ISSN 2237-8960 (Online)

CASE STUDY

A data-driven approach to maintenance-preventable causes of failure analysis in power distribution systems

Maycky Kennedy da Silva[1], Rodrigo Sampaio Lopes[2], Marcele Elisa Fontana[3], Jane Kelly Barbosa de Almeida[3]

[1] Federal University of Pernambuco (UFPE), Caruaru, PE, Brazil.
[2] Federal University of São João Del-Rei (UFSJ), São João Del-Rei, MG, Brazil.
[3] Federal University of Pernambuco (UFPE), Recife, PE, Brazil.

How to cite: Silva, M. K. da *et al.* (2025), "A data-driven approach to maintenance-preventable causes of failure analysis in power distribution systems", *Brazilian Journal of Operations and Production Management*, Vol. 22, No. 3, 20252702. https://doi.org/10.14488BJOPM.2702.2025

ABSTRACT

Goal: This study proposes to investigate the main causes of failures in a power distribution system (PDS) that can be mitigated through the implementation of best practices in maintenance management.

Design/Methodology/Approach: This research proposes a data-driven decision-making approach to aid the preventive maintenance management through the analysis of Maintenance-Preventable Causes (MPC) of failure in a real-life PDS. The proposed methodology is structured into three key steps: (1) Data collection and processing of 7,721 power distribution failure records over 12 months (all resolved within 6 hours); (2) Pattern detection using machine learning (ML) algorithms, specifically Association Rule Learning (ARL); and (3) Critical event identification via Social Network Analysis (SNA) with graph-based visualization.

Results: The results show that there was a reduction in the continuity indicators Equivalent Interruption Duration (EID) and Equivalent Interruption Frequency (EIF) by 8.54% and 6.26% respectively, taking as a basis only one MPC (vegetation). The model enables more assertive guidance for both resource planning and the execution of preventive maintenance actions in distribution networks.

Limitations of the investigation: The data used were limited to the southern region of Ceará, Brazil. Therefore, by applying the same methodological approach, other power distribution systems can also be analyzed.

Practical implications: A case study in the southern region of Ceará, Brazil, was conducted to demonstrate the practical applicability of the proposal. This study contributes to identifying variable dependency between failures associated with each MPC and the critical points with the highest impact on the distribution system.

Originality/Value: This study contributes to the academic literature by applying a model that aids in identifying and mitigating the primary failures occurring in power distribution systems, which result in financial losses associated with power supply interruptions, through the use of text mining techniques.

Keywords: Power Distribution System; Machine Learning; Association Rule Mining; Apriori Algorithm; Social Network Analysis.

1 INTRODUCTION

Failures in power distribution systems negatively impact system reliability, result in high costs for distribution utilities, and lead to negative customer experiences (Landegren et al., 2016; Duarte, Ribeiro and Costa, 2024). Gaining deeper knowledge about the root causes of failures and identifying significant variables associated with these causes can enable more

Financial support: This research was funded by Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES), grant number 001.
Conflict of interest: The authors have no conflict of interest to declare.
Corresponding author: marcele.elisa@ufpe.br
Received: 31 March 2025.
Accepted: 27 August 2025.
Editor: Osvaldo Luiz Gonsalves Quelhas.

effective decision-making for properly and efficiently restoring the system (Wesendrup et al., 2024; Campos, Ferreira and Freires, 2021).

Disruptions in power supply or destabilization of the entire system can arise from the malfunction of any component within the power network. Studies indicate that the primary causes of failures in electrical systems are related to vegetation, animals, and weather conditions, including lightning strikes (Xu and Mo, 2006; Xu et al., 2007; Silva and Saraee, 2019). These failures can generally be classified into two categories: Maintenance-Preventable Causes (MPC) and Maintenance-Non-Preventable Causes (MNPC) (National Electric Energy Agency - ANEEL, 2021). The primary interest of distribution utilities is to prevent avoidable failures, i.e., those related to MPC. A key approach in analyzing MPC is leveraging knowledge derived from past failure patterns.

However, due to the random nature of failures and the numerous contributing factors, predicting or identifying failures remains an extremely challenging task (Hasegawa et al., 2025). Moreover, the vast amounts of generated and stored data have made traditional data processing methods increasingly difficult and complex (Kumera et al., 2024). Doostan and Chowdhury (2017) demonstrated that failure prediction and identification tasks have been facilitated by the application of advanced data analytics techniques combined with decision models. According to Rezig et al. (2018), data mining primarily aims to extract the most relevant information from large datasets and, based on predefined criteria, create information and knowledge models. Consequently, several studies have been developed to analyze the characteristics of various failures, using statistical techniques and algorithms to explore failure databases (Doostan and Chowdhury, 2017; Ravi et al., 2019; Silva and Saraee, 2019).

In this context, analyzing maintenance failures in the power distribution system and proposing corrective actions is a time-consuming task that depends on the experience of maintenance personnel, the resources available for intervention, and the location of the failure event (Molęda et al., 2023). Therefore, the application of Machine Learning (ML) techniques presents an opportunity to address the complexity of decision-making, analyze large volumes of historical failure data, and identify patterns of recurring critical events to be mitigated (De Almeida, Lopes and Fontana, 2025). In this way, the use of machine learning algorithms enables the overcoming of human limitations by processing extensive datasets, uncovering hidden patterns, and supporting the proposition of effective intervention actions (Hamdan et al., 2024).

This research proposes a data-driven approach to support preventive maintenance management by analyzing Maintenance-Preventable Causes (MPC) of failures in a real-world power distribution system (PDS). While this study builds upon the approach adopted by Antomarioni et al. (2020, 2022), its primary distinction lies in the research locus. Antomarioni et al. (2020) applied their methodology in a hydroelectric power plant, whereas Antomarioni et al. (2022) focused on an onshore platform. In contrast, this study addresses PDS, which have inherently different operational characteristics, failure patterns, and maintenance challenges. Unlike power generation facilities, distribution networks involve a geographically dispersed infrastructure that is more susceptible to external environmental factors, making failure prediction and prevention even more complex.

This study aims to offer valuable insights into maintenance management practices by employing Machine Learning (ML) approaches. As its main practical contribution, the findings inform managers about the patterns associated with MPC failures, enabling a more efficient allocation of financial resources. The proposed study seeks to improve system availability by identifying correlations among preventable failures, ultimately reducing network disruptions. Given the increasing commercial pressure to minimize operational expenses, this research is justified by the need for cost reduction in distribution utilities (Novochadlo and Paladini, 2024).

In summary, this study contributes to reducing the number of customer complaints by decreasing failure occurrences in the network and mitigating potential costs and fines imposed by regulatory authorities due to future failures. Beyond bridging a gap in the literature and offering practical contributions to the power distribution sector, this research aligns with the Sustainable Development Goals (SDGs) in Brazil, specifically SDG-7 (Affordable and Clean Energy), by promoting energy efficiency through failure reduction and improved financial resource management in power networks.

2 LITERATURE OVERVIEW

To ensure a reliable energy supply to consumers, maintenance management in a power distribution system (PDS) is inherently complex. Operations are typically conducted manually and require territorial control over vast distances. Extensive geographic areas, which are susceptible to climatic effects, combined with the aging of electrical infrastructure, necessitate continuous monitoring to prevent frequent failures in the PDS (Du et al., 2024; Oboudi and Mohammadi, 2024; Liu et al., 2016).

Therefore, more advanced and efficient maintenance planning strategies are essential to ensure asset availability (Mi et al., 2024). The reliability of the PDS is critical for maintaining the uninterrupted supply of electricity to essential sectors, such as hospitals, major industrial centers, and communication hubs. In this context, preventive maintenance measures play a fundamental role (Luz et al., 2024; Yang et al., 2024). Preventive maintenance is an approach that enables planning maintenance actions in the PDS based on the failure probability analysis of system components, aiming to determine the optimal intervention period before a failure occurs (Du et al., 2024; Oboudi and Mohammadi, 2024; Liu et al., 2016).

The literature identifies three main objectives when implementing preventive maintenance: (i) minimizing costs while maintaining a predefined level of reliability, (ii) maximizing reliability within budgetary and time constraints, and (iii) minimizing overall risks (Dehghani et al., 2020). However, effective preventive maintenance requires identifying the primary causes of system failures, such as extreme weather events, bird collisions with transmission lines, line breakages, and failures in electrical components. Thus, preventive maintenance serves as a strategic tool for optimizing decision-making processes regarding scheduled maintenance intervals while ensuring energy availability (Molęda et al., 2023).

Various strategies for maintenance management have been proposed in the literature to establish optimal preventive policies that balance costs and asset availability within the power grid (Kammoun et al., 2022; Paiva et al., 2024). However, the inherent complexity of electrical systems presents challenges in developing rapid and efficient preventive maintenance routines (Yang, Yu, and Liu, 2022). Moreover, with the increasing dependence on energy systems, new tools, such as data mining, are being integrated to facilitate the identification of both critical and non-critical maintenance points (Al-Refaie and Hamdieh, 2024).

Data mining autonomously analyzes large datasets, enabling the extraction of meaningful patterns and relationships. In this context, machine learning (ML) algorithms play a crucial role in processing and learning from vast volumes of data (Jordan and Mitchell, 2015). ML algorithms can be categorized based on the type of learning they employ: (a) supervised learning, which relies on historical data with predefined response variables, and (b) unsupervised learning, where learning occurs through similarity or distance measures between observations (Ramasubramanian and Singh, 2017). Nowadays, there are other types of learning worth mentioning, such as reinforcement learning and transfer learning (Zhu et al., 2023).

Among the various ML algorithms available, this research focuses on Association Rule Learning (ARL), also known as Association Rule Mining (ARM). ARL is a type of unsupervised ML algorithm designed to uncover hidden interdependencies between variables and extract association rules from large databases (Lin et al., 2019; Sheng et al., 2018; Silva and Saraee, 2019). By identifying failure patterns within the explored dataset, ARL contributes to the development of more precise strategies for preserving the operational lifespan of monitored assets (Paiva et al., 2024). Despite its well-established presence in the literature, further research is needed to bridge the gap between mathematical methodologies and practical recommendations through real-world case studies.

3 METHODOLOGY

This research adopts a descriptive approach and employs quantitative methods to propose a data-driven framework for improving preventive maintenance management. The study focuses on analyzing maintenance-preventable causes (MPC) of failures in a power distribution system (PDS). The proposed methodology is structured into three main steps, based on Antomarioni et al. (2020; 2022):

1. **Data Collection and Management:** data on all failure causes are collected, and those with the highest frequency of occurrence are identified. Subsequently, the most

relevant MPC of failure is selected;
2. **Determination of Relevant Associations:** a machine learning algorithm, specifically Association Rule Learning (ARL), is applied to identify significant associations among the primary failure causes;
3. **Social Network Analysis (SNA):** this step is based on graph and network theory, enabling intuitive interpretation. Each node represents a distinct association rule, and arrows connect a cause or group of causes to a node (representing antecedents and consequents). If $x \rightarrow y$ it indicates that when cause x occurs y is also likely to occur. The node diameter represents the frequency of a failure cause in the dataset, while node color opacity indicates the relative lift value of the rule.

This information empowers maintenance managers to make more informed decisions to mitigate failure causes.

3.1 Association rule learning

ARL are expressed as A ⇒ B, where A, B ⊆ I and A ∩ B = ∅, with $I = (I_1, I_2, \ldots, I_m)$ representing a set of items, and T a set transaction such that $T \subseteq I$. The set transaction D comprises all possible transactions in the database. The components A and B are referred to as the antecedent and consequent, respectively, or the Left-Hand Side (LHS) and Right-Hand Side (RHS) of the rule. In simple terms, if antecedent A occurs, consequent B is also likely to occur. The Support (Sup) and Confidence ($Conf$) metrics are mathematically expressed in Equations (1) and (2), respectively (Wang et al., 2020):

$$Sup\ (\mathrm{A} \Longrightarrow \mathrm{B}) = \mathrm{P}(\mathrm{A} \cup \mathrm{B}) \tag{1}$$

$$Conf\ (\mathrm{A} \Longrightarrow \mathrm{B}) = \mathrm{P}(\mathrm{B}|\mathrm{A}) = \frac{\mathrm{P}(\mathrm{A} \cap \mathrm{B})}{\mathrm{P}(\mathrm{A})} \tag{2}$$

Support represents the probability of a given item appearing in a transaction, while confidence estimates the conditional probability of the occurrence of B given A (Liu et al., 2016). According to Doostan and Chowdhury (2017), support and confidence are key metrics for evaluating rule quality, as they indicate statistical significance and rule strength, respectively.

Another critical metric is the lift value, expressed in Equation (3):

$$lift\ (A \Longrightarrow B) = \frac{Sup\ (A \Longrightarrow B)}{Sup\ (B) \times Sup\ (A)} \tag{3}$$

The lift value measures the degree of independence between A and B. According to Yu et al. (2019), when $lift(A \Longrightarrow B) = 1$, the antecedent and consequent are independent, indicating no meaningful association. $lift(A \Longrightarrow B) > 1$, the occurrences of A and B are positively correlated, making the rule useful.

In this study, applying ARL first requires generating a set transaction from standardized failure causes, as illustrated in Table 1. The transactions are determined based on time intervals, which are defined according to the intended analysis. The time interval is established using event timestamps or by experts in maintenance management. It must be greater than the typical maintenance response time to ensure that managers can intervene in the power distribution network effectively.

Table 1 – Example of set transaction generated from failure causes

n	Interval	Set transaction
1	2019-01-01 00:00:00	instantaneous
2	2019-01-01 06:00:00	c("emergency shutdown", "instantaneous", " instantaneous")
3	2019-01-01 12:00:00	c("Large customer internal defect", "emergency shutdown", " fuse switch defect", "instantaneous")
4	2019-01-01 18:00:00	atmospheric discharges
5	2019-01-02 00:00:00	c("instantaneous", "temporary unidentified defect")

According to Chemweno et al. (2016), in the transaction generation process for a selected failure cause, all failure causes occurring before or after it are combined to form a transaction set. For instance, in Table 1, the second transaction initially records the failure cause “emergency shutdown,” followed chronologically by “instantaneous,” and then “instantaneous” again. This approach ensures that both preceding and subsequent failure events are considered during the transaction set generation process.

Next, the Apriori Algorithm is applied to extract frequent item subsets that can form association rules within the dataset through an iterative, layer-by-layer process (Wang et al., 2020). The Apriori-based data mining procedure consists of two main steps: (a) Identifying frequent item sets; (b) Generating strong association rules based on the frequent item sets (Tian et al., 2020; Xinchun et al., 2018). The algorithm terminates when no further frequent item sets can be found. This study follows the Apriori algorithm steps outlined by Tian et al. (2020). Different constraints can be applied when analyzing association rules, including:

- **Minimum Support**: Defines the threshold for identifying frequent itemset. The selection of this value depends on the characteristics of the dataset;
- **Minimum Confidence**: Determines rule strength. Increasing this threshold results in fewer but stronger rules;
- **Right-Hand Side (RHS)**: Since one objective of this research is to identify patterns leading to MPC failures, association rules with MPC failure causes as the RHS are prioritized.

Finally, the most relevant rules for MPC failures are identified based on their lift values. Rules with the highest lift values are selected and interpreted by maintenance managers. The SNA approach further enhances the analysis, providing a visual and structural representation of failure cause associations.

4 RESULT

The results will be presented following the three steps described in the methodology.

4.1 Data collection and management

This research was conducted in the Cariri region, located in the southern part of Ceará state, in northeastern Brazil. According to the Ceará Institute for Research and Economic Strategy (IPECE), the Cariri region comprises 29 municipalities. With a total land area of 17,298.35 km², the Brazilian Institute of Geography and Statistics (IBGE) estimated the region's population at 1,080,326 inhabitants in 2024. The local electricity distribution company supplies power to more than 400,000 consumer units.

For this study, failure incidence data were collected over a 12-month period, resulting in a total of 7,721 recorded failures. The data is stored in spreadsheets and manipulated by the R Studio program. Figure 1 presents the monthly distribution of failures. All figures and tables were created by the authors using the database provided by the case study company.

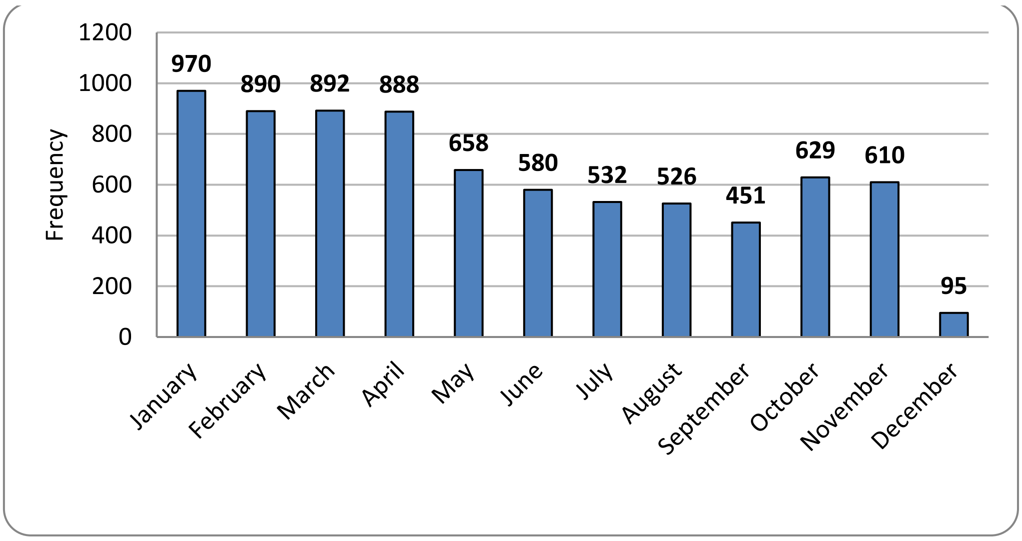


Figure 1 - Number of failures per month

As shown in Figure 1, the first four months account for 47% of the total failures. This is attributed to the region's rainy season, during which the average precipitation reaches 725 mm. Analyzing the failure causes, 54 different types were identified. Table 2 presents the most frequent failure causes.

Table 2 - Main causes identified

Type	Description	Is an MPC?	Frequency (%)
Atmospheric discharges	Refers to lightning strikes impacting the electrical network or its equipment. These events can lead to service interruptions and potential damage to system assets.	No	17,37
Load Transfer and Retransfer	This category is not associated with actual faults in the distribution network. It involves switching operations—either planned or emergency—to transfer loads between feeders, followed by restoration to the original configuration.	No	13,03
Instantaneous disturbances (voltage spikes)	Assigned to short-duration outages, typically lasting less than three minutes. These events are often caused by temporary contacts (e.g., tree branches touching the line), triggering automatic protective devices such as reclosers. Although service is quickly restored, such incidents signal the need for targeted inspections.	No	12,46
Bird contact	Occurs when birds create a short circuit by bridging energized components, such as transformers or switchgear, leading to localized outages. This cause is confirmed only when physical evidence of the bird is observed on site.	No	10,24
Emergency shutdown	Refers to deliberate disconnection requested by field crews upon identifying imminent risks (e.g., vegetation in contact with conductors), even if the service has already been automatically restored. The shutdown is performed to eliminate hazards prior to re-energization.	Yes	8,79
Unidentified temporary defect	Used when field personnel are unable to determine the root cause of a protection device operation. This may result from limited inspection time, lack of technical expertise, or difficulty in accessing the entire medium-voltage (MV) or low-voltage (LV) network segment.	Not defined	8,47
Vegetation interference	Applies to faults caused by excessive vegetation contact with overhead conductors. Restoration is achieved through vegetation management, such as pruning, by maintenance crews.	Yes	5,01
Critical Internal Day	Designation applied to operationally critical days—typically during severe weather events such as heavy rain. Faults classified under this cause are not included in continuity indices (e.g., Equivalent Interruption Duration (EID) and Equivalent Interruption Frequency (EIF)) for regulatory purposes.	No	4,14
Fuse switch defect	Malfunction of a fuse switch due to issues such as a broken base or disconnection of the fuse cartridge (link), which prevents normal operation. A temporary bypass is implemented until full replacement is completed.	Yes	3,42
Incident without impact	Refers to an event detected in the distribution system—either through monitoring systems, protection device activation, or field inspection—that did not result in any measurable disruption of energy supply to end-users.	Not defined	2,23
Third-party accidental damage	Faults caused by external agents, typically accidental, such as vehicle collisions with utility poles or construction activities that impact the distribution infrastructure.	No	1,98
Transformer defect	Internal failure of a distribution transformer, confirmed through no-load testing in the field, indicating that the unit is damaged and requires	Yes	1,68

	replacement.		
Large customer internal defect	Interruptions originating within a customer's internal electrical installation, particularly in large consumers or commercial/industrial facilities, and not attributable to the utility's network.	No	1,57
Normal condition	Assigned when service complaints are not confirmed upon field inspection, as the power supply has already been restored automatically before the crew's arrival.	No	1,30
Animal contact	General classification for faults caused by small animals (e.g., rodents or reptiles) making contact with energized components, potentially resulting in localized outages.	No	1,11
Other causes	Sum of all other causes with an incidence of less than 1%	Not defined	7,19

These insights provide maintenance managers with an overview of the months with the highest failure rates and a detailed breakdown of the most recurrent causes. For the analyzed case study, the maintenance-preventable causes (MPC) that imposed the highest demand on the maintenance sector were: Emergency shutdown, Vegetation interference, and Fuse switch defects.

4.2 Determination of the relevant associations

Based on these initial analyses, the model was applied to generate association rules to identify potential electrical system failures and recommend optimal mitigation strategies. The association rules were derived using a minimum support threshold of 0.0018, corresponding to at least 100 occurrences. This means that any itemset appearing more than 100 times is considered frequent. These rules can then be analyzed and interpreted to identify the most significant ones. It is worth noting that selecting a lower support threshold is generally preferable to choosing a higher one, as the latter may lead to the omission of potentially valuable rules.

The minimum confidence threshold was set at 70%. As previously mentioned, the selection of this value depends on the discretion of maintenance managers. To identify the most relevant association rules that reveal patterns for MPC failures, all rules were analyzed using the inspection function of the Apriori algorithm. This process involves classifying and identifying high-quality rules with significant lift values. In this study, rules with a lift value greater than 3.0 were selected for further analysis.

Tables 3, 4 and 5 summarize the results of the association rule mining process for failures related to emergency shutdowns, vegetation-related incidents, and fuse switch defects, respectively.

Table 3 - Association Rule Mining for Failures Related to "Emergency Shutdown"

LHS		RHS	Support	Confidence	Lift
{foreign object, large customer internal defect}	=>	{emergency shutdown}	0.00212	1.0000	3.5783
{temporary unidentified defect, bird, voltage level complaint}	=>	{emergency shutdown}	0.00282	1.0000	3.5783
{fallen tree, lightning strikes, vegetation}	=>	{emergency shutdown}	0.00212	1.0000	3.5783
{fallen tree, temporary unidentified defect, vegetation}	=>	{emergency shutdown}	0.00212	1.0000	3.5783
{fallen tree, load transfer and retransfer, vegetation}	=>	{emergency shutdown}	0.00212	1.0000	3.5783
{temporary unidentified defect, shutdown upon customer request, vegetation}	=>	{emergency shutdown}	0.00212	1.0000	3.5783
{temporary unidentified defect, shutdown upon customer request, load transfer and retransfer}	=>	{emergency shutdown}	0.00282	1.0000	3.5783
{shutdown upon customer request, instantaneous failure, load transfer and retransfer}	=>	{emergency shutdown}	0.00282	1.0000	3.5783
{animals, kite, load transfer and retransfer}	=>	{emergency	0.00212	1.0000	3.5783

		shutdown}			
{instantaneous failure, kite, load transfer and retransfer}	=>	{emergency shutdown}	0.00212	1.0000	3.5783
{instantaneous failure, bird, kite}	=>	{emergency shutdown}	0.00212	1.0000	3.5783
{insulator defect, normal condition, instantaneous failure}	=>	{emergency shutdown}	0.00282	1.0000	3.5783

Table 4 - Association Rule Mining Results for Failures Related to "Vegetation"

LHS		RHS	Support	Confidence	Lift
{fallen tree, lightning strikes, emergency shutdown}	=>	{vegetation}	0.00212	1.0000	4.4842
{fallen tree, temporary unidentified defect, emergency shutdown}	=>	{vegetation}	0.00212	1.0000	4.4842
{fallen tree, emergency shutdown, load transfer and retransfer}	=>	{vegetation}	0.00212	1.0000	4.4842
{material degradation, lightning strikes, bird}	=>	{vegetation}	0.00212	1.0000	4.4842
{large customer internal defect, instantaneous failure, bird's nest}	=>	{vegetation}	0.00212	1.0000	4.4842
{large customer internal defect, bird's nest, bird}	=>	{vegetation}	0.00212	1.0000	4.4842
{bird's nest, bird, accidental third-party interference}	=>	{vegetation}	0.00212	1.0000	4.4842
{fuse switch defect, lightning strikes, bird's nest}	=>	{vegetation}	0.00282	1.0000	4.4842
{lightning strikes, bird's nest, bird}	=>	{vegetation}	0.00494	1.0000	4.4842
{temporary unidentified defect, bird's nest, load transfer and retransfer}	=>	{vegetation}	0.00282	1.0000	4.4842
{temporary unidentified defect, bird's nest, bird}	=>	{vegetation}	0.00212	1.0000	4.4842
{bird's nest, bird, load transfer and retransfer}	=>	{vegetation}	0.00353	1.0000	4.4842

Table 5 - Association Rule Mining Results for Failures Related to "Fuse Switch Defect"

LHS		RHS	Support	Confidence	Lift
{customer internal defect, accidental third-party interference, load transfer and retransfer}	=>	{fuse switch defect}	0.00212	1.0000	6.1609
{customer internal defect, bird, accidental third-party interference, load transfer and retransfer}	=>	{fuse switch defect}	0.00212	1.0000	6.1609
{lightning strikes, emergency shutdown, broken jumper, bird, vegetation}	=>	{fuse switch defect}	0.00212	1.0000	6.1609
{lightning strikes, instantaneous failure, broken jumper, bird, vegetation}	=>	{fuse switch defect}	0.00212	1.0000	6.1609
{temporary unidentified defect, lightning strikes, instantaneous failure, broken jumper, bird}	=>	{fuse switch defect}	0.00212	1.0000	6.1609
{lightning strikes, emergency shutdown, broken jumper, bird, load transfer and retransfer, vegetation}	=>	{fuse switch defect}	0.00212	1.0000	6.1609
{temporary unidentified defect, lightning strikes, instantaneous failure, broken jumper, bird, vegetation}	=>	{fuse switch defect}	0.00212	1.0000	6.1609
{lightning strikes, instantaneous failure, broken jumper, bird, load transfer and retransfer, vegetation}	=>	{fuse switch defect}	0.00212	1.0000	6.1609
{temporary unidentified defect, lightning strikes, instantaneous failure, broken jumper, bird, load transfer and retransfer}	=>	{fuse switch defect}	0.00212	1.0000	6.1609
{temporary unidentified defect, lightning strikes, instantaneous failure, broken jumper, bird, load transfer and retransfer, vegetation}	=>	{fuse switch defect}	0.00212	1.0000	6.1609
{lightning strikes, broken jumper, bird, vegetation}	=>	{fuse switch defect}	0.00282	0.8000	4.9287
{instantaneous failure, broken jumper, bird, vegetation}	=>	{fuse switch defect}	0.00282	0.8000	4.9287

By analyzing Table 3, it is evident that fallen trees and vegetation are frequent causes of failures related to emergency shutdowns. Table 4 highlights that fallen trees and adverse weather conditions (lightning strikes) significantly impact failures related to vegetation. Finally, Table 5 indicates that the primary causes of fuse switch defects are broken jumpers, vegetation, and adverse weather conditions (lightning strikes).

4.3 Social network analysis

To enhance the visualization of these findings, social network analysis (SNA) was applied, allowing a clearer interpretation of the association rules related to MPC failures. Figures 2, 3 and 4 present the association rule mining results for emergency shutdowns, vegetation-related failures, and fuse switch defects, respectively.

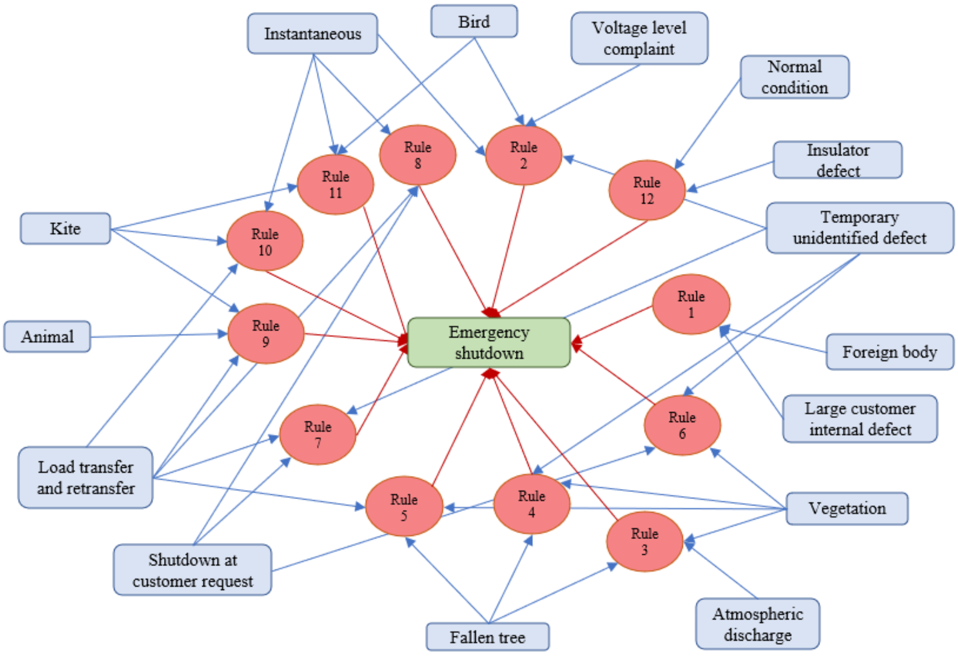


Figure 2 - SNA for failures related to “emergency shutdown”

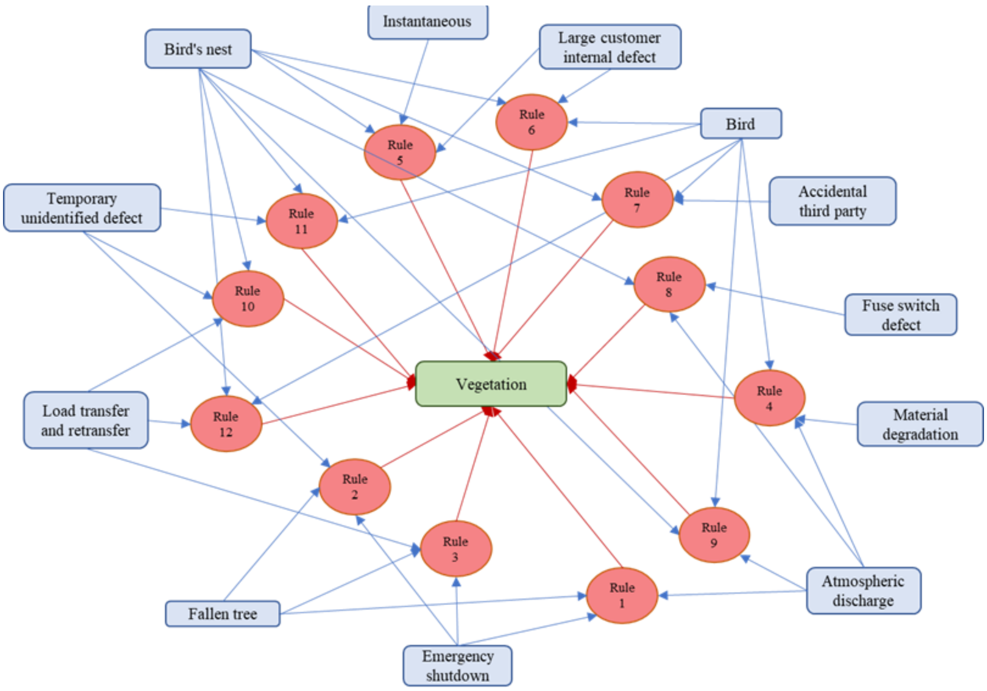


Figure 3 - SNA for failures related to “vegetation”

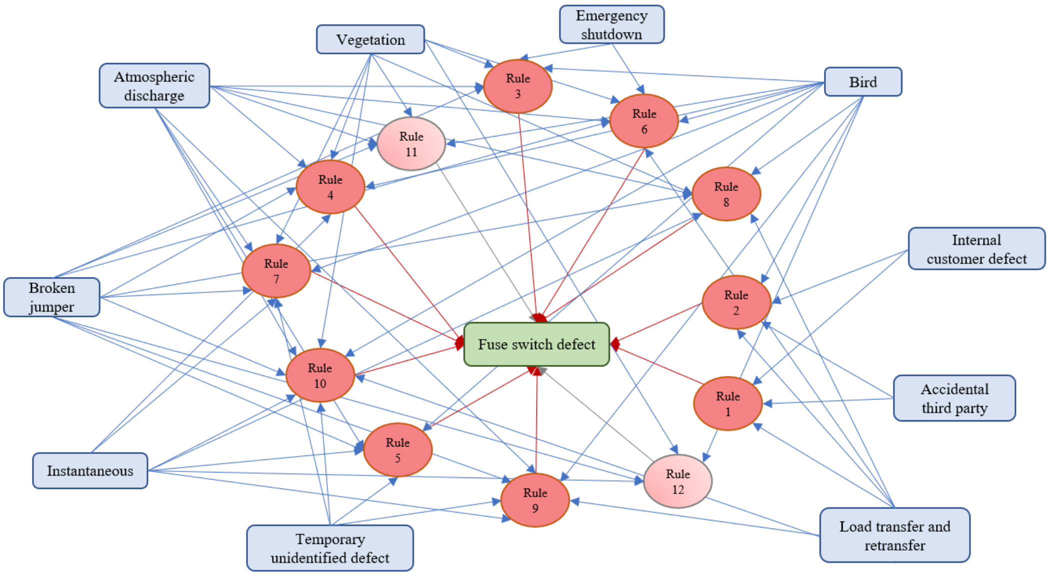


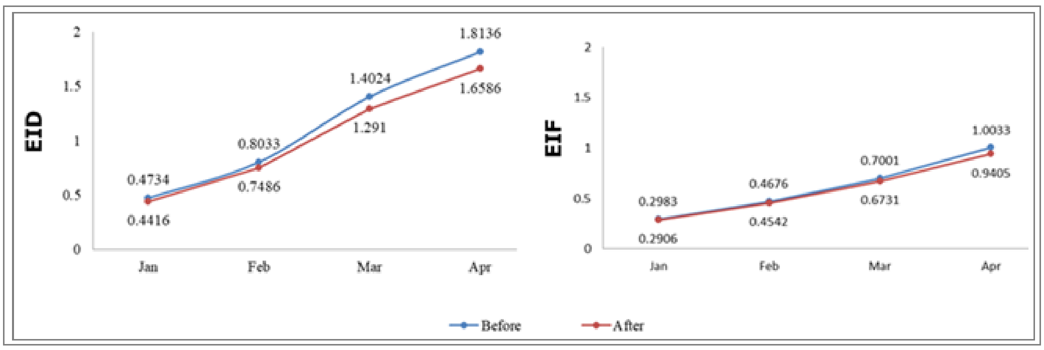
Figure 4 - SNA for failures related to "fuse switch defect"

In summary, the results indicate a strong correlation between failure types within each MPC category, with vegetation emerging as a predominant cause. Consequently, maintenance managers should implement preventive measures such as intensifying electrical grid inspections to identify critical points and schedule targeted interventions.

5 DISCUSSION

Based on the results of the association rules generated, maintenance managers decided to intensify on-site visual inspections across the entire network. Subsequently, preventive interventions were planned to proactively address failures, with a primary focus on the MPC category "vegetation." After one year of implementing a preventive maintenance plan specifically targeting vegetation-related failures, a comparative analysis of results (before and after implementation) was conducted.

In Brazil, the National Electric Energy Agency (ANEEL) monitors two key electricity reliability indicators to assess the quality of distribution services: Equivalent Interruption Duration per Consumer Unit (EID) – expressed in hours and hundredths of hours; and Equivalent Interruption Frequency per Consumer Unit (EIF) – expressed as the number of interruptions and hundredths of an interruption. Both EID and EIF are cumulative annual indicators that aggregate interruption data across all consumers in a given year. Thus, these indicators were used to evaluate our preventive maintenance plan by comparing the first four months (January-April) of two consecutive years: the baseline period before implementation (Year 1) and the same seasonal period after implementation (Year 2). This seasonal comparison controls for annual variation while isolating the maintenance plan's effects. Figures 5.a and 5.b show the performance of EID and EIF, respectively.



(a) (b)

Figure 5 - Performance of the indicators: (a) EID and (b) EIF

The results show a cumulative reduction of 8.54% and 6.26% in EID and EIF, respectively, considering only the MPC vegetation as the failure cause. This outcome demonstrates the benefits of a systemic approach to failure analysis, as proposed in this study. More specifically, the findings reveal several managerial implications relevant to maintenance management in electrical distribution systems.

The first implication concerns the prioritization of preventive actions based on data. Through Association Rule Mining, it was possible to identify critical patterns, such as the predominance of vegetation as the primary cause of failures. This enables managers to adopt a proactive approach, intensifying targeted inspections and planning preventive interventions before failures occur, thereby improving system reliability.

Furthermore, the application of computational analysis techniques and artificial intelligence in maintenance management underscores the importance of digital transformation in the electricity sector. The use of algorithms to identify failures and define strategic actions can be expanded to other causes beyond vegetation, making the maintenance system increasingly efficient and data-driven.

Additionally, reducing failures and power supply interruptions directly impacts consumer perception and regulatory indicators required by agencies such as ANEEL. This can result in an improved reputation for the utility company and a lower risk of regulatory penalties, corroborating the authors' claims (Rampini and Berssaneti, 2024) about the importance of risk management to implement decision-making that results in adequate control of maintenance operations. The reduction in interruption indicators (EID and EIF) confirms that a more targeted maintenance plan can optimize company resources (Dhewi et al., 2025). Instead of conducting generalized inspections and maintenance, the company can focus efforts on the most critical points, reducing operational costs and improving service performance.

The results suggest that this model can be integrated into the company's strategic planning, enabling the execution of short-, medium-, and long-term action plans. Moreover, there is potential to expand the use of this technique to analyze other failure causes (such as mechanical failures or equipment defects), thereby increasing the robustness of the maintenance system. However, implementing this model requires that managers and technical staff be prepared to interpret failure analysis results and apply the recommendations in maintenance planning. This implies the need for continuous training to ensure that data-driven strategies are correctly implemented.

6 CONCLUSION

A power distribution system is inherently vulnerable to failures, which impact utility companies responsible for ensuring high-quality power supply. Studies and analyses aimed at improving reliability with minimal costs and optimal performance indicators are essential for effective maintenance planning in power distribution companies. Therefore, it is crucial for electricity distribution companies to implement effective anomaly management strategies.

This study contributes to this objective by proposing a model capable of identifying variable dependency between failures related to each MPC and the critical points with the highest impact on the distribution system. Consequently, it was possible to strategically direct condition-based preventive interventions to the most critical segments within a given region.

From a theoretical perspective, the analysis of a dataset comprising more than 7,000 failure occurrences underscores the significance of data processing and analysis using computational techniques and Machine Learning (ML). This study contributes to this field by specifically addressing failures in power distribution networks.

From a practical perspective, applying the proposed model to a real-world distribution system resulted in a reduction of continuity indicators when compared to the performance of the existing network configuration. The model enabled more precise planning and execution of maintenance activities, improving maintenance management efficiency and mitigating the financial losses associated with power outages.

In summary, the results indicate that a data-driven approach not only improves reliability and operational efficiency but also generates competitive advantages, reduces costs, and strengthens maintenance management in the electrical distribution sector.

This study offers important insights into power distribution failures while acknowledging certain limitations. Our analysis specifically examined the most frequently occurring failure types during the study period, which means some rare but potentially significant causes may not have been captured. We selected the Apriori algorithm for its strengths in producing interpretable results, though we recognize that alternative approaches like FP-Growth or ECLAT could potentially reveal different patterns, especially when applied to more extensive datasets. For future research, important directions include investigating relationships between primary and secondary failure causes to improve prioritization, evaluating model performance over extended periods to assess robustness, comparing different association rule algorithms to optimize pattern discovery, and exploring the integration of environmental factors. These steps would further strengthen the practical application of these findings for utility companies.

ACKNOWLEDGMENTS

The authors would like to thank the Coordination for the Improvement of Higher Education Personnel (CAPES), and the National Council for Scientific and Technological Development (CNPq).

REFERENCES

Al-Refaie, A., & Hamdieh, B. A. (2024), “A data mining framework for maintenance prediction of faulty products under warranty,” *Journal of Advanced Manufacturing Systems, 23*(1), 35-59. https://doi.org/10.1142/S0219686724500021

Agência Nacional de Energia Elétrica (ANEEL). (2021), “Resolução 024,” Available at: https://www.aneel.gov.br/.

Antomarioni, S., Bellinello, M. M., Bevilacqua, M., Ciarapica, F. E., da Silva, R. F., & de Souza, G. F. M. (2020), “A data-driven approach to extend failure analysis: A framework development and a case study on a hydroelectric power plant,” *Energies, 13*(23), 6400. https://doi.org/10.3390/en13236400

Antomarioni, S., Ciarapica, F. E., & Bevilacqua, M. (2022), “Association rules and social network analysis for supporting failure mode effects and criticality analysis: Framework development and insights from an onshore platform,” *Safety Science, 150*, 105711. https://doi.org/10.1016/j.ssci.2022.105711

Campos, J. T. de G. A. e A., Ferreira, A. M. S., & Freires, F. G. M. (2021), “Time variability management and trade-off analysis of quality, productivity, and maintenance efficiency,” Brazilian Journal of Operations & Production Management, 18(4), 1–19. https://doi.org/10.14488/BJOPM.2021.018

Chemweno, P., Pintelon, L., Jongers, L., & Muchiri, P. (2016), “i-RCAM: Intelligent expert system for root cause analysis in maintenance decision making,” *2016 IEEE International Conference on Prognostics and Health Management (ICPHM)*. https://doi.org/10.1109/icphm.2016.7542830

DE ALMEIDA, Jane Kelly Barbosa; LOPES, Rodrigo Sampaio; FONTANA, Marcele Elisa. Predictive maintenance management of gear systems in the era of computer vision. International Journal of Quality & Reliability Management, 2025.

Dehghani, N. L., Darestani, Y. M., & Shafieezadeh, A. (2020), “Optimal life-cycle resilience enhancement of aging power distribution systems: A MINLP-based preventive maintenance planning,” *IEEE Access, 8*, 22324-22334. https://doi.org/10.1109/ACCESS.2020.2969997

Dhewi, R. M., Martunus, H., Hidayah, N. and Setiany, E. (2025), “Enterprise risk management manufacturing industry quality determinants: Malaysia and Indonesia context,” Brazilian Journal of Operations & Production Management, 22(1), p. 2238. https://doi.org/10.14488/BJOPM.2238.2025

Doostan, M., & Chowdhury, B. H. (2017), “Power distribution system fault cause analysis by using association rule mining,”. *Electric Power Systems Research, 152*, 140–147. https://doi.org/10.1016/j.epsr.2017.07.005

Du, W. L., et al. (2024), "An efficient nonlinear method for cascading failure analysis and reliability assessment of power distribution lines under wind hazard," *Reliability Engineering & System Safety, 245*, 109995. https://doi.org/10.1016/j.ress.2024.109995.

Hasegawa, H. L., Lima, R. S. de, Mota Junior, V. D. da and Teixeira, R. L. P. (2025), "Challenges in data collection for enhancing productivity in Brazilian industrial processes," Brazilian Journal of Operations & Production Management, 22(1), p. 2445. https://doi.org/10.14488/BJOPM.2445.2025

Duarte, T. E., Ribeiro, P. C. C., & Costa, H. G. (2024), "Indicators and performance requirements for suppliers' evaluation in the Brazilian electricity sector," Brazilian Journal of Operations & Production Management, 21(3), 1984. https://doi.org/10.14488/BJOPM.1984.2024

Jordan, M. I., & Mitchell, T. M. (2015), "Machine learning: Trends, perspectives, and prospects," *Science, 349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

Kammoun, M. A., Hajej, Z., & Rezg, N. (2022), "A multi-level selective maintenance strategy combined to data mining approach for multi-component system subject to propagated failures,". *Journal of Systems Science and Systems Engineering, 31*(3), 313-337. https://doi.org/10.1007/s11518-022-5525-9.

Kumera, D., Amentie, C. and Bali, N. (2024), "Effect of technological innovation on firm's performance: mediating effect of competitive advantage: a study on manufacturing firms operating in Ethiopian industrial parks," Brazilian Journal of Operations and Production Management, Vol. 21, No. 3, e20242146. https://doi.org/10.14488/BJOPM.2146.2024

Landegren, F. E., Johansson, J., & Samuelsson, O. (2016), "A method for assessing margin and sensitivity of electricity networks with respect to repair system resources," *IEEE Transactions on Smart Grid, 7*, 2880–2889. https://doi.org/10.1109/TSG.2016.2582080

Lin, H.-K., Hsieh, C.-H., Wei, N.-C., & Peng, Y.-C. (2019), "Association rules mining in R for product performance management in Industry 4.0," *Procedia CIRP, 83*, 699–704. https://doi.org/10.1016/j.procir.2019.04.099

Liu, X., et al. (2016), "Microgrids for enhancing the power grid resilience in extreme conditions," *IEEE Transactions on Smart Grid, 8*(2), 589-597. https://doi.org/10.1109/TSG.2016.2579999

Luz, T. J. da, Gapski, A. L., & Unsihuay-Vila, C. (2024), "Multiobjective optimization of maintenance applied in electric power distribution systems," *Brazilian Archives of Biology and Technology, 67*, e24230894. https://doi.org/10.1590/1678-4324-2024230894.

Mi, K., et al. (2024), "Research on path planning of intelligent maintenance robotic arm for distribution lines under complex environment," *Computers and Electrical Engineering, 120*, 109711. https://doi.org/10.1016/j.compeleceng.2024.109711.

Molęda, M., et al. (2023), "From corrective to predictive maintenance: A review of maintenance approaches for the power industry," *Sensors, 23*(13), 5970. https://doi.org/10.3390/s23135970.

Novochadlo, Y.M. and Paladini, E. P. (2024), "The application of real-time overall equipment efficiency indicator in a medium-sized company," Brazilian Journal of Operations and Production Management, Vol. 21, No. 2, e20242042. https://doi.org/10.14488/BJOPM.2042.2024

Oboudi, M. H., & Mohammadi, M. (2024), "Two-stage seismic resilience enhancement of electrical distribution systems," *Reliability Engineering & System Safety, 241*, 109635. https://doi.org/10.1016/j.ress.2023.109635

Paiva, R. G., Cavalcante, C. A., & Do, P. (2024), "Applying association rules in the maintenance and reliability of physical systems: A review," Computers & Industrial Engineering, 194, 110332. https://doi.org/10.1016/j.cie.2024.110332

Ramasubramanian, K., & Singh, A. (2017), “Machine learning using R: With time series and industry-based use cases in R,”.- 2nd ed.- New Delhi, India: Apress.

HAMDAN, Ahmad et al. (2024), “AI in renewable energy: A review of predictive maintenance and energy optimization,” International Journal of Science and Research Archive, v. 11, n. 1, p. 718-729. https://doi.org/10.30574/ijsra.2024.11.1.0112

Rampini, G.H.S. and Berssaneti, F.T. (2024), “Impact of critical success factors and risk management on organizational results,” Brazilian Journal of Operations and Production Management, Vol. 21, No. 1, e20241412. https://doi.org/10.14488/BJOPM.1412.2024

Ravi, N. N., Drus, S. M., Krishnan, P. S., & Ghani, N. L. A. (2019), “Substation transformer failure analysis through text mining,” *2019 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE)*. https://doi.org/10.1109/ISCAIE.2019.8743719

Rezig, S., Achour, Z., & Rezg, N. (2018), “Using data mining methods for predicting sequential maintenance activities,” *Applied Sciences, 8*, 2184. https://doi.org/10.3390/app8112184

Sheng, G., Hou, H., Jiang, X., & Chen, Y. (2018), “A novel association rule mining method of big data for power transformers state parameters based on probabilistic graph model,”. *IEEE Transactions on Smart Grid, 9*(2), 695–702. https://doi.org/10.1109/TSG.2016.2562123

Silva, C., & Saraee, M. (2019), “Understanding causes of low voltage (LV) faults in electricity distribution network using association rule mining and text clustering,” *2019 IEEE International Conference on Environment and Electrical Engineering and 2019 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*. https://doi.org/10.1109/EEEIC.2019.8783949

Tian, M., et al. (2020), “Data dependence analysis for defects data of relay protection devices based on Apriori algorithm,”. *IEEE Access, 8*, 120647-120653. https://doi.org/10.1109/ACCESS.2020.3006345

Xinchun, J., Haikuan, W., Minrui, F., Dajun, D., Qing, S., & Yang, T. C. (2018), “Anomaly behavior detection and reliability assessment of control systems based on association rules,” *International Journal of Critical Infrastructure Protection*. https://doi.org/10.1016/j.ijcip.2018.06.001

Xu, L., & Mo, Y. C. (2006), “A classification approach for power distribution systems fault cause identification,” *IEEE Transactions on Power Systems, 21*(1), 53-60. https://doi.org/10.1109/TPWRS.2005.861981

Xu, L., Chow, M., & Taylor, L. S. (2007), “Power distribution fault cause identification with imbalanced data using the data mining-based fuzzy classification E-algorithm,” *IEEE Transactions on Power Systems, 22*(1), 164-171. https://doi.org/10.1109/TPWRS.2006.888990

Wang, R., Tang, Z., Gao, J., Gao, Z., & Wang, Z. (2020), “Probabilistic model-checking based reliability analysis for failure correlation of multi-state systems,” *Quality Engineering*. https://doi.org/10.1080/08982112.2019.1692139

Wesendrup, K., Hellingrath, B. and Nikolarakis, Z. (2024), “A framework for conceptualizing integrated prescriptive maintenance and production planning and control models,” Brazilian Journal of Operations & Production Management, 21(3), p. 2172. https://doi.org/10.14488/BJOPM.2172.2024

Yang, J., et al. (2024), “A transformer maintenance interval optimization method considering imperfect maintenance and dynamic maintenance costs,” *Applied Sciences, 14*, 6845. https://doi.org/10.3390/app14156845.

Yang, X., Yu, M., & Liu, F. (2022), “Construction of power network operation and maintenance cost prediction model based on data information mining,” *2022 International Conference on Big Data, Information and Computer Network (BDICN)*. https://doi.org/10.1109/BDICN55575.2022.00031

Yu, S., Jia, Y., & Sun, D. (2019), "Identifying factors that influence the patterns of road crashes using association rules: A case study from Wisconsin, United States," *Sustainability, 11*(7), 1925. https://doi.org/10.3390/su11071925

Zhu, Zhuangdi et al. (2023), "Transfer learning in deep reinforcement learning: A survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 45, n. 11, p. 13344-13362. https://doi.org/10.48550/arXiv.2009.07888

Authors contributions: Conceptualization M.K da S. and R.S.L.; Methodology - M.K da S. and R.S.L.; Validation and formal analysis - M.K da S.; Writing—original draft preparation - M.K da S.; Writing—review and editing - M.K da S., J.K.B. de A., R.S.L. and M.E.F.; Project administration –R.S.L.; Funding acquisition - R.S.L. and M.E.F.