Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Balanced Attention-BiGRU Based Fault Detection and Classification in Electric Power System

Faiz Ahmad[1], Dharmendra Kumar Dheer[1], Ajit Kumar[1]
[1] *Electrical Engineering Department, National Institute of Technology, Patna, India*
E-mail: faiza.ph21.ee@nitp.ac.in

ARTICLE INFO

Received: 17 Oct 2024

Revised: 25 Nov 2024

Accepted: 20 Dec 2024

ABSTRACT

The modern era's expanding demand for electrical power, along with limited production and transmission capacity, has underlined the vital need for strong and rapid fault detection mechanisms inside electrical power networks. These intricate and dynamic systems, which rely primarily on transmission lines, are prone to disruptions and breakdowns induced by unanticipated factors. The work proposes a novel approach to addressing long-standing concerns in Artificial Intelligence (AI)-driven research, such as data imbalance, model complexity, computing cost, and over fitting issues. To solve these issues and improve the efficacy of AI-driven research, this study proposes an attention-based Bidirectional Gated Recurrent Unit (Bi-GRU) model designed specifically aimed for balanced datasets. The suggested approach initiates by preparing the dataset through normalization and one-hot encoding. Subsequently, the pre-processed dataset undergoes balancing using Synthetic Minority Oversampling Technique (SMOTE) before being inputted into the Attention-based Bi-GRU model. This model is designed to accommodate a fixed-length sequence of voltage and current values at its input layer. Finally, the batch normalization and dense layer with SoftMax were utilized to classify the types of defects in transmission lines. As a result, the suggested method classifies fault types with a high accuracy of 85.40% when compared to existing models.

Keywords: Fault Detection; AI; Transmission lines; Bi-GRU; Data Balance; SMOTE

1 INTRODUCTION

The size and complexity of the electrical power grids are expanding quickly in the modern era [1]. All sectors are involved in the growth of the energy industry, including distribution, transmission, and power generating [2]. The emergence of unexpected electrical problems in transmission or distribution lines is a common challenge faced by electrical power networks [3]. A basic definition of an electrical fault is the existence of higher current and voltage levels than those expected under normal operating environments, or an unusual deviation in current and voltage measurements [4]. The Electric Power Systems becomes unstable when a disturbance, like a short-circuit failure, happens because the line impedance varies. If fast corrective action is not performed in response to a disruption, there is a chance that failures could cascade and cause power outages throughout a significant area. Therefore, to preserve the power system and cause the least amount of harm possible, rapid fault type and location detection is essential [5]. Transmission lines may contain a variety of flaws, which can be classified as symmetrical or asymmetrical. Phase faults, including phase-to-ground, phase-to-phase, phase-to-ground, and three-phase faults, are several instances of the categories of flaws that transmission lines can experience. But the operation of the electricity system is unaffected by their presence. In addition, flaws such circuit faults, overlapping faults, and other problems are seen to be modest in comparison to the previously listed errors. These lines are typically preserved by using a megger device, which makes fault identification easier. Alternately, the lines [6] are physically examined to find any defects. Due to the length of time required by these two methods, nations all over the world are looking into innovative techniques to identify and fix defects quickly. The majority of the electrical power system defect diagnosis techniques rely on challenging mathematical calculations. To mitigate these constraints, recent research has attempted to use deep learning algorithms to detect and classify power quality interruptions [7]-[9]. Due to complex computations, one drawback of previous approaches is their lengthy calculating times, which could exacerbate harm by delaying the necessary decisions to safeguard the Electric Power System. On the other hand, a deep learning model is taught to increase the likelihood that it can quickly produce accurate findings from large amounts of data, and a well-trained model can produce outputs right away [10]. As a result, DL techniques have been frequently used to address problems that call for quick diagnosis, like power system breakdown [11]. Electric power transmission line defects have been identified using a variety of methods and procedures. Godse et al. presented a novel fault feature extraction method for real-time defect detection and classification in transmission lines that is based on mathematical morphology [12]. It is used to extract distinct fault features, which are then fed into a decision tree classifier to ascertain the type of the defect. However, it is sensitive to data noise, which reduces the precision of feature extraction and may result in false detections or incorrect classifications. For fault region identification (FRI), fault type classification (FTC), and fault location prediction (FLP), in the realm of Deep Learning (DL), Belagoune et al [13] developed three unique models for classification and regression based on Deep Recurrent Neural Networks (DRNN). The Imbalanced datasets, where the proportion of fault instances is considerably smaller than that of normal instances, can, however, skew LSTM models in favour of the majority class. In order to diagnose electrical power system faults, Ajagekar et al [14] presented a quality control-based hybrid deep learning framework that combines the deep network's adept categorization with the conditional restricted Boltzmann machine's feature extraction capabilities. However, qubit counts, error rates, and coherence periods are currently constrained in quantum computers. A novel diagnostic method for locating Power Quality Disturbances in transmission lines was presented by Yoon et al. [15]. This method involves representing real three-phase voltage and current waveforms using a convolutional neural network (CNN). An anomaly-based fault detection system was presented by Elmasry et al. in their paper [16]. In order to overcome the shortcomings and limitations of the present electrical fault detection systems, this system consists of four primary stages: data preprocessing, pre-training, training, and testing. The expansion of the feature space meant that the models in use were less and less effective as the amount of data pieces rose. In order to diagnose faults in transmission lines, Bon et al. [17] developed a combination method utilising the DWT, GoogleNet, and probabilistic neural network methods. The DWT method was used for detection, the GoogleNet method was used for classification, and the PNN method was suggested for fault location. However, combining several approaches can increase the complexity of data pre-processing, feature extraction, and model fusion. Elmasry et al. [18] proposed a novel Ensemble Deep Learning Approach for Electrical Fault Detection Systems (EDLA-EFDS), which uses a double Particle Swarm Optimisation (PSO) metaheuristic to select the best features and hyper parameters. This approach addresses the shortcomings of existing systems, such as overfitting, automation, and validation problems. Then, using three distinct deep learning techniques—Deep Neural Networks (DNN), Long Short-Term Memory Recurrent Neural Networks (LSTM-RNN), and Deep Belief Networks (DBN)—a bagging ensemble strategy is implemented. The ensemble system is used to generate the final choice, which is then followed by a majority voting engine. Bagging, on the other hand, necessitates the training of numerous models on various bootstrap samples, which can be computationally expensive. Using convolutional neural networks, Shakiba et al [19] established a reliable detection and identification system. For data feature extraction in this study, time-based voltage and current data will be used rather than FFT, WT, and ST. These techniques, meanwhile, are vulnerable to signal noise. Abed et al [20] concentrated on using the k-nearest neighbour technique to find faults in transmission lines. Due to its sensitivity to imbalanced datasets and the need to calculate the distance between each prediction's test instance and all training examples, this approach can be computationally expensive. To protect the safety and dependability of the power system, Goni et al. [21] presented innovative ELM for autonomous fault detection and classification system. In ELM, the connections between the input layer and the hidden layer are established randomly rather than intentionally. This randomness results in a loss of control over the concealed layer's organization and readability.

The preceding study makes it evident that issues with data imbalance, model complexity, computational expenditure, and the onset of over fitting are not effectively addressed by the current body of research. This research offers novel deep learning approaches to overcome these limitations. Despite the fact that they revealed encouraging findings, it is still difficult to accurately and effectively diagnose defects in the electric power system using an AI model. The main contribution of the proposed method is described below,

1. The input given to the model is first normalized to ensure consistent voltage and current levels before being separated into fixed-length sequences. To handle unbalanced features, fault labels are encoded using one-hot

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

encoding.

2. Second, in this study, SMOTE, which works by creating synthetic samples of the minority class in imbalanced dataset by interpolating between existing minority class instances and their nearest neighbours. Following dataset balancing, an Attention-based Bi-GRU model is implemented.

3. Finally, the Bi-GRU layer bidirectional analyses input sequences, capturing temporal dependencies. Attention mechanisms compute the relevance of hidden states, normalize their outputs, and aggregate them in a dense layer for final categorization. For multi-class fault classification, SoftMax creates fault class probability.

The structure of the research article is as follows: Section 2 offers an in-depth explanation of the suggested methodology. The performance parameters and the simulated outcomes of the suggested method are explained in Section 3. The suggested approach is summed up and concluded in the last section.

2 BALANCED ATTENTION-BIGRU

The demand for energy is increasing, but production and transmission capacity are not increasing at the same rate. Electrical power systems are complex and dynamic, and transmission lines are crucial for maintaining a reliable power supply. Researchers are working to identify failures to prevent financial losses. A dataset containing voltage and current values is pre-processed through normalization, division into fixed-length sequences, and encoding fault labels using one-hot encoding. The dataset is balanced using SMOTE. Subsequently, the balanced dataset is fed into the Attention-based Bi-GRU model. This model processes the input sequence bidirectionally to capture temporal dependencies. The hidden states are submitted to the attention mechanism, and the outcome is batch normalized for normalization. The dense layer combines the attention hidden states and performs the final classification. SoftMax is used for multi-class fault classification to generate fault class probabilities. The flow diagram for the proposed work has been shown in Fig 1.

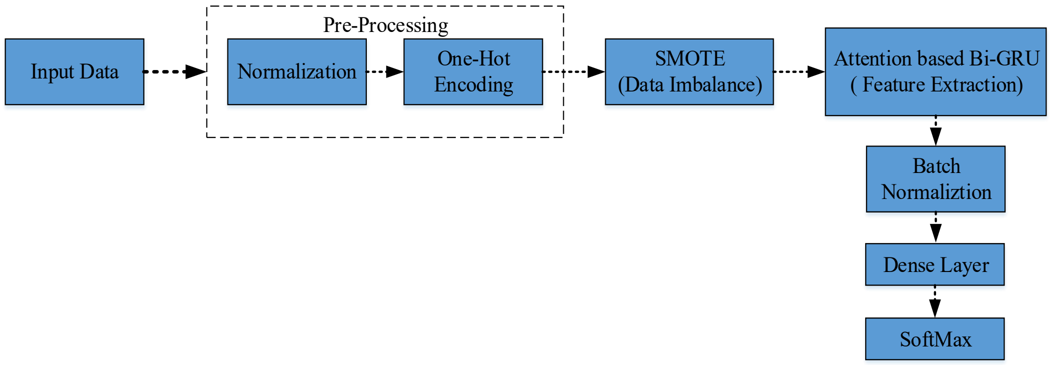


Fig. 1 Proposed work flow

2.1 PRE-PROCESSING STAGE

Data pre-processing is a crucial stage in any machine learning methods for classification. In the case of this study, the pre-processing stage is carried out to make sure the dataset is free of noisy data, such as data duplication to avoid redundancy, null data cleaning, etc. Data normalization is the next phase in the process once the initial steps of data interpretation and cleaning are finished. This stage entails scaling your dataset's numerical features to a fixed range, usually between 0 and 1 or -1 and 1. To prevent features with differing scales from predominating the training of machine learning models, normalization is crucial. Machine learning algorithms may readily analyze categorical data by converting it into a numeric format using a technique called one-hot encoding. When working with datasets that contain both numerical and categorical variables, this technique is often used after data normalization. For each category in the categorical feature, a binary "dummy" variable must be created. Each binary variable indicates whether a certain category is present or absent.

2.2 DATA BALANCING

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

A common issue in real-world deep learning applications is that certain classes have much more instances in the training set than others. When using such datasets, the learning algorithms are frequently biased towards the majority classes, resulting in a greater miss-classification rate for minority class cases. To solve this issue, different approaches can be used, including under sampling, oversampling, two-phase training, and cost-sensitive learning. This paper emphasizes the utilization of SMOTE, a technique employed to balance uneven datasets by generating synthetic instances of minority class.

SMOTE is an oversampling technique used to balance the class distribution of a dataset. The basic idea behind SMOTE is to create new minority class samples by taking small steps from one of the minority class samples to one of its K nearest neighbours in the feature space, where K is a parameter of the algorithm. The algorithm creates a new sample by randomly selecting one of the K nearest neighbours and then adding a small perturbation to the feature vector between the original sample and the selected neighbour. This creates new synthetic data that is similar to the minority class samples in the feature space but is not an exact copy of any of the existing samples. Figure 2 describes the SMOTE architecture.

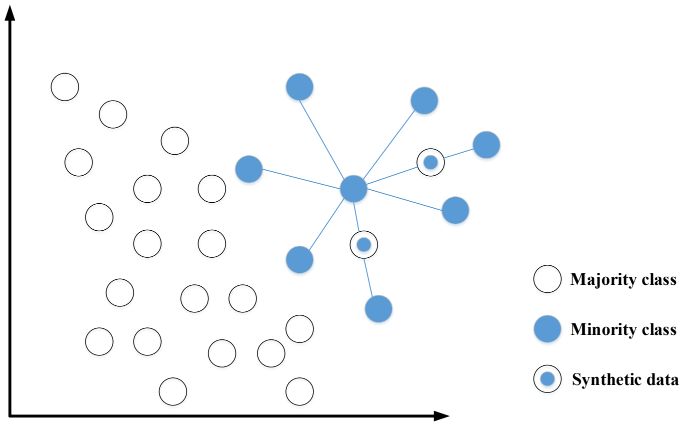


Figure 2 SMOTE Architecture.

The detailed process of data balancing using SMOTE is described in the algorithm as presented below.

Algorithm 1: Data balancing Using SMOTE

Step 1: Data Pre-processing and Class Separation

1. Open the dataset for Electrical Fault Detection and Classification.

2. Pre-process the dataset, including normalization and one-hot encoding.

3. Identify the minority class and majority class samples.

Step 2: Initialize SMOTE Components

4. Select a minority class samples from the original dataset.

5. Find its K nearest minority class neighbours in the feature space.

6. Randomly select one of the K nearest neighbours.

7. Generate a new synthetic sample by interpolating between the selected minority class sample and randomly selected neighbour

8. Repeat Steps 4-7 until the desired number of synthetic samples is generated.

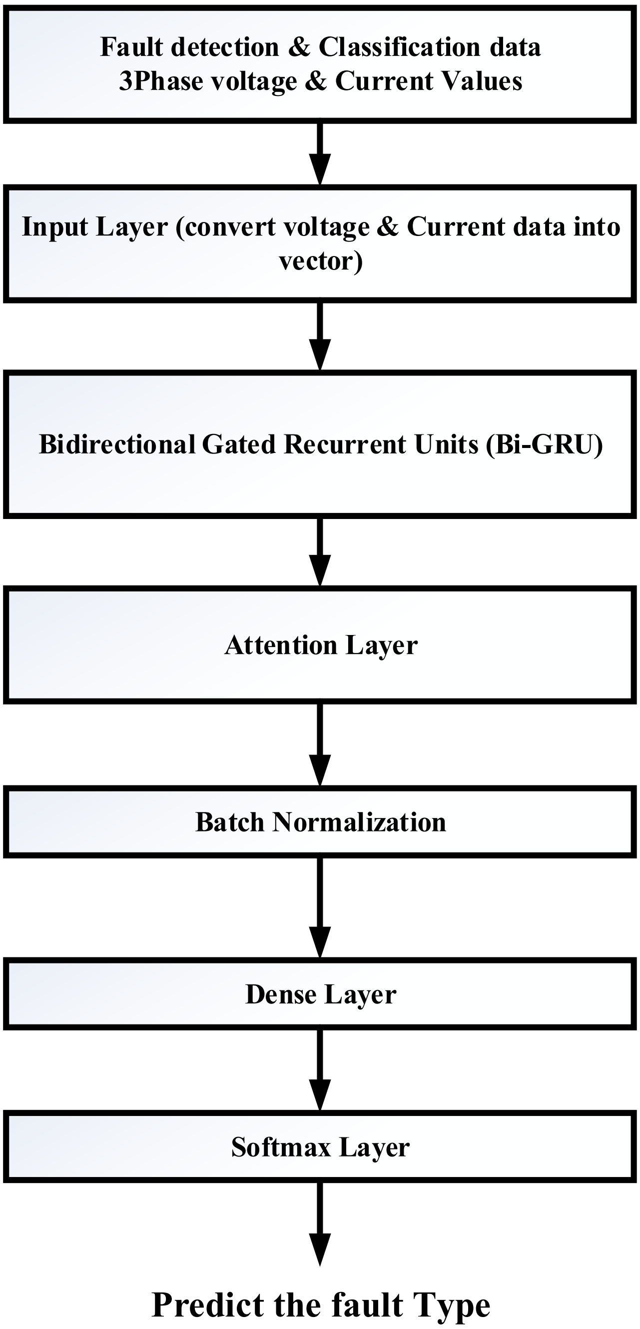
Step 3: Balancing the Dataset

9. Combine the synthetic sequences with the original dataset (including majority class samples) to create a balanced dataset.

This pre-processed balanced dataset is given into the Attention based Bi-GRU model which accepts the fixed length sequence of voltage and current values in the input layer.

2.3 ATTENTION BASED BI-GRU FEATURE EXTRACTOR:

The Electrical Fault detection and Classification dataset, which contains 3-phase voltage and current measurements, can be used to extract features using attention-based Bidirectional Gated Recurrent Units (Bi-GRU). Three main components make up the proposed model: the Input layer, the Bidirectional GRU layer with Attention, and the Output layer. The input layer used to convert the voltage and current data into continuous vector representations. The work flow for Bi-GRU has been shown in Fig.3.



Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Figure 3: Work Flow

Each sample as either normal or flawed is to be identified, and values for 3-phase voltage and current are contained in the dataset. There are a total of six input features per sample, with each sample having three voltage (V_1,V_2,V_3) and three current values(I_1,I_2,I_3) without explicitly considering time steps to detect and classify faults in an electrical power system. The basics Bi-GRU model has been represented in Fig.4.

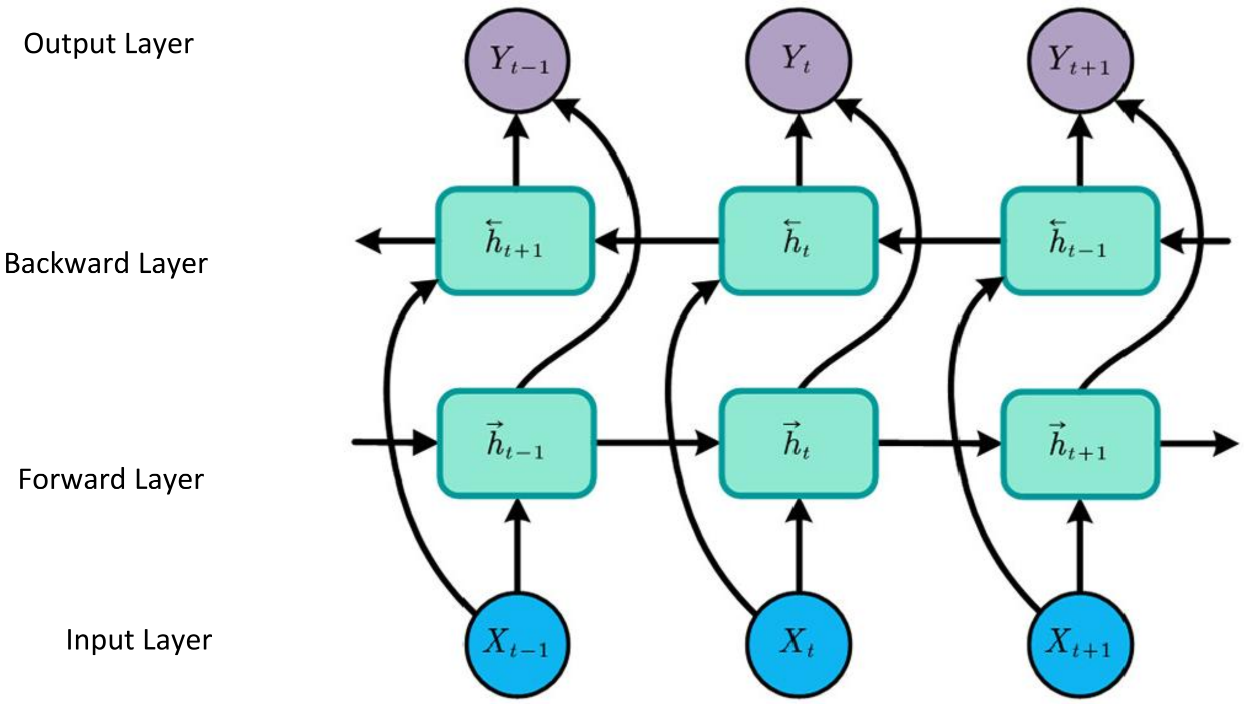


Figure 4 Bi-GRU Model

RNN has been effectively used in a variety of fields. Because they have hidden layers that are self-connected. To process time-related sequences, RNNs can make use of their internal states. The GRU is a type of RNN that can get around the problems with regular RNNs such gradient explosions and disappearing gradients. One important difference between a GRU and a traditional RNN is the former's capacity to capture both immediate and extended relationships. A forward and a backward GRU are both included in a bidirectional GRU, which allows it to process data in both directions. There are input and hidden layers in every GRU. To keep track of the state of the pattern, it requires a gating mechanism. The each GRU has two different kinds of gates: update gates and reset gates. In order to strike a balance between adding new information and keeping outdated information updated, the update gate is essential. In the meantime, the reset gate controls how much data from earlier iterations is added to the current candidate state. The proposed research aims to concentrate on pertinent features inside each sample without taking time steps into account.

Let's denote the input features for a single sample as $Y = [V_1, V_2, V_3, I_1, I_2, I_3]$. $H_i^{forward}$ and $H_i^{backward}$ denotes the hidden states of the forward and backward Bi-GRU units for the feature i, respectively. These units capture the dependencies and relationships with in the input features. The information from both the forward and backward hidden states is combined to create the context vector for feature i. It is determined as the weighted sum of the following states:

$$Cv_i = \sum_{j=1}^{6} \propto_{ij} * \{H_j^{forward}; H_j^{backward}\} \quad (1)$$

The 'i' will be ranges from 1 to 6 to cover all input features. $\propto_{ij}$ is the attention weights which are computed using softmax function. By ensuring that the weights add up to 1, the softmax function enables a weighted sum that

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

accurately reflects the relative value of each feature. To decide how much attention each feature should receive, the attention weights $\propto_{ij}$ are crucial. During the feature extraction process, these weights define the relative relevance of each feature and it is denoted as,

$$\propto_{ij}= \frac{\exp\,(e_{ij})}{\sum_{k=1}^{6} \exp\,(e_{ik})} \tag{2}$$

Where, e_{ij} is the attention score for feature i. This attention mechanism essentially assigns weights to each input feature based on how relevant they are to one another. It assists the model in concentrating on the input data's most instructive elements for later feature extraction. The attention scores used to measure the relative relevance of each feature throughout the feature extraction process are calculated using weight matrix and it can be calculated as follows:

$$e_{ij} = v_a.\tanh\,(W_a{}^*[\propto_{ij}*\,\{H_j^{forward}; H_j^{backward}]) \tag{3}$$

Where, v_a is the learnable parameter vector, W_a is the weight matrix used for the attention mechanism. The next critical step is featuring extraction, which comes after getting the context vector Cv_i for each of the six input characteristics. These context vectors are transformed during feature extraction into a feature representation that is appropriate for the next task, which is defect detection and classification. The context vector $Cv_i$is the input for the feature extraction process, which manipulates it to produce a feature representation. These context vectors are passed through the dense layer which consists of learnable parameters that allow the Bi-GRU model to learn complex relationships and representations from the input and it is represented by

$$F(Cv_i) = dense\ (Cv_i) \tag{4}$$

After features have been extracted, feature representations can be used to classify faults in a final task that involves fault detection. Before the classification process, this work uses a batch normalization to improve training stability and accelerate convergence. It could make your model more proficient at identifying and classifying errors. To classify different power system failures, this study uses a dense layer with a SoftMax activation function during the classification stage. The dense layer must be trained to combine the retrieved information and provide predictions.

The feature representations derived from feature extraction will be denoted as $F(Cv_1), F(Cv_2), F(Cv_3), F(Cv_4), F(Cv_5), F(Cv_6)$, where each $F(Cv_i)$ represents the feature representation for one of the six input features. These feature representations can be joined to create a single feature vector:

$$Feature\ Vector = [F(Cv_1), F(Cv_2), F(Cv_3), F(Cv_4), F(Cv_5), F(Cv_6)] \tag{5}$$

After that, a dense layer receives this feature vector:

$$Dense\ layer\ ouput = dense(Feature\ Vector) \tag{6}$$

By applying an activation function, such as ReLU, to each input feature, the dense layer creates nonlinearity and performs a calculation using the features. Finally, a SoftMax activation is applied to the dense layer's output, producing probabilities for various classes:

Class Probabilities = SoftMax (Output from Dense Layer) (7)

As the predicted fault class, the class with the highest likelihood is normally selected. In order to minimise an appropriate loss function (categorical cross-entropy), the Bi-GRU model learns the weights and biases of the dense layer during training using back propagation and optimisation algorithms (ADAM) for 500 epochs. By using the features that were retrieved, the network can produce precise predictions for fault detection and classification.

3 RESULT AND DISCUSSION

This section presents our simulation experiment's findings and provides a detailed examination of its performance characteristics. The proposed model has been implemented with a batch size of 32 and a learning rate of 0.001 on the Python platform.

3.1 TOOLS AND DATASET

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

The study makes use of the pre-processed Electrical Fault Detection and Classification dataset from Kaggle (https://www.kaggle.com/datasets/esathyaprakash/electrical-fault-detection-and-classification) for analysis and modelling. The system model is composed of four generators, each with a capacity of 11 kV. These generators are supported by two step-down transformers operating at 11 kV/400 V. The transmission line spans a length of 150 km. additionally, the system includes an RLC load with a phase voltage of 0.4 kV, a frequency of 50 Hz, and an active power output of 10 kW. To guarantee uniform scaling and seamless inclusion of categorical variables, normalization and one-hot encoding are used. Machine learning models may become biased in favour of the majority due to the dataset's intrinsic class imbalance. The SMOTE technique employs a generative adversarial network to generate synthetic data instances for underrepresented classes, aiming to enhance the accuracy and reliability of identifying and classifying electrical faults. The proposed model is implemented using the Python platform.

3.2 TRAINING AND VALIDATION

The training and validation graph shows the learning process of the machine learning model across 500 epochs. The large early losses point to the poor initial weights and parameters of the model. But both losses dropped as training progressed, suggesting better predictions. Around epoch 100, the training loss diverged, showing that the model was fitting the data effectively. When both losses stabilized at 0.2 at epoch 500, the model's convergence was achieved.

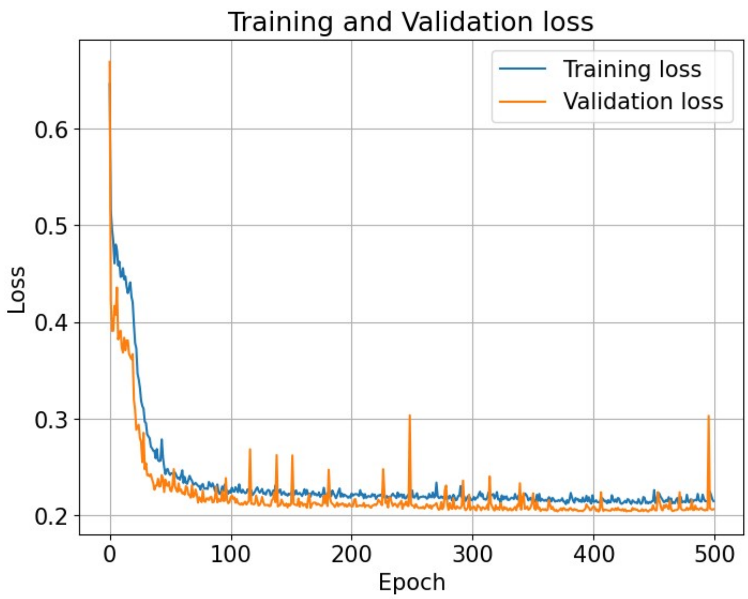


Figure 5 Training and Validation loss

As shown in Figure 5, the validation and training loss curves are remarkably similar. The training data (blue line) aids the model in learning, while the unobserved validation data (orange line) aids in generalization. This close alignment indicates that the model is learning in a fashion that works well for both datasets rather than over fitting. This similarity in loss curves is an indicator of the model's performance and generalization capacity since it finds a balance between learning from the training data and avoiding excessive complexity, which is often the desired outcome when training machine learning models.

3.3 PERFORMANCE METRICS

Electric power systems use Balanced Attention-BiGRU models for fault detection and classification, which improves performance metrics. More complex analysis is made possible by these models, which identify both local and global relationships in power system data. The model's accuracy, recall, and F1-score are performance indicators that assess how well the model can detect errors and categorize them. Recall assesses the completeness of the model, precision gauges the accuracy of positive predictions, and the F1-score offers a fair evaluation. AUC-ROC and accuracy gauge how well the model can distinguish between different fault classes. The accuracy and dependability of the model are

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

improved by ongoing assessment and optimization based on these indicators, which raises the robustness and effectiveness of electric power systems.

3.3.1 ACCURACY

The degree of correctness reached by a computing system or algorithm in image analysis, classification, or recognition tasks is referred to as accuracy in image processing. It computes the percentage of images in the whole collection that were successfully identified or processed. Accuracy is a key statistic for evaluating the success of image processing algorithms and models since it reveals how well they comprehend and interpret images. High accuracy ratings indicate correct object or pattern identification, whereas low scores indicate failures in recognition. The accuracy formula is depicted in equation 8.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

The ratio of correct forecasts to all predictions made is a more straightforward way to define accuracy. It is expressed as a percentage or a numeric value ranging from 0 to 1, with 1 representing perfect accuracy and 0 signifying the worst possible accuracy.

3.3.2 PRECISION

Precision is an important image processing evaluation statistic that measures the accuracy of positive predictions made by a classification or recognition method. It focuses on the proportion of genuine positive (TP) forecasts to total positive predictions (TP+FP). In other words, accuracy assesses how many of the expected positives are true positives. The precision formula is stated in equation 9.

$$\text{Precision} = \frac{TP}{TP+FP} \tag{9}$$

Where:

- True positives (TP) refer to instances where the model accurately predicted the presence of the positive class.
- False positives (FP) are situations where the true class was negative even though the model predicted a positive class.

3.3.3 RECALL

Recall is a key performance indicator in the assessment of image processing. It is also known as sensitivity or the true positive rate. It evaluates a classification or identification algorithm's ability to reliably identify each occurrence of a specific class within a dataset. The emphasis of this statistic is on the ratio of true positives (TP) or correct positive predictions to the total of false negatives (FN) or missed positive cases. The recall formula is stated in equation 10.

$$\text{Recall} = \frac{TP}{TP+FN} \tag{10}$$

Where:

- True positives (TP) are the cases where the model correctly predicted the positive class.
- False negatives (FN) are the cases where the model predicted the negative class but the actual class was positive.

The performance metrics for the proposed model has been shown in Fig.6.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

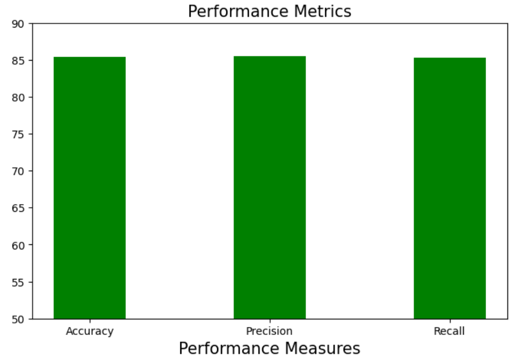


Figure 6 Performance Metric

3.4 COMPARATIVE ANALYSIS

Table 1 Comparison of different classifiers

	ACCURACY	PRECISION	RECALL
Logistic Regression	34.58	34.58	34.58
SVM	81.75	81.75	81.75
K-Nearest Neighbors	80.41	80.41	80.41
Gradient Boosting	84.04	84.04	84.04
Naïve Bayes	79.46	79.46	79.46
Ada Boost	77.36	77.36	77.36
XG Boost	81.30	81.30	81.30
Light GBM	81.11	81.11	81.11
Cat Boost	79.59	79.59	79.59
Proposed	85.40	85.48	85.31

Table 1 and Figure 7 present a comprehensive analysis of performance metrics across multiple models utilized within the domain of identifying and categorizing faults within electrical power systems. The evaluation of these models encompasses crucial performance indicators such as accuracy, precision, and recall. These metrics stand as pivotal benchmarks, profoundly influencing the efficacy of these models in effectively managing fault-related activities within power systems.

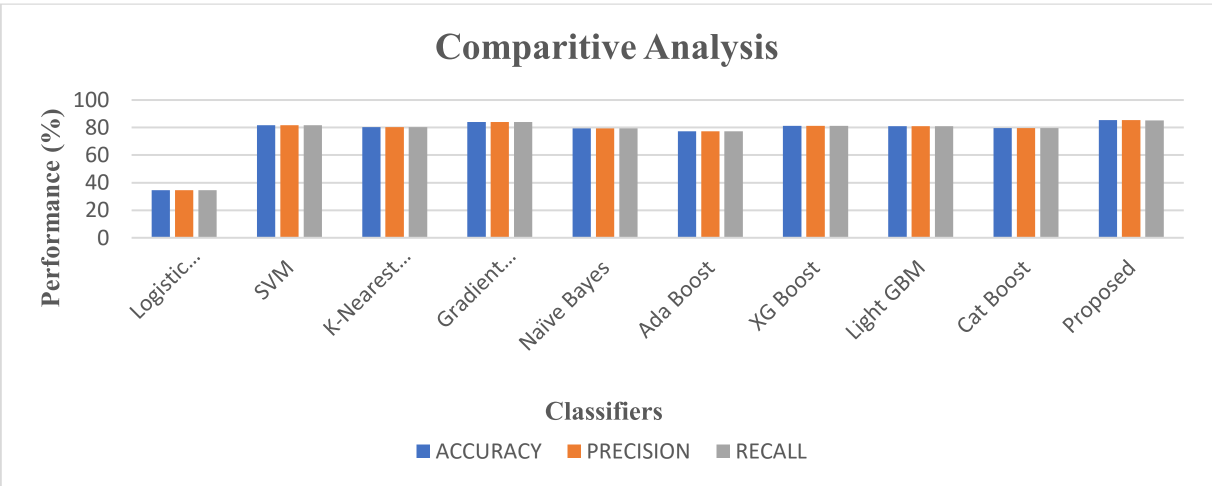


Figure 7 Comparative Analysis

Accurate fault detection and classification are integral facets of ensuring the reliability and safety of electrical power systems. The precision metric gauges the model's ability to correctly identify relevant instances among the total identified cases. Meanwhile, recall assesses the model's capacity to capture all relevant instances within the dataset. Both precision and recall contribute significantly to the overall accuracy of these fault-locating models, directly impacting their utility in practical fault management scenarios.

3.5 SUMMARY

In the evaluation of various machine learning models based on accuracy, precision, and recall, the performance metrics reveal distinctive characteristics. Logistic Regression exhibits the lowest values across all three metrics, with an accuracy, precision, and recall of 34.58%. Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and XG Boost share similar performance, demonstrating an accuracy, precision, and recall around 80-81%. Gradient Boosting performs slightly better, achieving 84.04% in all three metrics. Naïve Bayes, Ada Boost, Light GBM, and Cat Boost also present comparable results, with accuracy, precision, and recall scores ranging from 77.36% to 81.30%. Notably, the proposed model outperforms all others, boasting the highest accuracy at 85.40%, precision at 85.48%, and recall at 85.31%. These results suggest that the proposed model is the most effective in terms of overall predictive performance compared to the other models assessed in this study.

4. CONCLUSION

In conclusion, the limited production and transmission capacity in the current period, along with the rising demand for electricity, have highlighted the crucial relevance of effective defect detection methods in electrical power networks. These complex, dynamic systems rely largely on transmission lines, which leaves them open to unforeseen disruptions and failures. Even though many research work have used AI approaches, they have run into substantial problems with data imbalance, model complexity, cost of computing, and overfitting. This research has developed a novel strategy built around an attention-based Bi-GRU model with a balanced dataset to address these serious problems. The suggested approach begins with systematically preparing the dataset through procedures like one-hot encoding and normalization. To address data scarcity, SMOTE is applied, generating synthetic data for achieving better balance. This balanced dataset is fed into the Attention-based Bi-GRU model, specialized in handling fixed-length sequences of voltage and current values in the input layer, effectively processing them. Additionally, batch normalization and a dense layer with SoftMax activation are used in the research to categorize different fault types in transmission lines. The results of this work are quite encouraging; the proposed method shows a remarkable accuracy of 85.40% in classifying fault types. This performance outperforms current models and demonstrates the value of the developed strategy. Future efforts should focus on interdisciplinary collaboration, developing technologies, and resolving privacy and security concerns to uphold the constancy and efficiency of power grid operations.

REFERENCES

[1] M. Wadi and W. Elmasry, "Statistical analysis of wind energy potential using different estimation methods for Weibull parameters: a case study," Electrical Engineering, vol. 103, pp. 2573–2594, 2021.
[2] M. Wadi and W. Elmasry, "Modeling of wind energy potential in marmara region using different statistical distributions and genetic algorithms," in Proceedings of the 2021 International Conference on Electric Power Engineering – Palestine (ICEPE-P), pp. 1–7, Gaza, Palestine, March 2021.
[3] A. Raza, A. Benrabah, T. Alquthami, and M. Akmal, "A review of fault diagnosing methods in power transmission systems," Applied Sciences, vol. 10, no. 4, 1312 pages, 2020.
[4] M. Wadi and M. Baysal, "Reliability assessment of radial networks via modified RBD analytical technique," Sigma Journal of Engineering and Natural Sciences, vol. 35, pp. 717–726, 2017.
[5] A. P. S. Meliopoulos, G. J. Cokkinides, P. Myrda, Y. Liu, R. Fan, L. Sun, R. Huang, and Z. Tan, "Dynamic state estimation-based protection: Status and promise," IEEE Trans. Power Del., vol. 32, no. 1, pp. 320–330, Feb. 2017.
[6] N. A. Hussien, A. A. D. Al-Magsoosi, H. T. S. AlRikabi, and F. T. Abed, "Monitoring the consumption of electrical energy based on the internet of things applications," International Journal of Interactive Mobile Technologies (iJIM), vol.15, no. 07, p. 17, Apr. 2021.

Journal of Information Systems Engineering and Management
2024, 9(4s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[7] A. K. Ozcanli, F. Yaprakdal, and M. Baysal, “Deep learning methods and applications for electrical power systems: A comprehensive review,” Int. J. Energy Res., vol. 44, no. 9, pp. 7136–7157, Jul. 2020.
[8] M. Mishra, “Power quality disturbance detection and classification using signal processing and soft computing techniques: A comprehensive review,” Int. Trans. Elect. Energy Syst., vol. 29, no. 8, 2019, Art. no. e12008.
[9] G. S. Chawda, A. G. Shaik, M. Shaik, S. Padmanaban, J. B. Holm-Nielsen, O. P. Mahela, and P. Kaliannan, “Comprehensive review on detection and classification of power quality disturbances in utility grid with renewable energy penetration,” IEEE Access, vol. 8, pp. 146807–146830, 2020.
[10] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.
[11] W. Wang, H. Yin, C. Chen, A. Till, W. Yao, X. Deng, and Y. Liu, “Frequency disturbance event detection based on synchro phasors and deep learning,” IEEE Trans. Smart Grid, vol. 11, no. 4, pp. 3593–3605, Jul. 2020.
[12] Godse, R., & Bhat, S. (2020). Mathematical morphology-based feature-extraction technique for detection and classification of faults on power transmission line. IEEE Access, 8, 38459-38471.
[13] Belagoune, S., Bali, N., Bakdi, A., Baadji, B., & Atif, K. (2021). Deep learning through LSTM classification and regression for transmission line fault detection, diagnosis and location in large-scale multi-machine power systems. Measurement, 177, 109330.
[14] Ajagekar, A., & You, F. (2021). Quantum computing based hybrid deep learning for fault diagnosis in electrical power systems. Applied Energy, 303, 117628.
[15] Yoon, D. H., & Yoon, J. (2022). Deep learning-based method for the robust and efficient fault diagnosis in the electric power system. IEEE Access, 10, 44660-44668.
[16] Elmasry, W., & Wadi, M. (2022). Enhanced anomaly-based fault detection system in electrical power grids. International Transactions on Electrical Energy Systems, 2022.
[17] Bon, N. N. (2022). Deep Learning Method for Fault Diagnosis in High Voltage Transmission Lines: A Case of the Vietnam 220kV Transmission Line. International Journal on Electrical Engineering and Informatics, 14(2), 254-275.
[18] Elmasry, W., & Wadi, M. (2022). EDLA-EFDS: A novel ensemble deep learning approach for electrical fault detection systems. Electric Power Systems Research, 207, 107834.
[19] Shakiba, F. M., Shojaee, M., Azizi, S. M., & Zhou, M. Real-Time Sensing and Fault Diagnosis for Transmission Lines. International Journal of Network Dynamics and Intelligence. 2022, 1(1), 36–47.
[20] Abed, F. T., Abed, F. T., Al-Yasriy, H. F., & ALRikabi, H. T. S. (2023). Detection of power transmission lines faults based on voltages and currents values using K-nearest neighbors. International Journal of Power Electronics and Drive Systems (IJPEDS), 14(02), 1033-1043.
[21] Goni, M. O. F., Nahiduzzaman, M., Anower, M. S., Rahman, M. M., Islam, M. R., Ahsan, M., & Shahjalal, M. (2023). Fast and accurate fault detection and classification in transmission lines using extreme learning machine. e-Prime-Advances in Electrical Engineering, Electronics and Energy, 3, 100107.