Detection of Power System Anomalies Using a Fusion of Machine Learning & Deep Learning

1st Jishnu Teja Dandamudi
Amrita School of Artificial Intelligence
Amrita Vishwa Vidyapeetham
Coimbatore, Tamilnadu, India
djishnuteja2006@gmail.com

2nd Rupa Kandula
Amrita School of Artificial Intelligence
Amrita Vishwa Vidyapeetham
Coimbatore, Tamilnadu, India
rupakandula21@gmail.com

3rd Lekshmi C.R.
Amrita School of Artificial Intelligence
Amrita Vishwa Vidyapeetham
Coimbatore, Tamilnadu, India
cr_lekshmi@cb.amrita.edu

***Abstract*—Anomaly detection in Phasor Measurement Unit (PMU) data is crucial for ensuring the stability and security of modern power grids. However, the complexity of power system dynamics and diverse operational conditions pose significant challenges. This study proposes a hybrid anomaly detection framework that combines Isolation Forest (IF) and Long Short-Term Memory (LSTM) Autoencoders, leveraging the strengths of both machine learning (ML) and deep learning (DL) techniques. IF efficiently detects anomalies using unsupervised learning, while LSTM Autoencoders capture temporal dependencies in PMU data for sequential anomaly detection. A fusion strategy integrates both models to enhance detection accuracy. The proposed framework is trained and tested on real-world PMU datasets encompassing normal operations, missing data, power disturbances, and cyber-induced anomalies. To address class imbalance, we employ the Synthetic Minority Over-sampling Technique (SMOTE), improving the detection of rare events. Performance evaluation using standard classification metrics demonstrates that the fusion model outperforms individual techniques, offering a robust and adaptive anomaly detection solution. This approach enhances situational awareness, fault diagnosis, and grid resilience, contributing to the reliability of modern power systems.**

Dataset: Realistic Labelled PMU Data for Cyber-Power Anomaly Detection Using Real-Time Synchrophasor Testbed available in IEEE DataPortal.

***Index Terms*—Anomaly Detection, Machine Learning(ML) Metrics, Deep Learning(DL), PMU Data, File Handling.**

I. Introduction

Maintaining voltage stability in modern power systems is becoming increasingly complex due to the integration of renewable energy sources and fluctuating load demands. Traditional droop control methods, while commonly used, often struggle to adapt to nonlinear system behavior and rapidly changing operating conditions. With the development of smart grids, power networks now incorporate advanced monitoring and communication technologies, producing vast amounts of real-time data from Phasor Measurement Units (PMUs). However, detecting anomalies in such high-dimensional, time-dependent data remains a significant challenge, particularly due to noise, operational variability, and the imbalance between normal and anomalous events.

The proposed method leverages Isolation Forest, a decision-tree-based algorithm effective for unsupervised anomaly detection, and Long Short-Term Memory (LSTM) Autoencoders, which are well-suited for learning sequential dependencies in time-series data. By combining these techniques, the framework enhances the identification of anomalies in power system data, ensuring a more adaptive and effective detection process. Unlike conventional rule-based approaches that rely on predefined thresholds, this model learns patterns directly from data, enabling it to generalize across diverse scenarios, including system disturbances, missing data, and potential cyber threats.

The novelty of this study lies in the strategic fusion of ML and DL to enhance detection performance and robustness. Isolation Forest [1] effectively detects irregular patterns through recursive partitioning, while LSTM Autoencoders [7] reconstruct time-series sequences to highlight deviations from normal behavior. Additionally, to address the common issue of class imbalance—where anomalies occur far less frequently than normal conditions—the Synthetic Minority Over-sampling Technique (SMOTE) is employed to improve model sensitivity to rare events. Unlike traditional methods that may overlook infrequent anomalies, this approach ensures better detection rates while maintaining a low false alarm rate, making it more suitable for real-world applications.

Anomaly detection in power systems presents several key challenges. First, PMU data is high-dimensional, requiring efficient techniques to process and extract meaningful features. The dynamic nature of the grid, influenced by fluctuations in demand and renewable energy generation, further complicates the task of distinguishing between normal variations and true anomalies. Additionally, conventional classifiers often struggle with heavily imbalanced datasets, leading to poor performance in detecting rare events. Another major concern is the increasing vulnerability of smart grids to cyberattacks, making it crucial to develop anomaly detection methods that can differentiate between natural disturbances and malicious activities.

The proposed framework offers several advantages that make it well-suited for real-time deployment in smart grids. By integrating both ML and DL, the model achieves improved anomaly detection accuracy while maintaining scalability and adaptability. Its ability to handle multiple types of anomalies—including physical faults, sensor failures, and cyber-related anomalies—enhances its reliability in diverse operating conditions. Furthermore, its robustness to missing or noisy

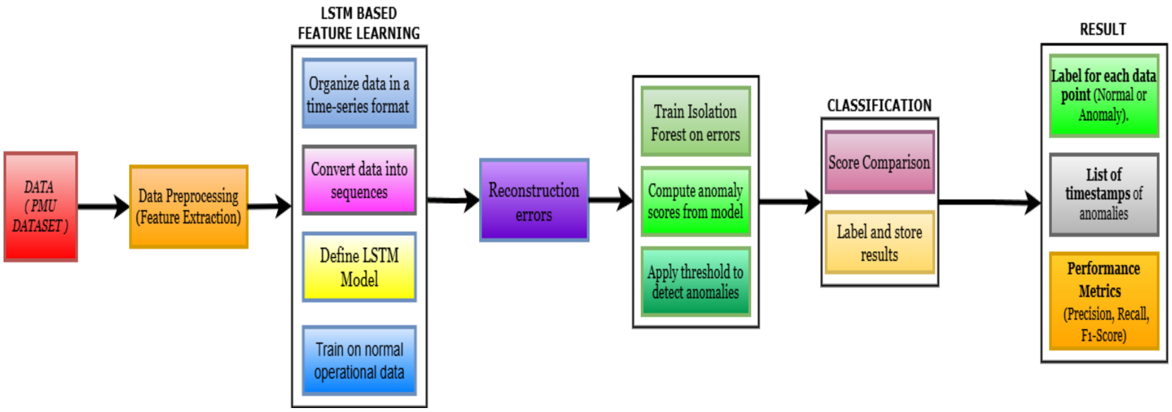
LSTM BASED
FEATURE LEARNING
RESULT
DATA
(PMU
DATASET)
Data Preprocessing
(Feature Extraction)
Organize data in a time-series format
Convert data into sequences
Define LSTM Model
Train on normal operational data
Reconstruction errors
Train Isolation Forest on errors
Compute anomaly scores from model
Apply threshold to detect anomalies
CLASSIFICATION
Score Comparison
Label and store results
Label for each data point (Normal or Anomaly).
List of timestamps of anomalies
Performance Metrics
(Precision, Recall, F1-Score)

Fig. 1. Block diagram of proposed method of detection of power system anomalies

data ensures consistent performance in practical applications. These features collectively contribute to improving grid stability, fault diagnosis, and overall system resilience.

The applications of this anomaly detection framework extend across various domains in smart grid operations. It can be used for continuous power grid monitoring to detect faults before they escalate into larger failures. Additionally, it plays a role in identifying cyber intrusions by recognizing deviations in grid behavior caused by unauthorized activities. In predictive maintenance, the framework helps utility providers anticipate potential failures, reducing downtime and improving operational efficiency. Furthermore, with the increasing integration of renewable energy, this approach assists in maintaining grid stability by monitoring fluctuations and ensuring reliable energy distribution.

In summary, this study presents a hybrid approach that integrates ML and DL to enhance the accuracy and efficiency of anomaly detection in power systems against the comparison done for many ML models. [14] By addressing critical challenges such as high data dimensionality, class imbalance, and real-time monitoring constraints, the proposed framework provides a reliable and scalable solution for improving the security and resilience of modern smart grids.

II. Related Work

According to Mustafa, Hussain M in their paper ”Realistic Synchrophasor Data Generation for Anomaly Detection using Cyber-Power Testbed,” the growing complexity of modern power systems necessitates advanced monitoring and anomaly detection to ensure stability, reliability, and efficiency. PMUs are critical for providing real-time, time-synchronized measurements of electrical quantities but are susceptible to anomalies from equipment malfunctions, cyberattacks, environmental factors, and data transmission errors. To address this, a Hardware-in-the-Loop (HIL) synchrophasor testbed has been developed to generate realistic data for training and validating anomaly detection algorithms. This testbed integrates components like a Real-Time Digital Simulator (RTDS), hardware/software PMUs, a Real-Time Automation Controller (RTAC), a network simulator (ns-3), and cloud-based data storage via PingThings. It emulates real-world scenarios, including normal operations, missing data, faults, and cyber-induced anomalies, generating high-fidelity, time-synchronized datasets for robust anomaly detection. A key feature is the introduction of realistic noise, bad data, and cyber disruptions. The datasets are used to evaluate anomaly detection algorithms like SyncAD, which combines statistical methods and clustering. Experimental results show that the testbed-generated data accurately reflects real-world PMU behavior, enabling precise anomaly detection with high precision and recall. The publicly available dataset was also used in an anomaly detection competition at the International Conference on Smart Grid Synchronized Measurements & Analytics (SGSMA), fostering innovation in smart grid monitoring. This testbed highlights the importance of high-quality, labeled datasets for advancing ML-based anomaly detection and strengthening cyber-physical security and grid resilience. [11]

In the paper ”Advanced Anomaly Detection in Energy Control Systems Using ML and Feature Engineering” according to the authors, the increasing integration of Industrial Control Systems (ICSs) with modern digital platforms has led to significant improvements in operational efficiency but has also introduced substantial cybersecurity risks. Traditional rule-based intrusion detection methods often struggle to detect sophisticated cyber threats, necessitating the use of advanced ML techniques for anomaly detection. [12] Their study presents a novel two-tiered anomaly detection framework specifically designed for ICSs in energy control environments, such as smart grids and renewable energy systems. The first tier is dedicated to fault detection, utilizing tree-based models to identify deviations from normal operational patterns with high precision. Among the models tested, XGBoost demonstrated the highest accuracy, achieving an outstanding 99.9% detection rate while maintaining minimal false positives and negatives. The second tier focuses on anomaly diagnosis, classifying specific types of cyber threats such as port scanning, replay attacks, Distributed Denial of Service (DDoS), and Man-In-The-Middle (MITM) attacks. The Random Forest classifier excelled in this phase, achieving near-perfect accuracy, correctly identifying all instances of port scan, replay, and DDoS attacks with minimal errors in detecting MITM threats. To ensure robustness, the framework employs a comprehensive preprocessing pipeline, including feature selection, normalization, and engineered features derived from ICS network communications. The dataset used for training was carefully curated, incorporating real and simulated ICS traffic, with labeled anomalies for both supervised and unsupervised learning approaches. A comparative analysis with previous studies highlights the superior performance of this approach, demonstrating that combining feature engineering with an optimized ML model selection significantly enhances detection capabilities. The proposed method is not only highly accurate but also computationally efficient, making it suitable for real-time ICS security monitoring. By proactively identifying and diagnosing cyber anomalies, this framework offers a critical advancement in securing industrial energy control systems, ultimately improving resilience against evolving cyber threats. [3]

Smart grids represent a transformative advancement in power generation, distribution, and consumption by integrating digital technologies and two-way communication for improved efficiency and resilience. However, this increased connectivity also introduces new cybersecurity challenges, making anomaly detection a crucial aspect of smart grid security according to the authors of ”Anomaly Detection in Smart Grids: A Survey From Cybersecurity Perspective”. Traditional power grids relied on one-way communication from power stations to consumers, but smart grids enable dynamic interactions through smart meters, distributed energy resources, and automated control systems. While these advancements enhance reliability and sustainability, they also expand the attack surface for cyber threats, including false data injection, denial-of-service attacks, and energy theft. To address these risks, advanced anomaly detection techniques leveraging ML and Artificial Intelligence (AI) [6], [13] have emerged as key solutions. Various supervised, unsupervised, and DL models have been developed to analyze power consumption patterns, network traffic, and state estimation data to detect anomalies in real time. Supervised learning methods, such as Support Vector Machines (SVM) [5] and Decision Trees, require labeled datasets for training and have shown success in identifying cyber intrusions. Unsupervised techniques, including IFs and Principal Component Analysis (PCA), help detect novel anomalies without predefined labels. DL approaches, such as Convolutional Neural Networks (CNN) [10] and Recurrent Neural Networks (RNN) [8], offer high accuracy in identifying complex attack patterns by analyzing time-series data. Furthermore, hybrid models combining multiple detection strategies have demonstrated superior performance in identifying and mitigating cyber threats. The integration of these AI-driven anomaly detection methods with real-time monitoring systems enhances the security and resilience of smart grids, ensuring reliable energy distribution while mitigating risks associated with cyber-attacks. As smart grids continue to evolve, ongoing research in advanced anomaly detection and cybersecurity measures will be essential to safeguard critical energy infrastructure from emerging threats. [2]

The integration of ML and DL in smart grids, as discussed in the paper ”HIL Testbed-based Auto Feature Extraction and Data Generation Framework for ML/DL-based Anomaly Detection and Classification,” has greatly enhanced anomaly detection and classification, improving power system resilience and efficiency. However, the effectiveness of these models relies on high-quality, diverse, and unbiased datasets, which traditional datasets often lack due to issues like data insufficiency, biases, and an inability to capture the complex behaviors of power systems. To address these challenges, the Auto Feature Extraction and Data Generation (AFEDG) framework combines a Python-based API with a Hardware-in-the-Loop (HIL) Cyber-Physical System (CPS) testbed, ensuring automated feature extraction, real-time data generation, and coverage of critical events, such as faults and cyberattacks. The framework uses protocols like IEEE C37.118 and IEC-61850 to collect high-resolution data from virtual sensors like PMUs and Intelligent Electronic Devices (IEDs). Monte Carlo simulations generate diverse fault scenarios, while OPAL-RT powers real-time simulation of power grid dynamics. This approach addresses data sufficiency, reduces biases, and improves ML/DL [4] training with balanced, event-specific data. Ultimately, this advancement enhances the security and intelligence of smart grid operations, with future work expanding to cyberattack and perturbation datasets, further strengthening system stability and security. [9]

III. Proposed Methodology

Anomaly detection in power systems requires a robust framework capable of distinguishing between normal operational variations and critical faults. This study introduces a fusion approach combining Long Short-Term Memory (LSTM) Autoencoder and Isolation Forest (IF) to enhance anomaly detection accuracy. The methodology as shown in Fig. 1 follows a structured pipeline consisting of data preprocessing, model training, anomaly detection, and evaluation.

A. Dataset

The dataset used in this study is specifically designed for anomaly detection in power systems and consists of realistic, labeled synchrophasor data. It contains time-series measurements from PMUs, capturing the dynamic behavior of the power grid under various conditions. Detailed description of the dataset is given in Table. I The dataset includes both normal and anomalous scenarios, covering a wide range of physical and cyber-induced disturbances.

The collected data consists of multiple features such as three-phase voltages and currents (both magnitude and angle), frequency, and the Rate of Change of Frequency (ROCOF). The dataset is divided into training and testing sets, ensuring a comprehensive evaluation of the proposed anomaly detection framework. Various types of events, including load changes, line faults, cyber-attacks, and combined cyber-physical faults,

TABLE I
DATASET INFORMATION

Category	Description
Features	- 3-phase voltages (magnitude and angle) - 3-phase currents (magnitude and angle) - Frequency and ROCOF
Training Data	- Collected over **90 minutes**, consisting of approximately **160,000 data points** - Covers **physical events** (e.g., load changes, line-to-ground faults) and **cyber events** (e.g., data drops)
Testing Data	- Two separate test sets, each spanning **25 minutes** (˜44,000 data points) - Includes events at different locations and of varying types (e.g., generation changes, cyber-physical faults)
Event Labels	- Categories: **Normal operation, fault events, line outages, generation changes, load changes, cyber faults,** and **combined cyber-physical faults**
Noise Integration	- **Signal-to-Noise Ratio (SNR) levels** introduced to simulate real-world conditions - Example: **75 dB** noise added to frequency measurements and **47 dB** noise added to voltage and current measurements

are labeled for supervised learning and performance benchmarking. To simulate real-world conditions, controlled noise is introduced at different Signal-to-Noise Ratio (SNR) levels, making the dataset more robust for practical deployment. This structured dataset enables the evaluation of anomaly detection models under diverse operating conditions, ensuring robustness and reliability in real-world applications.

B. Data Preprocessing

To ensure high-quality input data for model training, we apply several preprocessing steps:

1. Handling Missing Data: Missing values in PMU data are addressed using the forward-fill technique, where missing values are replaced with the last available observation, preserving temporal continuity.

2. Feature Scaling: Since DL models perform optimally on normalized inputs, all feature values are scaled between [0,1] using MinMax normalization using Eqn. 1:

$$X_{\text{scaled}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

3. Addressing Class Imbalance: Anomalies in power systems occur infrequently, leading to imbalanced datasets. To counter this, the Synthetic Minority Over-sampling Technique (SMOTE) is employed to generate synthetic samples for underrepresented anomaly classes, improving model generalization.

C. LSTM Autoencoder for Temporal Anomaly Detection

LSTM Autoencoder is employed to model the temporal dependencies in PMU data. Traditional feedforward neural networks are insufficient for capturing sequential patterns, as they lack memory mechanisms. LSTM networks, however, address this limitation by maintaining long-range dependencies through their gated memory structure. The Autoencoder architecture consists of an encoder that compresses input sequences into a latent representation and a decoder that reconstructs the input from this lower-dimensional space. The reconstruction error serves as a metric to identify anomalies, as deviations from expected patterns indicate potential faults in the power system.

The LSTM Autoencoder is formulated as shown in Eqn. 2 Let $X = \{X_1, X_2, \ldots, X_T\}$ be the input sequence with T time steps. The encoder maps this sequence into a lower-dimensional latent representation h_t at each time step t:

$$h_t = f(W_h X_t + U_h h_{t-1} + b_h) \tag{2}$$

where W_h and U_h are weight matrices, b_h is the bias term, and f represents the non-linear activation function (ReLU). The decoder reconstructs the input sequence $\hat{X}$ using the Eqn. 3:

$$\hat{X}_t = g(W_o h_t + b_o) \tag{3}$$

where W_o and b_o are the output weights and bias, respectively, and g represents the activation function.

The model learns to minimize the reconstruction loss, typically measured using Mean Squared Error (MSE) using Eqn. 4:

$$MSE = \frac{1}{T} \sum_{t=1}^{T} (X_t - \hat{X}_t)^2 \tag{4}$$

A high reconstruction error indicates that the input sequence significantly deviates from learned normal patterns, signifying an anomaly.

The LSTM Autoencoder is particularly suitable for PMU-based anomaly detection due to its ability to:

- Preserve temporal dependencies and detect subtle variations in time-series data.
- Encode and reconstruct complex patterns while filtering noise.
- Identify deviations in real-time, making it well-suited for power system monitoring.

Table II presents the architecture of the LSTM Autoencoder, detailing the layers and their corresponding parameters.

TABLE II
LSTM AUTOENCODER ARCHITECTURE

Layer	Units	Activation
LSTM Layer 1 (Encoder)	64	ReLU
Dropout	20%	-
LSTM Layer 2 (Encoder)	32	ReLU
Dense Layer (Decoder)	32	ReLU
Output Layer (Decoder)	Same as Input	-

Anomalies are detected when the reconstruction loss exceeds a predefined threshold, indicating a deviation from learned normal patterns.

D. *IF for Statistical Anomaly Detection*

While the LSTM Autoencoder detects temporal anomalies, the Isolation Forest (IF) is employed to enhance robustness by identifying statistical outliers in the feature space. Unlike traditional density-based methods, IF isolates anomalies by recursively partitioning the data and measuring how quickly observations become isolated within a tree structure.

The fundamental premise of IF is that anomalies, being distinct and infrequent, require significantly fewer splits to become isolated compared to normal data points. Consequently, anomalies exhibit shorter path lengths within the tree structure, making them easily distinguishable.

1) Training and Anomaly Classification: The IF is trained using the reconstruction error values obtained from the LSTM Autoencoder. This dual-stage approach ensures that both temporal deviations and statistical outliers are detected effectively.

- **Input:** The model is fed with reconstruction error values derived from the LSTM network.
- **Contamination Factor:** A contamination rate of 5% is assumed, meaning that approximately 5% of the dataset is expected to contain anomalous observations.
- **Output:** The model assigns binary labels based on anomaly scores:
  - Label -1: The observation is classified as an anomaly.
  - Label $+1$: The observation is classified as **normal**.

2) Threshold-Based Anomaly Refinement: In addition to the IF classification, an independent threshold-based approach is applied to further refine anomaly detection. The threshold is defined as the 95^{th} percentile of the reconstruction error distribution:

$$T = P_{95}(E), \tag{5}$$

where T represents the anomaly detection threshold and $P_{95}(E)$ denotes the 95^{th} percentile of the reconstruction error values E.

- If the reconstruction error of a data point surpasses T, it is classified as an anomaly.
- Otherwise, it is considered part of the normal data distribution.

Thus proposed methodology ensures a robust and adaptive framework for real-time power system anomaly detection, improving grid resilience and operational stability.

E. *Fusion of LSTM Autoencoder and Isolation Forest*

The integration of LSTM Autoencoder and IF combines the strengths of both techniques to enhance anomaly detection accuracy. The LSTM Autoencoder detects temporal anomalies by reconstructing time-series data and computing the reconstruction error, while the IF identifies statistical anomalies based on recursive partitioning of data points. To achieve a robust and accurate anomaly detection mechanism, a weighted fusion model is applied, where the final anomaly score $S_f(x)$ is computed as:

$$S_f(x) = \alpha \cdot S_{LSTM}(x) + \beta \cdot S_{IF}(x) \tag{6}$$

where $S_{LSTM}(x)$ is the anomaly score derived from the reconstruction error of the LSTM Autoencoder, and $S_{IF}(x)$ is the anomaly score assigned by the IF model. The coefficients α and β represent the respective weight contributions of each model, ensuring $\alpha + \beta = 1$. A threshold-based classification mechanism is then applied to categorize each data point as either normal or anomalous:

$$\text{Anomaly} = \begin{cases} 1, & S_f(x) > T \\ 0, & S_f(x) \leq T \end{cases} \tag{7}$$

where T represents the anomaly detection threshold. If the final anomaly score $S_f(x)$ exceeds T, the data point is classified as an anomaly; otherwise, it is considered part of the normal data distribution. This fusion approach effectively captures both temporal dependencies and statistical deviations, improving anomaly detection performance in complex time-series datasets.

Signal-to-Noise Ratio (SNR) is used to quantify the level of signal quality in a dataset. It helps understand how much useful information (signal) is present relative to unwanted disturbances (noise). Higher SNR means less noise, leading to better classification because anomalies and normal conditions are more distinguishable. Lower SNR increases uncertainty in the data, making it harder for a classifier to differentiate between normal and anomalous conditions.In real-world power systems (like PMU data with anomalies), noise can come from sensor errors, communication channels, or environmental interference. By simulating different SNR levels, you can test how robust your anomaly detection system is under noisy conditions. Noisy PMU data can disrupt anomaly detection, leading to incorrect system responses.

F. *Evaluation Metrics*

To evaluate the performance of the anomaly detection model, three key metrics are considered: **Precision, Recall, and F1-score**. These metrics provide insights into the model's ability to correctly identify anomalies while minimizing false detections.

Precision measures the proportion of correctly identified anomalies out of all instances classified as anomalies, ensuring that the detected anomalies are truly anomalous. It is defined as:

Algorithm 1 Fusion of LSTM Autoencoder and IF for Anomaly Detection

```
Require: Time-series data $X$
Ensure: Anomaly labels $Y$ (0: Normal, 1: Anomaly)
Preprocessing:
Normalize time-series data $X$
Create overlapping sequences as input for the LSTM Autoencoder
Train LSTM Autoencoder:
Define encoder with LSTM layers
Define decoder with Dense and output layers
Train using Mean Squared Error (MSE) loss:
    $$MSE = \frac{1}{N}\sum_{i=1}^{N}(X_i - \hat{X}_i)^2 \quad (8)$$
Compute reconstruction error $E$
Train IF:
Use reconstruction error $E$ as input
Train IF model
Compute anomaly scores $S_{IF}$
Fusion of Anomaly Scores:
Compute final anomaly score:
    $$S_f(x) = \alpha S_{LSTM}(x) + \beta S_{IF}(x) \quad (9)$$
    where $\alpha + \beta = 1$
Anomaly Classification:
Define threshold $T$ as the 95th percentile of $E$:
    $$T = P_{95}(E) \quad (10)$$
if $S_f(x) > T$ then
    Label as anomaly ($Y = 1$)
else
    Label as normal ($Y = 0$)
end if
Output anomaly labels $Y$
```

$$\text{Precision} = \frac{TP}{TP + FP} \tag{11}$$

where TP represents true positives (correctly detected anomalies), and FP denotes false positives (normal instances incorrectly classified as anomalies). [7]

Recall, also known as Sensitivity or True Positive Rate (TPR), evaluates the proportion of actual anomalies that were successfully detected. A higher recall value indicates fewer missed anomalies. It is given by:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{12}$$

where FN represents false negatives (actual anomalies that were not detected by the model). [7]

The **F1-score** serves as a harmonic mean between precision and recall, balancing both false positives and false negatives. It is computed as:

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{13}$$

These evaluation metrics are crucial for assessing the model's effectiveness in distinguishing between normal and anomalous instances, ensuring an optimal balance between detection accuracy and false alarm rates.

IV. Results and Analysis

The table III represents the precision, recall, and F1-score values also plotted in Fig. 2, which exhibit a decline as the number of training epochs increases. This decline can be attributed to the model's tendency to overfit normal data patterns. As training progresses, the LSTM autoencoder improves its ability to reconstruct normal sequences with high accuracy. However, this also means that it starts reconstructing anomalous sequences with lower errors, making them appear less distinct from normal data.

TABLE III
Performance Metrics

Metric	Value	Percentage (%)
Precision	0.8521	85.21
Recall	0.8529	85.29
F1-Score	0.8525	85.25

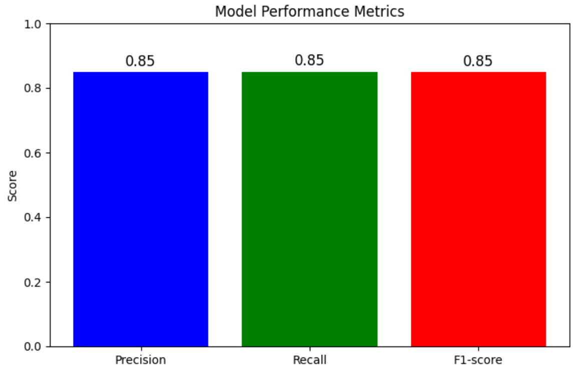

Fig. 2. Precision, Recall and F1 Plot

Consequently, fewer anomalies exceed the predefined reconstruction error threshold, leading to a drop in recall. Additionally, the decrease in reconstruction errors results in a less effective threshold for anomaly detection, causing the model to miss several true anomalies, thereby increasing false negatives and further reducing recall. Precision is also affected, as the model may still classify some normal instances as anomalies due to minor deviations, leading to false positives.

The plot of Fig. 3 represents the reconstruction error profile of an LSTM autoencoder-based anomaly detection system, with anomalous instances distinctly marked in red. The x-axis denotes the temporal evolution of the dataset, while the y-axis quantifies the magnitude of reconstruction error, serving

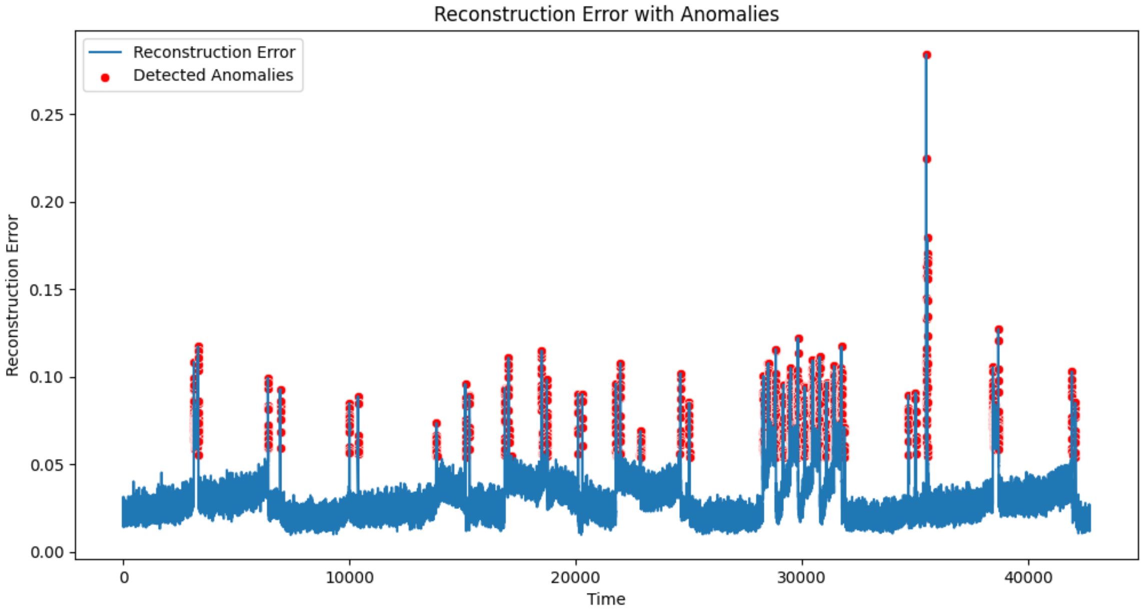

Fig. 3. Reconstruction Error with Anomalies

as an indicator of deviations between the model's learned normalcy and observed input patterns. The continuous blue curve encapsulates the reconstruction error trajectory, where relatively low values signify conformity to expected system behavior, whereas pronounced spikes indicate substantial deviations. The red markers correspond to data points classified as anomalous based on a predefined thresholding mechanism. The observed clustering of anomalies suggests transient perturbations, potentially linked to systemic fluctuations or operational inconsistencies within the monitored microgrid. Additionally, periodic excursions in reconstruction error hint at cyclical disturbances, possibly arising from recurring faults or transient instability within the system. The presence of extreme peaks further underscores the occurrence of substantial deviations, warranting a deeper investigation into their underlying causes, which could range from sensor anomalies to critical infrastructural faults.

The IF, which relies on statistical separation, is also influenced by changing reconstruction errors. Initially trained on higher reconstruction errors, its effectiveness diminishes as these errors shrink with more epochs. The shift in error distribution affects its ability to distinguish anomalies from normal instances, contributing to a lower anomaly detection rate. The combination of these factors leads to a decline in the overall F1-score, as both precision and recall are negatively impacted.

Fig. 4 illustrates the relationship between Signal-to-Noise Ratio (SNR) and classification accuracy, highlighting how noise levels impact model performance. At low SNR values (e.g., 100 dB), classification accuracy is around 90%, indicating that excessive noise interferes with the model's ability to distinguish normal and anomalous patterns. As the SNR increases, accuracy improves significantly, with a sharp rise observed between 100 dB and 400 dB, suggesting that noise reduction enhances feature clarity and improves anomaly detection. However, beyond a certain turning point in the curve, the improvement plateaus, indicating that excessively high SNR does not provide substantial additional benefits.

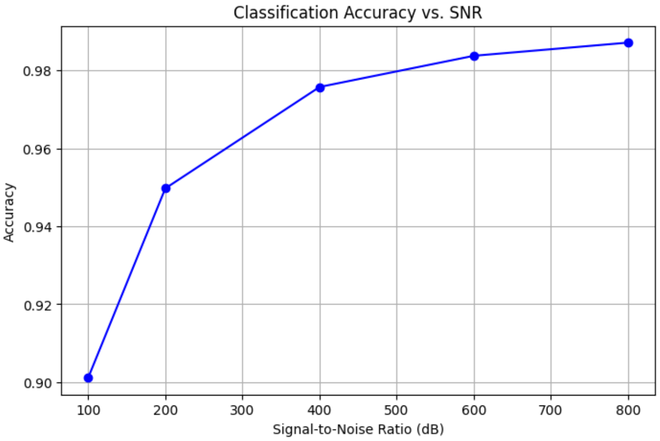

Fig. 4. Classification Accuracy vs. SNR

This saturation effect suggests that after a certain SNR threshold, the model has captured sufficient information for optimal classification, and further noise reduction primarily increases computational complexity rather than performance gains. The turning point in the curve helps determine the minimum SNR required to balance noise robustness and computational efficiency. If the SNR is too low, applying denoising techniques (e.g., filtering or preprocessing) may be necessary to improve data quality before feeding it into the anomaly detection model.

V. Conclusion & Future Works

The proposed hybrid anomaly detection framework combines an LSTM autoencoder for feature extraction with an IF for anomaly classification, effectively leveraging DL and ML for power system monitoring. The LSTM autoencoder captures temporal dependencies, while reconstruction errors guide the IF in isolating anomalies efficiently. This dual-stage approach enhances detection accuracy while minimizing false positives and negatives. Performance evaluation using precision, recall, and F1-score confirms its effectiveness for real-time power grid surveillance.

Future enhancements in anomaly detection can focus on adaptive thresholding using Bayesian optimization or reinforcement learning to improve classification accuracy. A multimodal DL approach integrating PMU data, weather conditions, and cybersecurity logs with Graph Neural Networks (GNNs) or Transformers could enhance detection robustness. *Self-supervised learning* methods like contrastive learning can address the challenge of limited labeled anomalies. Quantum ML techniques, such as Quantum-enhanced LSTMs, may further improve computational efficiency. Explainable AI (XAI) using SHAP or LIME can enhance model interpretability, while federated learning enables decentralized model training for secure data sharing. Additionally, adversarial robustness* through GAN-based anomaly simulation can strengthen cybersecurity. These advancements will lead to a more intelligent, secure, and scalable anomaly detection framework for real-time grid monitoring.

References

[1] Gandhimathinathan A, Ananthakrishnan C G, Lavanya R, R Jehadeesan, and Pidapa Raghava Reddy. Sensor anomaly detection in nuclear power plant using deep lstm denoising autoencoder and isolation forest. *IEEE Sensors Letters*, 8(12):1–4, 2024.

[2] Ahmad N Alkuwari, Saif Al-Kuwari, and Marwa Qaraqe. Anomaly detection in smart grids: A survey from cybersecurity perspective. In *2022 3rd International Conference on Smart Grid and Renewable Energy (SGRE)*, pages 1–7. IEEE, 2022.

[3] Zaid Allal, Hassan Noura, Ola Salman, and Ali Chehab. Advanced anomaly detection in energy control systems using machine learning and feature engineering. In *2024 8th Cyber Security in Networking Conference (CSNet)*, pages 15–21. IEEE, 2024.

[4] Raghavendra Chalapathy and Sanjay Chawla. Deep learning for anomaly detection: A survey. *arXiv preprint arXiv:1901.03407*, 2019.

[5] Ramin Ghiasi, Muhammad Arslan Khan, Danilo Sorrentino, Cassandre Diaine, and Abdollah Malekjafarian. An unsupervised anomaly detection framework for onboard monitoring of railway track geometrical defects using one-class support vector machine. *Engineering Applications of Artificial Intelligence*, 133:108167, 2024.

[6] Apoorv Joshi, Amrita, Rohan Sahai Mathur, Nitendra Kumar, and Padmesh Tripathi. Study of traditional, artificial intelligence and machine learning based approaches for moving object detection. *Mathematical Models Using Artificial Intelligence for Surveillance Systems*, pages 187–214, 2024.

[7] R Prasanna Kumar, Bharathi Mohan G, Elakkiya R, Charan Kumar M, and Rithani M. Automated sentiment classification of amazon product reviews using lstm and bidirectional lstm. In *2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT)*, pages 1–6, 2023.

[8] S Joshua Kumaresan, C Senthilkumar, Dinokumar Kongkham, Beenarani BB, and P Nirmala. Investigating the effectiveness of recurrent neural networks for network anomaly detection. In *2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)*, pages 1–5. IEEE, 2024.

[9] Aditya Akilesh Mantha, Arif Hussain, and Gelli Ravikumar. Hil testbed-based auto feature extraction and data generation framework for ml/dl-based anomaly detection and classification. In *2024 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pages 1–5. IEEE, 2024.

[10] Jose Mendoza-Bernal, Aurora Gonzalez-Vidal, and Antonio F Skarmeta. A convolutional neural network approach for image-based anomaly detection in smart agriculture. *Expert Systems with Applications*, 247:123210, 2024.

[11] Hussain M Mustafa, Vasavi Sivaramakrishnan, Vignesh VG Krishnan, and Anurag Srivastava. Realistic synchrophasor data generation for anomaly detection using cyber-power testbed. In *2024 56th North American Power Symposium (NAPS)*, pages 1–6. IEEE, 2024.

[12] SB Kavya Preetha, Jalakam Venu Madhava Sai, V Sowbaranic Raj, M Jayan Sekhar, and R Lavanya. Anomaly detection in satellite power system using deep learning. In *2024 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*, pages 1–5. IEEE, 2024.

[13] Ch Rahul AN Sharma, Dutta Swetchana, and Shinu M Rajagopal. Ai anomaly detection with decentralized identity management on blockchain. In *2024 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*, pages 1–6. IEEE, 2024.

[14] TS Siddhesh, Shinu M Rajagopal, and Sreebha Bhaskaran. Comparative analysis of machine learning algorithms for anomaly detection. In *2024 IEEE 9th International Conference for Convergence in Technology (I2CT)*, pages 1–5. IEEE, 2024.