Comparative analysis of tabular Generative Adversarial Network (GAN) models for generation and validation of power grid synthetic datasets

Darshana Upadhyay
Faculty of Computer Science
Dalhousie University
Halifax, NS, Canada, B3H1W5
darshana@dal.ca

Qiaodan Luo
Faculty of Computer Science
Dalhousie University
Halifax, NS, Canada, B3H1W5
qiaodan.luo@dal.ca

Jaume Manero
Computer Science Department
Universitat Politècnica de Catalunya
Barcelona, Spain, 08034
jaume.manero@upc.edu

Marzia Zaman
Research & Development Department
Cistel Technology
Ottawa, ON, Canada, K2E7V7
marzia@cistel.com

Srinivas Sampalli
Faculty of Computer Science
Dalhousie University
Halifax, NS, Canada, B3H1W5
srini@cs.dal.ca

IEEE EUROCON 2023 - 20th International Conference on Smart Technologies | 978-1-6654-6397-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/EUROCON56442.2023.10199093

***Abstract*—The demand for securing SCADA (Supervisory Control and Data Acquisition)-based power grid systems from cyber-attacks has been increasing significantly in the last few years. Current research trends widely adopt Machine Learning (ML) techniques to prevent attacks against such critical infrastructure. However, the efficiency of these techniques largely depends upon the availability of large datasets. Acquiring large data from such critical systems is not always feasible and this has inhibited the research progress in the development of advanced ML algorithms that can make a notable difference in the prediction of malicious events. Thus, there is a strong need for generating large synthetic yet realistic datasets from existing small datasets. This paper presents a comparative analysis of tabular Generative Adversarial Network (GAN) models for the generation and validation of synthetic datasets from existing datasets of power grids. Moreover, the synthetic datasets are validated using statistical analysis, and machine learning efficacy. These synthetic datasets open opportunities for the research community to explore advanced machine learning and deep learning methodologies for the protection of industrial systems.**

***Index Terms*—Power Grids, SCADA, GAN models, synthetic data, machine learning, cyber attacks, intrusion detection systems**

I. Introduction

Power grids are the basic infrastructure that supports our economies and livelihood by providing and maintaining a continuous supply of electricity. SCADA (Supervisory Control and Data Acquisition) systems play an important role in power grids by monitoring and controlling the system components such as motors, PMUs (Phase Measurement Units), actuators, sensors, etc., from remote locations. As illustrated in Figure 1, the typical SCADA architecture for a power system is divided into four major parts, namely, power stations, communication networks, sub-stations, and SCADA control center. The electrical operators and technicians have the flexibility to control and monitor the power station components remotely using a SCADA system.

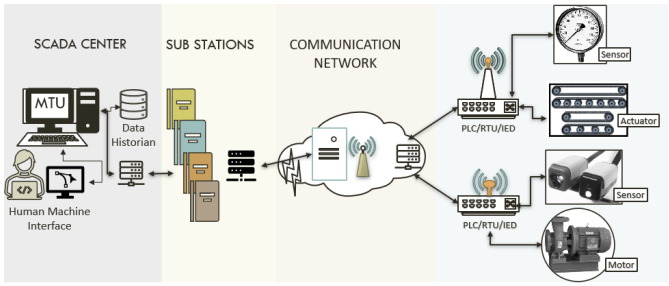


Fig. 1: SCADA System Architecture for Power Grids
Legend: PLC: Programmable Logic Controller, RTU: Remote Terminal Unit, IED: Intelligent Electronic Device, MTU: Master Terminal Unit

However, this advancement opens up power systems to the Internet and exposes them to various cyber attacks, such as Denial of Service (DoS), Man-In-the-Middle (MItM) and False Data Injection (FDI) attacks [1]. There have been several real cases of cyber attacks on power grids in recent years. For example, in 2015, owing to a cyber attack on a Ukrainian power plant, nearly 250,000 people were left without electricity for many hours [2], [3]. Similarly, in 2019, the operators of a power grid center in the USA lost communication with multiple substations for about 10 hours as intruders exposed a firewall vulnerability [4]. Such attacks can have catastrophic consequences. Enhancing performance along with the security of power grids has been one of the greatest challenges in the past few decades [5]. This includes various aspects such as incorporating efficient monitoring and controlling of the

978-1-6654-6397-3/23/$31.00 ©2023 IEEE

components at the power plants, adopting expansion plans, adding renewable resources, vulnerability analysis and risk assessment, and the deployment of intrusion detection and prevention systems [6]. Such protection mechanisms need large datasets to train advanced models for effective data analysis and monitoring of real network traffic. Therefore, real-time traffic monitoring by incorporating an advanced Intrusion Detection System (IDS) is critical to the safe operations of power grids.

In the past few years the application of Machine Learning (ML) and Deep Learning (DL) to protect SCADA systems from malicious events and intrusions has shown promising results [7], [8]. However, the development of these approaches faces a structural issue as they require large quantities of data for training. Moreover, acquiring a large amount of data is always not feasible for such critical systems and that leads the research community to generate a synthetic dataset.

Two major deep learning approaches have been used to generate the synthetic datasets, namely, Autoencoders (AE), and Generative Adversarial Network (GAN) Models. These techniques rely on neural network composition. Moreover, the quality of the synthetic dataset can be improved using hyperparameter tuning or adding new features methodology. The primary objective of this paper is to generate a synthetic dataset for a power grid system, (created at the Oak Ridge National Laboratories (ORNL) [9]) using four tabular GAN models and comparing the efficiency of GAN models by validating the quality of the synthetic data by statistical analysis and machine learning modeling. The main contributions of the paper are the following:

1) We use four GAN models to generate the synthetic dataset of power grid systems, namely, CopulaGAN, Conditional Tabular Generative Adversarial Network (CTGAN), Triplet-Based Variational Autoencoder Generative Adversarial Network (TVAE-GAN), and GaussianCopula GAN.
2) We apply grid search to identify hyper-tuning parameters for efficient modeling. Moreover, the quality of the synthetic dataset has been improved using post-processing by converting floating point zeros to zeros based on the type of features.
3) We compare the efficiency of four tabular GAN models by validating the quality of synthetic data using statistical methods and machine learning modeling, before and after hyper-tuning and post-processing.

The rest of the paper is structured as follows. Section II provides the background on GAN models and validation techniques. The proposed approach is described in Section III. Section IV covers the experimental results and discussions based on a comparative analysis of four tabular GAN models. Concluding remarks are provided in Section V.

II. Background

A. Generative deep learning

Deep learning can be applied to many other tasks, one of them being the generation of synthetic data, from tabular data to highly unstructured elements like images. Generative applications can be built using several well-known constructions, like autoencoders, variational autoencoders, or adversarial networks.

Generative Adversarial Networks (GAN) can be traced back to 2016 in the article by Radford et al. [10] and have shown robust results with images, which is a heavily unstructured datatype. However, the use of this deep learning architecture has been extended to other kinds of data, such as tabular structured data for the medical sector [11]. From this experience, other applications to new fields have been developed, and specifically for synthetic network data generation, an area where access to real data is always difficult [12], [13]. A GAN comprises two models, namely, generator G and discriminator D. The generator model generates the synthetic data, while the discriminator model is used to improve the training efficiency of the generator by comparing features of synthetic data with real data.

B. Validation Techniques

For proper validation of synthetic datasets and to measure the efficiency of tabular GAN models, we have mainly focused on two techniques, namely, statistical analysis, and machine learning modeling.

1) Statistical methods: The validation of the quality of the synthetic dataset of power grids is carried out using a total of six statistical techniques, namely, the Kolmogorov-Smirnov (KS) complement test, boundary adherence test, range coverage test, statistic similarity test, Chi-squared (CS) test, and pairwise similarity. A description of each of these statistical tests is given below. Researchers can explore [14] for more details.

- KS Complement Test: This test uses the Kolmogorov-Smirnov statistic property that converts numerical distribution into its cumulative distribution function (CDF) according to the given input data. The difference between the CDF of the real dataset and the synthetic dataset computes the quality of the dataset. The CDF is a value between 0 and 1. Synthetic Data Metrics (SDMetrics) represent the statistics by inverting the value of CDF that is $(1 - CDF)$ which indicates the score of this test. A high score reflects high-quality data.
- Boundary adherence test: This statistical metric measures whether the synthetic feature belongs to the minimum and maximum values of the same feature of the real dataset or not. This test defines the test boundary between 0 and 1, where 1 indicates the best case and 0 indicates the worst quality of the data.
- Range coverage test: This test covers the validation of the synthetic dataset by computing the full range of the values of the given data column. If the value is presented in the real column then the range coverage test is successfully passed.
- Statistic similarity test: This metric measures the similarity between a real column and a synthetic column by

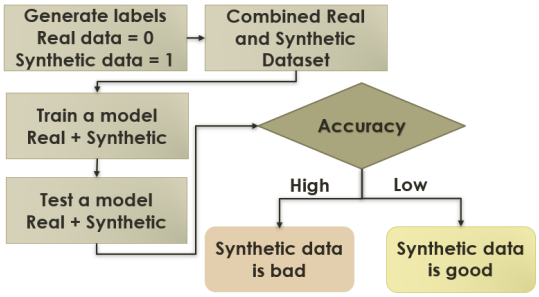


Fig. 2: Machine Learning Detection Model

comparing a summary statistic such as mean, median, and standard deviation.

- Chi-squared (CS) test: This metric is meant for categorical and boolean data types to assess whether the given synthetic data is having same frequencies as real data or not.
- Pairwise similarity: This test is performed to measure the correlation between a pair of numerical columns to compute the similarity between the real and synthetic data.

2) ML Detection model: The objective of the proposed ML modeling technique is to generate a detection score by comparing values of features of real and synthetic datasets. As shown in Figure 2, we created a dataset by merging real and synthetic datasets and added a feature, which labels the nature of the data, to the resulting dataset. The records of the real dataset are defined using the class label "0" and synthetic records by class label "1". After that, we trained classification models using the combined dataset to test whether the models were able to differentiate between real and synthetic data. Our goal is to achieve a test accuracy that is close to 50%. Such accuracy is a sign of a synthetic dataset with high quality because the classification models cannot differentiate between real and synthetic datasets.

III. Proposed Framework

This section describes the proposed framework and method implementation to measure the quality of the synthetic dataset of the power systems using various GAN Models. The entire process is divided into four steps, namely, preprocessing of the real dataset, synthetic data generation using default parameters of GAN Models, validation of synthetic dataset by comparing with a real dataset, and improving the quality of the data using hyper tunning and post-processing. Figure 3 shows the block diagram of the proposed approach to generate and validate the synthetic dataset for power grids using various GAN models. During the pre-processing of the real dataset, feature mapping and feature normalization are applied for streamed and sanitized data. Moreover, a data cleansing operation is done to remove the incorrect data. We have removed 8 rows (#51132 to #51139) of the dataset during this phase which reflects the incorrect format of the column R4:S as it can mislead the GAN models during the data generation process. Once the data is pre-processed, the synthetic dataset is generated using the GAN model with default parameters. We have generated four synthetic datasets using the four tabular GAN models to compare the techniques in terms of quality. The validation is carried out by comparing the synthetic dataset with the real dataset. Further, to improve the quality of the dataset, we have hyper-tuned the GAN models using grid search. Moreover, post-processing is applied by converting floating point zeros to zeros for some of the features to maintain the data similarity. In an electrical system during an anomalous event (due to a fault by natural or cyber attack), some measurements are likely to be absolutely zero, not a small number. However, it is not feasible to program the existing GAN to set constraints based on domain knowledge. This could be the future scope of the research. Once post-processing is done, we again compared the synthetic dataset with the real dataset using validation parameters and identified the efficient GAN model and final synthetic dataset.

TABLE I: Hyper-tuning parameters for tabular GAN Models

GAN Models	Hyper-tuning parameters
CopulaGAN, CTGAN, TVAE-GAN	epochs=500, batch_size=500, embedding_dim=256, generator_dim=(512, 512), discriminator_dim=(512, 512), generator_lr=0.0003, discriminator_lr=0.0003, discriminator_steps=5

IV. Experiments and Results

A. Power grids dataset

1) Real Dataset: The dataset used for this experimental study is the publicly available dataset generated at the ORNL laboratory on a small power grid testbed [9]. The features used in this dataset include application measurement and network-based parameters such as synchrophasor measurements, voltage and phase angles at various locations of the IEDs, relays, readings from snort and Syslog analyzer, etc. The binary dataset is used for this experiment which consists of a total of 128 features with 22,714 normal events and 55,663 attack events. The detailed description of the power systems dataset and its features are available in [15].

2) Synthetic Dataset: We have generated the synthetic datasets from the existing original dataset for the power system application using four GAN models, namely, CopulaGAN, CTGAN, TVAE-GAN, and GaussianCopula GAN. A similar implementation strategy has been followed to generate the synthetic data using each of these four GAN models. The Synthetic Data Vault (SDV) Python library provides the utility of a GAN-based DL data synthesizer for such tabular datasets. First, we created an instance of GAN and fit that instant for the given real dataset. The GAN model learns the statistical properties of the real dataset during the training process and accordingly generates synthetic data that captures the characteristics of the model. To increase the performance of the model, we have used hyper-tuning parameters of the GAN model. The grid search is applied to tune the parameters of

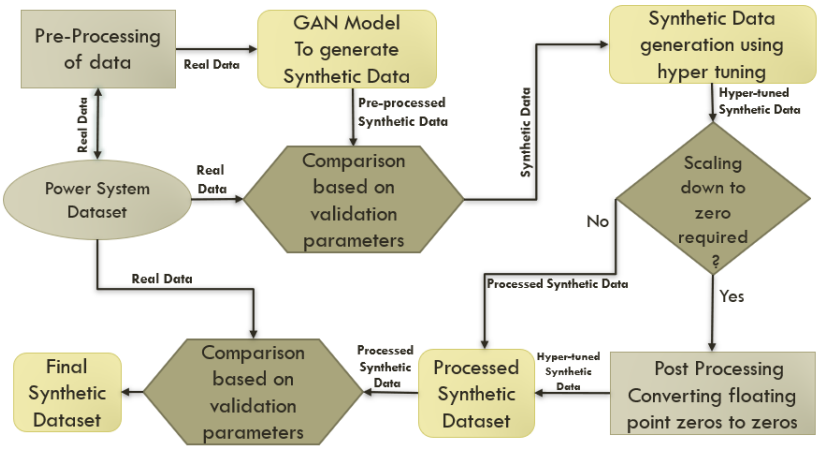


Fig. 3: Process Diagram of the proposed approach to generate and validate the synthetic data of power grids

the GAN model. Table I represents the hyper-tuning parameters that we used during the implementation. The utility of hyper tunning is available for three GAN models, namely, CopulaGAN, CTGAN, and TVAE-GAN. The GaussianCopula doesn't have the utility to set up the hyper tunning parameter. Furthermore, once synthetic data was generated, we processed the data by converting floating point zeros to zeros for some of the features to maintain the data similarity of a real and synthetic dataset. The final dataset of each of the four GAN models consists of 200,000 records. All the GAN models are trained and datasets were generated on Compute Canada DataBase (CCDB) [16]. This database offers facilities for Advanced Research Computing (ARC) to the research community for high computational tasks and for parallel processing. To train the multiple GAN models parallel for synthetic data generation, 8GB GPU and 60GB CPU were used in CCDB.

B. Statistical Validation

For effective assessment, we assessed the quality of the synthetic dataset using the SDV toolkit. Specifically, we have used the SDMetrics library of the SDV tool which consists of various statistical methods to evaluate the synthetic dataset generated by various GAN models. We configured the SDV tool in Anaconda distribution and used the utilities of statistical modeling in Jupyter Notebook. The tool offers various statistical tests to determine the similarity property of the given synthetic data with the real dataset. Each test assessed the quality of the dataset based on specific criteria by considering the statistical properties of the given data. Table II summarizes the validation score of various statistical methods along with the baseline score that was computed using a real dataset for all four tabular GAN models. We set the baseline by performing the evaluation on the real data only, by randomly sampling 50% as real and considering the rest as synthetic data. The purpose of setting up the baseline score is to efficiently map the validation score of synthetic data with a real dataset.

The results depicted in Table II compare the score of each method and GAN model with others after pre-processing and post-processing. According to the result analysis, the GaussianCopula model performed the lowest compared to other GAN models. The TVAE is at the third rank, CTGAN is in the second position, and Copula GAN seems to be the most promising among all in terms of the statistical properties of features of the synthetic dataset.

C. Validation using Machine Learning Modeling

For proper validation, experiments were computed by merging real and synthetic datasets for all four GAN models. First, we have dropped the output class label ("marker") from both datasets and created an additional output label ("cls") that depicts the data samples as real and synthetic respectively. Furthermore, observations were carried out to create a balanced dataset where we considered around 78,369 records for each of the datasets. For fair distribution and assessment, the dataset was split randomly into two subsets, training (70%) and testing (30%). The training data was used for the algorithm training and the testing data was used to test the extendability of the ML model to differentiate the real and synthetic datasets. To avoid selection bias in the datasets and to reduce overfitting, we sampled the synthetic dataset according to the size of the real dataset. We repeated the experiment again, by merging the entire synthetic dataset with a real dataset. We got promising results in both approaches. We set the baseline by performing the evaluation on the real data only by sampling 50% as real and considering the rest as synthetic data. However, in this case, we also make sure the class is balanced, i.e. both portions have 50% of attack vectors and 50% of normal vectors to avoid model bias due to an imbalanced class.

From this experiment, we found that linear-based classifiers are more suitable than other classifiers. Since there are limited numbers of different values for some of the columns of the real dataset, tree-based classifiers can differentiate the real and synthetic data by observing certain fixed values and that might lead the experiment in the wrong direction. Unlike tree-based models, linear-based classifiers do not focus on local

TABLE II: Comparative analysis of tabular GAN models by Validating the quality of synthetic dataset using statistical methods (⇑: high improvement, ↑: improvement →: No improvement, ↓: Decrement)

Statistical Analysis	COPULA GAN		CTGAN		TVAE GAN		GAUSSIAN GAN		BASELINE real dataset
Assessment Paramenters	Pre-Process	Post-Process	Pre Process	Post Process	Pre Process	Post Process	Pre Process	Post Process	50% random sampling
KS Complement	75.46%	**94.95%** ⇑	74.25%	89.74% ⇑	82.92%	87.66% ⇑	68.65%	71.15% ⇑	**99.59%**
Boundary Adherence	100%	100% →	100%	100% →	100%	100% →	100%	100% →	**100%**
Range Coverage	58.65%	59.69% ↑	69.68%	**71.20%** ↑	67.38%	63.55% ↓	58.18%	58.18% →	**91.71%**
Statistic Similarity	98.42%	**99.64%** ↑	98.02%	98.95% ↑	99.09%	98.19% ↓	99.02%	99.06% ↑	**99.95%**
Chi-squared test	99.04%	**99.22%** ↑	20.56%	23.01% ↑	95.95%	97.26% ↑	96.15%	96.15% →	**99.96%**
Pairwise similarity grid R1	95.13%	**96.89%** ↑	95.11%	96.34% ↑	95.74%	94.84% ↓	95.54%	95.62% ↑	**99.59%**
Pairwise similarity grid R2	96.11%	**97.13%** ↑	95.23%	96.56% ↑	96.80%	95.62% ↓	95.33%	95.34% ↑	**99.67%**
Pairwise similarity grid R3	96.54%	**97.32%** ↑	94.22%	96.27% ↑	96.50%	95.58% ↓	95.15%	95.81% ↑	**99.71%**
Pairwise similarity grid R4	95.64%	**96.85%** ↑	94.54%	96.06% ↑	96.09%	96.02% ↓	95.73%	95.81% ↑	**99.69%**

TABLE III: Comparative analysis of tabular GAN models by Validating the quality of synthetic dataset using ML classifiers (⇓: high improvement, ↓: improvement →: No improvement, ↑: Decrement)

ML Classifers	COPULA GAN		CTGAN		TVAE GAN		GAUSSIAN GAN		Baseline
Accuracy & AUC Score	Pre-processed	Post-processing	Pre-processed	Post-processed	Pre-processed	Post-processed	Pre-processed	Post-processed	50% split of real data
Logistic Regression (Accuracy)	95.02%	59.02% ⇓	95.71%	68.88% ⇓	59.60%	**58.31%** ↓	96.69%	95.95% ↓	**49.63%**
Logistic Regression (AUC)	93.67%	59.01% ⇓	94.57%	68.83% ⇓	59.84%	**58.31%** ↓	96.68%	95.94% ↓	**49.69%**
SGD Classifier (Accuracy)	79.44%	**49.90%** ⇓	94.40%	49.77% ⇓	61.79%	51.53% ↓	98.00%	98.42% ↑	**50.12%**
SGD Classifier (AUC)	82.83%	**50.14%** ⇓	95.05%	49.85% ⇓	50.51%	51.48% ↑	97.98%	98.43% ↑	**49.93%**
MLP Classifier (Accuracy)	61.10%	49.75% ↓	61.35%	50.36% ↓	61.09%	**49.81%** ↓	49.82%	49.75% ↓	**50.13%**
MLP Classifier (AUC)	49.57%	49.50% ↓	49.92%	50.28% ↑	49.88%	**49.76%**↓	49.51%	49.52% ↑	**49.94%**
Voting Classifier (Accuracy)	82.60%	**51.20%** ⇓	93.94%	68.99% ⇓	62.79%	58.40% ↓	96.22%	96.07% ↓	**50.17%**
Voting Classifier (AUC)	77.39%	**50.96%** ⇓	92.28%	68.94% ⇓	59.73%	58.39% ↓	96.20%	96.05% ↓	**49.98%**

differences. Hence, when it comes to detectability, linear-based classifiers can be a good fit to analyze the features of such datasets. We have used a total of four linear-based classifiers for this experiment, namely, logistic regression, Stochastic Gradient Descent (SGD) classifier, Multi-layer Perceptron (MLP) classifier, and Majority Vote (MV) ensemble method (this technique predicts the output label based on the prediction of a majority of the three linear-based classifiers that we have mentioned before). Table III depicts the accuracy and area under the ROC Curve (AUC) score of each of the classifiers that represent the detectability score to differentiate real and synthetic datasets. Ideally, according to the baseline, the accuracy of around 50% seems to be perfect which depicts no difference in real and synthetic datasets and that represents a given input data as a high-quality synthetic dataset. During analysis, we found that the synthetic dataset generated by CopulaGAN (after hyper tunning and post-processing) seems to be most promising than other GAN models. Moreover, the majority vote ensemble method is more effective compared to the other three classifiers.

D. Experimentation Summary

Based on the validation results from statistical analysis and the machine learning detectability method, we have identified CopulaGAN as the most promising model. Table IV depicts the validation results of synthetic data generated using CopulaGAN. As summarized, the results after applying fine-tuning and post-processing seem to be more accurate and make a significant difference in the quality of the synthetic dataset.

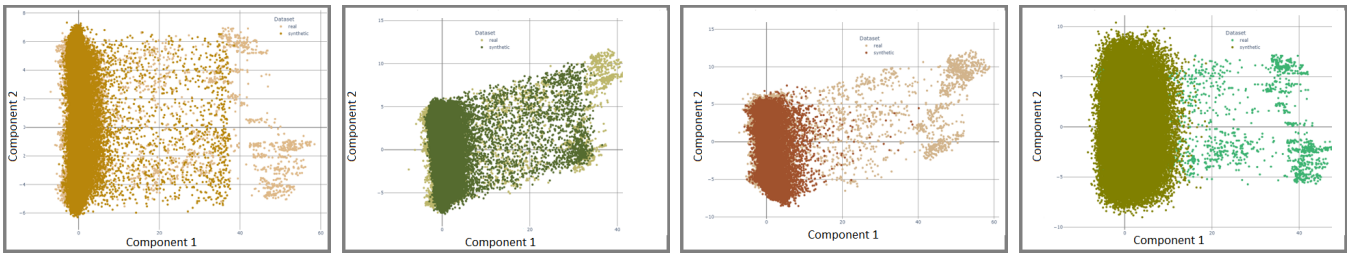
Furthermore, we have used a popular dimensionality reduction technique called Principal Component Analysis (PCA) for visualizing the synthetic and real datasets. Figure 4 represents the transformation of 2 components to the latent space. This figure shows a similarity between real and synthetic datasets by overlapped data points. This visualization also validates the quality of the synthetic dataset. We have uploaded all the programs that we have implemented for this research on the GitHub repository [17].

V. Conclusion

Advanced machine learning and data analytic methods perform effectively to protect SCADA-based power grids from various cyber-attacks. However, such advanced techniques need large datasets for efficient modeling and data representation. In this research, we have compared four tabular GAN models by generating and validating the quality of synthetic datasets of power grids. Furthermore, we have improved the quality of synthetic data by applying hyper tunning and post-processing techniques. The post-processing step is performed based on domain knowledge of the power system. The experimental results reveal that the CopulaGAN outperforms compares to other tabular GAN models. Moreover, the majority vote ensemble method seems promising while analyzing the similarity in the features of the synthetic and real datasets using ML modeling. Future research extends the machine learning detection model by taking a subset of normal and attack events separately to circumvent the inadequacy of bias performance assessment.

TABLE IV: Validation score to evaluate the quality of synthetic dataset generated using CopulaGAN

Preformace Analysis	COPULA GAN	Default Parameters		Finetuned Parameters		Baseline
		before post processing	after post processing	before post processing	after post processing	50% split of real data
Statistical Analysis	KSComplement	75.46%	88.59%	83.17%	**94.95%**	**99.59%**
	BoundaryAdherence	100.00%	100.00%	100.00%	**100.00%**	**100.00%**
	RangeCoverage	58.65%	58.65%	59.69%	**59.69%**	**91.71%**
	StatisticSimilarity	98.42%	98.42%	99.64%	**99.64%**	**99.95%**
	CS test	99.04%	99.04%	99.22%	**99.22%**	**99.69%**
	Pairwise similarity grid R1	95.13%	96.25%	96.62%	**96.89%**	**99.59%**
	Pairwise similarity grid R2	96.11%	96.65%	97.03%	**97.13%**	**99.67%**
	Pairwise similarity grid R3	96.54%	96.89%	97.33%	**97.32%**	**99.71%**
	Pairwise similarity grid R4	95.64%	96.66%	96.50%	**96.85%**	**99.69%**
ML Classifiers	Logistic Regression (Accuracy)	95.02%	93.67%	63.10%	**59.02%**	**49.63%**
	Logistic Regression (AUC Score)	93.67%	93.70%	63.12%	**59.01%**	**49.69%**
	SGD Classifier (Accuracy)	79.44%	96.08%	50.93%	**49.90%**	**50.12%**
	SGD Classifier (AUC Score)	82.83%	96.10%	50.57%	**50.14%**	**49.93%**
	MLP Classifier (Accuracy)	61.10%	49.44%	49.16%	**49.75%**	**50.13%**
	MLP Classifier (AUC Score)	49.57%	49.65%	49.52%	**49.50%**	**49.94%**
	Majority Voting (Accuracy)	82.60%	93.59%	62.75%	**51.20%**	**50.17%**
	Majority Voting (AUC Score)	77.39%	93.61%	62.77%	**50.96%**	**49.98%**



(a) CopulaGAN (b) CTGAN (c) TVAE GAN (d) GaussianCopula GAN

Fig. 4: Representation of power grid's data consist of the real and synthetic dataset (Fine-tuned post processed synthetic dataset)

ACKNOWLEDGMENT

The authors gratefully acknowledge the support in part by the Natural Sciences and Engineering Research Council (NSERC), Canada through a Collaborative Research Grant.

REFERENCES

[1] D. Upadhyay, S. Sampalli, and B. Plourde, "Vulnerabilities' assessment and mitigation strategies for the small linux server, Onion Omega2," *Electronics*, vol. 9, no. 6, p. 967, 2020.

[2] B. Kesler, "The vulnerability of nuclear facilities to cyber attack," *Strategic Insights*, vol. 10, no. 1, pp. 15–25, spring 2011.

[3] D. Upadhyay and S. Sampalli, "Scada (supervisory control and data acquisition) systems: Vulnerability assessment and security recommendations," *Computers & Security*, vol. 89, p. 101666, 2020.

[4] D. Upadhyay, J. Manero, M. Zaman, and S. Sampalli, "Gradient boosting feature selection with machine learning classifiers for intrusion detection on power grids," *IEEE Transactions on Network and Service Management*, 2020.

[5] ——, "Intrusion detection in scada based power grids: Recursive feature elimination model with majority vote ensemble algorithm," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 3, pp. 2559–2574, 2021.

[6] H. Xu, Y. Lin, X. Zhang, and F. Wang, "Power system parameter attack for financial profits in electricity markets," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3438–3446, 2020.

[7] L. A. Maglaras and J. Jiang, "Intrusion detection in scada systems using machine learning techniques," in *2014 Science and Information Conference*, 2014, pp. 626–631.

[8] S. Ghosh and S. Sampalli, "A survey of security in scada networks: Current issues and future challenges," *IEEE Access*, vol. 7, pp. 135 812–135 831, 2019.

[9] U. Adhikari, S. Pan, T. Morris, R. Borges, and J. Beaver, "Power system datasets," datasets used in the experimentation. [Online]. Available: https://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data-sets

[10] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in *4th International Conference on Learning Representations, San Juan, Puerto Rico*, Y. Bengio and Y. LeCun, Eds., 2016. [Online]. Available: http://arxiv.org/abs/1511.06434

[11] E. Choi, S. Biswal, B. A. Malin, J. Duke, W. F. Stewart, and J. Sun, "Generating multi-label discrete electronic health records using generative adversarial networks," *CoRR*, vol. abs/1703.06490, 2017. [Online]. Available: http://arxiv.org/abs/1703.06490

[12] M. H. Shahriar, N. I. Haque, M. A. Rahman, and M. Alonso, "G-ids: Generative adversarial networks assisted intrusion detection system," in *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, 2020, pp. 376–385.

[13] L. D. Manocchio, S. Layeghy, and M. Portmann, "Flowgan - synthetic network flow generation using generative adversarial networks," in *2021 IEEE 24th International Conference on Computational Science and Engineering (CSE)*, 2021, pp. 168–176.

[14] M. D. to AI Lab, "Synthetic data metrics (sdmetrics)," statistical methods. [Online]. Available: https://docs.sdv.dev/sdmetrics/

[15] S. Pan, T. Morris, and U. Adhikari, "Industrial control system (ics) cyber attack datasets," binary dataset used in the experimentation. [Online]. Available: http://www.ece.uah.edu/~thm0009/icsdatasets/PowerSystem_Dataset_README.pdf

[16] C. C. F. (CCF), "Compute canada database by digital research alliance of canada," synthetic dataset generation. [Online]. Available: https://ccdb.computecanada.ca/security/login

[17] Q. Lui, D. Upadhyay, J. Manero, M. Zaman, and S. Sampalli, "Implementation of synthetic data for power grids," github repo for source code. [Online]. Available: https://github.com/qiaodan97/PowerGridSyntheticData