Electrical Substation Fault Analysis Using 1D-CNNs

Anand Rinchinjugder
M.S. Data Science
Anderson College of Computing
Regis University, Denver, CO, USA
arinchinjugder@regis.edu

Abstract

Substations are easily susceptible to failure and the cost of such a failure is substantial. For instance, a severe ice storm in China in 2008 caused widespread power outages that imposed an economic loss exceeding 2.2 billion USD on the country, prompting a reassessment of its electrical grid infrastructure in response to the snow and ice storms encountered that year [1]. There are no comprehensive real solutions to address this vulnerability effectively. Short term outages, referred to as momentary outages, typically have causes that are not recorded in outage datasets. However, repeated momentary outages can escalate into long term permanent outages, thereby substantially diminishing overall system reliability [2]. This proposal focuses on predicting substation failures using incomplete, noisy, and sensitive data from electrical substations using modern machine learning techniques.

I. Introduction/Background

Electricity serves as a crucial driver of any nation's economy, and ensuring reliability and quality in power supply is essential for satisfying demand and supporting sustained economic growth. Power plants generate electricity at low voltages ranging from 5 to 34.5 kilovolts (kV) and substations perform all the remaining critical functions of supplying electricity. Identifying failures requires the collection of relevant data to train predictive models. However, voltage data are usually continuous and require careful processing to extract meaningful signals from the noise. This difficulty extends beyond data collection and requires using modern machine learning end to end training pipelines to introduce novelty in solving the problem.

II. Problem Statement

According to a 2013 report, the United States experienced average annual costs from power interruptions due to severe weather events estimated between 18 billion USD and 33 billion USD [3]. Causes of failure in power systems include natural factors such as adverse weather, vegetation contact, and animal activity, which all contribute to outages [4]. When substation faults occur, they trigger protective devices such as breakers or medium voltage fuses. After the activation of the protection system, operators dispatch a maintenance team to locate and fix the outage involving equipment malfunction or damage from natural factors. However, these protective devices are not reliable and there are people always checking electrical substation state of operation. Therefore, it is important to predict possible failures by collecting substation data, preparing the time series through cleaning, and demonstrating the detection of failures.

III. Related Work

Failures at substations are traditionally analyzed through manual or ”tacit” knowledge which involves long manual review of data to diagnose issues or physical site visits to the substation. However, there has been some progress has been made toward digitalization and the application of machine learning. For example, Ahmad and Chen [5] utilized regression based models and random forest techniques to forecast load in distribution systems, while other approaches incorporated backpropagation artificial neural networks. Niazazari and Livani [6] applied classification methods with phasor measurement unit data to detect disruptive events in distribution systems. Sun, Zhou, and Yang [7] employed ensemble clustering to identify patterns in household load consumption. Support vector machines have been used by Hosseini, Mahoor, and Khodaei [8] to diagnose causes of power outages. In summary, many studies have tried to improve distribution system reliability using classical machine learning methods. In this proposal, the intent is to explore more advanced machine learning techniques and determine the feasible boundaries for advancement. So although substation failure problems are usually modeled with gradient boosting, random forests, or topology graphs, this paper proposes to train one dimensional Convolutional Neural Network (1D-CNN) models.

IV. Methodology

A. Data Source

Electrical substations have specialized real time data collection devices called Power Communications Monitors (PCMs) that collect receive and transmit signal data. The PCMs produce millions of rows of continuous timestamped values. Therefore, data from these PCMs will be utilized to predict substation failures. We use a dataset spanning 2018-2025 with a total of 25 million records across 268 substations. The data is stored in SQL database; the PCM metadata table includes PCM information such as failure threshold, PCM name, and measurement type. the raw measurement log table contains the 12 features including signal level, margin level, and standing wave ratio for 5 channels.

Fig. 1. Power Communications Monitor Device. Source: https://powercommsolutions.com/products/5350

B. Problem Framing

We frame the problem as a binary classification to predict substation failure events in advance using a 1D-CNN model, given earlier days of data from a measurement device. 1D-CNN models requires segmentation. Therefore, we use sliding window strategy to segment the raw sequential data by creating a window size. More formally, Let,

$i \in \{1, \cdots, N\}$ = Index of measurement devices,
$t \in \mathbb{Z}$ = Integer time steps (hourly),
$x_{i,t} \in \mathrm{R}^f$ = Measurement device logs, (the feature)
τ_i^u = Upper failure threshold for each device, i
τ_i^l = Lower failure threshold for each device i,
W=$\Delta T = [x_{i,t-w+1}, \cdots, x_{i,t}]$ = Sliding window of length W,

and we define a binary target $y \in \{1\}$ if $x_{i,t} \leq \tau_i^l \mid \tau_i^u \leq x_{i,t}$, else $y \in \{0\}$ predicting whether a substation is failing or normal.

V. Data Description/Data Analysis

A. Exploratory Data Analysis

We conduct basic exploratory descriptive analysis processes such as summary statistics, checking for null values, and checking for outliers. The biggest discovery during the EDA process was the imbalance of positive cases of failure.

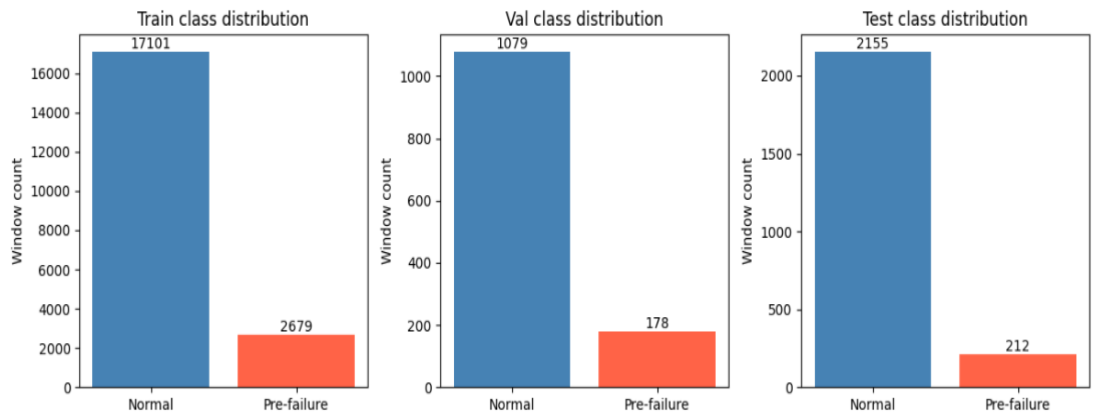

Fig. 2. Imbalance of Normal and Failure Data

B. Data cleaning and pre-processing

The raw dataset contains multiple errors. We handle errors such as no identifiers of measurements, SQL table joins, and unformatted timestamps. We also remove redundant data by dropping values over 3 standard deviations and resampling to 1 hour grid.

Pre-processing was done per-PCM to keep device-specific features. Temporal split was used to prevent data leakage. To handle class imbalance, we split the data into windows based on PCM failures

VI. Model Architecture and Training

A. 1D-CNN Models

1. Simple 1D-CNN

1D-CNN is a low computational complexity deep learning architecture that can learn challenging 1D signals using relatively small number of layers and neurons [9]. According to Kiranyaz et al., 1D forward propagation is defined as:

$$x_k^l = b_k^l + \sum_{i=1}^{N_{l-1}} conv1D(w_{ik}^{l-1}, s_i^{l-1})$$

where x_k^l is defined as the input, b_k^l is defined as the bias of the k^{th} neuron at layer l, s_i^{l-1} is the output of the i^{th} neuron at layer $l-1$, w_{ik}^{l-1} is the kernel from the i^{th} neuron at layer $l-1$ to the k^{th} neuron at layer l.

2. Resnet

Residual neural networks are deep learning models that allow training very deep networks by solving the vanishing gradients problem using skip connections [10]. He et al. defines the fundamental concept as:

$$y = \mathcal{F}(x, \{W_i\}) + W_s x$$

where x and y are the input and output vectors of the layers considered. The function $\mathcal{F}(x, \{W_i\})$ represents the residual mapping to be learned.

3. TCN (Temporal CNN)

TCNs are causal convolutional networks that outperform recurrent networks on sequential tasks [11]. Bai et al. introduce the concepts of causal convolutions as ”convolutions where an output at time t is convolved only with elements from time t and earlier in the previous layer.” Therefore, a simple TCN = 1D Fully convolutions + causal convolution and dilated convolution.

Using these architectures, we train (in increasing order of complexity): 1. 38K parameter simple 1D-CNN with 3 convolutional stages each with BatchNorm, and Maxpool after first 2 stages 2. 221K parameter

1D Resnet with 3 residual blocks and skip connections 3. 287K parameter TCN with 8 temporal blocks and 1, 2, 4, 8, 16, 32, 64, and 128 dilations

B. Training

We use AdamW optimizer [12] on Binary Cross Entropy loss function with logits and OneCycleLR scheduler [13] for convergence. In fig 1, we search the best hyperparameters for each model. To fix class imbalance during training, we use WeightedRandomSampler and give more weights to positive failure cases. For robustness, we used/compared different lead times (day, week, 2 weeks, month, quarter, 6 month, and year).

C. Evaluation

Model performance was evaluated primarily using area under the precision–recall curve (AUC-PR) [14], which is more informative than ROC-based metrics under class imbalance. The model has early stopping based on AUC-PR.

D. Baseline models

To compare the performance of the modern model, we implement 3 classical ML models. Logistic regression . Random forest [15]: using 300 trees and class balanced weighting and XGBoost [16] with depth-6 500 trees.

VII. Expected Outcomes

The objective is to identify substation failures from PCM data using 1D-CNN. However, the noisy characteristics of electrical substation data, influenced by environmental factors such as electromagnetic interference from switching transients, impose inherent limitations [17]. Our assumption is that more complex models will perform better compared to simpler models.

VIII. Timeline

Even though an 8 week period is limited for developing an end to end model, the goal is to extend beyond the knowledge acquired in the MSDS program. Following proposal submission in week 1 and data collection by week 3, the initial phases of code implementation will be presented in week 4. By week 8, the aim is to deliver the full predictive model.

IX. Results

The first table shows the training results of 1D-CNN models with different complexity. All 3 were evaluated on AUC-PR. Simple 1D-CNN achieved 0.81 and Resnet1D 0.79, while TCN performed at 0.74. In the second table, we compare our best performing model with the baseline models. Simple 1D-CNN performs best on AUC-PR than traditional ML models, but all models performs similarly on AUC-ROC.

TABLE I
Best Hyperparameters

Model	Best LR	Best Channels	Best Dropout	AUC-PR
Simple 1D-CNN	0.001	32	0.3	0.81
Resnet1D	0.0005	32	0.3	0.79
TCN	0.001	32	0.4	0.74

TABLE II
BASELINE COMPARISON

Model	AUC-PR	AUC-ROC
Simple 1D-CNN	0.81	0.80
XGBoost	0.67	0.83
Logistic Regression	0.66	0.80
Random Forest	0.63	0.81

X. CONCLUSION

This project addressed the challenge of predicting failures in critical energy infrastructure by developing 3 types of 1D-CNN in increasing order of complexity. Using 25 million rows of data from PCM devices spanning from 2018 to 2025, the simple 1D-CNN successfully predicted electrical substation failures with AUC-PR of 0.81. We expected the TCN to perform better than the simple 1D-CNN and other baseline models, but the results indicate that local patterns outperform full window TCN. Our best explanation is that electrical substation failures are sudden bursts, not gradual trends.

A. Limitations and Future Work

The target labeling strategy for this project is a binary classification. Even though this is a good way to approach the problem in the absence of enough data, future research will focus on obtaining failure event type data with logs and training a multi-label classification model to compare with the existing 1D-CNN models results. We can further extend the problem by assessing the impact of weather on substation failures. In addition, future work can generalize this problem for different geographical locations to compare the performance of 1D-CNN models.

XI. ACKNOWLEDGMENT

I would like to thank the faculty of Anderson College of Computing at Regis University and Professor. Mujtaba for supporting me during the implementation of this research.

REFERENCES

[1] Q. Xie and R. Zhu, "Earth, wind, and ice," *IEEE Power and Energy Magazine*, vol. 9, no. 2, pp. 28–36, 2011.

[2] M. R. Dehbozorgi, M. Rastegar, and A. Sami, "Data mining-based cause identification of momentary outages in power distribution systems," *Sustainable Cities and Society*, vol. 77, p. 103587, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210670721008520

[3] M. Z. Çelik, Sağlam, and B. Oral, "Transmission line failures due to high-impact, low-probability meteorological conditions," *Applied Sciences*, vol. 16, no. 1, 2026. [Online]. Available: https://www.mdpi.com/2076-3417/16/1/379

[4] M. Doostan and B. Chowdhury, "Power distribution system fault cause analysis by using association rule mining," *Electric Power Systems Research*, vol. 152, pp. 140–147, Nov. 2017, publisher Copyright: © 2017 Elsevier B.V.

[5] T. Ahmad and H. Chen, "Nonlinear autoregressive and random forest approaches to forecasting electricity load for utility energy management systems," *Sustainable Cities and Society*, vol. 45, pp. 460–473, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210670718310254

[6] I. Niazazari and H. Livani, "Attack on grid event cause analysis: An adversarial machine learning approach," in *2020 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2020, pp. 1–5.

[7] L. Sun, K. Zhou, and S. Yang, "An ensemble clustering based framework for household load profiling and driven factors identification," *Sustainable Cities and Society*, vol. 53, p. 101958, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210670719309424

[8] Z. Hosseini, M. Mahoor, and A. Khodaei, "Ami-enabled distribution network line outage identification via multi-label svm," *IEEE Transactions on Smart Grid*, vol. PP, pp. 1–1, 06 2018.

[9] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1d convolutional neural networks and applications: A survey," *Mechanical systems and signal processing*, vol. 151, p. 107398, 2021.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[11] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *arXiv preprint arXiv:1803.01271*, 2018.

[12] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *arXiv preprint arXiv:1711.05101*, 2017.
[13] L. N. Smith and N. Topin, "Super-convergence: Very fast training of neural networks using large learning rates," in *Artificial intelligence and machine learning for multi-domain operations applications*, vol. 11006. SPIE, 2019, pp. 369–386.
[14] J. Davis and M. Goadrich, "The relationship between precision-recall and roc curves," vol. 06, 06 2006.
[15] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, p. 5–32, Oct. 2001. [Online]. Available: https://doi.org/10.1023/A:1010933404324
[16] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.
[17] C. M. Wiggins and S. E. Wright, "Switching transient fields in substations," *IEEE Transactions on Power Delivery*, vol. 6, pp. 591–600, 1991. [Online]. Available: https://api.semanticscholar.org/CorpusID:110896331