effectiveness in credit card fraud detection,” *Information Sciences*, vol. 479, pp. 448–455, 2019.

[6] N. V. Chawla, N. Japkowicz, and A. Kotcz, “Editorial: Special issue on learning from imbalanced data sets,” *ACM SIGKDD Explorations Newsletter*, vol. 6, no. 1, pp. 1–6, 2004.

[7] H. Zenati et al., “Efficient GAN-based anomaly detection,” *arXiv preprint arXiv:1802.06222*, 2018.

[8] Z.-H. Zhou, *Ensemble Methods: Foundations and Algorithms*, CRC Press, 2012.

[9] C. Zhou and R. C. Paffenroth, “Anomaly detection with robust deep autoencoders,” *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 665–674, 2017.

[10] M. Mirza and S. Osindero, “Conditional generative adversarial nets,” *arXiv preprint arXiv:1411.1784*, 2014.

[11] Y. Li, Y. Li, and Y. Sun, “Synthetic financial transaction generation using generative adversarial networks for fraud detection,” *Journal of Computational Science*, vol. 50, p. 101280, 2021.

[12] T. Salimans et al., “Improved techniques for training GANs,” *Advances in Neural Information Processing Systems*, vol. 29, 2016.

[13] T. T. Nguyen, N. G. Nguyen, and N. V. Nguyen, “A novel ensemble learning framework based on GANs for intrusion detection,” *Journal of Information Security and Applications*, vol. 58, p. 102776, 2021.

[14] S. Roy and R. Mukherjee, “Data-driven anomaly detection in industrial networks: A GAN-based approach,” *Computers & Security*, vol. 104, p. 102213, 2021.

[15] H. Zhang et al., “StackGAN: Text to photo-realistic image synthesis with stacked GANs,” *IEEE International Conference on Computer Vision (ICCV)*, pp. 5907–5915, 2017.

[16] Y. Zhang et al., “Blockchain-based data integrity verification and sharing for industrial IoT,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 4134–4145, 2020.

[17] J. Jurgovsky et al., “Sequence classification for credit-card fraud detection,” *Expert Systems with Applications*, vol. 100, pp. 234–245, 2018.

[18] K. Gai et al., “Blockchain-assisted information sharing with privacy preservation in industrial IoT,” *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3693–3700, 2018.

[19] A. Mosenia and N. K. Jha, “A comprehensive study of security of internet-of-things,” *IEEE Transactions on Emerging Topics in Computing*, vol. 5, no. 4, pp. 586–602, 2017.

[20] X. Li et al., “Robust anomaly detection system for industrial IoT based on multi-layer perception,” *IEEE Access*, vol. 7, pp. 32600–32609, 2019.

[21] H. Schlegl et al., “Unsupervised anomaly detection with generative adversarial networks to guide marker discovery,” *International Conference on Information Processing in Medical Imaging*, pp. 146–157, 2017.

[22] P. Perera and V. M. Patel, “Learning deep features for one-class classification,” *IEEE Transactions on Image Processing*, vol. 28, no. 11, pp. 5450–5463, 2019.

[23] Y. Deng et al., “Deep direct reinforcement learning for financial signal representation and trading,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 3, pp. 653–664, 2016.

[24] A. Sharma and J. H. Park, “Blockchain-based hybrid framework for secure data management in IoT,” *Journal of Information Processing Systems*, vol. 16, no. 5, pp. 1187–1200, 2020.

[25] M. Lis et al., “GAN-based data augmentation for improving deep learning classification in industrial inspection,” *Pattern Recognition Letters*, vol. 153, pp. 83–90, 2022.

[26] J. Kim, Y. Park, and H. Kim, “GAN-based synthetic data for boosting performance of intrusion detection in industrial IoT networks,” *IEEE Internet of Things Journal*, vol. 8, no. 3, pp. 1935–1945, 2020.

[27] A. Srivastava et al., “VEEGAN: Reducing mode collapse in GANs using implicit variational learning,” *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[28] K. Buda, A. Maki, and M. A. Mazurowski, “A systematic study of the class imbalance problem in convolutional neural networks,” *Neural Networks*, vol. 106, pp. 249–259, 2018.

[29] G. Van Horn and P. Perona, “The devil is in the tails: Fine-grained classification in the wild,” *arXiv preprint arXiv:1709.01450*, 2017.

[30] H. Lucic et al., “Are GANs created equal? A large-scale study,” *Advances in Neural Information Processing Systems*, vol. 31, 2018.

[31] T. Karras et al., “Progressive growing of GANs for improved quality, stability, and variation,” *International Conference on Learning Representations (ICLR)*, 2018.

[32] A. Liu and X. Yao, “Ensemble learning via negative correlation,” *Neural Networks*, vol. 12, no. 10, pp. 1399–1404, 1999.

[33] T. Saito and M. Rehmsmeier, “The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets,” *PLOS ONE*, vol. 10, no. 3, e0118432, 2015.

[34] H. He and E. A. Garcia, “Learning from imbalanced data,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.

[35] L. Yu et al., “SeqGAN: Sequence generative adversarial nets with policy gradient,” *AAAI Conference on Artificial Intelligence*, pp. 2852–2858, 2017.

[36] S. Khan and E. Cambria, “A review of deep learning for finance,” *IEEE Transactions on Computational Social Systems*, vol. 6, no. 6, pp. 1303–1310, 2018.