LSTM-Based Approach for Predictive Link Monitoring and RRU Anomaly Detection in 4G/5G Networks

Pannatorn Kijphitayarit[1] and Settawit Poochaya[2*]

[1,2]School of Telecommunication Engineering, Institute of Engineering, Suranaree University of Technology, Nakhon Ratchasima

settawit@sut.ac.th

Received 25 April 2025

Revised 13 June 2025

Accepted 24 June 2025

Abstract

The increasing complexity of 4G and 5G networks has intensified the need for reliable and efficient telecommunication infrastructures. Remote Radio Units (RRUs) are crucial components responsible for ensuring seamless signal transmission between user devices and the core network. However, anomalies in RRU voltage and temperature can cause signal degradation, network inefficiencies, and potential failures. Traditional rule-based fault detection systems often struggle to adapt to dynamic network conditions, necessitating the integration of advanced deep learning models for proactive anomaly detection. Long Short-Term Memory (LSTM) networks have demonstrated superior capabilities in analyzing time-series data, making them well-suited for detecting performance anomalies in RRUs. However, a significant challenge in deploying these models is the class imbalance problem, where normal operational conditions vastly outnumber rare fault instances, leading to biased predictions and poor recall for minority-class anomalies. This research aims to determine the effectiveness of RandomOverSampler in improving the performance of LSTM-based anomaly detection models when applied to imbalanced RRU datasets in the context of 4G/5G network monitoring. To address this, a resampling strategy utilizing RandomOverSampler is implemented to balance the dataset, which consists of 5,000 time-series samples with two key features: voltage and temperature, ensuring improved detection of rare failures without introducing synthetic noise. The proposed framework processes sequential RRU voltage and temperature data, capturing temporal dependencies to improve failure predictions. Performance evaluations show that the minority-class recall improved from 33% to 99%, and the F1-score increased from 32% to 99% after resampling, effectively addressing a major limitation of conventional machine learning-based anomaly detection systems. The model also achieves an

overall accuracy of 99%, demonstrating its robustness and suitability for real-world deployment in mobile network monitoring. Future work will focus on extending this framework to predict Radio Link Failure (RLF) based on RAW RRU data performance patterns.

Keywords: 5G Technology, 4G Technology, LSTM, Imbalanced Data, Time Series

1. Introduction

The rapid expansion of 4G and 5G networks has significantly increased the demand for stable and reliable mobile communication services. Remote Radio Units (RRUs) play a critical role in these networks, serving as the primary interface for signal transmission between user devices and the core network. Ensuring the operational stability of RRUs is essential for maintaining network performance and Quality of Service (QoS). However, voltage fluctuations and temperature anomalies within RRU systems can lead to signal degradation, reduced efficiency, or even complete network failure. According to field observations shared by a global Telco vendor, a key vendor collaborating with SUT under a memorandum of understanding (MOU), such anomalies frequently contribute to unexpected service instability in real-world 4G/5G deployments. As a result, effective anomaly detection mechanisms are crucial for preventing disruptions and ensuring seamless mobile network operations.

Advancements in machine learning and deep learning have enabled the development of automated anomaly detection systems for mobile network monitoring. Traditional rule-based and threshold-based approaches, while widely used, often struggle to adapt to dynamic network environments. Machine learning models, particularly those based on classification algorithms, have demonstrated improved accuracy in fault detection by identifying patterns in historical data. More recently, Long Short-Term Memory (LSTM) networks have gained attention due to their ability to process sequential data and capture temporal dependencies in network performance metrics, making them well-suited for detecting anomalies in time-series data.

Despite the effectiveness of deep learning models in anomaly detection, a key challenge remains: imbalanced datasets. In real-world RRU monitoring, normal operating conditions vastly outnumber instances of actual anomalies, making rare faults difficult to detect. Standard LSTM models trained on such datasets tend to favor the majority class, leading to low recall for minority-class anomalies. This bias severely limits the practical utility of anomaly detection systems, as critical but infrequent failures may go undetected, posing risks to network stability. Moreover, this study is constrained by the limited sample size in certain minority classes, which prevents the effective use of synthetic oversampling techniques such as SMOTE. As a result, resampling methods that avoid synthetic data generation were adopted to preserve data integrity while mitigating class imbalance.

Several techniques have been proposed to mitigate class imbalance in machine learning, including oversampling and undersampling methods. However, conventional synthetic oversampling techniques often generate artificial data that may not accurately represent real-world anomalies, potentially introducing noise and leading to overfitting. An alternative approach involves balancing the dataset while preserving the original data distribution, allowing the model to learn more effectively from minority-class samples without distorting their characteristics.

This study aims to develop a robust LSTM-based anomaly detection framework designed to effectively monitor RRU operational stability under imbalanced data conditions. To achieve this, the proposed approach integrates a resampling strategy to ensure a more balanced class distribution, thereby enhancing the model's ability to detect anomalies in minority classes. To this end, the study specifically focuses on preprocessing and resampling imbalanced RRU operational data using the RandomOverSampler to balance the dataset without introducing synthetic noise, developing an LSTM-based model optimized for time-series anomaly detection that captures temporal dependencies in RRU voltage and temperature fluctuations, and evaluating the model's performance before and after resampling to assess the impact of balancing techniques on precision, recall, and F1-score.

2. Background

The rapid evolution of 4G/5G mobile networks has introduced new challenges in ensuring network reliability and efficient operation and maintenance (O&M) [1]. As mobile communication services become more critical, maintaining the performance of Radio Remote Units (RRUs) is essential to prevent service degradation and failures. Traditional reactive fault detection methods often fail to provide timely insights, leading to prolonged downtimes and reduced network efficiency [2]. The increasing complexity of large-scale distributed antenna systems further necessitates advanced fault detection mechanisms that can accurately identify network anomalies [3].

To enhance fault detection and anomaly prediction, machine learning (ML) and deep learning (DL) techniques have gained attention due to their ability to process large volumes of network data and detect patterns that indicate potential failures [4]. Among these approaches, Long Short-Term Memory (LSTM) networks have demonstrated strong capabilities in analyzing time-series network data, enabling real-time failure prediction [5]. However, applying ML-based failure detection models in real-world network environments is challenging due to the class imbalance problem, where network failure events occur significantly less frequently than normal operational data [6].

Addressing class imbalance is crucial for improving the accuracy of ML models in network anomaly detection. Various techniques, such as undersampling, oversampling, and the Synthetic Minority Oversampling Technique (SMOTE), have been proposed to improve model robustness in handling imbalanced datasets [7]. Additionally, the presence of class overlap further complicates the classification process, making it difficult for models to distinguish between normal and failure states [8].

To overcome these limitations, this study proposes an LSTM-based predictive maintenance framework tailored for 4G/5G network fault detection. By integrating advanced data preprocessing techniques and deep learning-based time-series analysis, this approach aims to enhance failure prediction accuracy, support proactive maintenance strategies, and contribute to the resilience of modern mobile communication networks [9].

3. Related Works

The increasing complexity of 5G networks has driven significant advancements in fault detection, resource management, and predictive analytics to maintain network stability and service reliability. Researchers have explored various approaches, including machine learning (ML), deep learning (DL), and artificial intelligence (AI)-based methods, to enhance fault prediction and network optimization. One of the key challenges in 5G network maintenance is radio link failure (RLF) prediction, which is influenced by multiple environmental factors. Several studies have demonstrated that integrating historical weather data and link parameters can improve the accuracy of failure predictions, allowing operators to anticipate and mitigate disruptions more effectively [10] . The impact of environmental conditions, including humidity, wind speed, and temperature fluctuations, has also been examined, revealing a strong correlation between adverse weather and increased radio link degradation [11]. These findings highlight the importance of real-time environmental monitoring in predictive network maintenance.

Advancements in ML-based fault prediction have further expanded the capabilities of network anomaly detection. Ensemble learning techniques have been studied extensively, demonstrating their effectiveness in enhancing fault detection accuracy in telecommunication networks [12] . Additionally, ML-based fault detection in large-scale distributed antenna systems has been explored through time-series similarity analysis, enabling early identification of system anomalies . A broader review of ML applications in heterogeneous telecommunication networks underscores the growing role of AI-driven methodologies in predictive maintenance [13].

Beyond traditional ML models, deep learning techniques, particularly Long Short-Term Memory (LSTM) networks, have shown promise in optimizing 5G resource allocation. These models effectively process time-series data, capturing long-term dependencies to enhance network efficiency and reduce latency [14]. LSTM-based methods have also been applied in Quality of Service (QoS) prediction, particularly for connected and automated mobility (CAM) applications, ensuring proactive resource allocation under dynamic network conditions [15].

Further innovations in predictive analytics have emerged with the use of Graph Neural Networks (GNNs), which leverage spatiotemporal dependencies within network datasets to improve cellular Key Performance Indicator (KPI) prediction [16]. This approach has been found to enhance network performance forecasting, allowing for smarter and more adaptive network management strategies. One of the critical challenges in network fault prediction is the class imbalance problem, where failure events constitute only a small fraction of total network data. Addressing this issue, researchers have explored various data resampling techniques, such as undersampling, oversampling, and Synthetic Minority Oversampling Technique (SMOTE), to improve model accuracy in detecting rare faults [17] . Comparative studies have suggested that integrating sampling strategies with ensemble learning can significantly enhance prediction performance, ensuring a more balanced approach to network anomaly detection.

Among oversampling techniques, SMOTE [18] and ADASYN [19] have been widely adopted to generate synthetic samples and enhance class representation. SMOTE applies interpolation between existing minority-class samples, while ADASYN adaptively generates synthetic samples based on the distribution of difficult-to-classify instances. However, both methods require a sufficient number of minority-class instances to function effectively. Given the dataset constraints and specific requirements set by the vendor, these methods were not applicable in this study. Instead, RandomOverSampler was employed as a practical alternative to balance the dataset while ensuring compliance with vendor requirements and maintaining data integrity. This approach allowed the model to learn from a more evenly distributed dataset while avoiding potential biases introduced by synthetic sample generation.

Recent studies have demonstrated that despite its simplicity, RandomOverSampler (ROS) remains a robust method for handling imbalanced datasets, often achieving competitive accuracy compared to more advanced sampling techniques [20]. Unlike SMOTE and ADASYN, which generate synthetic samples, ROS preserves the original feature distribution by duplicating existing minority instances rather than synthesizing new data points. Given its computational efficiency and practical implementation benefits, ROS was chosen as the primary resampling technique in this study. This decision was further reinforced by vendor constraints, which restricted the introduction of synthetic data and required strict preservation of the original dataset structure.

Multiple Sampling Methods (MSM) have been introduced to handle class imbalance by integrating different resampling techniques, such as combining oversampling with undersampling to improve classifier performance. Previous studies have demonstrated that MSM-based approaches can enhance predictive accuracy in imbalanced datasets, particularly in network anomaly detection and fault prediction tasks [21]. However, these methods typically require careful tuning to prevent overfitting and may introduce synthetic data distributions that deviate from real-world conditions. In this study, RandomOverSampler was employed as the primary technique to address class imbalance, ensuring compliance with vendor constraints while maintaining the original data distribution. Although MSM-based approaches offer advantages in general cases, the dataset constraints in this study limited the feasibility of synthetic oversampling techniques, making RandomOverSampler the most suitable choice.

In addition to ML and deep learning approaches, recent studies have explored alarm log-based fault prediction models that utilize entropy-based layered analysis to predict hardware failures [22]. By analyzing historical alarm data, these models improve predictive maintenance capabilities, reducing downtime and enhancing overall network reliability. Furthermore, research in multi-scenario

KPI prediction has focused on deep learning-based optimizations, which help in improving network stability and fault recovery mechanisms [23].

The collective findings from these studies provide a strong foundation for developing advanced predictive maintenance frameworks that integrate LSTM-based fault detection, imbalanced data handling, and KPI-driven optimization. As 5 G networks continue to evolve, leveraging AI-driven methodologies will be essential for ensuring scalable, resilient, and self-optimizing telecommunication infrastructures.

4. Research Method

4.1 Data Preprocessing

4.1.1 Data Collection and Cleaning

The dataset used in this study consists of 5,000 samples of Remote Radio Unit (RRU) operational parameters, specifically voltage and temperature readings. The data were extracted from structured Excel files and underwent a thorough cleaning process to ensure reliability. Non-numeric entries and missing values were identified and removed. This preprocessing step ensured that the data were accurate, consistent, and suitable for subsequent analysis.

4.1.2 Label Assignment

Each sample in the dataset was categorized into one of four predefined classes, based on voltage and temperature thresholds:

• Normal (1): Voltage between 40–60 V and temperature between -25°C and 60°C.

• Abnormal Voltage (0.1): Voltage outside the acceptable range (40–60 V), while temperature remained within -25°C to 60°C.

• Abnormal Temperature (0.2): Voltage within the acceptable range (40–60 V), but temperature outside -25°C to 60°C.

• Both Abnormal (0): Both voltage and temperature outside their respective thresholds.

This labeling process ensured that critical RRU operational states were captured comprehensively. A significant class imbalance was observed, with the majority of samples labeled as Normal (1), while the other three categories were underrepresented.

4.1.3 Addressing Class Imbalance

Due to specific constraints set by the vendor, the implementation of synthetic oversampling techniques such as SMOTE and ADASYN was not feasible for this dataset. Instead, RandomOverSampler was employed to address class imbalance while ensuring compliance with vendor requirements. This approach maintains data integrity and avoids potential issues associated with synthetic sample generation.

The RandomOverSampler method was applied to duplicate instances of the underrepresented classes until a balanced class distribution was achieved. This ensured that all categories had sufficient representation during training. Unlike synthetic oversampling techniques, which generate artificial samples, RandomOverSampler preserves the original feature distribution by using actual instances from the minority classes. This decision was supported by previous studies demonstrating that RandomOverSampler can enhance model robustness while maintaining computational efficiency.

The overall workflow of the proposed anomaly detection framework is illustrated in Fig. 1.

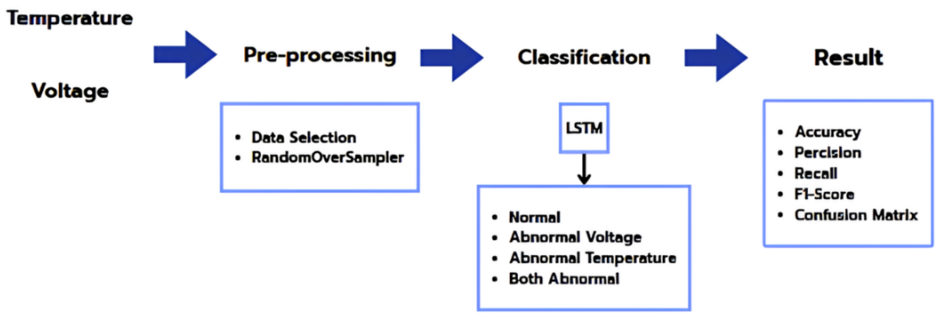

Fig. 1 Overview of the proposed RRU anomaly detection framework.

4.2 Model Development

4.2.1 Feature Scaling and Sequence Creation

Voltage and temperature values were normalized to a range of [0, 1] using Min-Max Scaling to standardize the features. This ensured that both features contributed equally during training and avoided bias toward one feature over the other. To capture temporal dependencies in the data, sequences of three consecutive time steps were generated for each sample. This sequence creation step allowed the model to learn patterns over time and improved its ability to detect anomalies.

A fixed sequence length of 3 timesteps was selected to balance model simplicity with computational efficiency. This window size allows the model to learn short-term temporal patterns without overfitting or incurring excessive training cost. Future work will explore longer sequence lengths to capture more extended temporal dependencies and enhance detection of delayed anomalies. Only two features voltage and temperature were used in this study. This selection was informed by domain knowledge shared by the vendor, which identified these parameters as the most frequent and critical indicators of RRU/AAU faults. Although this is a limited feature set, it reflects real-world operational focus. Future studies will consider additional KPIs, such as RSRP, SINR, and RSSI, to enhance predictive accuracy and generalizability.

4.2.2 LSTM Model Architecture

The model was designed using TensorFlow/Keras and featured a Long Short-Term Memory (LSTM) architecture optimized for time-series data. The key components of the model are as follows:

• Input Layer: Accepting sequences with a shape of (3, 2), corresponding to three time steps and two features (voltage and temperature).

• First LSTM Layer: 64 memory units with ReLU activation and return_sequences=True to pass the sequence to the next LSTM layer.

• Dropout Layer: A dropout rate of 30% to reduce overfitting.

• Second LSTM Layer: 32 memory units with ReLU activation.

• Second Dropout Layer: A dropout rate of 20% to further mitigate overfitting.

• Output Layer: A dense layer with softmax activation, producing probabilities for the four output classes.

The model was compiled using the Adam optimizer and categorical cross entropy as the loss function. This architecture was designed to balance performance and complexity while effectively capturing temporal patterns in the data.

The model was trained over 50 epochs using a batch size of 32. The learning process was monitored using a 20% validation split extracted from the training data. Class weights were incorporated to address residual class imbalance, computed using an inverse frequency strategy via scikit-learn's compute_class_weight function. These training hyperparameters and preprocessing steps are summarized in Table 1 to support reproducibility. These values were selected based on initial experiments to balance overfitting risk and training efficiency.

Table 1 Summary of model architecture and training parameters

Component	Parameter	Value
Input Shape	(timesteps × features)	3 timesteps, 2 features
LSTM Layer 1	Units	64 (ReLU)
Dropout Layer 1	Rate	0.30
LSTM Layer 2	Units	32 (ReLU)
Dropout Layer 2	Rate	0.20
Output Layer	Activation	Softmax (4 classes)
Optimizer	-	Adam
Loss Function	-	Categorical Crossentropy
Epochs	-	50
Batch Size	-	32
Validation Split	-	20% from training
Class Weighting	-	Inverse frequency (scikit-learn)

4.2.3 Training and Validation

The dataset was split into training and testing sets with an 80/20 ratio. The model was trained over 50 epochs with a batch size of 32, using a validation split of 20% to monitor performance during training. GPU acceleration was utilized to improve training efficiency. Additionally, class weights were applied during training to further balance the learning process and ensure that minority classes were adequately represented. Class weights were computed using the compute_class_weight function from the scikit-learn library, which applies an inverse-frequency strategy. This approach ensures that minority classes receive higher importance during training, thus reducing model bias toward overrepresented classes.

4.3 Performance Evaluation

4.3.1 Evaluation Metrics

The model's performance was evaluated using several standard metrics:

- Accuracy: The proportion of correctly predicted samples across all classes.
- Precision: The proportion of true positive predictions among all positive predictions for each class.
- Recall: The proportion of true positives among all actual positives for each class.
- F1-Score: The harmonic mean of precision and recall, providing a balanced performance measure.

4.3.2 Confusion Matrix and Classification Report

Confusion matrices were used to visualize classification performance, showing the number of correct and incorrect predictions for each class. Detailed classification reports provided insights into precision, recall, and F1-scores for individual classes, highlighting strengths and weaknesses in the model's predictions.

4.3.3 Comparison of Pre and Post Resampling Performance

The model's performance was compared before and after applying the RandomOverSampler. The initial training on the imbalanced dataset resulted in high overall accuracy (95%) but poor recall and F1-scores for minority classes. After resampling, the model achieved significant improvements, with precision, recall, and F1-scores for minority classes exceeding 99%, as shown in the results section. These improvements demonstrate the importance of addressing class imbalance for robust anomaly detection.

5. Results and Discussion

5.1 Performance Before Resampling

The confusion matrix before resampling Fig. 2 demonstrates that the model achieves high accuracy for the majority class (Normal), but fails to correctly classify minority classes (Abnormal Voltage and Abnormal Temperature). The significant misclassification of these rare events indicates the model's bias toward the dominant class, which is a common issue in imbalanced datasets.

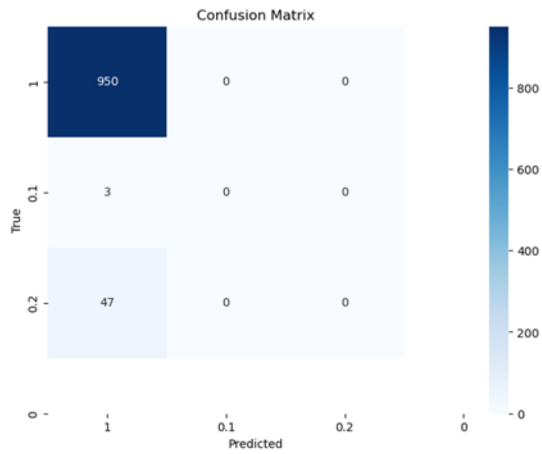

Fig. 2 The confusion matrix before resampling.

Such bias occurs because the model learns patterns predominantly from the majority class while failing to capture the characteristics of underrepresented categories. This results in low recall for minority classes, leading to a high rate of false negatives, where anomalies remain undetected. For real-world 4G/5G network monitoring, this is a serious issue because an undetected RRU fault could lead to signal degradation, service interruptions, or even complete network failure.

Table 2 The classification report before resampling.

	Precision	Recall	F1-Score	Support
0	0.95	1.00	0.97	950
1	1.00	0.00	0.00	3
2	1.00	0.00	0.00	47
Accuracy			0.95	1000
Macro avg	0.98	0.33	0.32	1000
Weighted avg	0.95	0.95	0.93	1000

The classification report before resampling Table 2 highlights the imbalance problem. While the model achieves an overall accuracy of 95%, this metric alone is misleading because the recall for minority classes is near zero. This means that although the model correctly classifies most Normal cases, it fails to detect critical anomalies.

A recall value close to zero suggests that the model prioritizes accuracy for the majority class at the expense of identifying minority-class instances. In scenarios where anomaly detection is crucial such as predictive maintenance for RRU systems low recall for rare failures makes the model ineffective. Consequently, accuracy alone is not a sufficient measure of performance for imbalanced datasets. Instead, metrics such as recall and F1-score provide a more reliable evaluation.

Without addressing this imbalance, the model is not viable for deployment in real-world telecommunication networks, as it cannot reliably detect RRU failures that require immediate maintenance.

5.2 Impact of RandomOverSampler

The confusion matrix after applying the RandomOverSampler Fig. 3 shows a significant improvement in classification performance. Unlike Fig. 2, where minority classes were largely misclassified, the resampled dataset allows the model to learn a more balanced distribution, ensuring that all classes receive proper attention.

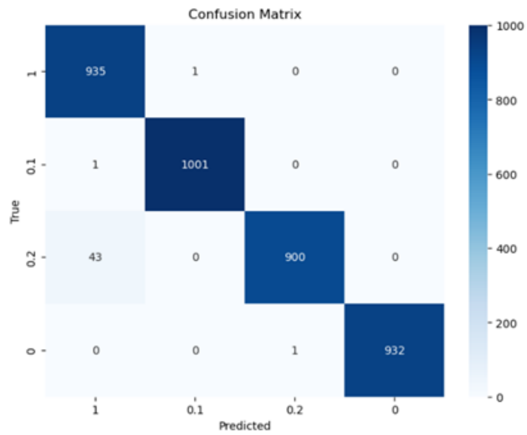

Fig. 3 The confusion matrix with oversampling.

Oversampling artificially increases the number of minority-class samples by duplicating existing instances, thereby preventing the model from being biased toward the majority class. This technique does not alter the fundamental characteristics of the dataset but ensures that the model receives sufficient exposure to all classes during training. As a result, it improves the model's ability to generalize across different class distributions, reducing the false-negative rate for anomaly detection.

Table 3 The classification report with oversampling.

	Precision	Recall	F1-Score	Support
0	0.96	1.00	0.97	936
1	1.00	1.00	1.00	1002
2	1.00	0.95	0.98	943
3	1.00	1.00	1.00	933
Accuracy			0.99	3814
Macro avg	0.99	0.99	0.99	3814
Weighted avg	0.99	0.99	0.99	3814

The classification report after oversampling Table 3 further demonstrates the effectiveness of addressing class imbalance. Precision, recall, and F1-scores for all classes have increased significantly, with recall values improving from near-zero to over 95%. This enhancement confirms that resampling plays a critical role in improving the reliability of anomaly detection in RRU systems.

By balancing the dataset, the model now correctly recognizes previously underrepresented classes, which is crucial for proactive network maintenance. With a higher recall rate, the model significantly reduces the likelihood of missing actual RRU failures, ensuring timely intervention and minimizing service disruptions. Another important observation is that precision remains high, meaning that the model does not overcompensate by misclassifying normal instances as anomalies. This balance between precision and recall is what makes the resampling technique effective.

5.3 Comparative Analysis

The comparative bar chart Fig. 4 provides a clear visualization of the improvements in model performance after applying the RandomOverSampler. The most notable improvements are in the recall and F1-score for minority classes, which increased from 33% and 32% to over 99%. Meanwhile, the overall accuracy improved from 95% to 99%.

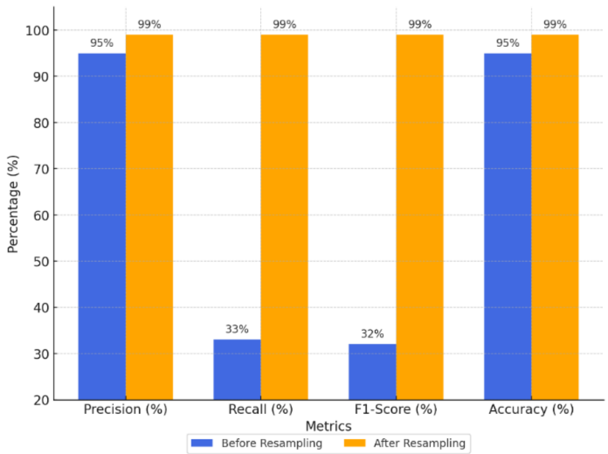

Fig. 4 Performance metrics before and after resampling.

These results confirm that resampling is crucial for achieving a well-balanced and effective anomaly detection system. However, while RandomOverSampler successfully addresses the class imbalance, it may introduce redundant samples in the dataset, potentially increasing training time.

5.4 Comparative Analysis

Improved Minority-Class Detection: By addressing class imbalance through RandomOverSampler, the model demonstrated significant improvements in detecting rare but critical anomalies, such as abnormal voltage and temperature conditions.

Maintained Generalization: The model preserved its generalization ability, achieving high accuracy and avoiding overfitting, even with the resampled dataset. Applicability to Real-World Scenarios: The proposed framework is highly applicable to real-world settings where imbalanced data distributions are common, such as monitoring RRU systems in 4G/5G networks. The ability to reliably detect anomalies ensures improved network stability and reliability.

6. Conclusions

This study presents an LSTM-based framework for detecting anomalies in Remote Radio Units (RRUs) within 4G/5G networks, under imbalanced data conditions. The integration of RandomOverSampler significantly improved the model's ability to detect minority-class events, with recall and F1-score increasing from under 33% to over 99%. These results highlight the necessity of addressing data imbalance in real-world telecom applications.

However, this study has several limitations. In accordance with the Telco vendor's specifications, the model was trained using only two features: voltage and temperature. Although these parameters are relevant, they may not comprehensively capture all factors affecting RRU performance. Additionally, the use of a short 3-step sequence, though computationally efficient, may limit the model's ability to capture longer-term temporal patterns. Another key limitation is that the evaluation was conducted only on static, historical datasets; no real-time system testing or adaptation to concept drift was performed. There is also a potential risk of overfitting due to the small feature space and the use of oversampling.

Despite these limitations, the results demonstrate that the proposed framework is a strong candidate for anomaly detection in network monitoring scenarios. Nonetheless, further validation is essential before deployment, particularly under live network conditions where data dynamics can change unpredictably. Addressing these issues will be vital to ensure robust performance and generalizability.

7. Future Work

To enhance the applicability of this research in operational environments, several extensions are planned. First, we will expand the input feature set by incorporating additional RRU performance me trics, such as RSRP, SINR, and RSSI, to improve the model's predictive power. Second, we will experiment with longer sequence lengths to better capture temporal dependencies that may influence fault detection accuracy. Third, to assess generalization, the model will be tested on external datasets from different network environments. If consistent performance is observed, the framework will be integrated into a real-time monitoring prototype. This deployment will also explore challenges such as handling streaming input, retraining models periodically, and adapting to concept drift. These steps are essential to address the variability and dynamics in live 4G/5G networks.

Lastly, we will apply statistical significance tests, including confidence intervals and McNemar's test, to strengthen the reliability of the reported metrics and support comparative evaluations with baseline models such as SVM or Random Forest. These improvements aim to enhance the framework's robustness, scalability, and practical deployment potential.

8. Acknowledgement

This work was supported by Suranaree University of Technology and an Telco industry partner.

9. References

[1] Y. F. Li, C. Jia, J. Ye, and B. Xu, "On the reliability of 4G/5G mobile telecommunication networks from the perspective of operation & maintenance," in *2021 Annual Reliability and Maintainability Symposium (RAMS)*, Orlando, United States, May 24-27, 2021, pp. 1-7.

[2] J. M. de Oliveira Pereira, *"A framework for 5G network data analytics function with emphasis on anomaly detection,"* M.S. thesis, Graduate Program in Computer Science, Federal University of Minas Gerais, Belo Horizonte, Brazil, 2024.

[3] Y. Wang, S. Meng, Y. Song, and D. Liu, "Fault detection for large scale indoor distributed antenna system based on time series similarity," in *2022 Prognostics and Health Management Conference (PHM-2022 London)*, London, United Kingdom, May 27-29, 2022, pp. 269-275.

[4] S. Wang, J. F. Balarezo, S. Kandeepan, A. Al-Hourani, K. G. Chavez, and B. Rubinstein, "Machine learning in network anomaly detection: A survey," *IEEE Access*, vol. 9, pp. 152379-152396, November. 2021.

[5] K. Ghosh, C. Bellinger, R. Corizzo, P. Branco, B. Krawczyk, and N. Japkowicz, "The class imbalance problem in deep learning," *Machine Learning*, vol. 113, no. 7, pp. 4845-4901, December. 2022.

[6] P. Vuttipittayamongkol, E. Elyan, and A. Petrovski, "On the class overlap problem in imbalanced data classification," *Knowledge-Based Systems*, vol. 212, January. 2021.

[7] T. Wongvorachan, S. He, and O. Bulut, "A comparison of undersampling, oversampling, and SMOTE methods for dealing with imbalanced classification in educational data mining," *Information*, vol. 14, no. 1, January. 2023.

[8] K. R. Balmuri, S. Konda, W.-C. Lai, P. B. Divakarachari, K. M. V. Gowda, and H. K. Lingappa, "A long short-term memory network-based radio resource management for 5G network," *Future Internet*, vol. 14, no. 6, June. 2022.

[9] Y. Wei, J. Jang-Jaccard, W. Xu, F. Sabrina, S. Camtepe, and M. Boulic, "LSTM-autoencoder-based anomaly detection for indoor air quality time-series data," *IEEE Sensors Journal*, vol. 23, no. 4, pp. 3787-3800, February. 2023.

[10] S. Aktaş, H. Alemdar, and S. Ergüt, "Towards 5G and beyond radio link diagnosis: Radio link failure prediction by using historical weather, link parameters," *Computers and Electrical Engineering*, vol. 99, April. 2022.

[11] S. K. Agarwal, S. Banerjee, and R. Mahapatra, "Prediction and recovery of radio link failure caused by environmental factors," in *2022 IEEE 19th India Council International Conference (INDICON)*, Kochi, India, November 24-26, 2022, pp. 1-6.

[12] W. Zhang, Y. Li, M. Wen, and R. He, "Comparative study of ensemble learning methods in just-in-time software defect prediction," in *2023 IEEE 23rd International Conference on Software Quality, Reliability, and Security Companion (QRS-C)*, Chiang Mai, Thailand, October 22-26, 2023, pp. 83-92.

[13] K. Murphy, A. Lavignotte, and C. Lepers, "Fault prediction for heterogeneous telecommunication networks using machine learning: a survey," *IEEE Transactions on Network and Service Management*, vol. 21, no. 2, pp. 2515-2538, April. 2024.

[14] A. Pourmahboubi and H. Tabrizchi, "LSTM-based framework for 5G resource allocation prediction," in *2024 11th International Symposium on Telecommunications (IST)*, Tehran, Iran, October 9-10, 2024, pp. 295-302.

[15] S. Barmpounakis, L. Magoula, N. Koursioumpas, R. Khalili, J. M. Perdomo, and R. P. Manjunath, "LSTM-based QoS prediction for 5G-enabled connected and automated mobility applications," in *2021 IEEE 4th 5G World Forum (5GWF)*, Montreal, Canada, October 13-15, 2021, pp. 436-440.

[16] J. Lin, et al., "Multi-scenario cellular KPI prediction based on spatiotemporal graph neural network," *IEEE Transactions on Automation Science and Engineering*, vol. 22, pp. 5131-5142, June. 2024.

[17] F. Nhita and I. Kurniawan, "Performance and statistical evaluation of three sampling approaches in handling binary imbalanced data sets," in *2023 International Conference on Data Science and Its Applications (ICoDSA)*, Bandung, Indonesia, August 9-10, 2023, pp. 420-425.

[18] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321-357, June. 2002.

[19] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: adaptive synthetic sampling approach for imbalanced learning," in *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, Hong Kong, China, June 1-8, 2008, pp. 1322-1328.

[20] F. Kamalov, H. H. Leung, and A. K. Cherukuri, "Keep it simple: random oversampling for imbalanced data," in *2023 Advances in Science and Engineering Technology International Conferences (ASET)*, Dubai, United Arab Emirates, February 20-23, 2023, pp. 1-4.

[21] G. E. A. P. A. Batista, R. C. Prati, and M. C. Monard, "A study of the behavior of several methods for balancing machine learning training data," *ACM SIGKDD Explorations Newsletter*, vol. 6, no. 1, pp. 20-29, June. 2004.

[22] A. Massaro, D. Kostadinov, A. Silva, A. O. Guzman, and A. Aghasaryan, “Predicting network hardware faults through layered treatment of alarms logs,” *Entropy*, vol. 25, no. 6, June. 2023.

[23] L. Sana, et al., “Securing the IoT cyber environment: Enhancing intrusion anomaly detection with vision transformers,” *IEEE Access*, vol. 12, pp. 82443-82468, May. 2024.