Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

DIGITAL TWIN BASED LEARNING FRAMEWORK FOR ADAPTIVE FAULT DIAGNOSIS IN MICROGRIDS WITH AUTONOMOUS RECONFIGURATION CAPABILITIES

Temitope Runsewe
Abdurrahman Yavuz
Nurcin Celik

University of Miami
Department of Industrial and Systems Engineering
1251 Memorial Drive
Coral Gables, FL 33146, USA

ABSTRACT

The world is increasingly reliant on energy systems, making them a critical infrastructure for various essential services. However, this also makes them vulnerable to attacks, which can result in significant disruptions and damage. Microgrid (MG) monitoring systems play a crucial role in ensuring the safety and reliability of energy systems. However, traditional fault diagnosis techniques are limited to already established faults due to the use of only historical data, making it challenging to keep up with the increasing demand for safety and reliability. This paper proposes a digital twin based machine learning (DTML) framework for fault diagnosis in MG monitoring systems, with a focus on assessing the resilience of MG end-to-end systems to potential disruptions from adversaries. The proposed framework utilizes digital twin based random forest (RF) and support vector machine (SVM) and logistic regression (LR) model and shows that the RF based model outperforms other models with an accuracy of 95%.

1 INTRODUCTION

Energy sector has been designated as a critical infrastructure because of its enabling function for all infrastructure sectors. Several recent weather-related events and cyberattacks have placed the energy sector's resilience at the forefront of national research priorities. In the event of a natural or human-induced disaster, electricity may be lost, resulting in large financial losses that can affect a wide range of sectors and service types (Eskandarpour et al. 2017). About 83% of all reported power outages from 2000-2021 can be attributed to a weather-related event. Beyond weather events, human-induced outages are also a concern (Ramirez 2022). The electric grid has been described as an "attractive target" for domestic violent extremists in the US. In 2020, intelligence analysts saw major uptick in online chatter focused on attacking the power grid (U.S. EIA 2023). The average annualized costs caused by cybercrimes worldwide in 2018 is 13.77 million U.S. dollars (Accenture 2019). Due to natural and human-induced threats and the complexity of power systems, building resilience in energy infrastructure is challenging. There has been a number of studies on resilience in both physical and cyber layers of power systems.

Bie et al. (2017) presents an extensive review of the studies focusing on assessment of the power system resilience and emphasizes the vital importance of recognizing the threats and their potential effects in order to build power systems that are swift in responding to the events such as natural disasters and cyber-attacks. Once the root cause, magnitude and scope of a contingency that is causing disruption in the power system are accurately identified, then the problem becomes clearer and can be addressed by approaches such as optimal management of electricity storage units (Tavakoli et al. 2018), load curtailment (Mehrjerdi 2020), reformation of the network (Gilani et al. 2020), operation scheduling and reconfiguration (Damgacioglu et al. 2022) or stochastic tuning of the distributed generation units (Yavuz et al. 2023).

The research on diagnosis of a disruptive event in both cyber and physical layers of power systems (e.g., natural disaster, cyber-attack) has gathered significant attention from both academia and industry. Badr et al. (2015) points out that prioritizing resiliency only during the design phase is not sufficient to achieve end-to-end resiliency of systems and proposes a cloud-based real-time monitoring and runtime crisis management framework to improve the resiliency of control mechanisms deployed for maintaining efficient system operations. Jin et al. (2017) emphasizes on the pervasive utilization of decentralized control mechanisms employed in grid control, thus the exacerbating need for cyber resilient communication infrastructures to govern the supervisory control and data acquisition operations. The authors propose a software-defined networking-based communication network architecture to sustain functionality against cyber-attacks aiming to hinder efficient control of grid operations. Mishra et al. (2020) presents a comprehensive review of studies centering around the analysis of cyber threats posed on microgrids and the precautionary measures proposed to achieve enhanced resiliency and mitigate the impact of such interferences. The authors also highlight the importance of having pre-, during- and after-event recovery strategies for continuous supply of electricity. Mohammadpourfard et al. (2021) proposes a long-short-term memory recurrent neural network to detect an attack that targets to distort measurements collected from measurement units. A novel cyber-attack type that imitates the cascading impact of physical component failure by manipulating data from not only a single source but from multiple sources is introduced and the proposed fault detection approach is tested over extensive simulations of the system. The authors also underline the increasing challenges of fault detection when the power distribution network has the flexibility to adjust its configurations and make topological changes dynamically.

Another vein of research has focused on analyzing the threats posed by natural disasters and the preventative and corrective measures that can mitigate the impact of such events on the power network resiliency. Wang et al. (2015) reviews the studies focusing on modeling the effect of natural disasters on electric power systems and exploring ways to prepare and recover the grid. The authors point out the fact that although statistical models can be effective in damage assessment, they heavily rely on the availability of the data. Since the preventative measures should be determined before the event happens, i.e., when we have little or no information on the operating conditions that the system will experience, the use of statistical models in determining proactive and corrective actions is limited. To this end, the use simulation techniques (e.g., Monte Carlo simulations, agent-based simulations) in conjunction with statistical inferences drawn from the initial observations of both the grid and the event are utilized to estimate the effect and scope of the natural disasters (e.g., wildfires, earthquakes, floods) (Wang et al. 2015; Waseem et al. 2020). Huang et al. (2022) proposes the use of simulation software specifically designed for power networks (i.e., OpenDSS) to express the physical properties of the system and embeds the simulation model in a reinforcement learning environment that is designed to find out set of decisions (i.e., policies) related to optimal formation of the power system. The application of digital-twinning in monitoring power system conditions is presented by Moutis et al. (2020) on tracking the voltage and current at sub-cycle detail. The authors report significant accuracy of the designed digital twin when compared to field data obtained for medium voltage-low voltage distribution transformers. Darville et al. (2022) utilizes machine learning techniques as well as the simulation models to overcome the data sparsity in problem of fault detection. The generated data is then used in the training process of binary classification models aiming to distinguish faulty measurements collected from the real-world microgrid. Besides the tasks such as monitoring and tracking the system condition, digital twin of the electricity distribution systems has been extensively benefited for various purposes including but not limited to electric utility resource planning (Senz et al. 2012), assessment of communicational (Yavuz et al. 2022) or computational infrastructure (Yavuz et al. 2020) and load dispatching (Thanos et al. 2013).

Digital twinning has also found itself an application area in fault detection problems in various domains. Xie et al. (2023) employs the digital representation of the heating, ventilation, and air conditioning (HVAC) system for fault detection and diagnosis processes. The authors propose a lightweight artificial intelligence technique to identify the most informative sensory measurements within a building and prevent the model from overfitting by eliminating the uninformative measurements. The insights gained via proposed approach are then utilized in dynamic asset management problem. Nguyen et al. (2022) employs a physics-informed digital twin of thermal-hydraulic components of a nuclear reactor and include virtual sensors for improved diagnosis and fault detection capabilities. Jain et al. (2019) proposes the use of digital twin in estimating the photovoltaic energy conversion unit outputs. The estimation is then compared with the real-world measurements and the analysis of the residuals (i.e., the difference between estimation and the measurement) are carried out to detect faults. Saad et al. (2020) addresses the problem of mitigating ethe impact of coordinated cyber-attacks on networked microgrids within Internet of Things (IoT)-based digital twin framework. The authors tested their framework against coordinated false data injection and denial of service cyber-attacks.

In this study, we address the challenges brought by the increasing flexibility and automation of electricity distribution systems such as i) dynamically and drastically changing network topologies as a result of embedded reformation capabilities, and ii) adaptive system reconfiguration via centralized/decentralized smart control mechanisms within a digital twin-based learning framework that can facilitate learning for data-driven algorithms. The proposed framework addresses the challenges of data sparsity and performance deterioration of the models by incorporating digital twin of the model into the learning (i.e., model development) process to provide it with i) the data for the out of sample network structures formed as a response to a natural disaster or a cyber-attack, ii) continuous learning environment with the flexibility to encompass the decisions made by dynamic control mechanisms (e.g., load shedding, unit commitment).

The remainder of this paper is organized as follows. First, we outline the digital twin based machine learning (DTML) framework and detail the microgrid that is subject to experimentation. Subsequently, in Section 3, we present the comparative analysis of the ML models benefiting from the DTML framework. Finally, in Section 4, we conclude with the summary of findings obtained from the study.

2 METHODOLOGY

In this paper, a microgrid is modeled based on a modified IEEE 30-node distribution system connected with renewable energy and other non-renewable energy sources. The model is simulated in OpenDSS. These simulations include normal and faulty operations during grid-connected mode of operation of the microgrid.

2.1 Digital Twin based Machine Learning (DTML) Framework

The utilization of digital twins as a basis for machine learning is a cyclical system throughout the engineering tasks for energy systems (Min et al. 2019). This research utilizes two crucial interfaces for enhancing the resilience in microgrids: one for "data-driven modeling or model improvements" to transfer knowledge extracted from the tangible energy system, and the other for an "AI environment" that characterizes the digital twin as a milieu for third-party providers of learning algorithms. By amalgamating these interfaces, two significant features can be furnished:

I. Independent problem-solving using the digital twin.

II. Model extraction/improvements by means of data-oriented learning approaches.

Once a digital twin has been developed for a particular use case, model-based approaches like machine learning can be utilized to address specific tasks such as programming control logic. Once the physical energy system is in place, data-driven learning methods can improve the model foundation for both current and future systems. The engineering process for an energy system follows a step-by-step problem-solving approach where each engineering phase involves collecting requirements and shaping them into a specified task (Ritto et al. 2021; Hossain et al. 2022). Digital twins have become an increasingly popular tool in the field of microgrids, providing a virtual replica of the physical system that serves as an input parameter for problem-solving and optimization. This enables subsequent engineering phases in the life cycle of the microgrid. By providing a detailed and accessible model, digital twins facilitate communication across multiple domains and offer several benefits. One such advantage is the ability to conduct many problem-solving and optimization steps in earlier stages, such as programming control systems before the physical energy system is assembled. However, despite the use of digital twins, programming still requires an engineer for refinements and adjustments. Overall, the use of digital twins provides a powerful and cost-effective means for designing, testing, and operating microgrids. (Bazmohammadi et al. 2021). In this scenario, the digital twin serves as a test framework. Figure 1 depicts the digital twin model used in this study.

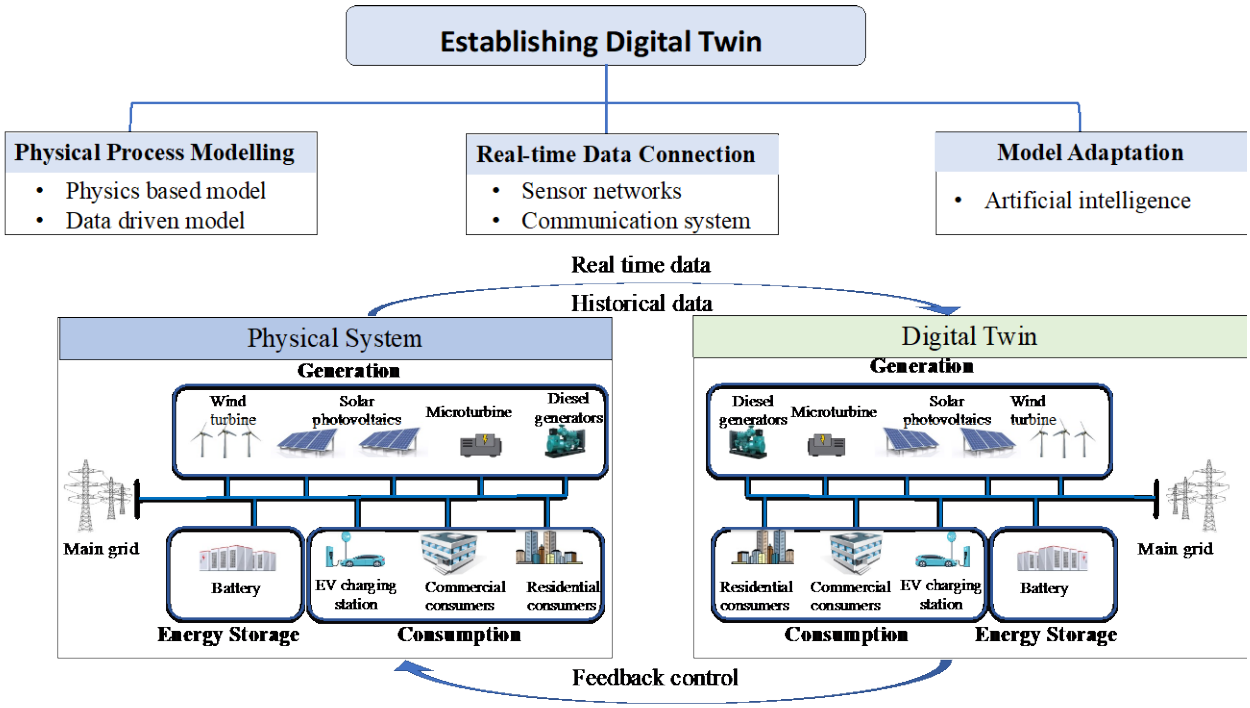

Figure 1: Digital twin of a MG.

The digital twin as a foundation for machine learning provides independent problem-solving in the engineering context. While executing an independent problem solver with ML, the input values necessitate adjustment to a machine-readable function. All other information required for resolving or refining the problem is present in the digital twin itself (Danilczyk et al. 2021). This study exemplifies the utilization of a problem description and a digital twin independently to program control logic, while incorporating advanced simulation models to effectively tackle complex problem-solving. In this regard, we propose the adoption of a digital twin-based machine learning model (DTML) as a suitable approach for data-oriented learning. The DTML model utilizes historical and real-time data from the digital twin simulation to train a machine learning model specifically designed for detecting anomalies or faults within the microgrid system. By doing so, the DTML model acts as a preventive measure against cascading effects within the microgrid.

Maintaining a continuous information flow between the digital twin and the training model becomes crucial in preventing model degradation and minimizing false predictions. This bidirectional information exchange ensures that the machine learning model remains updated and accurately reflects the behavior of the microgrid system. Figure 2 provides a visual representation of the DTML framework, illustrating the dynamic nature of this approach.

The integration of digital twins and data-driven methods, as exemplified by the DTML framework, showcases the powerful potential of digital twins in microgrid control and optimization. By harnessing the vast amounts of data available within the digital twin, coupled with machine learning techniques, we can effectively address and solve complex problems encountered within microgrids. This approach holds great promise for improving the efficiency, reliability, and performance of microgrid systems.

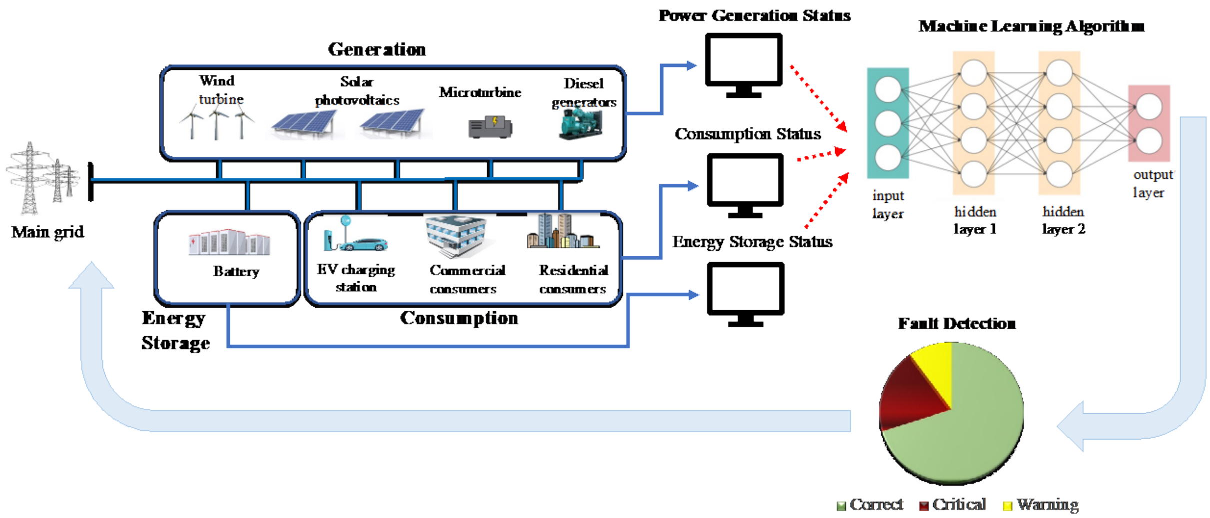

Figure 2: DTML framework.

2.2 Digital Twin Modelling

The microgrid model implemented using OpenDSS is based on the modified IEEE 30-node system, which operates at 132 kV and 60 Hz, and emulates the common characteristics of a distribution system (see Figure 3). The model integrates renewable energy sources such as solar PV, and wind generator. Additionally, a diesel generator system is also included. The model also simulates different faults on the distribution line that connects buses (1-2,10-17,14-15,15-18 and 21-22) of the microgrid. However simultaneous faults are not considered in this paper.

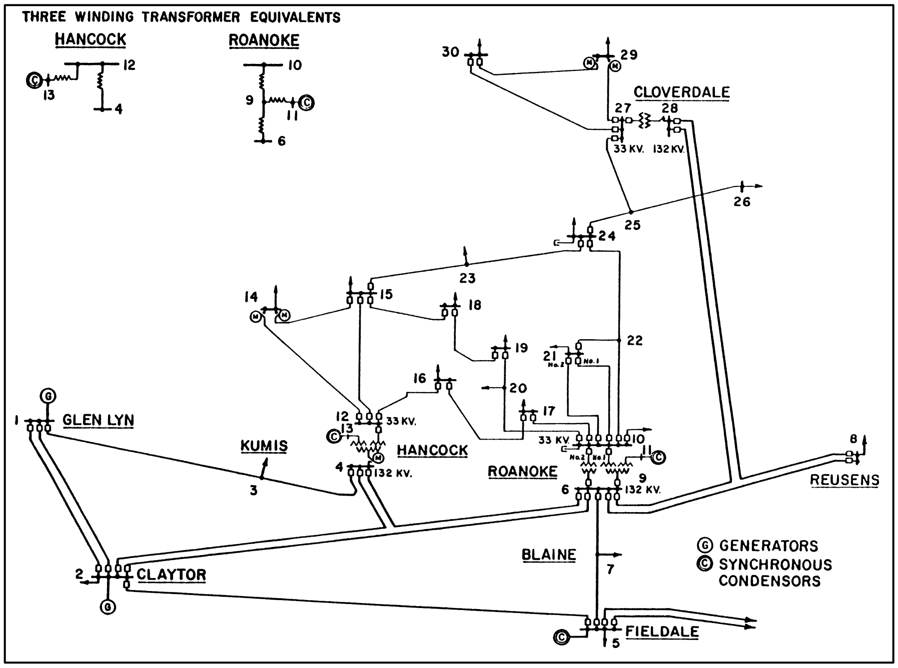

Figure 3: Simulated IEEE 30-node system.

2.3 Machine Learning Model Building

The data collected comprises training and test datasets which were split into 80% and 20% respectively. Each case of the training dataset consists of a baseline scenario and fault scenario. The baseline scenario serves as a benchmark to characterize the behavior of the system under normal operation. For this scenario, we collect data under expected normal circumstances, with no faults while working continuously. Whereas the fault scenario consists of the microgrid operating under different types of faults (see Figure 4).

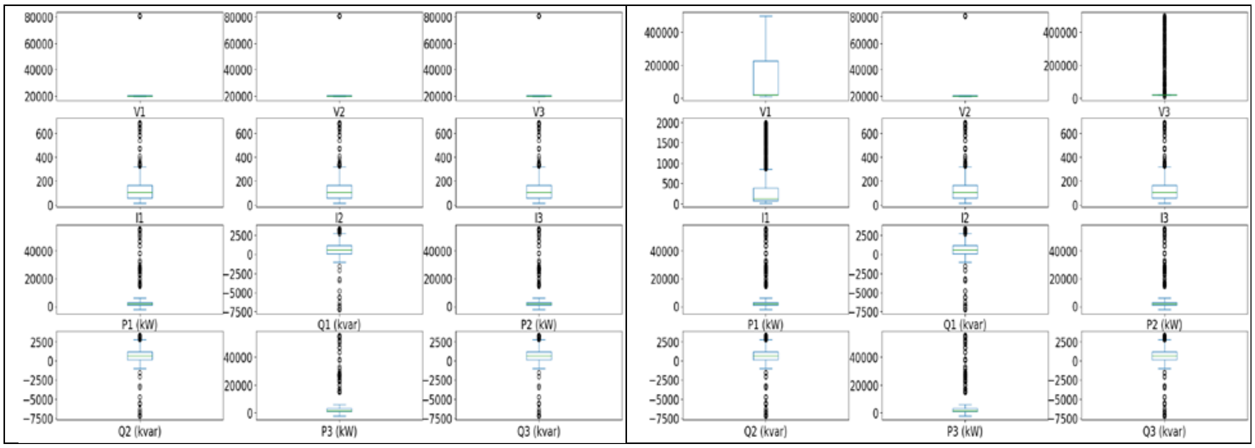

Figure 4: Boxplots of variables operating under normal operations (left) and under fault (right).

Since the data primarily used for our machine learning model was from the digital twin, there were no missing points in the data. However, a summary of the dataset was performed, and the distribution of the normal and fault data was checked to ensure the classes were not imbalanced. Here, if an imbalance in the class distributions is observed, additional steps such as over-sampling of the minority class observations or under-sampling of the majority class observations may be required (Japkowicz, 2000). Figure 5 shows a pictorial depiction of the current, voltage and powers of the lines.

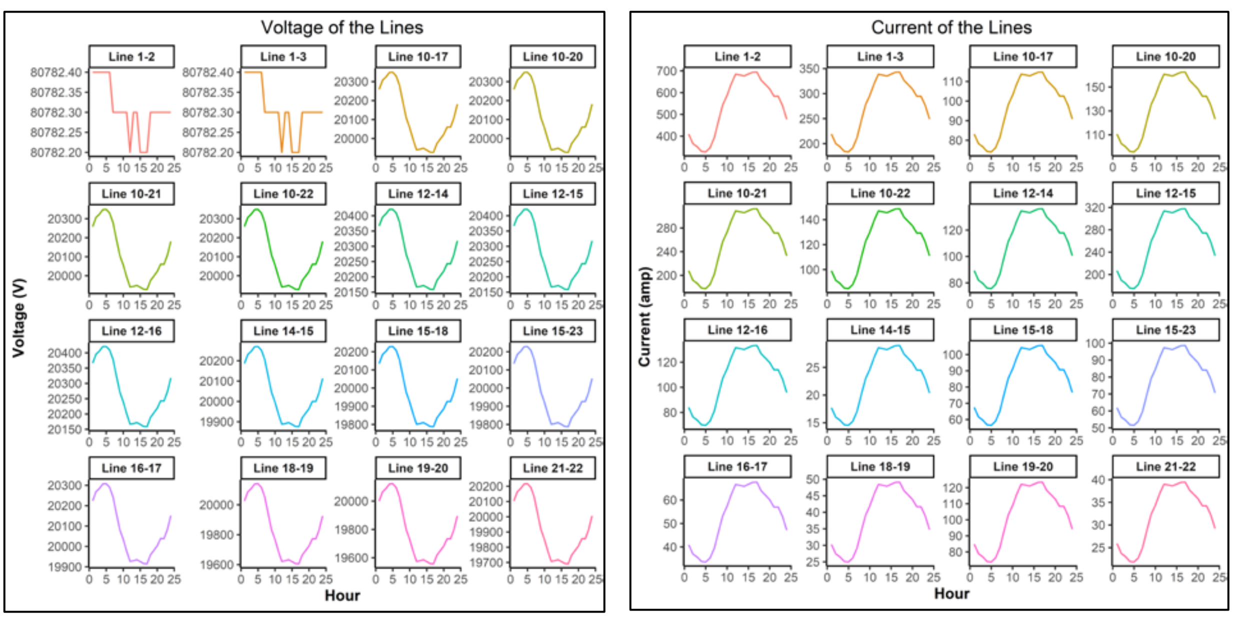

Figure 5: Current and voltage flowing through the line during a 24-hour period.

ML models for binary classification such as random forest (RF), support vector machine (SVM) and logistics regression (LR) is developed for the DTML framework. The developed ML models are tested with both training and test datasets. LR is a binary classification model (i.e., it only contains data classified as 1 or 0, which in our case refers to an MG operating under normal or fault that is positive or negative operation) that estimates the probability of the dependent variable taking on a particular outcome, given the values of the independent variables (Pampel 2020). The LR model is based on Equation (1) below.

$$p = \frac{1}{1 + e^{-z}} \tag{1}$$

Where p is the predicted probability of the dependent variable taking on a particular outcome and z is the linear combination of the independent variables and their coefficients.

SVM is a supervised machine learning algorithm that aims to find an optimal hyperplane to separate data points of different classes. The decision function of SVM can be represented as:

$$f(x) = sign(w^T x + b)$$

where x represents the input vector, w denotes the weight vector perpendicular to the hyperplane, b is the bias term, and sign (·) is the sign function that assigns the positive class (+1) or the negative class (-1) based on the predicted value. The SVM algorithm solves an optimization problem to determine the optimal values of w and b by minimizing the objective function, which includes the hinge loss function to penalize misclassified samples.

RF is an ensemble learning method that combines large number of decision trees. In RF, each decision tree is trained on a randomly selected subset of the data and a randomly selected subset of the independent variables, which helps to reduce overfitting and improve the accuracy of the model. The decision trees in a RF are constructed using a process called "bagging", which involves repeatedly resampling the data with replacement and building decision trees on the resampled data. Once the decision trees are constructed, the RF algorithm combines their outputs by taking a majority vote for classification tasks or averaging their predictions for regression tasks. This ensemble approach helps to reduce the variance of the model and improve its overall performance (Breiman 2001; Dhanaraj et al. 2021). RF model is based on Equation 2 below. Where $\hat{y}_k$ is the predicted class label for decision tree k and J is the number of class labels.

$$\hat{y} = argmax\left(\sum [\hat{y}_k = j]\right) \quad for\ j = 1, 2, \dots, J\ and\ k = 1\ to\ K \tag{2}$$

2.3.1 Performance Assessment

This research evaluates the performance of the proposed ML models using common criteria employed in fault identification and classification problems for microgrids, such as accuracy, precision, recall, and F1 score, as this is a binary classification problem (Iwendi et al. 2020). The accuracy (ACC) is calculated using the Equation (3):

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \tag{3}$$

where ACC is the classification accuracy, true positive (TP) is the number of instances that are correctly classified as positive, true negative (TN) is the number of instances that are correctly classified as negative, false positive (FP) is the number of instances that are incorrectly classified as positive, and false negative

(FN) is the number of instances that are incorrectly classified as negative. The accuracy alone may not provide a comprehensive assessment of a classifier model's performance. The precision (PRE) is another parameter used to evaluate model performance. The PRE is calculated using the following equation:

$$PRE = \frac{TP}{TP + FP} \tag{4}$$

A good model has a higher precision, with a maximum value of one. Another metric, known as recall (REC) or sensitivity, is used for classifier model evaluation and calculated as follows:

$$REC = \frac{TP}{TP + FN} \tag{5}$$

A good classifier model has a maximum recall value of one. It can be observed that PRE is dependent on FP and REC is dependent on FN values from equation (4) and (5). As a result, the weighted average of precision and recall is used as another performance evaluation metric, known as the F1 score, which is defined as in Equation (6):

$$F1\ score = 2 * \frac{(PRE * REC)}{(PRE + REC)} \tag{6}$$

3 RESULTS AND DISCUSSION

In the DTML framework proposed in this study, we evaluated the performance of each machine learning (ML) model in detecting faults during MG operations. All ML models were developed using Python 3.9 and executed on a computer equipped with an Intel Xeon® 2.90 GHz processor and 64 GB RAM, in order to ensure replicability of results. Prior to implementation, we assessed and addressed the assumptions necessary for each ML model. In the case of the LR model, we investigated the potential issue of multicollinearity. To test for collinearity between variables, we employed a correlation matrix, where coefficients approaching 1 indicate co-dependence between variables, as depicted in Figure 6.

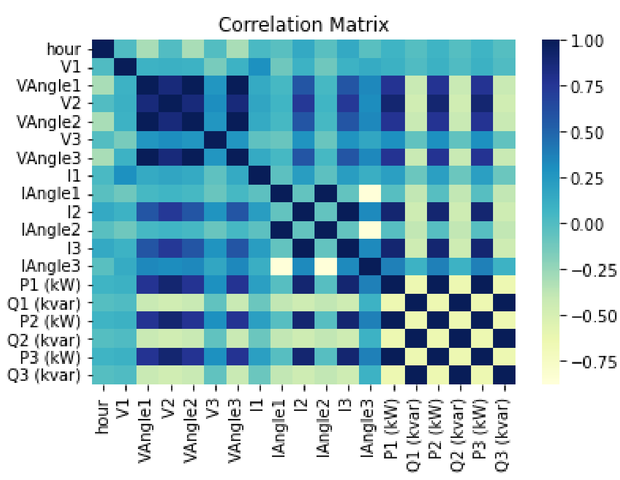

Figure 6: Correlation matrix of model features.

In Figure 4, the presence of multicollinearity among variables in the MG Faults dataset is evident. However, these correlations are inherent and not a result of creating new variables from existing ones. Therefore, the approach to address multicollinearity should be informed by domain knowledge of the system and data collection methods. In this study, data was collected from a DT system of an MG that produces three-phase power, where power is a function of voltage and current at each phase. Notably, the voltage and current at one phase are independent, while there is a dependence between the voltage at one phase and the current at the other two phases. This dependence is expected as three symmetrical single-phase components operate simultaneously to achieve a balanced three-phase loading within an MG (Darville et al. 2022). Therefore, such correlations should not be manipulated or eliminated to enhance the reliability of regression coefficients, as they are inherent to MG systems. Other variables such as (P2, P3, Q2 and Q3) which possessed high correlation were removed to avoid multicollinearity.

3.1 Training Performance of the ML Model

This section presents an analysis of the training and testing performance of the ML models used in this study. Figure 7a displays the confusion matrix, where the rows indicate the predicted class, and the columns represent the true class. The trained models were tested using the same training datasets. To further assess the overall performance of the proposed ML model, an offline trained ML model was evaluated using various test datasets. Specifically, four different faults were introduced during the peak usage period of hours 8-16, resulting in testing datasets that were not used during model training. These simulated faults generated 3080 windows of data samples comprising instantaneous voltage and current waveforms. Among these windows, 785 represent no fault cases, while 2295 correspond to fault cases. The confusion matrix obtained from this testing is depicted in Figure 7b.

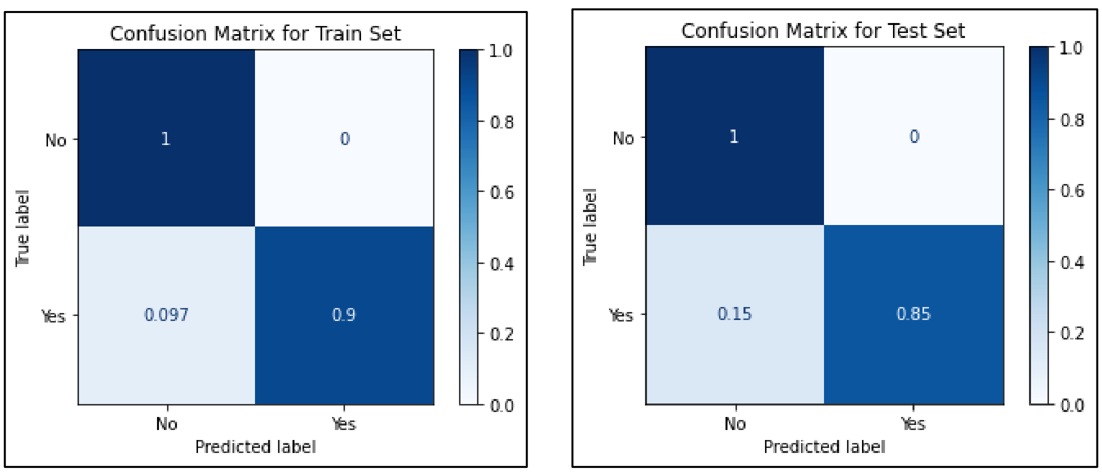

Figure 7: Confusion matrix of RF based model for training set (a) and test set (b).

As shown in Table 1 below, the RF model exhibited significantly higher accuracy compared to the LR and SVM model. Moreover, the classification accuracy indicates that the developed RF model is capable of detecting and classifying most of the microgrid faults, with only a few instances of misclassification.

Table 1: Training performance of RF based model for training set.

	Precision	Recall	F1-score	Accuracy
LR	0.87	0.75	0.81	0.80
RF	1	0.91	0.95	0.95
SVM	0.95	0.85	0.92	0.92

Table 2 below shows that the RF model outperforms the LR and SVM model in terms of all performance metrics considered in this study. This shows that the DT based RF model performs better in detecting faults in constantly changing microgrid topographies.

Table 2: Training performance of RF based model for test set.

	Precision	Recall	F1-score	Accuracy
LR	0.81	0.80	0.80	0.80
RF	1	0.87	0.93	0.93
SVM	0.95	0.87	0.90	0.90

Upon analyzing the performance metrics of both the training and testing data sets, we discovered that the RF model had achieved high accuracy values of 0.95 and 0.93, respectively, while the SVM model had slightly lower values of 0.92 and 0.90, respectively and LR model had obtained the lowest values with 0.80 and 0.80, respectively. However, relying solely on accuracy to assess the effectiveness of a model can be misleading, as it does not provide a complete picture of the model's performance. Therefore, we also evaluated other metrics such as precision, recall, and F1-score, and found that all models demonstrated high values in these measures for both the training and testing data sets. This is an important indication that the developed ML models did not overfit or overtrain, and can generalize well to new, unseen data.

4 CONCLUSION

This paper contributes to the literature on machine learning, digital twin, and control optimization, by proposing a theoretical framework for digital twin-based control. This novel framework helps to reduce the dependency of control management decision-making on expert experience and domain knowledge and avoid excessive influences of single machine learning results. This paper introduces the usage of real time data and provides methodological insights on building digital twin using machine learning, training digital twin-based models using historical and real time data. As the concept of a digital twin has gained traction, an increasing emphasis has been placed on the interaction between the physical system and digital twins. Therefore, the results of this study provide guidelines for the efficient and effective integration of digital twins and physical systems. The framework and approach proposed in this paper, including model development, data processing, and model training, contribute to the methodological research on the applications of machine learning and digital twin to solving control decisions in virtual representations of real-life scenarios. The effectiveness of these approaches is proved by applying them to an IEEE 30-node system. The modeling processes of digital twins proposed in this paper based on machine learning also shows that the digital twin-based RF outperformed other models.

In conclusion, this paper has presented a comprehensive digital twin-based adaptive fault diagnosis framework for enhancing fault diagnosis performance in microgrids with autonomous reconfiguration capabilities. Analyzing data during off-peak hours and predicting faults during this time can provide valuable insights into the effectiveness and robustness of the framework in different operating conditions. Additionally, assessing and evaluating scenarios with concurrent faults can enhance the framework's capabilities in handling complex fault scenarios. These research directions have the potential to enhance the scope and applicability of the framework, contributing to improved fault diagnosis in microgrids.

ACKNOWLEDGMENTS

This study was supported by the Air Force Office of Scientific Research Award No: FA9550-18-1-0075.

REFERENCES

Accenture. 2019. Annual Global Industry Losses Caused by Cyber Crime 2018. Statista. https://www.statista.com/statistics/474928/average-annual-costs-caused-by-cyber-crime-worldwide/, accessed 16th June 2023.

Badr, Y., S. Hariri, A.L.-N. Youssif, and E. Blasch. 2015. "Resilient and Trustworthy Dynamic Data-Driven Application Systems (Dddas) Services for Crisis Management Environments". *Procedia Computer Science* 51: 2623-2637.

Bazmohammadi, N., A. Madary, J.C. Vasquez, H.B. Mohammadi, B. Khan, Y. Wu, and J.M. Guerrero. 2021. "Microgrid Digital Twins: Concepts, Applications, and Future Trends". *IEEE Access* 10: 2284-2302.

Bie, Z., Y. Lin, G. Li, and F. Li. 2017. "Battling the Extreme: A Study on the Power System Resilience". *Proceedings of the IEEE* 105 (7): 1253-1266.

Breiman, L. 2001. "Random Forests". *Machine Learning* 45: 5-32.

Damgacioglu, H., and N. Celik. 2022. "A Two-Stage Decomposition Method for Integrated Optimization of Islanded Ac Grid Operation Scheduling and Network Reconfiguration". *International Journal of Electrical Power & Energy Systems* 136: 107647.

Danilczyk, W., Y.L. Sun, and H. He. 2021. "Smart Grid Anomaly Detection Using a Deep Learning Digital Twin". In *2020 52nd North American Power Symposium (NAPS)*, 1-6.

Darville, J., T. Runsewe, A. Yavuz, and N. Celik. 2022. "Machine Learning Based Simulation for Fault Detection in Microgrids". In *Proceedings of the 2022 Winter Simulation Conference*, edited by B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C.G. Corlu, L.H. Lee, E.P. Chew, T. Roeder, and P. Lendermann, 701-712. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Dhanaraj, R.K., V. Ramakrishnan, M. Poongodi, L. Krishnasamy, M. Hamdi, K. Kotecha, and V. Vijayakumar. 2021. "Random Forest Bagging and X-Means Clustered Antipattern Detection from Sql Query Log for Accessing Secure Mobile Data". *Wireless Communications and Mobile Computing* 2021: 1-9.

Eskandarpour, R., A. Khodaei, and A. Arab. 2017. "Improving Power Grid Resilience through Predictive Outage Estimation". In *2017 North American Power Symposium (NAPS)*, 1-5.

Gilani, M.A., A. Kazemi, and M. Ghasemi. 2020. "Distribution System Resilience Enhancement by Microgrid Formation Considering Distributed Energy Resources". *Energy* 191: 116442.

Hossain, M.M., and S.R. Kolla. 2022. "Fault Detection and Classification in Microgrid Using Long Short-Term Memory". In *2022 IEEE International Conference on Electro Information Technology (eIT)*, 019-026.

Huang, Y., G. Li, C. Chen, Y. Bian, T. Qian, and Z. Bie. 2022. "Resilient Distribution Networks by Microgrid Formation Using Deep Reinforcement Learning". *IEEE Transactions on Smart Grid* 13 (6): 4918-4930.

Iwendi, C., A.K. Bashir, A. Peshkar, R. Sujatha, J.M. Chatterjee, S. Pasupuleti, R. Mishra, S. Pillai, and O. Jo. 2020. "Covid-19 Patient Health Prediction Using Boosted Random Forest Algorithm". *Frontiers in Public Health* 8: 357.

Jain, P., J. Poon, J.P. Singh, C. Spanos, S.R. Sanders, and S.K. Panda. 2019. "A Digital Twin Approach for Fault Diagnosis in Distributed Photovoltaic Systems". *IEEE Transactions on Power Electronics* 35 (1): 940-956.

Japkowicz, N. 2000. "The Class Imbalance Problem: Significance and Strategies." In *Proceedings of the Intertnational Conference on Artificial Intelligence*, 56: 111-117.

Jin, D., Z. Li, C. Hannon, C. Chen, J. Wang, M. Shahidehpour, and C.W. Lee. 2017. "Toward a Cyber Resilient and Secure Microgrid Using Software-Defined Networking". *IEEE Transactions on Smart Grid* 8 (5): 2494-2504.

Mehrjerdi, H. 2020. "Resilience Improvement with Zero Load Curtailment by Multi-Microgrid Based on System of Systems". *IEEE Access* 8: 198494-198502.

Min, Q., Y. Lu, Z. Liu, C. Su, and B. Wang. 2019. "Machine Learning Based Digital Twin Framework for Production Optimization in Petrochemical Industry". *International Journal of Information Management* 49: 502-519.

Mishra, S., K. Anderson, B. Miller, K. Boyer, and A. Warren. 2020. "Microgrid Resilience: A Holistic Approach for Assessing Threats, Identifying Vulnerabilities, and Designing Corresponding Mitigation Strategies". *Applied Energy* 264: 114726.

Mohammadpourfard, M., A. Khalili, I. Genc, and C. Konstantinou. 2021. "Cyber-Resilient Smart Cities: Detection of Malicious Attacks in Smart Grids". *Sustainable Cities and Society* 75: 103116.

Moutis, P., and O. Alizadeh-Mousavi. 2020. "Digital Twin of Distribution Power Transformer for Real-Time Monitoring of Medium Voltage from Low Voltage Measurements". *IEEE Transactions on Power Delivery* 36 (4): 1952-1963.

Nguyen, T.N., R. Ponciroli, P. Bruck, T.C. Esselman, J.A. Rigatti, and R.B. Vilim. 2022. "A Digital Twin Approach to System-Level Fault Detection and Diagnosis for Improved Equipment Health Monitoring". *Annals of Nuclear Energy* 170: 109002.

Pampel, F.C. 2020. *Logistic Regression: A Primer*. Vol. 132: Sage publications.

Ramirez, R. 2022. Power Outages Are on the Rise, Led by Texas, Michigan and California. Here's What's to Blame.

Ritto, T., and F. Rochinha. 2021. "Digital Twin, Physics-Based Model, and Machine Learning Applied to Damage Detection in Structures". *Mechanical Systems and Signal Processing* 155: 107614.

Saad, A., S. Faddel, T. Youssef, and O.A. Mohammed. 2020. "On the Implementation of Iot-Based Digital Twin for Networked Microgrids Resiliency against Cyber Attacks". *IEEE Transactions on Smart Grid* 11 (6): 5138-5150.

Senz, J.P., N. Celik, S. Asfour, and Y.-J. Son. 2012. "Electric Utility Resource Planning Using Continuous-Discrete Modular Simulation and Optimization (Codimoso)". *Computers & Industrial Engineering* 63 (3): 671-694.

Tavakoli, M., F. Shokridehaki, M.F. Akorede, M. Marzband, I. Vechiu, and E. Pouresmaeil. 2018. "Cvar-Based Energy Management Scheme for Optimal Resilience and Operational Cost in Commercial Building Microgrids". *International Journal of Electrical Power & Energy Systems* 100: 1-9.

Thanos, A.E., X. Shi, J.P. Sáenz, and N. Celik. 2013. "A Dddams Framework for Real-Time Load Dispatching in Power Networks". In *Proceedings of the 2013 Winter Simulations Conference*, edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, M. E.Kuhl, 1893-1904. Washington, District of Columbia: Institute of Electrical and Electronics Engineers, Inc.

U.S. EIA. 2023. Monthly Energy Review. U.S. Energy Information Administration.

Wang, Y., C. Chen, J. Wang, and R. Baldick. 2015. "Research on Resilience of Power Systems under Natural Disasters—a Review". *IEEE Transactions on Power Systems* 31 (2): 1604-1613.

Waseem, M., and S.D. Manshadi. 2020. "Electricity Grid Resilience Amid Various Natural Disasters: Challenges and Solutions". *The Electricity Journal* 33 (10): 106864.

Xie, X., J. Merino, N. Moretti, P. Pauwels, J.Y. Chang, and A. Parlikad. 2023. "Digital Twin Enabled Fault Detection and Diagnosis Process for Building Hvac Systems". *Automation in Construction* 146: 104695.

Yavuz, A., and N. Celik. 2022. "Effect of Communication Network Properties on Data-Driven Decision-Making in Large-Scale Interconnected Microgrids". In *Proceedings of the IIE Annual Conference Proceedings*, 1-6. Seattle, WA.

Yavuz, A., N. Celik, C.-H. Chen, and J. Xu. 2023. "A Sequential Sampling-Based Particle Swarm Optimization to Control Droop Coefficients of Distributed Generation Units in Microgrid Clusters". *Electric Power Systems Research* 216: 109074.

Yavuz, A., J. Darville, N. Celik, J. Xu, C.-H. Chen, B. Langhals, and R. Engle. 2020. "Advancing Self-Healing Capabilities in Interconnected Microgrids Via Dynamic Data Driven Applications System with Relational Database Management". In *Proceedings of the 2020 Winter Simulation Conference*, edited by K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, 2030-2041. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

AUTHOR BIOGRAPHIES

TEMITOPE RUNSEWE is a Ph.D. candidate in the Department of Industrial and Systems Engineering at the University of Miami. She obtained her bachelor's degree in chemical engineering from Covenant University, Nigeria. She then received her master's degree in industrial engineering from Texas State University, San Marcos. She was also awarded the Texas State University graduate college scholarship and student government scholarship. Her research interest includes data analytics, simulation, and optimization of systems. She also received the 2020 REMADE Institute Rising Star Spotlight and the ISE Outstanding Graduate Student Award for Research in 2022. Her email address is tor9@miami.edu. Her website is https://simlab.coe.miami.edu/team/.

ABDURRAHMAN YAVUZ is a Ph.D. candidate in the Department of Industrial and Systems Engineering at the University of Miami. He received the B.Sc. and M.Sc. degrees in industrial engineering from the TOBB University of Economics and Technology, Ankara, Turkey, in 2017 and 2019, respectively. He is currently pursuing a Ph.D. degree in industrial engineering at the University of Miami. His current research effort is the development of a computational method to make optimal decisions under uncertain conditions. He also received the University of Miami Outstanding Graduate Research Award in 2023. His email address is axy286@miami.edu. His website is https://simlab.coe.miami.edu/team/.

NURCIN CELIK is a Professor in the Department of Industrial and Systems Engineering at the University of Miami (UM). She received her M.S. and Ph.D. degrees in Systems and Industrial Engineering from the University of Arizona with magna cum laude. Her reserch interests are on the development of integrated modeling and decision-making methodologies for large-scale, complex, and dynamic systems with learning capabilities. Notably, Dr. Celik received the Presidential Early Career Award for Scientists and Engineers from the White House in 2017. The PECASE Award is the highest honor bestowed by the U.S. government on outstanding scientists and engineers beginning their independent careers. She has served in the editorial board of academic journals and conference organizations, and has been an ad-hoc reviewer for more than (15) journals. Her email address is celik@miami.edu. Her website is https://simlab.coe.miami.edu/team/.