



UNIVERSIDADE FEDERAL DE PERNAMBUCO

CENTRO ACADÊMICO DO AGRESTE

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

MAYCKY KENNEDY DA SILVA

**APRENDIZADO DE MÁQUINA APLICADO A SISTEMAS DE DISTRIBUIÇÃO DE
ENERGIA: um estudo para auxiliar na gestão de manutenção do sistema**

Caruaru

2022

MAYCKY KENNEDY DA SILVA

**APRENDIZADO DE MÁQUINA APLICADO A SISTEMAS DE DISTRIBUIÇÃO DE
ENERGIA: um estudo para auxiliar na gestão de manutenção do sistema**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de mestre em Engenharia de Produção.

Área de concentração: Otimização e Gestão da Produção.

Orientador: Prof. Dr. Rodrigo Sampaio Lopes

Caruaru

2022

Catálogo na fonte:
Bibliotecário – Raul César de Melo - CRB/4 - 1735

S586a Silva, Maycky Kennedy da.
Aprendizado de máquina aplicado a sistemas de distribuição de energia: um estudo para auxiliar na gestão de manutenção do sistema / Maycky Kennedy da Silva. – 2022.
69 f.; il.: 30 cm.

Orientador: Rodrigo Sampaio Lopes.
Dissertação (Mestrado) – Universidade Federal de Pernambuco, CAA, Mestrado em Engenharia de Produção, 2022.
Inclui Referências.

1. Energia elétrica – Distribuição. 2. Aprendizado do computador. 3. Manutenção.
I. Lopes, Rodrigo Sampaio (Orientador). II. Título.

CDD 658.5 (23. ed.)

UFPE (CAA 2022-008)

MAYCKY KENNEDY DA SILVA

APRENDIZADO DE MÁQUINA APLICADO A SISTEMAS DE DISTRIBUIÇÃO DE ENERGIA: um estudo para auxiliar na gestão de manutenção do sistema

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de mestre em Engenharia de Produção.

Área de concentração: Otimização e Gestão da Produção.

Aprovada em: 24/02/2022.

BANCA EXAMINADORA

Prof. Dr. Rodrigo Sampaio Lopes (Orientador)
Universidade Federal de Pernambuco

Prof.^a Dr.^a Marcele Elisa Fontana (Examinadora Interna)
Universidade Federal de Pernambuco

Prof. Dr. Thiago Poletto (Examinador Externo)
Universidade Federal do Pará

Dedico a minha família, em especial a minha mãe Ana Lucia, por toda a base moral cedida.

A minha esposa Tatiana Sousa pelo apoio incondicional aos meus objetivos.

AGRADECIMENTOS

Em primeiro lugar, gostaria de agradecer a Deus por me dá força e sabedoria. Minha mãe Ana Lucia que é a pessoa que mais me inspiro, devido a sua força de vontade e determinação para me dar sempre o melhor.

Agradeço também a minha esposa Tatiana por sempre acreditar em mim e apoiar meus objetivos de todas as formas.

Agradeço ao amigo Saulo por toda a sua disponibilidade e paciência em ajudar com a parte do desenvolvimento computacional, pois sem sua ajuda com certeza minha luta teria sido bem mais árdua.

Agradeço ao meu orientador Dr. Rodrigo Sampaio Lopes, a paciência em acreditar que eu conseguiria realizar essa pesquisa, pelo conhecimento compartilhado e pela orientação dada.

A Universidade Federal de Pernambuco – UFPE - CAA pela oportunidade de aprendizado e crescimento profissional em uma instituição de grande renome.

RESUMO

A Agência Nacional de Energia Elétrica é responsável pela fiscalização e regularização da qualidade de fornecimento de energia do sistema elétrico brasileiro. Sendo assim, a busca por técnicas de manutenções a fim de reduzir as possíveis falhas no sistema de distribuição tem sido comum pelas distribuidoras de energia elétrica. Adquirir conhecimentos mais profundos, sobre as causas primárias das falhas e identificar variáveis significativas relacionadas a essas causas, pode proporcionar tomada de decisões mais eficazes para restaurar o sistema de forma adequada e com eficiência. As mudanças na estrutura do setor elétrico levam à necessidade de um maior controle de qualidade de energia elétrica fornecida aos consumidores finais. Portanto, as concessionárias procuram encontrar soluções práticas destinadas a prevenir falhas específicas. Tendo em vista tais questões, o objetivo do presente estudo é compreender as causas de falhas de média tensão (MT) no sistema de distribuição de energia, na região sul do Ceará, usando a mineração de dados com regras de associação, para dar suporte a análise e gestão da manutenção. Logo, uma aplicação da metodologia proposta demonstra como a abordagem pode ser usada para minerar regras de associação para causas das falhas PEM. Com base nos resultados das regras de associação geradas os gestores de manutenção tomaram a decisão de intensificar as inspeções visuais em toda a rede *in loco* e posteriormente foram planejadas intervenções preventivas com a finalidade de atuar de forma antecipada à falha. Através da aplicação da metodologia proposta neste estudo, foi possível um direcionamento mais assertivo tanto para o planejamento quanto para a execução das ações de manutenções nas redes de distribuição. Analisando os resultados encontrados, o modelo proposto obteve desempenho melhor quando comparado com o desempenho da configuração atual, houve a redução dos indicadores de continuidade DEC e FEC de 8,54% e 6,26% respectivamente, tudo isso atrelado apenas a causa PEM vegetação.

Palavras-chave: sistema de distribuição elétrica; aprendizado de máquina; mineração de regras de associação; algoritmo apriori; análise de causa de falha.

ABSTRACT

The National Electric Energy Agency is responsible for inspecting and regulating the quality of energy supply in the Brazilian electrical system. Therefore, the search for maintenance techniques in order to reduce possible failures in the distribution system has been common by electricity distributors. Acquiring deeper knowledge about the root causes of failures and identifying significant variables related to these causes can provide more effective decision-making to restore the system properly and efficiently. Changes in the structure of the electricity sector lead to the need for greater quality control of electricity supplied to final consumers. Therefore, utilities seek to find practical solutions aimed at preventing specific failures. In view of these issues, the objective of the present study is to understand the causes of medium voltage (MV) failures in the power distribution system, in the southern region of *Ceará*, using data mining with association rules, to support maintenance analysis and management. Therefore, an application of the proposed methodology demonstrates how the approach can be used to study association rules for causes of PEM failures. Based on the results of the association rules generated, maintenance managers made the decision to intensify visual inspections across the entire network on site. Subsequently, preventive interventions were planned in order to act in advance of failure. Through the application of the methodology proposed in this study, a more assertive direction for planning was possible. In addition to guiding the execution of maintenance actions in the distribution networks. Analyzing the results found, the proposed model performed better when compared to the performance of the current configuration. There was a reduction in the continuity indicators DEC (Equivalent Duration of Interruption per Consumer Unit) and FEC (Equivalent Frequency of Interruption per Consumer Unit) of 8.54% and 6.26% respectively, all linked only to the PEM vegetation cause.

Keywords: electrical distribution systems; machine learning; association rule learning; apriori algorithm; root cause failure.

LISTA DE FIGURAS

Figura 1 -	Rede de colaboração por país.....	36
Figura 2 -	Proposta de um sistema de gestão da manutenção para sistemas de distribuição elétrica.....	43
Figura 3 -	Mapeamento da inspeção visual na rede elétrica.....	44
Figura 4 -	Modelo do planeamento de anomalias.....	44
Figura 5 -	Delimitação da área de estudo.....	46
Figura 6 -	Visualização das regras de associação no plano de dispersão.....	52
Figura 7 -	15 principais regras de associação com relação à sua confiança.....	53
Figura 8 -	Rede de nós de associação para falhas relacionadas à desligamento emergencial.....	58
Figura 9 -	Rede de nós de associação para falhas relacionadas à vegetação.....	58
Figura 10 -	Rede de nós de associação para falhas relacionadas à defeito em chave fusível.....	59
Figura 11 -	Mapeamento das inspeções visuais na rede elétrica.....	60
Figura 12 -	Planejamento de anomalias de uma equipe referente ao mês de março/2020.....	60

LISTA DE FLUXOGRAMAS

Fluxograma 1 -	Algoritmo Apriori.....	26
Fluxograma 2 -	Processo de revisão sistemática da literatura.....	28
Fluxograma 3 -	Etapas do processo de filtragem da pesquisa.....	31
Fluxograma 4 -	Processo metodológico usado no estudo.....	40

LISTA DE GRÁFICOS

Gráfico 1 -	Número de artigos avaliados por ano de publicação.....	32
Gráfico 2 -	Número de citações de artigos avaliados por ano de citação.....	33
Gráfico 3 -	Periódicos mais relevantes.....	34
Gráfico 4 -	Produção por país.....	35
Gráfico 5 -	Autores e sua produção ao longo do tempo.....	37
Gráfico 6 -	Produtividade científica.....	37
Gráfico 7 -	Quantidade de falhas por mês.....	48
Gráfico 8 -	Top 10 causas mais recorrentes.....	49
Gráfico 9 -	Top 10 equipamentos que mais falharam.....	49
Gráfico 10 -	Comparação de desempenho DEC.....	61
Gráfico 11 -	Comparação de desempenho FEC.....	62

LISTA DE QUADROS

Quadro 1 - Taxonomia simples dos métodos de aprendizado máquina.....	23
Quadro 2 - Termos de busca utilizados.....	30
Quadro 3 - Amostra de dados tratados.....	42
Quadro 4 - Conjuntos de transações gerados a partir das causas de falhas.....	42
Quadro 5 - Explicação de alguns atributos do conjunto de dados.....	47
Quadro 6 - Lista de regras fortes de associação para dados transacionais de 6 horas.....	51

LISTA DE TABELAS

Tabela 1 -	Distribuição das áreas de pesquisa dos artigos avaliados.....	33
Tabela 2 -	Distribuição do comprimento das regras de associação.....	50
Tabela 3 -	Distribuição do comprimento das regras de associação.....	53
Tabela 4 -	Resultados da mineração de regras de associação para falhas relacionadas à desligamento emergencial.....	56
Tabela 5 -	Resultados da mineração de regras de associação para falhas relacionadas à vegetação.....	56
Tabela 6 -	Resultados da mineração de regras de associação para falhas relacionadas à defeito em chave fusível.....	57

LISTA DE SIGLAS

ABRADEE	Associação Brasileira de Distribuidores de Energia Elétrica
ANEEL	Agencia Nacional de Energia Elétrica
DEC	Duração Equivalente de Interrupção por Unidade Consumidora
DNAEE	Departamento Nacional de Águas e Energia Elétrica
FEC	Frequência Equivalente de Interrupção por Unidade Consumidora
FUNCEME	Fundação Cearense de Meteorologia e Recursos Hídricos
IBGE	Instituto Brasileiro de Geografia e Estatística
IPECE	Instituto de Pesquisa e Estratégia Econômica do Ceará
ISI	Institute for Scientific Information
IT	Instrução de Trabalho
MT	Média tensão
NPEM	Não são Possíveis de Evitar com Manutenção
NPL	Processamento de Linguagem Natural
OS	Ordem de Serviço
PEM	Possíveis de Evitar com Manutenção
SGD	Sistema de Gestão de Desligamento

SUMÁRIO

1	INTRODUÇÃO.....	16
1.1	JUSTIFICATIVA.....	17
1.2	OBJETIVOS.....	18
1.2.1	Objetivo geral.....	18
1.2.2	Objetivos específicos.....	18
1.3	ESTRUTURA DO TRABALHO.....	18
2	REFERENCIAL TEÓRICO.....	20
2.1	VISÃO GERAL DA ANÁLISE DE FALHAS NO SISTEMA DE DISTRIBUIÇÃO ELÉTRICA.....	20
2.2	INDICADORES DE CONTINUIDADE.....	21
2.3	APRENDIZADO DE MÁQUINA.....	22
2.3.1	Mineração de regras de associação.....	23
2.3.1.1	Algoritmo apriori.....	25
3	REVISÃO SISTEMÁTICA DA LITERATURA.....	27
3.1	MÉTODO.....	27
3.2	PLANEJAMENTO DA REVISÃO.....	28
3.3	CONDUÇÃO DA REVISÃO.....	29
3.4	DOCUMENTAÇÃO DA REVISÃO.....	31
4	METODOLOGIA.....	39
4.1	CLASSIFICAÇÃO DA PESQUISA EM ENGENHARIA DE PRODUÇÃO.....	39
4.2	DELINEAMENTO COM AS ETAPAS DA PESQUISA	39
4.2.1	Delimitação da área de estudo.....	40
4.2.2	Estudo da literatura.....	41
4.2.3	Coleta, tratamento e padronização dos dados.....	41
4.2.4	Mineração de regras de associação.....	42
4.2.5	Gestão da manutenção para o sistema de distribuição elétrica.....	43
5	APLICAÇÃO DA METODOLOGIA PROPOSTA.....	46
5.1	LOCAL DE ESTUDO.....	46
5.2	BASE DE DADOS.....	47

5.3	DADOS INICIAIS.....	47
5.4	REGRAS DE ASSOCIAÇÃO PARA CAUSAS DAS FALHAS.	50
5.5	REGRAS DE ASSOCIAÇÃO PARA CAUSAS DAS FALHAS PEM.....	54
5.6	GESTÃO DA MANUTENÇÃO PARA O SISTEMA DE DISTRIBUIÇÃO ELÉTRICA.....	59
6	CONSIDERAÇÕES FINAIS.....	63
6.1	CONTRIBUIÇÕES DA PESQUISA.....	63
6.2	TRABALHOS FUTUROS.....	64
	REFERÊNCIAS.....	65

1 INTRODUÇÃO

A agência Nacional de Energia Elétrica (ANEEL) é responsável pela fiscalização e regularização da qualidade de fornecimento de energia do sistema elétrico brasileiro. Sendo assim, a busca por técnicas de manutenções a fim de reduzir as possíveis falhas no sistema de distribuição tem sido recorrente pelas distribuidoras de energia elétrica.

Falhas nos sistemas de distribuição de energia influenciam negativamente na confiabilidade do sistema, resulta em altos custos para as concessionárias de distribuição, além de proporcionar experiências negativas para os clientes (LANDEGREN *et al.*, 2016). Adquirir conhecimentos mais profundos, sobre as causas primárias das falhas e identificar variáveis significativas relacionadas a essas causas, pode proporcionar tomada de decisões mais eficazes para restaurar o sistema de forma adequada e com eficiência.

Perturbações no fornecimento ou desestabilizações de todo o sistema pode ser causada por qualquer mau funcionamento de componentes no sistema de energia. De acordo com a Associação Brasileira de Distribuidores de Energia Elétrica (ABRADEE) a parte que contém maior grau de fiscalização e regularização é a fase da distribuição (ABRADEE, 2021).

Para a agência distribuidora de energia elétrica, as falhas podem ser divididas em falhas PEM, que são as causas possíveis de evitar com manutenção e NPEM, as causas que não são possíveis evitar com manutenção (ENEL, 2021). Uma ação a ser tomada ao analisar as causas PEM é levar como base o conhecimento adquirido das falhas que ocorreram no passado. O interesse das concessionárias de distribuição de energia geralmente está em prevenir falhas evitáveis.

Com base no histórico de ocorrências de um sistema de distribuição de energia, é possível através da adoção de metodologia e estudos descobrir os fatores associados que contribuem para a falha na rede de distribuição elétrica. Tendo em vista, que para ter a qualidade no atendimento das metas definidas pelo órgão regulador, é necessário saber quando atuar de forma assertiva, para que interrupções sejam evitadas.

Em estudos recentes, as principais causas para as falhas nos sistemas elétricos estão relacionadas à vegetação, animais e efeitos climáticos, assim como descarga atmosférica (XU e MO, 2006; XU ET AL., 2007; SILVA e SARAEE; 2019).

Devido à natureza aleatória das falhas e os inúmeros fatores contribuintes, prevenir ou identificar falhas são atividades extremamente desafiadoras. Por motivo da quantidade de dados

gerados e armazenados, o processamento de dados usando métodos tradicionais tornou-se mais difícil e complexo. Doostan e Chowdhury (2017) mostram que nos últimos anos as tarefas de previsão e identificação de falhas tem sido facilitada pela aplicação de técnicas avançadas de análise de dados combinadas com modelos de decisão.

De acordo Rezig *et al.*, (2018) a mineração de dados tem como objetivo principal extrair de uma grande quantidade de dados as informações com maior relevância e, desta forma, com base em critérios fixos criar modelos de informações e conhecimentos.

Deste modo, estudos têm sido desenvolvidos a fim de analisar características de várias falhas, utilizando as técnicas estatísticas e o uso de algoritmos para analisar o banco de dados (DOOSTAN e CHOWDHURY, 2017; RAVI et al, 2019; SILVA e SARAEE; 2019). Neste trabalho buscamos contribuir para literatura utilizada técnicas de mineração de dados e propondo uma análise para o sistema de distribuição de energia elétrica.

1.1 JUSTIFICATIVA

As mudanças na estrutura do setor elétrico levam à necessidade de um maior controle de qualidade de energia elétrica fornecida aos consumidores finais. As distribuidoras, portanto, trabalham em ambientes regulados necessitando justificar investimentos e tarifas, e ainda, obter eficiência econômica reduzindo custos de investimento. Portanto, as concessionárias procuram encontrar soluções práticas destinadas a prevenir falhas específicas ou tentam tomar medidas eficazes para restaurar o sistema de forma adequada e rápida após a ocorrência de falhas.

Tanto do ponto de vista da distribuidora em estudo quanto do consumidor, a compreensão profunda da causa da falha na rede de distribuição elétrica é de suma importância. O estudo de mineração de dados proposto tem por objetivo, melhorar o nível de disponibilidade do sistema, identificando correlação nas falhas relacionadas a cada causa PEM e reduzindo as falhas da rede. Além disso, há uma necessidade comercial de reduzir os gastos operacionais nos departamentos de engenharia. Portanto, esse estudo pode ajudar reduzir no número de reclamações devido a menos falhas na rede e prevenção de custos e multas em potencial associadas a falhas futuras na rede por parte dos reguladores.

1.2 OBJETIVOS

1.2.1 Objetivo geral

O presente trabalho tem como objetivo geral compreender as causas de falhas de média tensão (MT) no sistema de distribuição de energia, usando a mineração de dados com regras de associação, para dar suporte a análise e gestão da manutenção.

No próximo tópico serão abordados os objetivos específicos necessários para o alcance deste objetivo geral definido.

1.2.2 Objetivos específicos

Nesse contexto, o presente trabalho tem como objetivos específicos:

- Realizar uma pesquisa exploratória sobre mineração de regras de associação e sobre os temas relacionados a sua problemática;
- Explorar a possibilidade de aprimorar o ganho de conhecimento das regras de associação geradas;
- Explorar a aplicabilidade das regras de associação para causas das falhas PEM;
- Desenvolver e explorar a aplicabilidade do modelo de gestão da manutenção proposto na área de estudo;
- Realizar análises e discussões dos resultados obtidos.

1.3 ESTRUTURA DO TRABALHO

Este trabalho está dividido em seis capítulos: introdução, referencial teórico, revisão sistemática da literatura, metodologia, estudo de caso e considerações finais. Cada um dos capítulos cumpre um determinado objetivo, a saber:

- Capítulo 1: são apresentados os aspectos iniciais da pesquisa, contendo uma introdução que sintetiza o tema a ser explorado e os objetivos geral e específicos deste trabalho;
- Capítulo 2: é apresentado o referencial teórico, no qual se aborda os principais conceitos para o entendimento do estudo.

- Capítulo 3: é apresentada uma revisão da literatura sobre mineração de regras de associação aplicada a manutenção de sistemas de distribuição elétrica
- Capítulo 4: é apresentado a metodologia geral utilizada no estudo desta dissertação, ou seja, a diretriz seguida desde os estudos iniciais até a apresentação dos resultados. Define-se também, a caracterização e o delineamento da pesquisa de forma que o leitor possa entender melhor como foi realizada a construção desta pesquisa.
- Capítulo 5: é apresentado a aplicação da metodologia proposta em um caso real. Analisam-se e apresentam-se os resultados obtidos e compara-se com os resultados vivenciados na região nos dias atuais.
- Capítulo 6: são apresentadas as considerações finais da pesquisa, além da perspectiva para o desenvolvimento de trabalhos futuros.

Por fim, são apresentadas as referências utilizadas para a elaboração desta dissertação.

2 REFERENCIAL TEÓRICO

Nesse capítulo é apresentado os conceitos essenciais para o desenvolvimento desta dissertação. No primeiro momento é exposto uma visão geral da análise de falhas em sistemas de distribuição de energia. Em seguida será exibido os conceitos sobre aprendizado de máquina, mineração de regras de associação e suas aplicações.

2.1 VISÃO GERAL DA ANÁLISE DE FALHAS NO SISTEMA DE DISTRIBUIÇÃO ELÉTRICA

Atualmente, os sistemas de distribuição elétrica estão em constante crescimento, em consequência do aumento dos centros populacionais e industriais. Este crescimento é um desafio para as concessionárias, que devem fornecer energia elétrica aos seus consumidores de maneira sustentável, confiável e com qualidade. (SOUZA, 2008).

De acordo com Xu (2006), os sistemas de distribuição elétrica estão geograficamente dispersos e em vários ambientes operacionais dinâmicos, podem ser significativamente afetados por várias falhas em termos de confiabilidade de energia e qualidade dos serviços.

Segundo Lout (2015) os componentes da rede de distribuição elétrica são vulneráveis a falhas frequentes, que podem ocorrer em componentes ou subcomponentes. Já para Silva (2019), os sistemas de distribuição elétrica são propensos a falhas, que podem ocorrer em qualquer uma das unidades geradoras, transformadores, meios de distribuição de energia, como cabos aéreos e subterrâneos.

Todos os sistemas de distribuição elétrica estão sujeitos a problemas e falhas em sua operação (LIMA, 2013). Os sistemas de distribuição, por apresentar uma topologia predominantemente radial com inúmeras ramificações, está mais susceptível a falhas provocadas de forma direta ou indireta. As falhas podem ser causadas por animais que entram em contato com a rede energizada, queimadas, tempestades, pela ação do homem, vandalismo, por problemas elétricos, queima/explosão de transformadores, excesso de carga e subdimensionamento de componentes (TONELLI NETO, 2015).

Para resolver adequadamente os problemas causados por interrupções nos sistemas de distribuição elétrica, é necessária a identificação da causa. Adquirir conhecimento da causa da interrupção imediatamente após a falha é extremamente benéfico em termos de redução da

duração da interrupção, pois dá à equipe de operação e manutenção a chance de encontrar a evidência da falha mais rapidamente e permite que as concessionárias enviem a equipe certa com o equipamento específico necessário para reparar ou substituir os componentes danificados (DOOSTAN e CHOWDHURY, 2017).

Vários estudos foram realizados para identificar as principais causas de falhas nos sistemas de distribuição elétrica. Os autores Xu e Mo (2006), consideram seis características como fatores mais influentes para falhas de vegetação e animais. Além disso, os autores Xu *et al.*, (2007), desenvolveram um classificador para identificar três causas de interrupções: vegetação, animais e descarga atmosférica. Segundo os autores Silva e Saraee (2019), as falhas que geralmente ocorrem nos sistemas de distribuição elétrica são causadas predominantemente por vegetação, animais e efeitos climáticos.

Conseqüentemente, caracterizar falhas de acordo com suas causas e identificar variáveis significativas que impactam fortemente a frequência de falha são extremamente relevantes, pois permitem encontrar soluções para restringir causas específicas e dar uma resposta apropriada para falhas inevitáveis (DOOSTAN e CHOWDHURY, 2017).

2.2 INDICADORES DE CONTINUIDADE

Para as concessionárias, licenciadas de serviço público de distribuição e de transmissão de energia elétrica, é importante manter a continuidade no serviço de distribuição de energia elétrica, pois existem metas estabelecidas pela ANEEL a serem cumpridas nos enfoques de Duração Equivalente de Interrupção por Unidade Consumidora DEC e Frequência Equivalente de Interrupção por Unidade Consumidora FEC, para conjuntos de unidades consumidoras (CUNHA, 2011).

A Regulação da qualidade de energia no Brasil foi iniciada através da Portaria 46/1978 em 17 de abril de 1978, pelo já extinto Departamento Nacional de Águas e Energia Elétrica (DNAEE), onde ficaram estabelecidos os indicadores de continuidade do fornecimento de energia elétrica a serem observados pelas concessionárias de serviços públicos para o Brasil. Dessa forma, a Portaria determinou que os indicadores DEC e FEC deveriam ser apurados para cada conjunto da unidade consumidora, considerando as interrupções com duração superior a 3 minutos (BERNARDO, 2013).

Conjunto de unidades consumidoras é um grupo de unidades consumidoras da mesma área de concessão definidos pela concessionária. Podendo abranger mais de um município ou dividir o município em mais de um conjunto (ANEEL, 2021).

Para os indicadores citados, DEC representa o tempo que, em média, no período de observação, cada unidade consumidora ficou sem energia elétrica e FEC é número de interrupções ocorridas, em média, no período de observação (ANEEL, 2021).

Com as crescentes fiscalizações da ANEEL, bem como a criticidade dos clientes atendidos pelas distribuidoras, se faz necessária uma verificação mais apurada sobre qual valor de indicador DEC e FEC seria ideal para a empresa manter sem ter uma perda em seu investimento (RAMOS, 2021).

2.3 APRENDIZADO DE MÁQUINA

Um dos primeiros conceitos sobre Aprendizado de Máquina foi publicado por Samuel (1959). De acordo com o autor, aprendizado de máquina é o processo de ensinar máquinas através da experiência. Segundo Shalev-Shwartz e Bem-David (2013), o aprendizado de máquina é o processo de converter experiência em expertise ou conhecimento. Ainda segundo os autores, aprendizado de máquina tem como objetivo programar computadores de modo que eles aprendam através de *inputs*.

Com o aumento da disponibilidade de grandes volumes de dados em todas as áreas do conhecimento humano provocou-se uma demanda pelo entendimento de algoritmos de aprendizado de máquina (JORDAN e MITCHELL, 2015). Ainda segundo Jordan e Mitchell (2015), a área de estudo desse tópico está concentrada entre ciência da computação, estatística e uma variedade de outras disciplinas.

Os algoritmos de aprendizado de máquina podem ser classificados segundo o tipo de aprendizagem. Aprendizado supervisionado utiliza dados históricos com variáveis repostas previamente identificadas. Já no aprendizado não supervisionado os dados não possuem variável resposta, o aprendizado ocorre com base em medidas de similaridade ou distância entre as observações. (RAMASUBRAMANIAN e SINGH, 2017). O Quadro 1 mostra uma taxonomia simples dos métodos de aprendizado e os algoritmos mais populares.

Quadro 1 - Taxonomia simples dos métodos de aprendizado máquina

Método de Aprendizagem	Algoritmos populares
Supervisionado	<i>Classification and Regression Trees, ANN, SVM, Genetic Algorithms, Decision Trees, ANN/MLP, Rough Sets, Genetic Algorithms, Linear/Nonlinear Regression.</i>
Não supervisionado	<i>OneR, ZeroR, Eclat, Expectation Maximization Apriori Algorithm, Graph-Based Matching, Apriori Algorithm, FP-Growth technique, K-means, ANN/SOM, Expectation Maximization (EM).</i>

Fonte: O Autor (2021)

Nota: Adaptado de Turban *et al.*, (2005)

Como pudemos identificar no Quadro 1, existe grande quantidade de algoritmos de aprendizado de máquina, do qual será abordado apenas um deles na presente pesquisa.

2.3.1 Mineração de regras de associação

Para Li *et al.*, (2015), as regras de associação tentam descrever as relações potenciais entre itens em um banco de dados. Segundo Sheng *et al.*, (2018), o objetivo da mineração de regras de associação é encontrar o relacionamento de associação entre diferentes termos em um banco de dados. De acordo com Silva e Saraee (2019), é uma técnica de aprendizado não supervisionado, usada para descobrir a associação de itens.

A mineração de regras de associação é uma técnica usada para identificar a interdependência desconhecida dos dados e descobrir as regras entre esses itens (LIN *et al.*, 2019). Ainda segundo os autores, compreende em encontrar as regras que podem conter associações e objetos causais (antecedentes e consequentes) entre conjuntos de itens.

Regras de associação é uma técnica de mineração de dados bem reconhecida aplicada para identificar relacionamentos ou associações em uma grande quantidade de dados. A abordagem da mineração de regras de associação tem como objetivo identificar regras robustas entre associações incorporadas em dados dos quais diferentes medidas de interesse são usadas, sendo as mais comuns as de suporte, confiança e elevação (CHEMWENO *et al.*, 2016).

Segundo Wang *et al.*, (2020), as regras de associação são expressas como $A \Rightarrow B$, com duas condições, onde $A, B \subseteq I$ e $A \cap B = \emptyset$, sendo que $I = (I_1, I_2, \dots, I_m)$ é um conjunto de itens e T um conjunto de transações, onde $T \subseteq I$, e o conjunto de transações D é uma coleção de todas as transações possíveis no banco de dados. Em palavras simples, se o conjunto antecedente A ocorre, então o conjunto consequente B também é provável que ocorra.

O suporte e a confiança podem ser matematicamente expressos nas Equações (1) e (2), respectivamente (WANG *et al.*, 2020).

$$\text{Suporte } (A \Rightarrow B) = P(A \cup B) \quad (1)$$

$$\text{Confiança } (A \Rightarrow B) = P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (2)$$

Para Chemweno *et al.*, (2016), a medida de suporte indica a proporção de transações nos dados que contêm os conjuntos de itens antecedentes A e consequentes B , e a medida de confiança avalia a probabilidade condicional em que as transações percentuais nos dados que contêm os conjuntos de itens antecedentes A e também contêm os conjuntos de itens consequentes B são medidos. Enquanto que Liu *et al.*, (2018), considera o suporte como a probabilidade de um determinado item aparecer em uma transação e a confiança pode ser interpretada como a estimativa da probabilidade condicional.

De acordo com Doostan e Chowdhury (2017), suporte e confiança são as principais métricas para avaliar a qualidade das regras geradas pelo algoritmo, pois indicam a significância estatística e a força de uma regra, respectivamente. Além disso, há outras métricas, uma das quais é o valor de elevação de uma regra. O valor de elevação pode ser expresso na Equação (3).

$$\text{Elevação } (A \Rightarrow B) = \frac{\text{Suporte } (A \Rightarrow B)}{\text{Suporte } (B) \times \text{Suporte } (A)} \quad (3)$$

O valor de elevação indica a força de uma regra sobre a coocorrência aleatória do antecedente e do consequente (FORGHANI, 2014). Já para Moharana *et al.*, (2019), o valor de elevação mede até que ponto A e B são independentes um do outro. Segundo Sun *et al.*, (2019), quando o valor de elevação de uma regra é um, a probabilidade de ocorrência do antecedente e da classe alvo são consideradas independentes uma da outra, ou seja, não existe associação. Se a elevação for maior do que um, será confirmado que essas duas ocorrências são dependentes uma da outra, tornando-se esta regra mais útil para futuras previsões.

Além do mais, vários algoritmos para mineração de regras de associação são propostos na literatura, dos quais os diferentes algoritmos incorporam diferentes medidas de interesse e, como resultado, podem gerar regras de associação variadas.

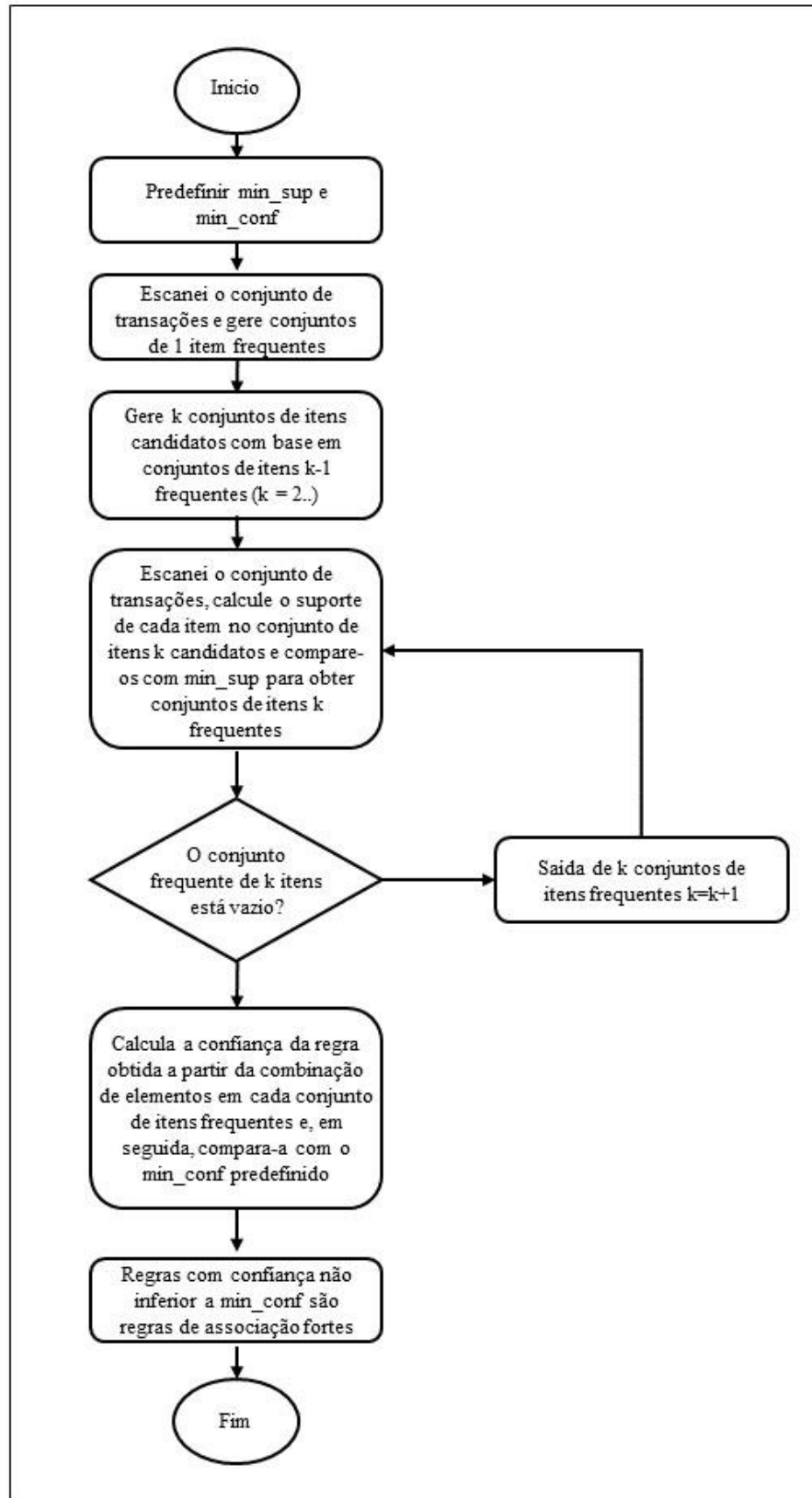
2.3.1.1 Algoritmo apriori

O algoritmo apriori foi proposto pela primeira vez por Agrawal e Srinikant (1994), com o objetivo de descobrir regras de associação entre itens em um grande banco de dados de transações de vendas. Segundo Yu *et al.*, (2011) o algoritmo Apriori é um dos algoritmos mais usados na mineração de conjuntos de itens frequentes de regras de associação. Já segundo Wang *et al.*, (2020) a ideia básica do algoritmo Apriori é encontrar o maior conjunto de itens processando itens em várias etapas.

O procedimento de mineração de dados baseado no algoritmo Apriori inclui duas etapas: encontrar conjuntos de itens frequentes e gerar regras de associação fortes com base nos conjuntos de itens frequentes (XINCHUN *et al.*, 2018)

Segundo Tian *et al.*, (2020) o algoritmo Apriori extrai subconjuntos de itens que podem representar as regras dos conjuntos de dados através da iteração camada por camada. A implementação básica do método é descobrir o conjunto de itens frequente de acordo com a combinação de todos os itens diferentes e, em seguida, fornece as regras de associação. O algoritmo encerra quando o conjunto de itens frequentes ficar vazio, o Fluxograma 1 correspondente é mostrado abaixo.

Fluxograma 1 - Algoritmo Apriori



Fonte: O Autor (2021)

Nota: Adaptado de Tian *et al.*, (2020)

3 REVISÃO SISTEMÁTICA DA LITERATURA

Este capítulo apresenta a revisão sistemática da literatura, que tem como objetivo identificar trabalhos científicos publicados em periódicos que abranjam o tema dessa pesquisa. Após a busca, todos os trabalhos foram analisados com o intuito de responder as questões de pesquisa previamente definidas.

3.1 MÉTODO

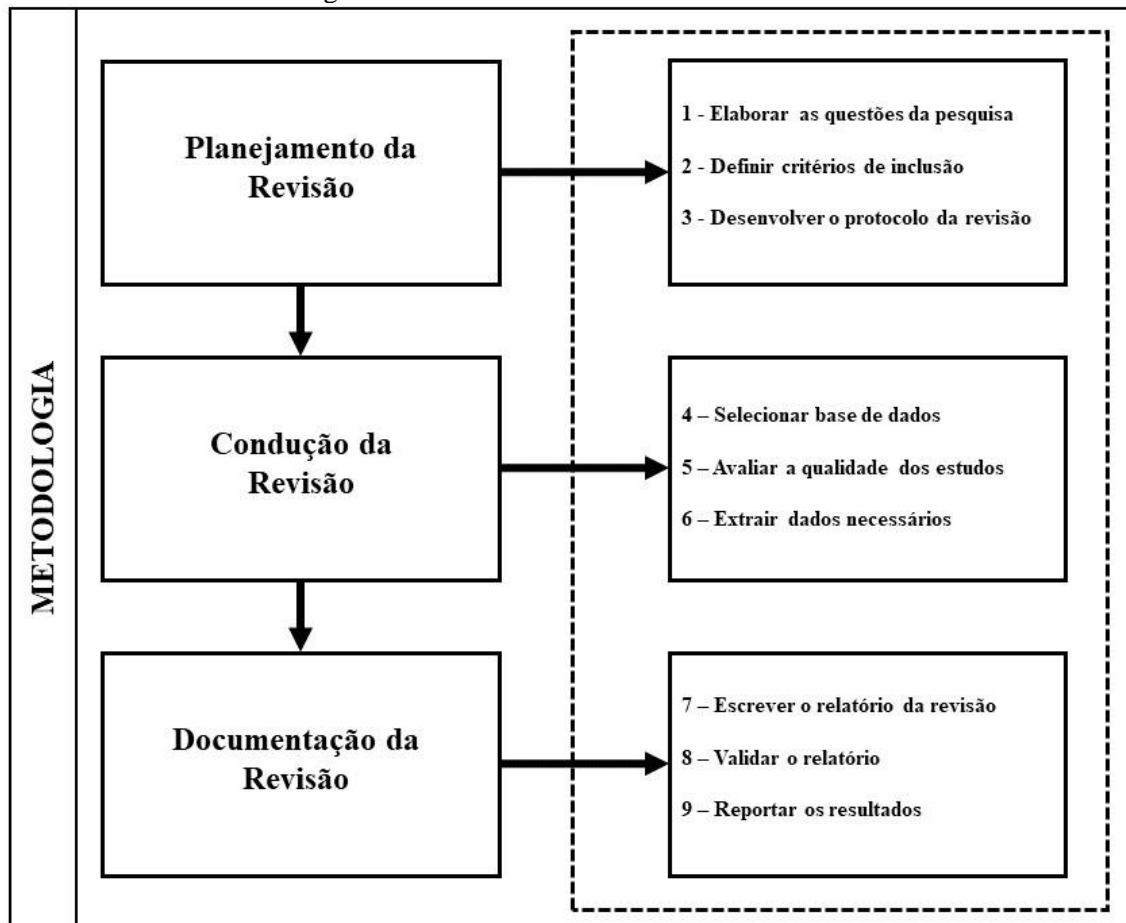
Para o desenvolvimento dessa seção, foi realizado uma revisão sistemática da literatura com base nas diretrizes propostas por Brereton *et al.* (2007). A abordagem de Brereton *et al.* (2007), sugere a realização de uma revisão sistemática da literatura envolvendo várias atividades discretas, que podem ser agrupadas em três fases principais: planejamento da revisão; condução da revisão; e documentação da revisão. Método esse, que é amplamente abordado na literatura quando se deseja realizar uma revisão sistemática da literatura mais estruturada.

Uma revisão sistemática da literatura é um meio de avaliar e interpretar todas as pesquisas disponíveis relevantes para uma questão de pesquisa específica, uma área de estudo ou um fenômeno de interesse (KITCHENHAM; CHARTERS, 2007). Gupta *et al.* (2018) afirmam que para uma revisão ser sistemática, deve-se responder a questões específicas e aplicar uma metodologia clara para a avaliação de todas as informações disponíveis.

A metodologia utilizada para essa pesquisa tomou como base o modelo de Brereton *et al.* (2007), como mostra o Fluxograma 1.

Apesar da forma sequencial na qual a metodologia adotada foi apresentada, vale ressaltar que o processo é interativo e podem ocorrer interações fora da ordem pré-estabelecida. Nas próximas subseções, cada uma das etapas apresentadas no Fluxograma 2 serão detalhadas.

Fluxograma 2 - Processo de revisão sistemática da literatura



Fonte: O Autor (2021)

Nota: Adaptado Brereton *et al.*, (2007)

3.2 PLANEJAMENTO DA REVISÃO

Para iniciar uma revisão sistemática da literatura, o escopo da problemática da pesquisa deve ser bem definido para assim guiar os primeiros passos. Kitchenham e Charters, (2007) sugerem que as questões da pesquisa devam ser elaboradas informando a população a ser estudada, o meio de intervenção, a forma de comparação e as saídas a serem analisadas. Para essa pesquisa, a população a ser estudada é a técnica de aprendizado de máquina, regra de associação, a intervenção é o campo de estudo referente a manutenção em sistemas de distribuição elétrica. A forma de comparação será entre os artigos selecionados para o estudo e a saída, será definida em cada questão de pesquisa.

Assim sendo, este estudo busca responder às seguintes questões de pesquisa:

- **RQ1:** Como a técnica regra de associação aplicada a manutenção de sistemas de distribuição elétrica evoluiu em termos de número de artigos publicados e citações?

- **RQ2:** Em quais áreas de pesquisa a técnica regra de associação é predominantemente aplicada na manutenção de sistemas de distribuição elétrica?
- **RQ3:** Quais são os principais periódicos onde encontra-se artigos sobre a técnica regra de associação aplicada a manutenção de sistemas de distribuição elétrica?
- **RQ4:** Quais são os países que estão à frente nas pesquisas relacionadas a técnica regra de associação aplicada a manutenção de sistemas de distribuição elétrica?
- **RQ5:** Quais autores com maior número de publicações e citações relacionada a técnica regra de associação aplicada a manutenção de sistemas de distribuição elétrica?

Essas questões, serão respondidas durante a análise e elaboração do relatório de revisão na terceira fase.

Segundo Gupta *et al.* (2018), os critérios de inclusão para as revisões sistemáticas devem claramente definir os tipos de estudos, participantes, testes independentes, condições-alvo e padrões de referência a serem incluídos na revisão. Os critérios de inclusão adotados nessa pesquisa são: somente artigos publicados em periódicos; somente artigos escritos em língua inglesa; e somente artigos publicados até 31/12/2020.

Para finalizar a primeira fase da metodologia, foi definido o protocolo de pesquisa visando minimizar qualquer viés no estudo. Para Gupta *et al.* (2018), um protocolo de revisão bem elaborado facilita o gerenciamento da revisão, além de evitar redundâncias e artigos irrelevantes.

3.3 CONDUÇÃO DA REVISÃO

Após a finalização do planejamento da revisão é introduzido a condução da revisão, onde, nessa fase é definido uma plataforma para se explorar as publicações sobre o tema regra de associação aplicada a manutenção de sistemas de distribuição elétrica. A base de dados escolhida para esta pesquisa foi *Web of Science core Collection*, que é uma plataforma bastante utilizada no meio acadêmico, possuindo mais de vinte mil periódicos disponíveis, todas as referências são indexadas e fundamentada em assinaturas que são originalmente produzidas pelo *Institute for Scientific Information (ISI)* depois alimentado pela *Clarivate Analytics*. Para realizar a busca, primeiro foram definidos os termos de busca e por fim, a busca inicial foi realizada no mecanismo disponibilizado.

Para que a pesquisa tivesse um embasamento forte e de qualidade relevante, foi definido que apenas seriam escolhidos os artigos completos e publicados em repositórios de grande impacto para área acadêmica. Dessa forma, a revisão sistemática da literatura iria possibilitar uma ampla visão dos principais pontos que estão sendo trabalhados pelos principais autores.

Foram definidos dois grupos de termos de busca baseados na literatura específica, que podem ser visualizados no Quadro 2. O primeiro, foi utilizado para especificar o tema principal a ser explorado. O segundo, referente a termos geralmente utilizados para representar a manutenção em sistemas de distribuição elétrica.

Quadro 2 - Termos de busca utilizados

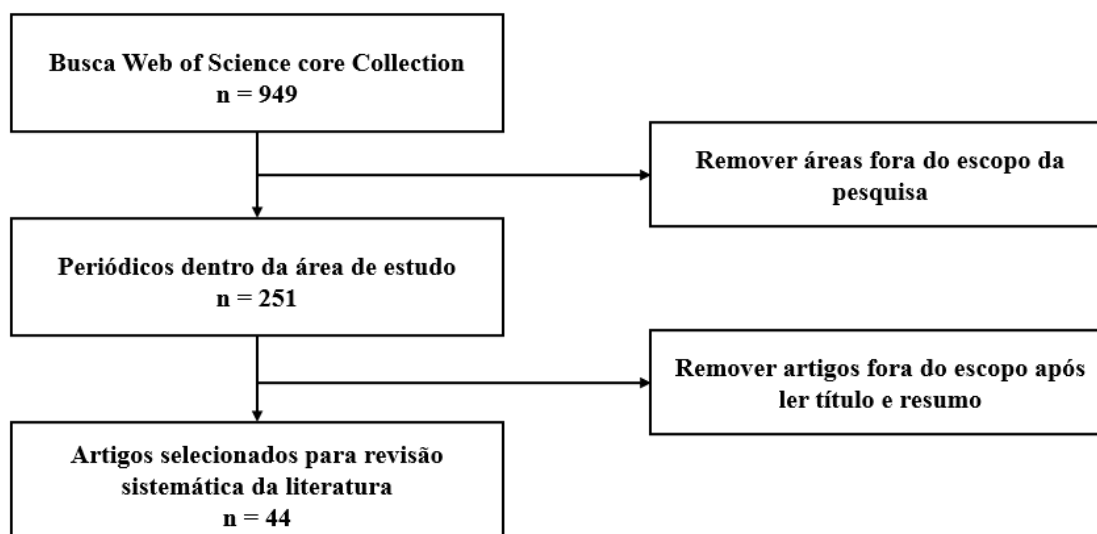
Palavras-Chaves E Operadores	
Grupo 1	<i>“Mining Association rules”</i> OR <i>“Association rules”</i> OR <i>“Apriori algorithm”</i>
Grupo 2	<i>“MT Faults”</i> OR <i>“Predictive maintenance”</i> OR <i>“Failure analysis”</i> OR <i>“Maintenance data”</i> OR <i>“Maintenance”</i> OR <i>“Causal mapping”</i> OR <i>“Fault correlation”</i> OR <i>“Maintenance decision-making”</i> OR <i>“Electricity Distribution Network”</i> OR <i>“Distribution System”</i> OR <i>“Distribution network”</i> OR <i>“Transmission System”</i>

Fonte: O Autor (2021)

Uma vez definido os termos de busca, serão usados os operadores booleanos OR (ou) e AND (e), para construção das estratégias avançadas de busca, onde OR equivale à união e AND equivale à intersecção. Todas as buscas foram realizadas nos campos título, resumo, palavras chaves e palavras chaves mais citadas.

Enfim, as buscas foram realizadas, resultando em 949 trabalhos. Após essa etapa, foram aplicados todos os critérios estabelecidos, sendo possível definir a amostra utilizada no estudo. Todas as etapas do processo de filtragem da pesquisa estão expostas no Fluxograma 3.

Fluxograma 3 - Etapas do processo de filtragem da pesquisa



Fonte: O Autor (2021)

Logo, alguns filtros foram usados para refinar a seleção do conjunto final de artigos, chegando em um número final de 44 artigos.

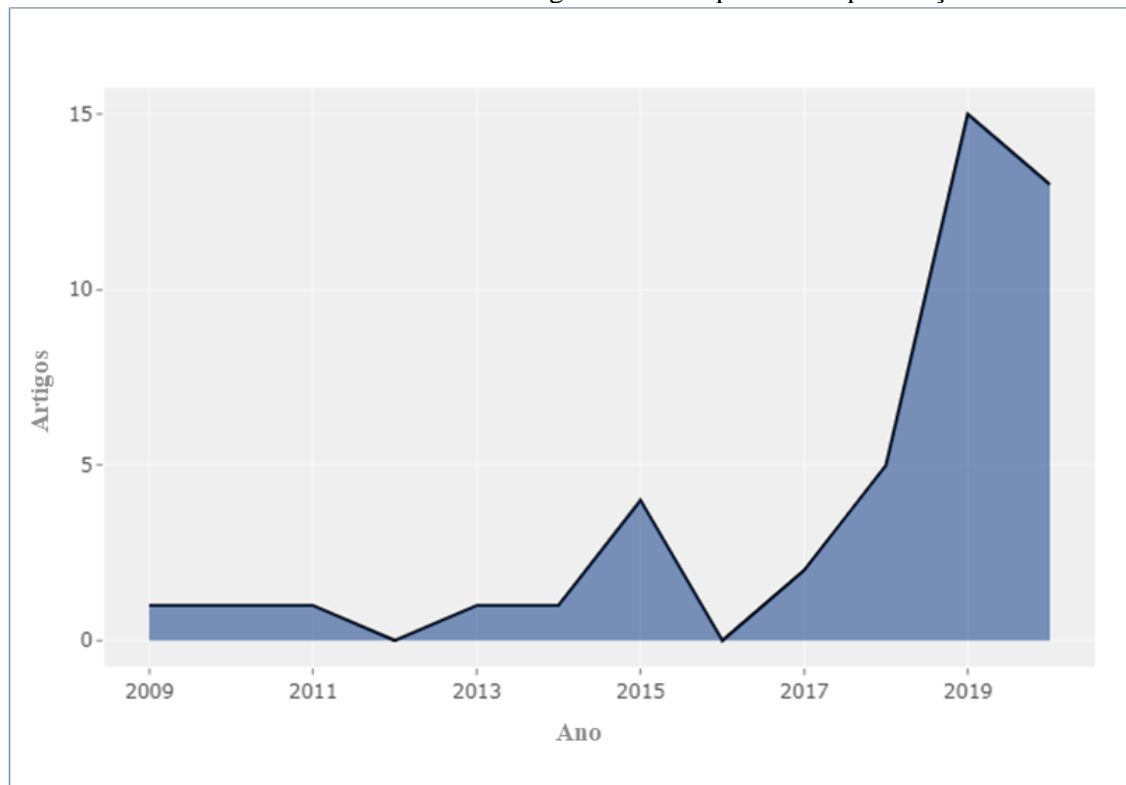
A extração dos dados teve o auxílio do pacote desenvolvido na linguagem R: Bibliometrix Aria e Cuccurullo (2017). Tal pacote fornece uma vasta quantidade de informações sobre os estudos preliminares de forma rápida e confiável.

3.4 DOCUMENTAÇÃO DA REVISÃO

A presente seção irá apresentar o relatório da revisão, validação do relatório e por fim será reportados os resultados. Um resumo da literatura sobre a técnica regra de associação aplicada a manutenção de sistemas de distribuição elétrica publicado de 2009 a 2020 é apresentada. Uma amostra de 44 artigos foi avaliada usando uma revisão sistemática da literatura. O objetivo desta análise é representar o estado atual da pesquisa e identificar as direções da pesquisa sobre esse tópico.

O número de publicações apresentou uma taxa de crescimento anual de 32,98% mostrando a relevância e atualidade do tema, como é mostrado na Gráfico 1, fato ocorrido provavelmente pela larga popularização das técnicas de aprendizado de máquina, bem como o surgimento de tecnologia que possibilitaram a análise de grandes volumes de dados.

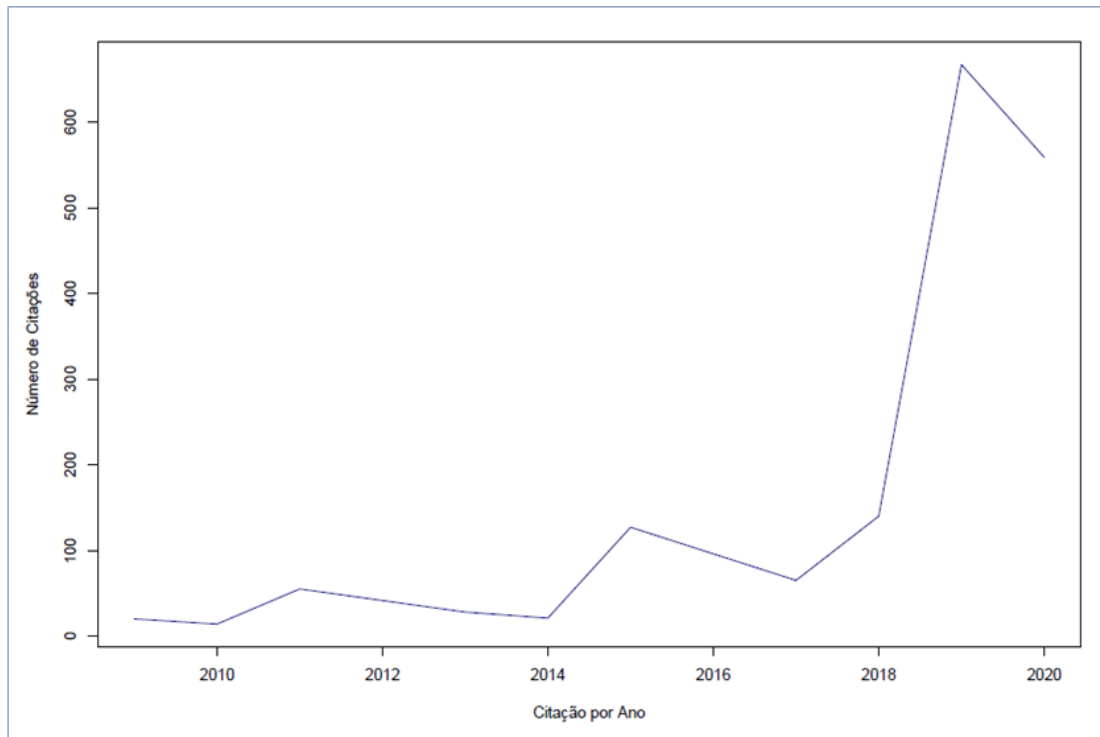
Gráfico 1 - Número de artigos avaliados por ano de publicação



Fonte: O Autor (2021)

Nos últimos anos houve um crescimento constante nas áreas de NPL e análise preditiva. De acordo com Wipo (2019), o crescimento acentuado nas aplicações de inteligência artificial começou nos últimos 7 anos, movido principalmente pelo crescimento do poder computacional e a grande conectividade com grandes volumes de dados compilados e compartilhados. Apesar de ser importante ressaltar que a relevância de um assunto possa ser evidenciada em termos da evolução do número de artigos por ano, o número de citações por ano também pode ser utilizado como outra métrica importante para este fim. A coleção estudada foi citada por 1696 artigos no período entre 2009 e 2020 de acordo com o banco de dados da *Web of Science Core Collection*. Na Gráfico 2 podemos observar o crescimento do número de citações por ano, onde a média de citação por artigo é 6,14.

Gráfico 2 - Número de citações de artigos avaliados por ano de citação



Fonte: O Autor (2021)

A pesquisa em questão está restrita à 16 áreas, que são classificadas com base em dados extraídos do banco de dados da *Web of Science*, sendo que mais de uma área de pesquisa pode ser designada para um artigo, a distribuição dos trabalhos por área de pesquisa é apresentada na Tabela 1.

Tabela 1 - Distribuição das áreas de pesquisa dos artigos avaliados

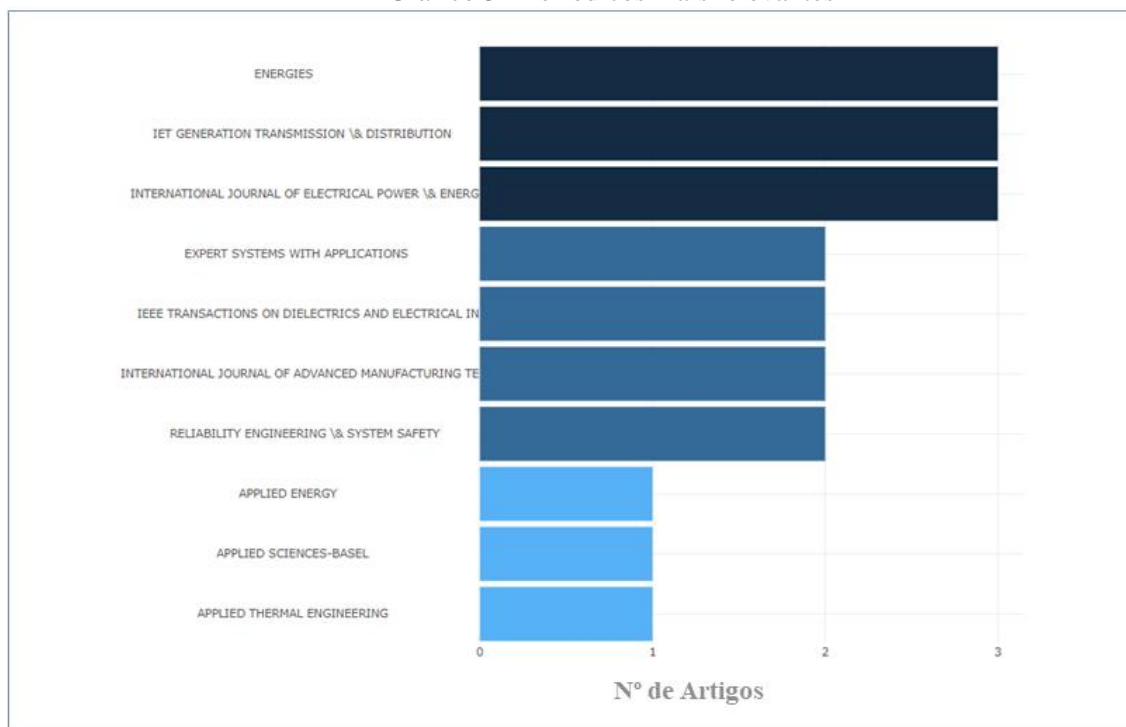
Áreas de pesquisa	Contagem registrada	% de 44
Engenharia	35	79,54
Energia e Combustíveis	10	22,72
Ciência da Computação	7	15,90
Pesquisa Operacional e Ciência de Gestão	5	11,36
Construção e Tecnologia de Construção	4	9,09
Mecânica	3	6,81
Física	3	6,81
Termodinâmica	3	6,81
Sistemas de Controle e Automação	2	4,54
Economia de Negócios	2	4,54
Matemática	2	4,54
Acústica	1	2,27
Química	1	2,27
Ciência de Materiais	1	2,27
Ciência e Tecnologia - outros tópicos	1	2,27
Telecomunicações	1	2,27

Fonte: O Autor (2021)

A porcentagem de artigos por áreas de pesquisa definidas nas etapas da revisão sistemática da literatura mostra que algumas áreas fizeram uso mais frequente do que outras da técnica regra de associação aplicada a manutenção de sistemas de distribuição elétrica.

Já na Gráfico 3, temos a relação dos dez periódicos com maior número de publicações, e os artigos aqui discutidos estão distribuídos em 20 periódicos diferentes.

Gráfico 3 - Periódicos mais relevantes

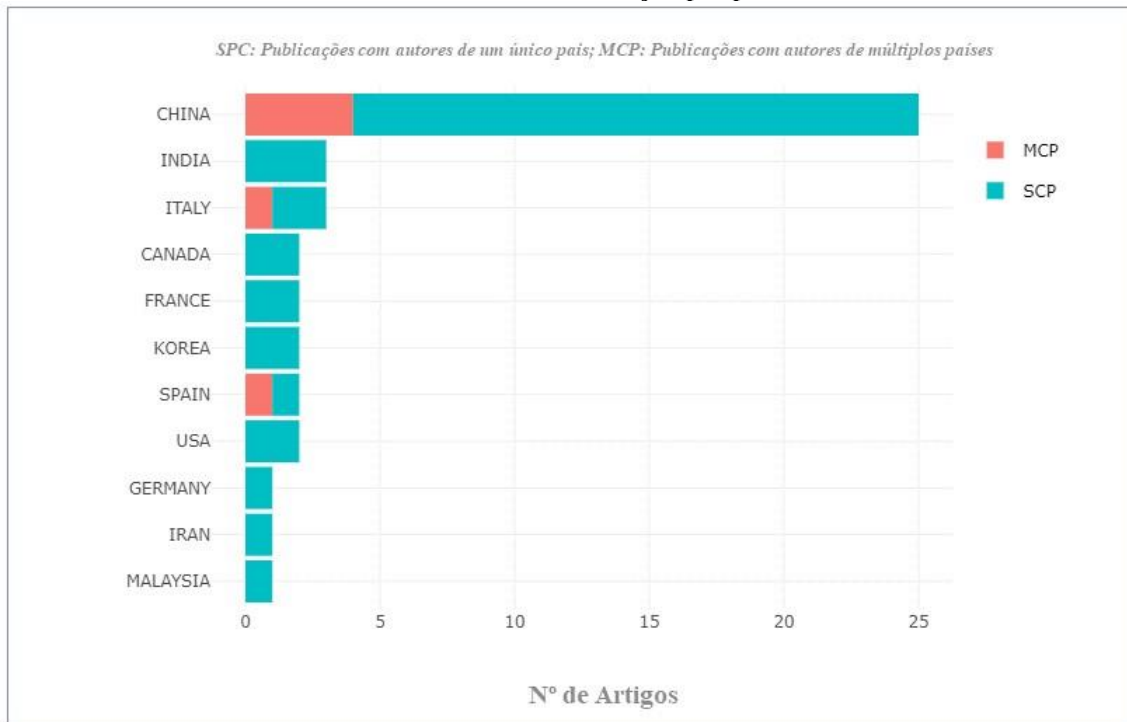


Fonte: O Autor (2021)

Segundo a lei de Bradford, Machado Junior *et al.* (2016), os periódicos com maior produtividade no tema em questão são: *Energies*, com 3 artigos; *Iet Generation Transmission \& Distribution*, com 3 artigos; *International Journal Of Electrical Power \& Energy Systems*, com 3 artigos; *Expert Systems With Applications*, com 2 artigos e; *Ieee Transactions On Dielectrics And Electrical Insulation*, com 2 artigos.

Para definir quem são os países que estão à frente das pesquisas foram analisados os dados da amostra, tomando como base o número de documentos produzidos em cada país, a China foi o país com maior número de publicações, seguida por Índia e Itália. Além disso, podemos observar que China, Itália e Espanha publicam artigos com colaboração de autores de outros países, como mostra o Gráfico 4.

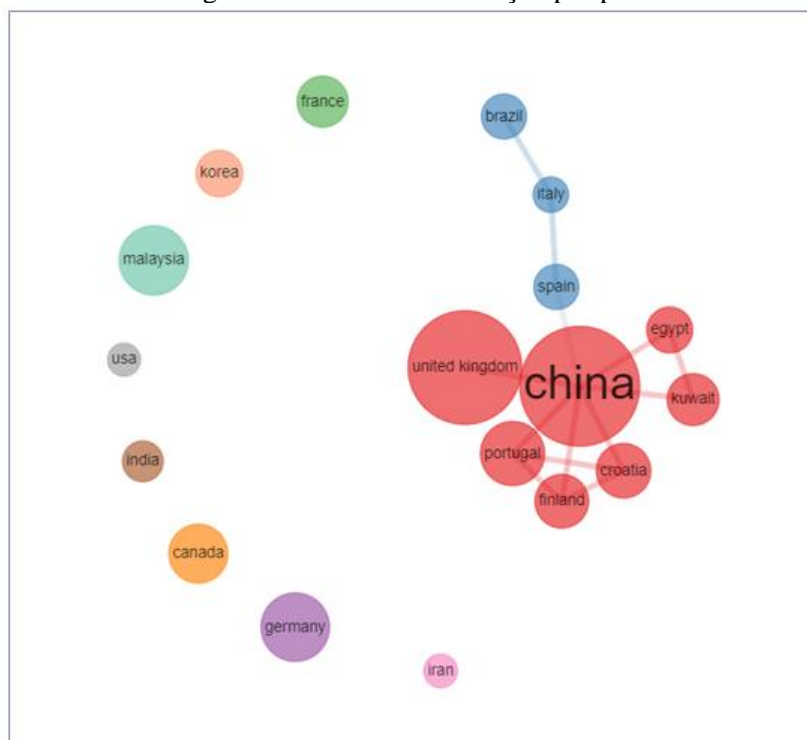
Gráfico 4 - Produção por país



Fonte: O Autor (2021)

A Figura 1 nos fornece uma análise complementar mostrando a rede de colaboração dos países, pois ilustra a densidade de estudos publicados por meio do tamanho dos círculos. Dessa forma, percebe-se a formação de um cluster principal (em vermelho) formado pelo país com maior colaboração de autores de outros países.

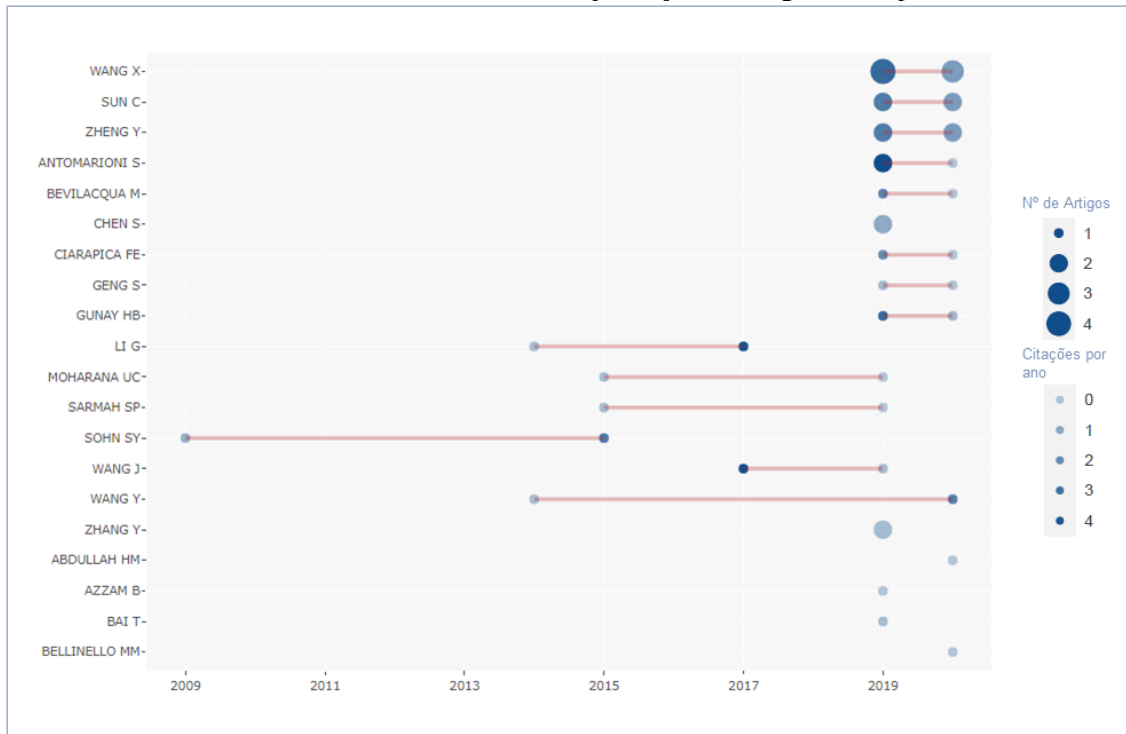
Figura 1 - Rede de colaboração por país



Fonte: O Autor (2021)

Para entender o comportamento dos principais autores na área em estudo foi desenvolvida uma análise temporal sobre suas publicações. No Gráfico 5 estão expostos os autores mais produtivos do portfólio bibliográfico analisado, relacionando o total de publicações com a quantidade de citações recebidas por cada autor. Foram identificados 146 autores diferentes, o que significa 3,32 autores por documento. Destaque para Wang X, que entre 2019 e 2020 publicou 7 artigos.

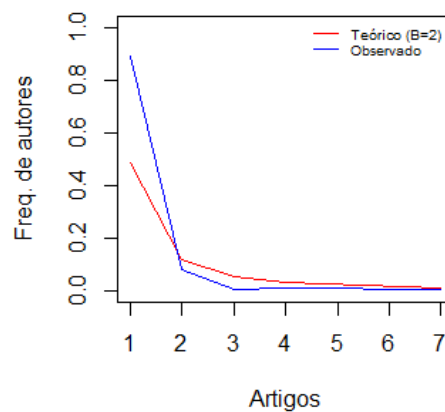
Gráfico 5 - Autores e sua produção ao longo do tempo



Fonte: O Autor (2021)

Analisando a produtividade segundo a Lei de Lotka Machado Junior *et al.* (2016), a amostra apresentou coeficiente de Lotka igual a 2,29, constatando que os autores possuem produtividade acima do padrão teórico. O teste de Kolmogorov-Smirnov de duas amostras entre a distribuição empírica e teórica retornou valor-p de 0,8186. Significando que a diferença entre o teórico e o observado não é estatisticamente significativa, mostrando que a produção dos autores se assemelha ao padrão teórico estabelecido, que pode ser observado na Gráfico 6.

Gráfico 6 - Produtividade Científica



Fonte: O Autor (2021)

No presente capítulo foi possível responder as questões de pesquisa estabelecidas com dados coletados em trabalhos científicos, fornecendo uma base sólida e confiável para comparação, com o objetivo de auxiliar novos estudos na área.

4 METODOLOGIA

Esse capítulo tem como propósito explicar todos os procedimentos metodológicos utilizados nesta pesquisa, apresentando todas as fases, desde o método de pesquisa, a caracterização do estudo, bem como expondo uma aplicação em um caso real em uma empresa de distribuição de energia elétrica.

4.1 CLASSIFICAÇÃO DA PESQUISA EM ENGENHARIA DE PRODUÇÃO

Uma pesquisa na área de engenharia de produção pode ser classificada em sua natureza, objetivos, abordagem e o método utilizado (TURRIONI e MELLO, 2012). Do ponto de vista da sua natureza, este estudo pode ser classificado em pesquisa aplicada. Segundo Silva e Menezes (2005), uma pesquisa aplicada tem o intuito de gerar conhecimentos para aplicações práticas e dirigidas a soluções de problemas reais.

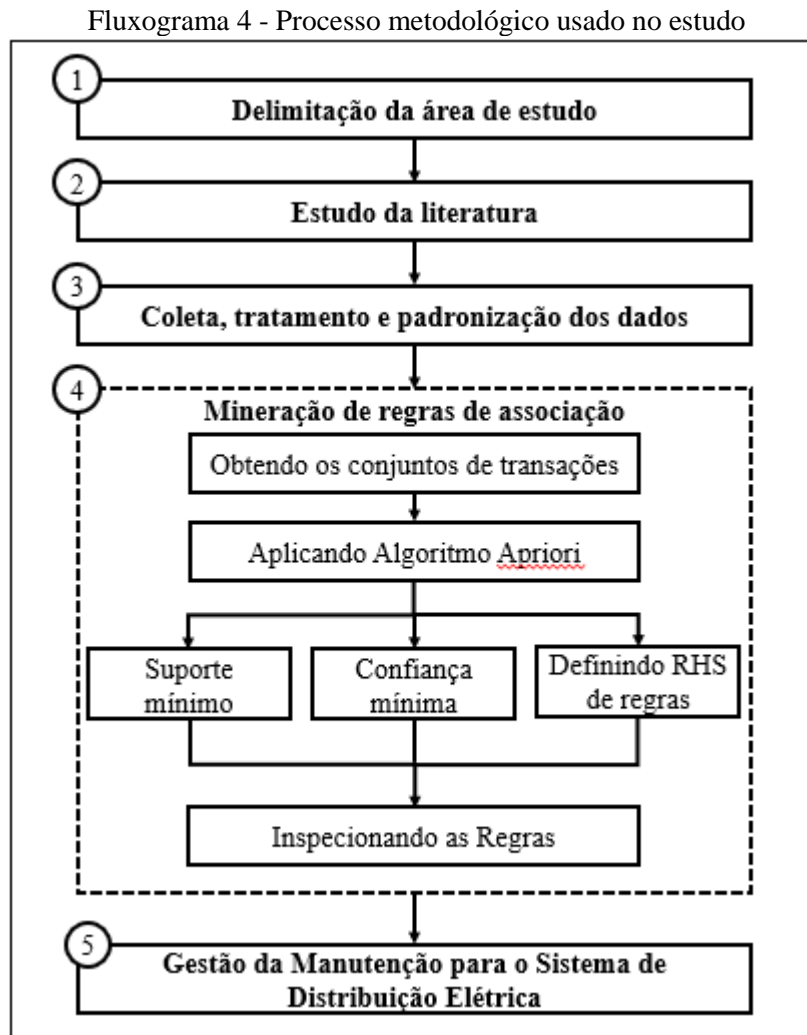
Com relação ao objetivo, segundo Gil (2002), o atual estudo pertence ao grupo de pesquisa exploratória, que visa proporcionar maior familiaridade ao problema, com o intuito de torná-lo mais explícito. Já a classificação da abordagem dessa pesquisa é quantitativa. Turrioni e Mello (2012), considera que tudo pode ser quantificável e conseqüentemente utilizado para classificá-las e analisá-las.

Por fim, o método utilizado para essa pesquisa é o estudo de caso. De acordo com Silva e Menezes (2005), o estudo de caso envolve um estudo minucioso de alguns objetos do problema, de maneira que se permita um amplo e detalhado conhecimento.

4.2 DELINEAMENTO COM AS ETAPAS DA PESQUISA

A metodologia utilizada no presente trabalho consiste em cinco etapas principais, resumidas no Fluxograma 4. Na primeira etapa, a delimitação da área estudada foi escolhida. Na segunda etapa, a literatura disponível sobre o assunto foi analisada. Na terceira etapa, os dados necessários para atingir os objetivos do estudo foram coletados. Na quarta etapa, é extraída as regras de associação. Finalmente, na quinta etapa é desenvolvido uma proposta de gestão da manutenção para sistemas de distribuição elétrica.

De acordo com Gil (2002), o procedimento em formato de fluxograma recebe o nome na literatura de Delineamento da Pesquisa. O delineamento da pesquisa refere-se ao planejamento da pesquisa em sua dimensão mais ampla. Nesse sentido, as próximas subseções descrevem a importância de cada etapa.



Fonte: O autor (2021)

4.2.1 Delimitação da área de estudo

O local de estudo que foi escolhido para esta pesquisa, foi a região do cariri, localizada no sul do estado do Ceará. A empresa de distribuição de energia elétrica que atua na região, distribui energia para mais de 400.000 mil unidades consumidoras. Essa área foi escolhida por existir uma necessidade comercial de entender com precisão qualquer mal funcionamento no sistema de distribuição de energia elétrica.

4.2.2 Estudo da literatura

A realização da análise da literatura busca um profundo conhecimento em problemas relacionados a técnica regra de associação aplicada a manutenção de sistemas de distribuição elétrica. Segundo Marconi e Lakatos (2017), o levantamento de um referencial teórico permite que o pesquisador obtenha uma visão geral do problema, se tornando indispensável para a realização das discussões e apresentações dos pontos mais relevantes da pesquisa.

Nesse sentido, para a construção do referencial teórico desta pesquisa foram apresentados no capítulo 3 como principais fontes para a coleta de artigos, jornais e revistas, os bancos de dados do portal de periódicos CAPES e os da *Web of Science core Collection*.

4.2.3 Coleta, tratamento e padronização dos dados

Nesta etapa, os dados de incidências de manutenção da empresa de distribuição de energia elétrica que atua na região do cariri foram utilizados na análise. Os dados foram coletados em um período de um ano e aspectos relacionados à manutenção, como descrição da causa da falha, data de início da ocorrência da falha, entre outros estão contidos no banco de dados.

Uma vez feita a coleta dos dados de incidência de manutenção, é preciso saber como tratar e analisar os dados. Segundo Turban *et al.*, (2005), informações brutas dificilmente são uteis, sendo necessário realizar um tratamento desses dados para que sejam possíveis manipula-los.

As ferramentas utilizadas para processamento dos dados foram pacotes disponíveis na linguagem R R Core Team (2021). Na linguagem R, foram utilizados pacotes para tratamento, padronização, visualização e modelagem dos dados. Foi utilizado o modelo de *machine learning* não supervisionado regras de associação.

Em seguida, os dados foram tratados e padronizados, parte das ações realizadas incluiu a exclusão de alguns atributos desnecessários, omissão de causas de falhas redundantes e melhoria do conteúdo, removendo elementos indesejados dos dados brutos. Isso foi além de estruturar os dados, de uma maneira que facilitou o processo de mineração de regras de associação. Uma amostra dos dados tratados e padronizados está representado no Quadro 3.

Quadro 3 - Amostra de dados tratados

	Data Inicio	Causa
1	2019-01-01 02:18:30	instantanea
2	2019-01-01 10:03:17	desligamento emergencial
3	2019-01-01 11:56:56	instantanea
4	2019-01-01 11:56:56	instantanea
5	2019-01-01 12:06:33	grande cliente defeito interno

Fonte: O Autor (2021)

4.2.4 Mineração de regras de associação

Esta etapa é ainda decomposta em três sub etapas distintas, como mostrado no Fluxograma 4. A primeira sub etapa envolve a geração dos conjuntos de itens de transações a partir das causas de falhas padronizadas. As transações nesse estudo são determinadas por intervalos de tempo, sendo que o intervalo é definido pela análise pretendida. O Quadro 4 ilustra um exemplo dos conjuntos de transações de itens gerados a partir das causas de falhas, em que a segunda transação consiste em três itens de causa de falha.

Segundo Chemweno *et al.*, (2016), o processo de geração das transações, para uma causa de falha selecionada, os itens de causa de falha que ocorrem antes ou após são combinados para formar um conjunto de itens. Por exemplo, na segunda transação a causa da falha “desligamento emergencial” foi observado pela primeira vez, seguido cronologicamente por “instantanea”, em seguida, “instantanea” novamente. Essa abordagem de seleção de transação de item garante que os eventos de causa de falha cronológica anteriores ou que ocorram após a causa de falha de interesse, sejam levados em consideração durante o processo de geração dos conjuntos de itens de transação.

Quadro 4 - Conjuntos de transações gerados a partir das causas de falhas

Intervalo	Conjunto de transações
2019-01-01 00:00:00	instantanea
2019-01-01 06:00:00	c("desligamento emergencial", "instantanea", "instantanea")
2019-01-01 12:00:00	c("grande cliente defeito interno", "desligamento emergencial", "defeito em chave fusivel", "instantanea")
2019-01-01 18:00:00	descargas atmosfericas
2019-01-02 00:00:00	c("instantanea", "defeito temporario nao identificado")

Fonte: O Autor (2021)

Na segunda sub etapa, o algoritmo Apriori foi aplicado ao conjunto de transações geradas a partir das causas de falhas. Para realizar a análise das regras de associação, diferentes restrições devem ser determinadas, tais como:

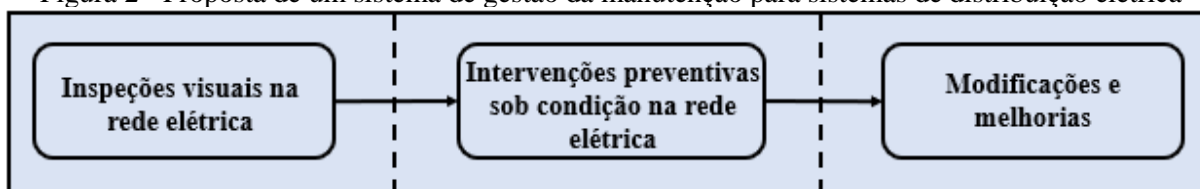
- **Suporte mínimo:** Para encontrar conjuntos de itens frequentes, o valor mínimo de suporte deve ser definido. Em geral, a seleção desse valor depende das características do conjunto de dados.
- **Confiança mínima:** Para fins de determinação da força das regras, o valor mínimo de confiança deve ser definido. Semelhante ao valor mínimo de suporte, aumentar o valor mínimo de confiança resulta em encontrar regras mais fortes.
- **Definindo RHS de regras:** Ao aplicar o algoritmo Apriori, várias regras podem ser encontradas, no entanto, uma vez que um dos objetivos da pesquisa é encontrar os padrões para as causas das falhas PEM, essas regras são desejadas cujo RHS é a causa da falha alvo.

Na terceira sub etapa, busca-se identificar as melhores regras, que mostram os padrões para a causa das falhas PEM, elas devem ser classificadas pelo valor de elevação. Posteriormente, aquelas com o valor de elevação mais altos são selecionados e interpretados pelo gestor de manutenção.

4.2.5 Gestão da manutenção para o sistema de distribuição elétrica

Nesta etapa, é desenvolvido uma proposta de sistema de gestão da manutenção para sistemas de distribuição elétrica, que objetiva a gestão dos recursos organizacionais de forma eficiente, assumindo papel fundamental para a empresa. Ressalta-se ainda que a manutenção contribui para o incremento da capacidade de inovação em relação aos seus concorrentes. A proposta de um sistema de gestão da manutenção para sistemas de distribuição elétrica divide-se em três sub etapas conforme podemos ver na imagem abaixo.

Figura 2 - Proposta de um sistema de gestão da manutenção para sistemas de distribuição elétrica

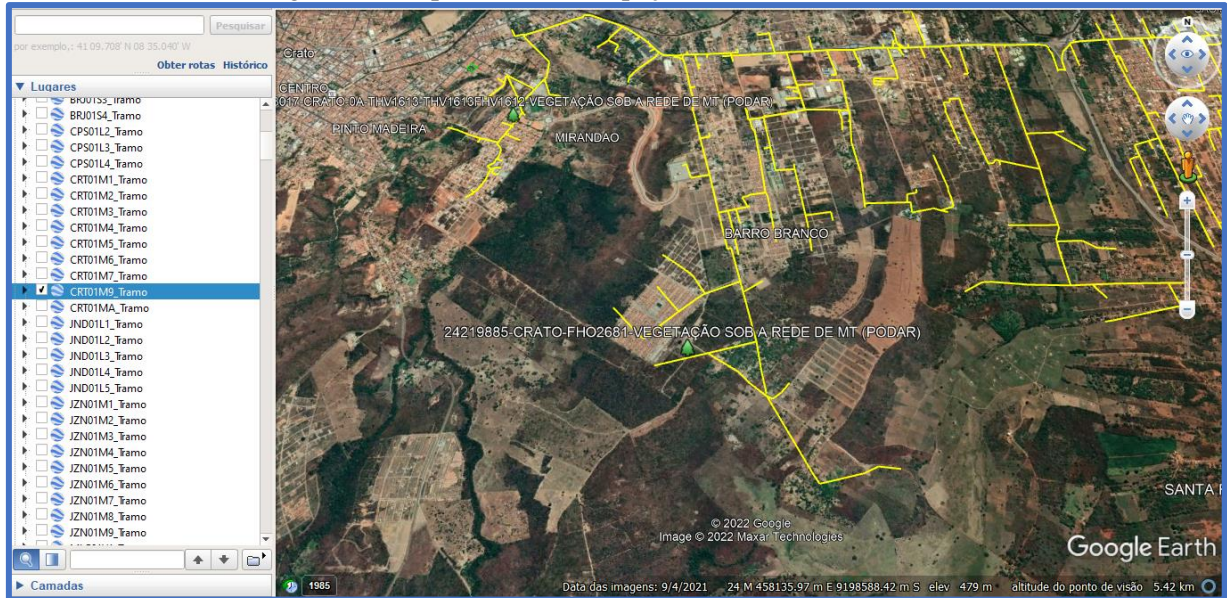


Fonte: O Autor (2021)

Na primeira sub etapa são executadas inspeções visuais na rede elétrica tomando como base a identificação das melhores regras de associação, que mostram os padrões para a causa das falhas PEM, com isso ajudará a nortear a inspeção tornando-a bem mais eficiente. A Figura 3 ilustra

um exemplo de como serão mapeadas e geradas as ordens de serviço a partir das inspeções visuais na rede elétrica para posteriormente realizar as ações preventivas sob condição.

Figura 3 – Mapeamento da inspeção visual na rede elétrica



Fonte: O Autor (2021)

Nota: Adaptado imagem do Google Earth (2021)

Com o tratamento das informações levantadas pelas as inspeções visuais, deverá seguir para segunda sub etapa, nela serão planejadas ações mantenedoras (intervenções preventivas sob condição na rede elétrica) advindas das informações levantadas pelas inspeções visuais, afim de atuar de forma antecipada à falha. Com isso se constitui uma política de manutenção importante para um desempenho adequado do sistema de distribuição elétrica. A Figura 4 exemplifica o modelo de como serão planejadas as ações mantenedoras.

Figura 4 – Modelo do planejamento de anomalias

TURMA/ Tur	Componentes (Qtz)	Superf	DATA	DI	OS	MUNICIPIO	SUBSTACAO	TRAMO DA ANAMOL	TIPO SGD	HORARIO	DESCRIÇÃO DA ANOMALIA	PREVISÃO DE EXECUÇÃO	Coordenada da anomalia	Coordenada da anomalia
			01/03/2020	DOM										
			02/03/2020	SEG										
			03/03/2020	TER										
			04/03/2020	QUA										

Fonte: O Autor (2021)

No planejamento de anomalias possui várias colunas como: turma de execução da anomalia, OS, definição de endereço da obra, coordenadas geográficas, tipo de SGD, descrição da anomalia, previsão de conclusão, dentre outras colunas.

Na terceira sub etapa, busca-se modificações e melhorias, pois os mantenedores podem e devem propor alterações e melhorias nas ações planejadas, afim de sempre obter os melhores resultados.

Nesse sentido, por meio da análise dos resultados é possível gerar um conjunto de informações cruciais para uma possível tomada de decisão dos gestores de manutenção. Em outras palavras os resultados buscam informar para os decisores os padrões para as causas das falhas PEM, e com isso possibilitar a gestão dos recursos organizacionais de forma eficiente.

5 APLICAÇÃO DA METODOLOGIA PROPOSTA

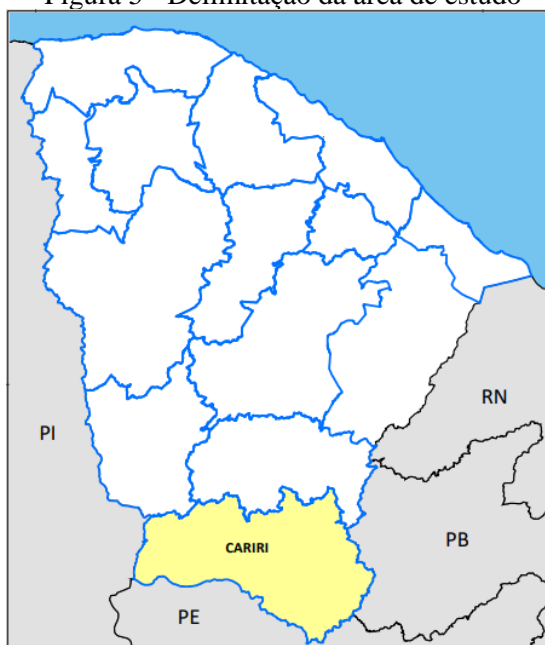
Este capítulo tem como objetivo apresentar em detalhes a aplicação de regras de associação nos dados de incidência de manutenção de média tensão (MT), de uma empresa de distribuição de energia elétrica, na região sul do Ceará. A abordagem é aplicada ao problema de prognóstico de falha em redes de distribuição de energia elétrica de escala real.

O estudo é baseado na aplicação do algoritmo de associação Apriori. Os resultados obtidos são ilustrados em tabelas, gráfico, figuras e posteriormente discutidos.

5.1 LOCAL DE ESTUDO

A pesquisa foi desenvolvida na região do cariri, localizada no sul do estado do Ceará, situado na região Nordeste do país, conforme exposto na Figura 5. Segundo o Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE), a região do cariri é formada por 29 municípios, com uma área territorial de 17.298,35 km², a região tem população estimada para 2019 de mais de 990 mil pessoas com densidade populacional de 57,76 hab/km².

Figura 5 - Delimitação da área de estudo



Fonte: IPECE/IBGE (2021)

A empresa de distribuição de energia elétrica que atua na região do cariri distribui energia para aproximadamente 400.000 mil unidades consumidoras. Portanto, existe uma necessidade

comercial de entender com precisão qualquer mal funcionamento no sistema de distribuição de energia elétrica, que causem perturbações significativas no fornecimento dos clientes. Dessa forma, a região é uma excelente candidata para o estudo de caso que será desenvolvido nas seções subsequentes.

5.2 BASE DE DADOS

A base de dados utilizada nessa pesquisa foi obtida na empresa de distribuição de energia elétrica atuante na região do cariri cearense. Devido à natureza da sensibilidade comercial dos dados, utilizou-se apenas para fins de pesquisa. Esses dados permitirão que a mineração de dados seja realizada com mais facilidade. Os atributos dos dados são de natureza numérica e categórica. Alguns atributos foram transformados para simplificar a análise, o Quadro 5 mostra a explicação dos principais atributos do conjunto de dados.

Quadro 5 - Explicação de alguns atributos do conjunto de dados

Atributos	Descrição
Data inicio	Data de início da falha
Causa da falha	Causa direta
Número de cliente	Quantidade de clientes afetados
Equipamentos	Equipamento envolvido
Município	Local onde ocorreu a falha

Fonte: O Autor (2021)

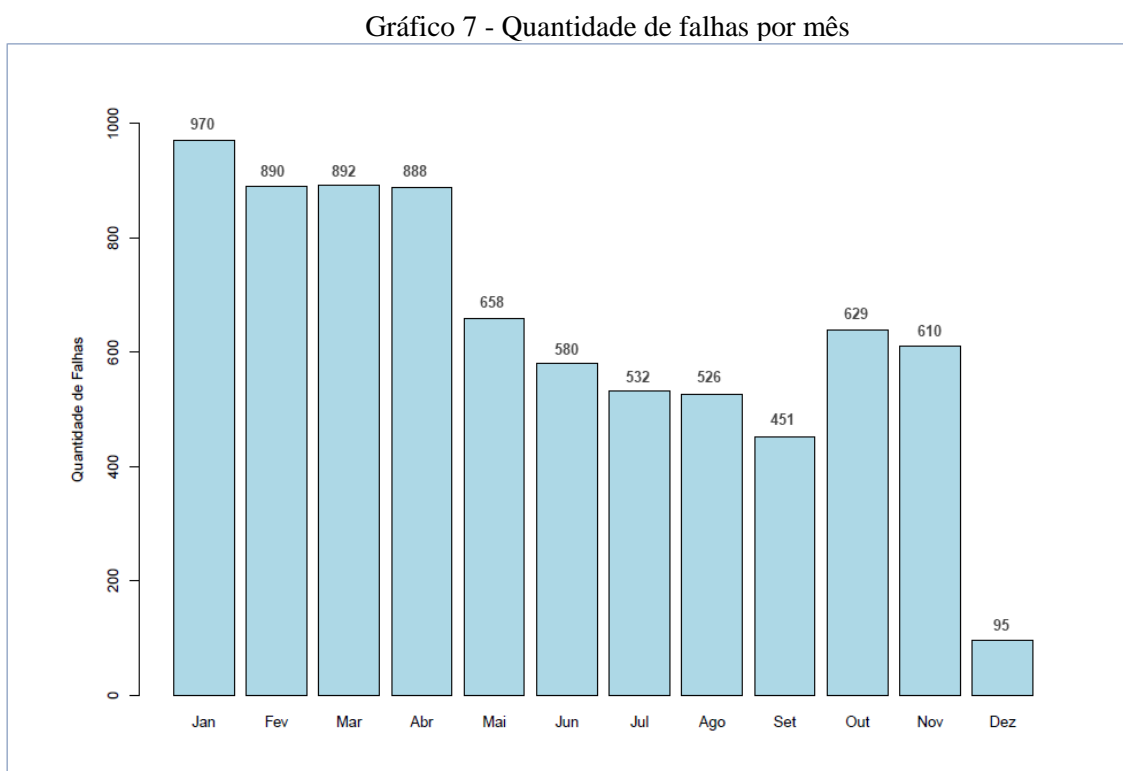
- Causa da falha: A causa da falha é usada principalmente para descrever o motivo que causou a interrupção, existem 54 causas diretas no conjunto de dados;
- Número de clientes: Quantidade de clientes que ficaram sem fornecimento de energia elétrica;
- Equipamentos: Equipamento afetado pela falha;
- Município: Localidade que ocorreu a falha no sistema elétrico.

5.3 DADOS INICIAIS

Para o desenvolvimento desta pesquisa, foram coletados os dados de incidências de manutenção da empresa de distribuição de energia elétrica que atua na região do cariri. As coletas foram realizadas no período de um ano (de 01 de janeiro à 30 de dezembro de 2019). No total, 7721

falhas foram registradas, totalizando um tamanho de 2,55 MB. Os dados são armazenados no arquivo Excel e manipulados pelo R Studio.

Nesse sentido, o Gráfico 7 apresenta a quantidade de falhas ocorridas em cada mês do ano. Foram identificados nos quatro primeiros meses 3640 falhas representando 47% do total.

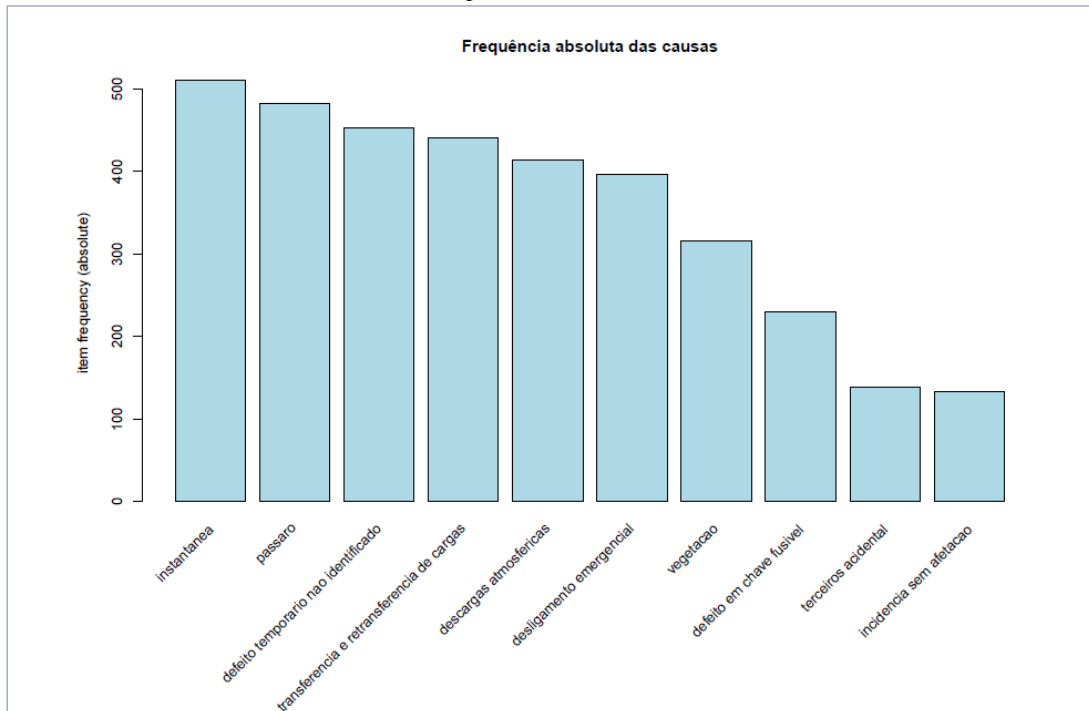


Fonte: O Autor (2021)

Uma das explicações para essa concentração de falhas nos primeiros meses do ano é por conta do período chuvoso na região. Segundo a Fundação Cearense de Meteorologia e Recursos Hídricos (FUNCEME) a precipitação pluviométrica média da região do cariri nesses meses foram de 725 mm.

Prosseguindo com a apresentação estatística dos resultados, adquirido através do banco de dados, o Gráfico 8 demonstra as dez causas de falha que mais ocorreram, ou seja, as causas mais recorrentes dentro do espaço de tempo observado. As causas de falha se dividem em causas possíveis evitar com manutenção - PEM e as não possíveis evitar com manutenção – NPEM. Dentre as causas PEM, as que mais impactam o setor de manutenção são desligamento emergencial, vegetação e defeito em chave fusível.

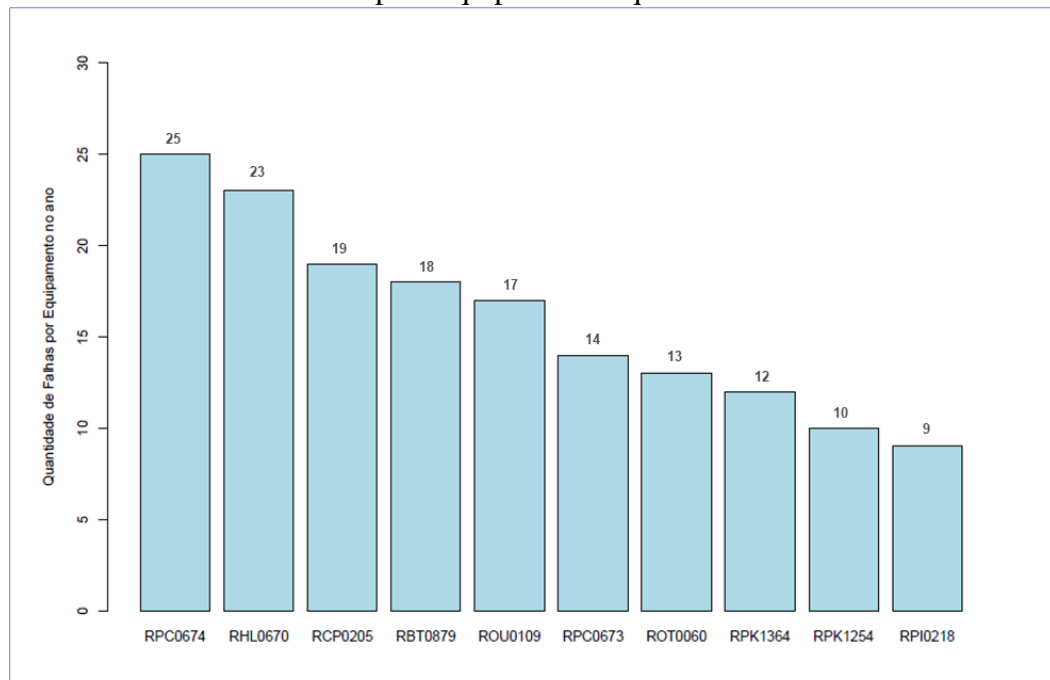
Gráfico 8 - Top 10 causas mais recorrentes



Fonte: O Autor (2021)

No Gráfico 9, é apresentado os dez equipamentos do sistema elétrico em estudo que mais falharam durante o período observado, os mesmos são identificados por códigos, como por exemplo o equipamento que mais falhou foi o RPC0674.

Gráfico 9 - Top 10 equipamentos que mais falharam



Fonte: O Autor (2021)

Com essas informações, os gestores de manutenção conseguem ter uma visão geral dos meses que ocorrem mais falhas, um detalhamento das maiores causas de falha, conseguindo identificar as causas PEM que implica na maior demanda do setor de manutenção que são desligamento emergencial, vegetação, defeito em chave fusível e os equipamentos que mais falharam no período estudado.

Com a utilização do modelo proposto, foi possível realizar e extrair esse estudo inicial e estatístico, que fornece dados relevantes para os gestores de manutenção. Logo, a partir desses estudos iniciais, é possível fazer uso do modelo para gerar regras de associações afim de prever possíveis falhas no sistema elétrico.

5.4 REGRAS DE ASSOCIAÇÃO PARA CAUSAS DAS FALHAS

O conjunto de dados consiste em mais de 7000 falhas, este grande número de falhas evidencia a importância do estudo e tratamento de dados através de técnicas computacionais, no primeiro experimento vamos descobrir os fatores associados que contribuem para a falha na rede de distribuição elétrica. Para conseguir isso, tratamos as causas das falhas dentro de determinado intervalo de tempo como uma transação. O intervalo de tempo é determinado pelo registro de data e hora dos eventos ou por especialistas da área de manutenção. O intervalo de tempo de interesse deve ser maior que o tempo de espera das ações de manutenção, para que os gestores de manutenção possam intervir na rede de distribuição elétrica a tempo. Para obter mais informações sobre as relações entre as causas das falhas, foi definido junto aos gestores da empresa o intervalo de tempo de 6 horas.

Para os dados transacionais de 6 horas, o comprimento das regras de associação é apresentado na Tabela 2 e as fortes regras de associação são evidenciadas no Quadro 6. Logo, o conjunto de regras foi obtido com limite mínimo de suporte definido como 0,01 e o limite mínimo de confiança como 0,7. No total, essas configurações geraram 23 regras.

Tabela 2 - Distribuição do comprimento das regras de associação

Comprimento	2 Item	3 Item	4 Item	5 Item
Nº de Regras	3	5	10	5

Fonte: O Autor (2021)

Quadro 6 - Lista de regras fortes de associação para dados transacionais de 6 horas

Antecedentes (lhs)	Consequentes (rhs)	Confiança
jump partido	instantanea	0.700
jump partido	transferencia e retransferencia de cargas	0.700
dia critico interno int	descargas atmosfericas	0.840
defeito em para raios, desligamento emergencial	instantanea	0.739
estava normal, vegetacao	transferencia e retransferencia de cargas	0.800
desligamento emergencial, dia critico interno int	instantanea	0.772
grande cliente defeito interno, terceiros acidental	passaro	0.761
incidencia sem afetacao, terceiros acidental	passaro	0.761
instantanea, terceiros acidental, vegetacao	passaro	0.772
defeito temporario nao identificado, terceiros acidental, transferencia e retransferencia de cargas	passaro	0.714
desligamento emergencial, incidencia sem afetacao, instantanea	descargas atmosfericas	0.739
defeito em chave fusivel, desligamento emergencial, vegetacao	passaro	0.718
defeito em chave fusivel, descargas atmosfericas, vegetacao	passaro	0.750
defeito em chave fusivel, defeito temporario nao identificado, vegetacao	passaro	0.735
defeito em chave fusivel, transferencia e retransferencia de cargas, vegetacao	passaro	0.705
defeito em chave fusivel, defeito temporario nao identificado, desligamento emergencial	passaro	0.722
defeito em chave fusivel, desligamento emergencial, transferencia e retransferencia de cargas	passaro	0.729
defeito em chave fusivel, defeito temporario nao identificado, transferencia e retransferencia de cargas	passaro	0.787
defeito temporario nao identificado, descargas atmosfericas, desligamento emergencial, vegetacao	transferencia e retransferencia de cargas	0.714
defeito temporario nao identificado, desligamento emergencial, instantanea, vegetacao	transferencia e retransferencia de cargas	0.769
defeito temporario nao identificado, descargas atmosfericas, desligamento emergencial, transferencia e retransferencia de cargas	instantanea	0.724
defeito temporario nao identificado, descargas atmosfericas, desligamento emergencial, passaro	instantanea	0.730
desligamento emergencial, instantanea, passaro, transferencia e retransferencia de cargas	defeito temporario nao identificado	0.729

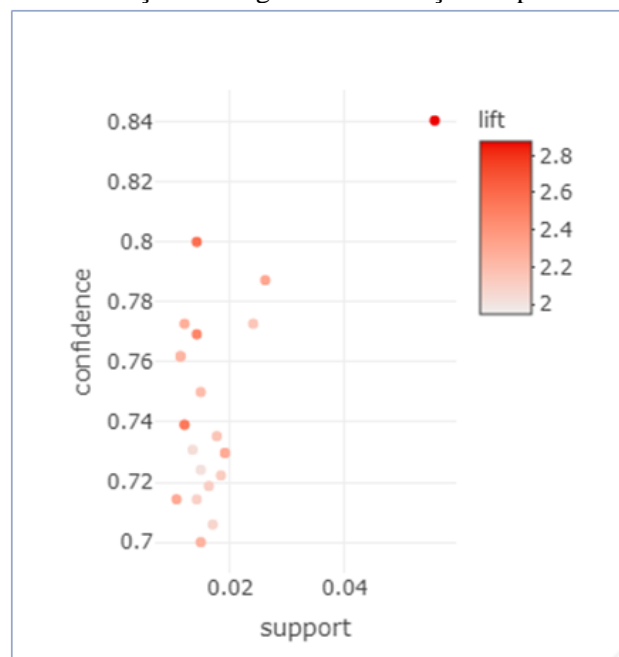
Fonte: O Autor (2021)

Na Tabela 2 podemos observar que o comprimento de 4 itens tem a maioria das regras e comprimento de 2 itens tem o menor número de regras. No Quadro 6 apresenta as fortes regras de associação com confiança maior que 0,7. É possível observar que a regra mais significativa é {dia critico interno int} => {descargas atmosféricas}, com confiança 0,840, o que implica, dado a causa da falha {dia critico interno int} ocorra, a causa da falha {descargas atmosféricas} ocorrerá dentre 6 horas com uma probabilidade 0,840. Outro ponto de fácil observação é que a grande maioria das causas das falhas são ocasionadas por animais e efeitos climáticos que são causas NPEM.

Através deste resultado e um estudo junto com os gestores da empresa foi possível identificar uma falha em projeto, onde, foi possível identificar que as falhas associadas a descargas atmosféricas, tem causa raiz a inexistência de conjunto de para-raios a cada 3 km de rede de MT e nos pontos de mudança de seção de condutores da rede de MT, descumprindo a Instrução de Trabalho no.60/2018 (IT0060). Caso a IT0060 estivesse sendo respeitada as falhas associadas a descarga atmosféricas teriam impactado bem menos a rede elétrica.

Um gráfico de dispersão pode ser usado para visualizar as regras de associação geradas. Suporte e confiança aparecem nos eixos X e Y, além disso a opacidade dos pontos indica o valor relativo à elevação da regra. Na Figura 6 mostra que a regra com maior elevação também é a com maior confiança.

Figura 6 - Visualização das regras de associação no plano de dispersão



Fonte: O Autor (2021)

As redes de nós de associação permitem uma fácil interpretação. Cada nó representa uma regra de associação distinta, as setas conectam uma causa ou um grupo de causas a um nó (anteriores e consequentes), por exemplo, no conjunto de dados, quando a causa da falha “estai quebrado” ocorre, a causa da falha “desligamento emergencial” também irá ocorrer. O diâmetro dos nós indica uma frequência de causa de falha maior no conjunto de dados analisados, e a opacidade dos nós indica o valor relativo da elevação da regra.

Observou-se que a ordenação de regras de associação com relação a seus valores de confiança ajudou a identificar causa das falhas raras no conjunto de dados, causas essas que podem melhorar a decisão dos gestores de manutenção, logo que estarão de posse de informações futuras, com isso podem prever a falha baseado na sua causa raiz.

5.5 REGRAS DE ASSOCIAÇÃO PARA CAUSAS DAS FALHAS PEM

A seguir, são apresentados e discutidos os resultados da mineração de regras de associação para causas das falhas PEM. O conjunto de regras foi obtido com limite mínimo de suporte definido como 0,0018, o que equivale a 100 ocasiões. Desta maneira se um conjunto de itens aparecer mais de 100 vezes, este é considerado frequente. Essas regras podem ser investigadas e interpretadas para selecionar as mais importantes. Vale ressaltar que selecionar um valor baixo é sempre uma opção melhor em comparação com selecionar um suporte mínimo mais alto, pois a última opção pode resultar na perda de possíveis regras importantes.

O limite de confiança mínima foi definido para ser 70%. Conforme mencionado anteriormente, a seleção desse valor depende da decisão dos gestores de manutenção. Afim de encontrar as melhores regras, que mostram os padrões para a causa das falhas PEM, todas são inspecionadas aplicando a função de inspeção definida no pacote Apriori. O procedimento consiste em classificar e identificar todas as regras de alta qualidade cujos valores de elevação são significativos. Neste estudo, aquelas regras cuja a elevação é maior 3,0 são selecionadas e apresentadas.

- (1) Falhas relacionadas à desligamento emergencial: As regras mais fortes alcançadas para falhas relacionadas à desligamento emergencial são fornecidos na Tabela 4. De acordo com os resultados, pode-se entender que arvore tombada e vegetação são causas frequentes para as falhas de desligamento emergencial. Com o intuito de eliminar esse tipo de falha faz-se necessário intensificar as inspeções na rede elétrica, para identificar

os pontos críticos e assim poder programa a manutenção preventiva para podar essa vegetação em excesso.

- (2) Falhas relacionadas à vegetação: A Tabela 5 resume as regras com maior intensidade para falhas relacionadas à vegetação. Com base nesta tabela, percebe-se que arvore tombada têm um grande impacto nas falhas de vegetação. De acordo com o restante dos resultados, entende-se que as condições climáticas (descarga atmosféricas) desempenha um papel crucial para ocorrência dessa falha. Além disso, religadores de linha e seccionadores são os principais dispositivos que eliminam as falhas relacionadas a vegetação.
- (3) Falhas relacionadas à defeito em chave fusível: As regras obtidas para falhas de equipamentos são fornecidas na Tabela 6. No entanto, com base na tabela, entende-se que jump partido, vegetação e condições climáticas (descarga atmosféricas) têm um grande impacto nas falhas em defeito em chave fusível.

Através dos resultados foi possível encontrar uma correlação nas falhas relacionadas a cada causa PEM, em todas a causa vegetação se destaca, fazendo com que os gestores de manutenção busquem mitigar as falhas relacionadas a esse tipo de causa, usando as seguintes medidas, intensificar as inspeções na rede elétrica, para identificar os pontos críticos e assim poder programa a intervenção preventiva para podar essa vegetação em excesso, quando não for possível fazer a poda fazer a substituição de condutores nu para condutores multiplexo.

Tabela 4 - Resultados da mineração de regras de associação para falhas relacionadas à desligamento emergencial

LHS	RHS	Support	Confidence	Lift
{corpo estranho, grande cliente defeito interno}	=> {desligamento emergencial}	0,00212	1,0000	3,5783
{defeito temporario nao identificado, passaro, reclamacao nivel de tensao}	=> {desligamento emergencial}	0,00282	1,0000	3,5783
{arvore tombada, descargas atmosfericas, vegetacao}	=> {desligamento emergencial}	0,00212	1,0000	3,5783
{arvore tombada, defeito temporario nao identificado, vegetacao}	=> {desligamento emergencial}	0,00212	1,0000	3,5783
{arvore tombada, transferencia e retransferencia de cargas, vegetacao}	=> {desligamento emergencial}	0,00212	1,0000	3,5783
{defeito temporario nao identificado, deslgrreligamento por solicitacao do cliente, vegetacao}	=> {desligamento emergencial}	0,00212	1,0000	3,5783
{defeito temporario nao identificado, deslgrreligamento por solicitacao do cliente, transferencia e retransferencia de cargas}	=> {desligamento emergencial}	0,00282	1,0000	3,5783
{deslgrreligamento por solicitacao do cliente, instantanea, transferencia e retransferencia de cargas}	=> {desligamento emergencial}	0,00282	1,0000	3,5783
{animais, pipa, transferencia e retransferencia de cargas}	=> {desligamento emergencial}	0,00212	1,0000	3,5783
{instantanea, pipa, transferencia e retransferencia de cargas}	=> {desligamento emergencial}	0,00212	1,0000	3,5783
{instantanea, passaro, pipa}	=> {desligamento emergencial}	0,00212	1,0000	3,5783
{defeito em isolador, estava normal, instantanea}	=> {desligamento emergencial}	0,00282	1,0000	3,5783

Fonte: O Autor (2021)

Tabela 5 - Resultados da mineração de regras de associação para falhas relacionadas à vegetação

LHS	RHS	Support	Confidence	Lift
{arvore tombada, descargas atmosfericas, desligamento emergencial}	=> {vegetacao}	0,00212	1,00000	4,484177
{arvore tombada, defeito temporario nao identificado, desligamento emergencial}	=> {vegetacao}	0,00212	1,00000	4,484177
{arvore tombada, desligamento emergencial, transferencia e retransferencia de cargas}	=> {vegetacao}	0,00212	1,00000	4,484177
{degradacao material, descargas atmosfericas, passaro}	=> {vegetacao}	0,00212	1,00000	4,484177
{grande cliente defeito interno, instantanea, ninho de passaro}	=> {vegetacao}	0,00212	1,00000	4,484177
{grande cliente defeito interno, ninho de passaro, passaro}	=> {vegetacao}	0,00212	1,00000	4,484177
{ninho de passaro, passaro, terceiros acidental}	=> {vegetacao}	0,00212	1,00000	4,484177
{defeito em chave fusivel, descargas atmosfericas, ninho de passaro}	=> {vegetacao}	0,00282	1,00000	4,484177
{descargas atmosfericas, ninho de passaro, passaro}	=> {vegetacao}	0,00494	1,00000	4,484177
{defeito temporario nao identificado, ninho de passaro, transferencia e retransferencia de cargas}	=> {vegetacao}	0,00282	1,00000	4,484177
{defeito temporario nao identificado, ninho de passaro, passaro}	=> {vegetacao}	0,00212	1,00000	4,484177
{ninho de passaro, passaro, transferencia e retransferencia de cargas}	=> {vegetacao}	0,00353	1,00000	4,484177

Fonte: O Autor (2021)

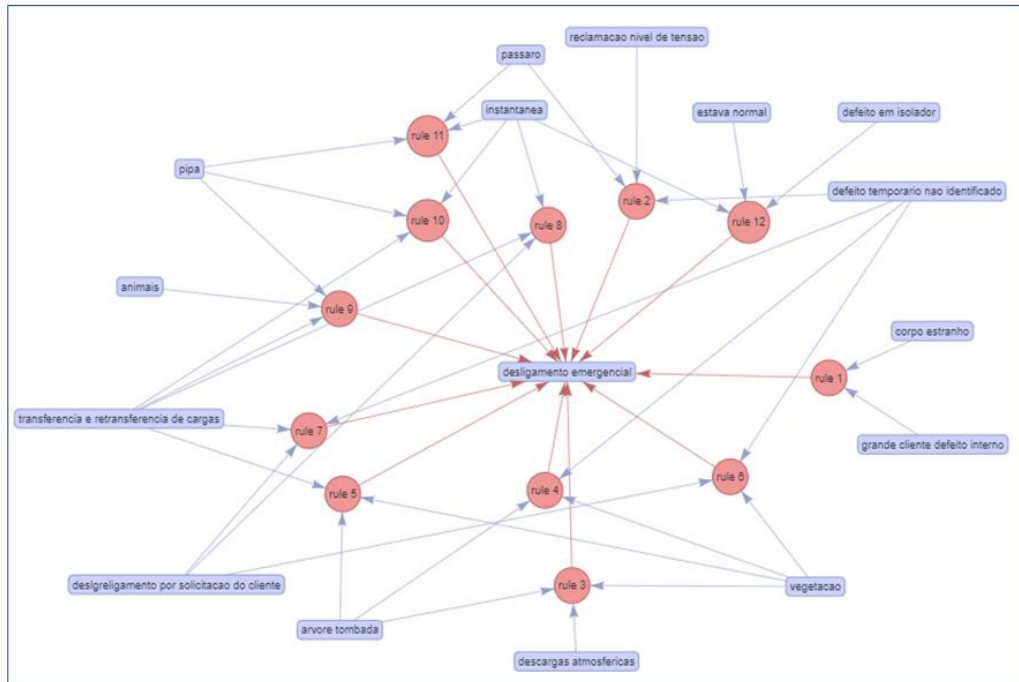
Tabela 6 - Resultados da mineração de regras de associação para falhas relacionadas à defeito em chave fusível

LHS	RHS	Support	Confidence	Lift
{defeito interno cliente, terceiros acidental, transferencia e retransferencia de cargas}	=> {defeito em chave fusivel}	0,00212	1,00000	6,16087
{defeito interno cliente, passaro, terceiros acidental, transferencia e retransferencia de cargas}	=> {defeito em chave fusivel}	0,00212	1,00000	6,16087
{descargas atmosfericas, desligamento emergencial, jump partido, passaro, vegetacao}	=> {defeito em chave fusivel}	0,00212	1,00000	6,16087
{descargas atmosfericas, instantanea, jump partido, passaro, vegetacao}	=> {defeito em chave fusivel}	0,00212	1,00000	6,16087
{defeito temporario nao identificado, descargas atmosfericas, instantanea, jump partido, passaro}	=> {defeito em chave fusivel}	0,00212	1,00000	6,16087
{descargas atmosfericas, desligamento emergencial, jump partido, passaro, transferencia e retransferencia de cargas, vegetacao}	=> {defeito em chave fusivel}	0,00212	1,00000	6,16087
{defeito temporario nao identificado, descargas atmosfericas, instantanea, jump partido, passaro, vegetacao}	=> {defeito em chave fusivel}	0,00212	1,00000	6,16087
{descargas atmosfericas, instantanea, jump partido, passaro, transferencia e retransferencia de cargas, vegetacao}	=> {defeito em chave fusivel}	0,00212	1,00000	6,16087
{defeito temporario nao identificado, descargas atmosfericas, instantanea, jump partido, passaro, transferencia e retransferencia de cargas}	=> {defeito em chave fusivel}	0,00212	1,00000	6,16087
{defeito temporario nao identificado, descargas atmosfericas, instantanea, jump partido, passaro, transferencia e retransferencia de cargas, vegetacao}	=> {defeito em chave fusivel}	0,00212	1,00000	6,16087
{descargas atmosfericas, jump partido, passaro, vegetacao}	=> {defeito em chave fusivel}	0,00282	0,80000	4,928696
{instantanea, jump partido, passaro, vegetacao}	=> {defeito em chave fusivel}	0,00282	0,80000	4,928696

Fonte: O Autor (2021)

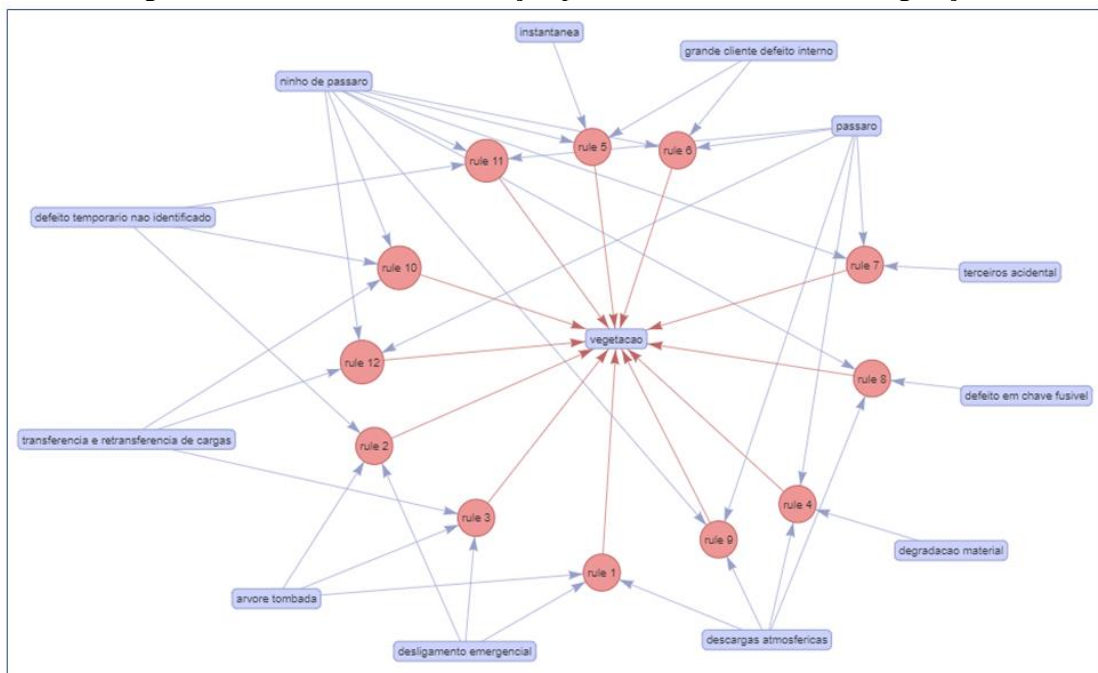
As redes de nós de associação permitem uma interpretação das regras de associação relacionadas as causas das falhas PEM. Nas Figuras 8, 9 e 10 podemos ver uma representação dos resultados da mineração de regras de associação para falhas relacionadas a cada causa das falhas PEM estudadas.

Figura 8 – Rede de nós de associação para falhas relacionadas à desligamento emergencial



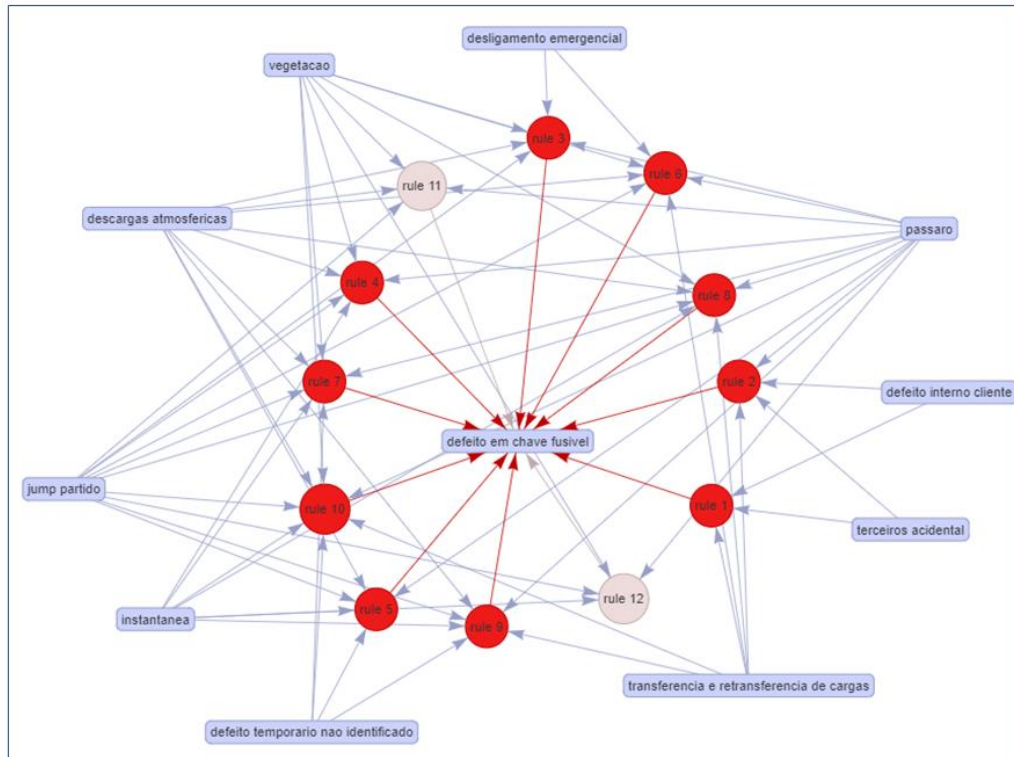
Fonte: O Autor (2021)

Figura 9 - Rede de nós de associação para falhas relacionadas à vegetação



Fonte: O Autor (2021)

Figura 10 - Rede de nós de associação para falhas relacionadas à defeito em chave fusível



Fonte: O Autor (2021)

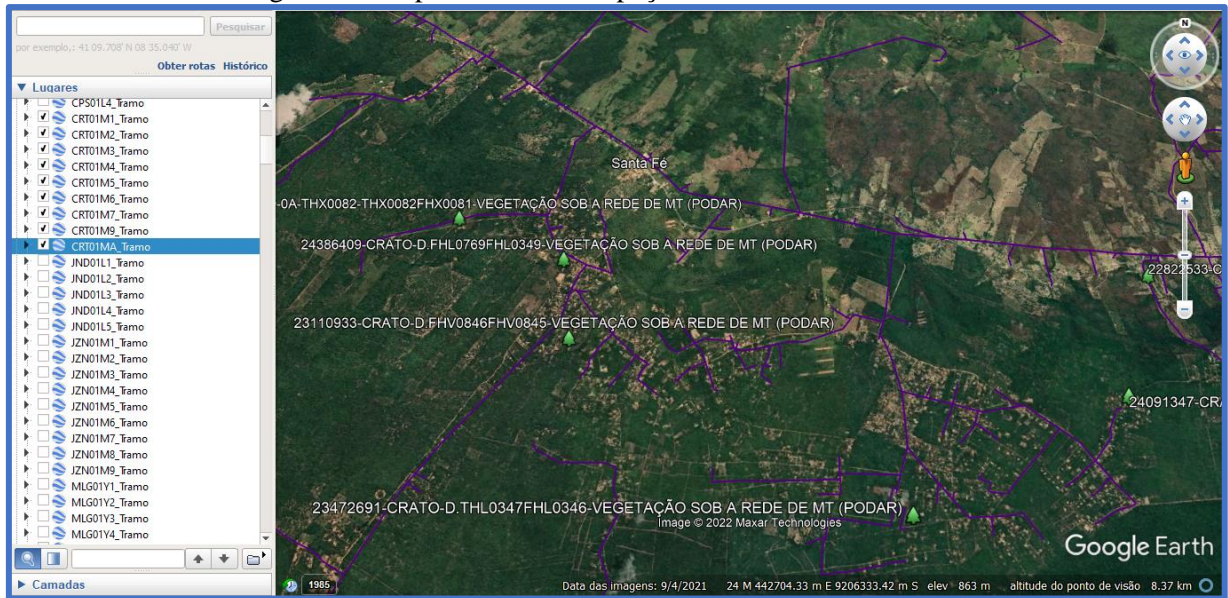
5.6 GESTÃO DA MANUTENÇÃO PARA O SISTEMA DE DISTRIBUIÇÃO ELÉTRICA

Após a mineração de regras de associação para causas das falhas PEM foi possível identificar uma correlação relacionada a cada causa PEM, em todas, a causa vegetação se destaca. Com base nos resultados das regras de associação geradas os gestores de manutenção tomaram a decisão de intensificar as inspeções visuais em toda a rede *in loco*, com o intuito de mapear todos os possíveis pontos que a vegetação se encontra em excesso, após identificação desses pontos foram planejadas e programadas intervenções preventivas com a finalidade de atuar de forma antecipada à falha. Na Figura 11 podemos ver alguns dos pontos identificados a partir das inspeções visuais na rede, em cada ponto consta a OS, cidade, estrutura mais próxima e observação.

Com o mapeamento de todos os pontos em que a vegetação está em excesso e geradas as ordens de serviços para cada ponto, inicia-se a análise de todas as OS para serem planejadas durante todo o mês. Após as informações necessárias para o planejamento coletadas, o gestor realiza o planejamento de anomalias que compõe todas as programações relacionadas ao mês, nele constam as anomalias a serem corrigidas por cada turma com informações de datas, horários,

tipo de SGD, descrição da anomalia, município e etc. Na Figura 12 conseguimos ver o planejamento de anomalias de uma equipe referente ao mês de março/2020.

Figura 11- Mapeamento das inspeções visuais na rede elétrica



Fonte: O Autor (2021)

Figura 12 – Planejamento de anomalias de uma equipe referente ao mês de março/2020

DATA	DI	OS	SUBESTACAO	TRAMO DA ANAMOL	TIPO SGD	HORARIO	DESCRIÇÃO DA ANOMALIA	PREVISÃO DE EXECUÇÃO
01/03/2020	DOM	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA
02/03/2020	SEG	23127497	MLG	FOZ1128	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	02/03/2020
03/03/2020	TER	24173861	MLG	FOZ0609	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	03/03/2020
04/03/2020	QUA	23578609	CPS	FCX0029	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	04/03/2020
05/03/2020	QUI	23233765	ANN	SPG1214	FOLGA	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	05/03/2020
06/03/2020	SEX	24096567	NVO	0A-TPE2267FPE2275	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	06/03/2020
07/03/2020	SÁB	23643379	CPS	0A-TPB4727FPB4824	PROX MT	08:00 AS 12:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	07/03/2020
08/03/2020	DOM	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA
09/03/2020	SEG	22838515	NVO	TPE2666-TPE2666FPE2	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	09/03/2020
10/03/2020	TER	23391261	NVO	D.TPD2059FPD2075	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	10/03/2020
11/03/2020	QUA	23064369	BLN	TMR4979	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	11/03/2020
12/03/2020	QUI	22705757	BLN	TBT1730-TBT1730FBT1	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	12/03/2020
13/03/2020	SEX	23908459	BLN	TBS0867-TBS0867FBS0	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	13/03/2020
14/03/2020	SÁB	23908459	BLN	TBS0867-TBS0867FBS0	PROX MT	08:00 AS 12:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	14/03/2020
15/03/2020	DOM	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA
16/03/2020	SEG	22617355	CRT	D.THX0119FHX0118	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	16/03/2020
17/03/2020	TER	23042927	CRT	THX7522-THX7522SHX7	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	17/03/2020
18/03/2020	QUA	23043921	BBL	D.TCP0180D.TCP1152	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	18/03/2020
19/03/2020	QUI	23956119	BBL	CO0039-TCO0039FCO0	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	19/03/2020
20/03/2020	SEX	23645311	BBL	D.TCL2580FCN0172	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	20/03/2020
21/03/2020	SÁB	24341137	BBL	FCL1367	PROX MT	08:00 AS 12:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	21/03/2020
22/03/2020	DOM	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA
23/03/2020	SEG	24254417	CRT	D.FHX1154D.THX2036	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	23/03/2020
24/03/2020	TER	23664305	BBL	CL0045-TCL0045D.TCL	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	24/03/2020
25/03/2020	QUA	23095207	BBL	CO2489-TCO2489SCO2	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	25/03/2020
26/03/2020	QUI	23095207	BBL	CO2489-TCO2489SCO2	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	26/03/2020
27/03/2020	SEX	24133237	CRT	THX7522	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	27/03/2020
28/03/2020	SÁB	23016613	BRJ	D.TOZ5771SOZ5763	PROX MT	08:00 AS 12:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	28/03/2020
29/03/2020	DOM	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA	FOLGA
30/03/2020	SEG	23932027	MLG	D.TOU0113FOU0015	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	30/03/2020
31/03/2020	TER	23038857	MLG	D.FOZ0190D.TOZ3426	PROX MT	08:00 AS 12:00 - 13:00 AS 17:00	VEGETAÇÃO SOB A REDE DE MT (PODAR)	31/03/2020

Fonte: O Autor (2021)

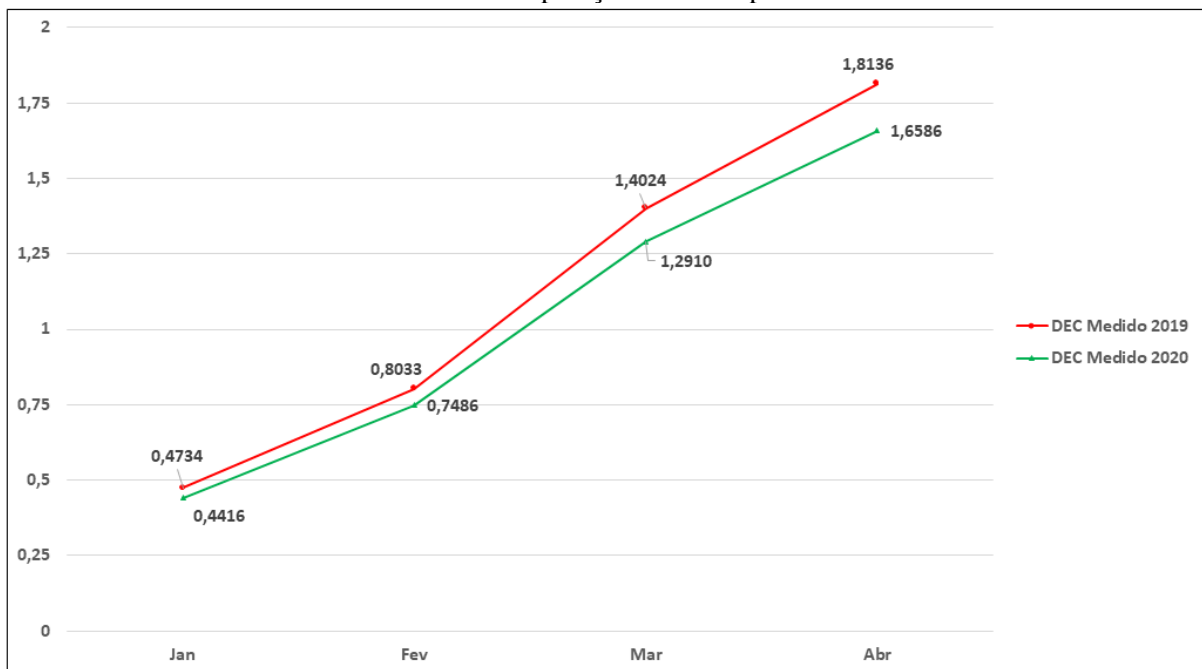
Após o término do planejamento de anomalias, encaminha-se para as equipes as programações dos serviços para que realizem as intervenções preventivas, que tem como objetivo principal podar o excesso de vegetação que se encontra sob a rede de MT. Tomando como base toda a pesquisa os gestores de manutenção passaram a saber o ponto exato onde atuar e qual causa estava impactando no sistema elétrico fazendo com que houvesse uma melhoria nos indicadores de continuidade DEC e FEC.

Após o desenvolvimento e aplicação da gestão da manutenção para o sistema de distribuição elétrica, foi possível comparar o cenário dos indicadores de continuidade do ano 2019 com o ano de 2020. Nesse sentido, com essa comparação, é possível verificar o desempenho de ambos os cenários e da mesma forma sugerir uma conclusão sobre o sistema de gestão proposto.

Logo, para desenvolvimento e comparação de ambos os cenários se tomou como base o DEC e FEC acumulado medido dos quatro primeiros meses de ambos os anos. Em que, é ilustrado a comparação do desempenho do sistema de gestão proposto no ano de 2020 com relação ao cenário que estava no ano de 2019.

No Gráfico 10, é apresentado a melhoria do indicador de continuidade DEC, ou seja, a partir do momento que começou-se usar o sistema de gestão proposto houve um ganho em relação ao cenário que não o utilizava, representando uma redução acumulada no DEC em 8,54%.

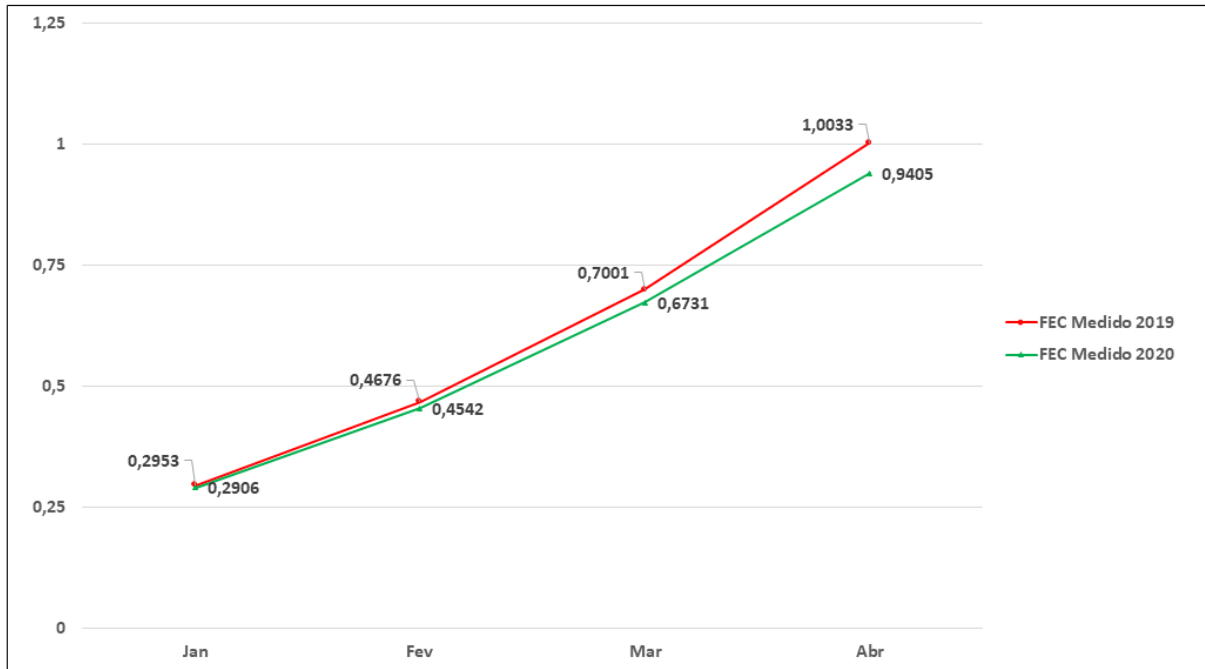
Gráfico 10 – Comparação de desempenho DEC



Fonte: O Autor (2021)

Seguindo com as análises comparativas, agora por perspectiva do indicador de continuidade FEC, é ilustrado no Gráfico 11. No qual, é demonstrado uma redução acumulada significativa no FEC de 6,26% em relação ao cenário que estava no ano de 2019.

Gráfico 11 – Comparação de desempenho FEC



Fonte: O Autor (2021)

Eventualmente, torna-se notório o desempenho do sistema de gestão proposto, permitindo um melhor desempenho nos indicadores de continuidade DEC e FEC.

6 CONSIDERAÇÕES FINAIS

Neste capítulo serão apresentadas as considerações finais relacionadas a esta pesquisa. Onde serão expostas as contribuições da pesquisa e as sugestões para trabalhos futuros.

6.1 CONTRIBUIÇÕES DA PESQUISA

Um sistema de distribuição elétrica está vulnerável a falhas. As falhas afetam as concessionárias que devem garantir boa qualidade no fornecimento de energia. Estudos e análises para obtenção de melhoria de confiabilidade, com o menor custo possível e bons indicadores de desempenho são essenciais para o planejamento de manutenção das concessionárias de energia. Portanto, a atuação das empresas distribuidoras de energia elétrica precisa ser efetiva no planejamento de anomalias.

No primeiro experimento desse estudo foi descoberto os fatores associados que contribuem para todas as falhas na rede de distribuição elétrica. Para conseguir isso, foi obtido um conjunto de regras de associação com limite mínimo de suporte definido como 0,01 e o limite mínimo de confiança como 0,7, essas configurações geraram 23 regras, dentre essas a mais significativa é {dia crítico interno int} => {descargas atmosféricas}. De posse dessa informação os gestores de manutenção identificaram a inexistência de conjunto de para-raios a cada 3 km de rede de MT e nos pontos de mudança de seção de condutores da rede de MT, descumprindo a Instrução de Trabalho no.60/2018 (IT0060). Caso a IT0060 estivesse sendo respeitada as falhas associadas a descarga atmosféricas teriam impactado os indicadores de continuidade.

A presente dissertação também apresentou uma metodologia capaz de identificar uma correlação nas falhas relacionadas a cada causa PEM associados aos pontos críticos e de maior impacto no sistema de distribuição elétrica. Desta forma, foi possível direcionar as intervenções preventivas sob condição para os trechos de maior impacto de uma determinada região. A técnica usada para analisar os dados foi a mineração de regras de associação.

Foi analisado um sistema de distribuição elétrica real, com o objetivo de demonstrar a capacidade da metodologia proposta de encontrar as soluções condizentes do ponto de vista técnico. Os resultados apontam uma redução dos indicadores de continuidade, colaborando assim na eficaz gestão da manutenção para o sistema de distribuição elétrica e na minimização dos prejuízos associados às interrupções de energia.

Através da metodologia proposta nesta dissertação, foi possível um direcionamento mais assertivo tanto para o planejamento quanto para a execução das ações de manutenções nas redes de distribuição. Analisando os resultados encontrados, o modelo proposto obteve desempenho melhor quando comparado com o desempenho da configuração atual, houve a redução dos indicadores de continuidade DEC e FEC de 8,54% e 6,26% respectivamente, tudo isso atrelado apenas a causa PEM vegetação.

6.2 TRABALHOS FUTUROS

Dessa forma, a metodologia adotada nesta pesquisa, bem como a gestão da manutenção para o sistema de distribuição elétrica que foi desenvolvido, mostrou-se adequados para auxiliar a compreender as causas de falhas de MT no sistema de distribuição de energia, usando a Mineração de Regras de Associação. Principalmente pelos resultados que foram encontrados e comprovados através da pesquisa.

Sugere-se, entretanto, recomendações futuras que podem auxiliar no planejamento e na gestão das falhas:

- Aplicação de uma nova metodologia abrangendo todas as causas PEM fazendo a combinação com outras técnicas de aprendizado de máquina;
- Execução de um plano de ação a curto, médio e longo prazos considerando o histórico de ocorrência de falhas e identificando os pontos considerados críticos sempre levando em consideração as causas PEM;
- Capacitação do gestores e técnicos responsáveis com relação à operabilidade da metodologia aplicada, com o intuito de expandir para outras áreas.

REFERÊNCIAS

- ABRADEE. Setor Elétrico. www.abradee.org.br, 2021. Disponível em: <<https://www.abradee.org.br/setor-eletrico/redes-de-energia-eletrica/>>. Acesso em: 2021.
- Agência Nacional de Energia Elétrica – ANEEL. Resolução 024 Disponível em: <<https://www.aneel.gov.br/>>. Acesso em 2021.
- Aria, M. & Cuccurullo, C. (2017) *bibliometrix: An R-tool for comprehensive science mapping analysis*, Journal of Informetrics, 11(4), pp 959-975, Elsevier.
- Bastos, Pedro; Lopes, Isabel; Pires, Luís (2012). **A maintenance prediction system using data mining techniques**. In World Congress on Engineering. London, p. 1448-1453. ISBN 978-988-19252-2-0
- BERNARDO, N. **Evolução da Gestão da Qualidade de Serviço de Energia**. Repositório da Politécnica, 2013. Disponível em: <<http://repositorio.poli.ufrj.br/monografias/monopoli10006255.pdf>>. Acesso em: 2021.
- Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). *Lessons from applying the systematic literature review process within the software engineering domain*. *Journal of Systems and Software*, 80(4), 571–583. doi:10.1016/j.jss.2006.07.009
- C. Silva and M. Saraee, "**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), Genova, Italy, 2019, pp. 1-6.
- 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). doi:10.1109/icphm.2016.7542830
- Cunha, R. F. C. e “**Operação do Sistema Elétrico de Distribuição do Estado do Ceará**”, Universidade Federal do Ceará – UFC, 2011, 70p.

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

Enel Distribuição Ceará. OPM-OMBR-OeM-18-0100-EDCE - Coleta e Apuração dos Indicadores de Atendimento Emergencial. Disponível em: <<https://www.eneldistribuicao.com.br/CE/NormasTecnicas.aspx>>. Acesso em: 2021

GIL, A. C. **Como elaborar projetos de pesquisa/Antônio Carlos Gil**. - 4. ed. - São Paulo: Atlas, 2002

Gupta, S., Rajiah, P., Middlebrooks, E. H., Baruah, D., Carter, B. W., Burton, K. R., ... Miller, M. M. (2018). *Systematic Review of the Literature: Best Practices. Academic Radiology*. doi:10.1016/j.acra.2018.04.025

JORDAN, M. I.; MITCHELL, T. M. **Machine learning: Trends, perspectives, and prospects. Science**, v. 349, n. 6245, p. 255–260, 2015.

Kitchenham, B. & Charters, S. (2007), *Guidelines for performing Systematic Literature Reviews in Software Engineering*. 2.3 ed. UK: Keele University and Durham University: EBSE Technical Report, 2007. v. 2

L. Xu and Mo-Yuen Chow, "**A classification approach for power distribution systems fault cause identification**," in *IEEE Transactions on Power Systems*, vol. 21, no. 1, pp. 53-60, Feb. 2006, doi: 10.1109/TPWRS.2005.861981.

L. Xu, M. Chow and L. S. Taylor, "**Power Distribution Fault Cause Identification With Imbalanced Data Using the Data Mining-Based Fuzzy Classification E-Algorithm**," in *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 164-171, Feb. 2007, doi: 10.1109/TPWRS.2006.888990.

LANDEGREN, F.E.; JOHANSSON, J.; SAMUELSSON, O. **A method for assessing marginand sensitivity of electricity networks with respect to repair systemresources**. *IEEE Trans. Smart Grid* 7 (November) (2016) 2880–2889.

Li, Lee & Yong, Cheng & Long-Jun, Xie & Li-Qiu, Jiang & Ning, Ma & Ming, Lu. (2015). **An Integrated Method of Set Pair Analysis and Association Rule for Fault Diagnosis of Power Transformers An Integrated Method of Set Pair Analysis and Association Rule for Fault**

Diagnosis of Power Transformers. IEEE Transactions on Dielectrics and Electrical Insulation.

LIMA, Fernando Parra dos Anjos. **Análise de distúrbios de tensão em sistemas de distribuição de energia elétrica baseada em sistemas imunológicos artificiais.** 2013. 169 f. Dissertação (mestrado) - Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Engenharia de Ilha Solteira, 2013. Disponível em: <<http://hdl.handle.net/11449/87159>>.

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. doi:10.1016/j.procir.2019.04.099

LOUT, K. (2015). **Development of a fault location method based on fault induced transients in distribution networks with wind farm connections** (Unpublished master's thesis). University of Bath.

M. Doostan and B. H. Chowdhury, "Power distribution system equipment failure identification using machine learning algorithms," 2017 IEEE Power & Energy Society General Meeting, 2017, pp. 1-5, doi: 10.1109/PESGM.2017.8274109.

M. Tian et al., "Data Dependence Analysis for Defects Data of Relay Protection Devices Based on Apriori Algorithm," in IEEE Access, vol. 8, pp. 120647-120653, 2020, doi: 10.1109/ACCESS.2020.3006345.

Machado Junior, C., Souza, M. T. S. de Parisotto, I. R. D. S., & Palmisano, A. (2016). *As Leis da Bibliometria em Diferentes Bases de Dados Científicos.* Revista de Ciências Da Administração, 1(1), 111. doi:10.5007/2175-8077.2016v18n44p111

MARCONI, M. A.; LAKATOS, E. M. **Metodologia do trabalho científico: projetos de pesquisa / pesquisa bibliográfica/ teses de doutorado, dissertações de mestrado, trabalhos de conclusão de curso.** – 8. ed. – São Paulo: Atlas, 2017.

R CORE TEAM. **R: A language and environment for statistical computing** Vienna, Austria **R Foundation for Statistical Computing**, 2021. Disponível em: <<https://www.r-project.org/>> Acesso em: 12 abr. 2021.

R. Agrawal and R. Srikant. **Fast Algorithms for Mining Association Rules.** In Proc. of the 20th Int'l Conference on Very Large Databases, Santiago, Chile, September 1994.

RAMASUBRAMANIAN, K.; SINGH, A. **Machine Learning Using R: With Time Series and Industry-Based Use Cases in R**. 2a ed. New Delhi: Apress, 2017.

RAVI, N.N.; DRUS, S.M.; KRISHNAN, P.S.; GHANI, N.L.A. **Substation Transformer Failure Analysis Through Text Mining**. 2019 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE).

REZIG, S.; ACHOUR, Z.; REZG, N. (2018). **Using Data Mining Methods for Predicting Sequential Maintenance Activities**. Applied Sciences. 8. 2184. 10.3390/app8112184.

Rongxi Wang, Zezhou Tang, Jianmin Gao, Zhiyong Gao & Zhen Wang (2020): **Probabilistic model-checking based reliability analysis for failure correlation of multi-state systems**, Quality Engineering, DOI: 10.1080/08982112.2019.1692139

Rongxi Wang, Zezhou Tang, Jianmin Gao, Zhiyong Gao & Zhen Wang (2020): **Probabilistic model-checking based reliability analysis for failure correlation of multi-state systems**, Quality Engineering, DOI: 10.1080/08982112.2019.1692139

S. Kazemi, M. Lehtonen and M. Fotuhi-Firuzabad, "**Impacts of Fault Diagnosis Schemes on Distribution System Reliability**," in IEEE Transactions on Smart Grid, vol. 3, no. 2, pp. 720-727, June 2012, doi: 10.1109/TSG.2011.2176352.

SAMUEL, A. L. **Some Studies in Machine Learning Using the game of Checkers**. IBM Journal, v. 3, n. 3, p. 534–554, 1959.

SHALEV-SHWARTZ, S.; BEN-DAVID, S. **Understanding machine learning: From theory to algorithms**. [s.l: s.n.]. v. 9781107057

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. doi:10.1109/tsg.2016.2562123

SILVA, E. L. DA; MENEZES, E. M. **Metodologia da Pesquisa e Elaboração de Dissertação**. 4. ed. Florianópolis: UFSC, 2005

SOUZA, F. A. **Detecção de falhas em sistemas de distribuição de energia elétrica usando dispositivos programáveis**. 2008. Dissertação (Mestrado em Engenharia Elétrica) –Faculdade de Engenharia, Universidade Estadual Paulista, Ilha Solteira, 2008.

TONELLI NETO, Mauro de Souza. **Desenvolvimento de um sistema inteligente para o diagnóstico e prognóstico de falhas em sistemas de distribuição de energia elétrica.** 2015. 111 p. Tese (doutorado) - Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Engenharia, 2015. Disponível em: <<http://hdl.handle.net/11449/126378>>.

TURBAN, E.; ARONSON, J. E.; LIANG, T.-P. **Decision Support System and Intelligent System.** Prentice Hall, 7th ed., 2005.

TURRIONI, J. B.; MELLO, C. H. P. **Metodologia de pesquisa em engenharia de produção: estratégias, métodos e técnicas para condução de pesquisas quantitativas e qualitativas.** Itajubá: UNIFEI, 2012.

WIPO. WIPO Technology Trends 2019: Artificial Intelligence. Genebra, Suíça: World Intellectual Property Organization, 2019.

Xinchun Jie, Haikuan Wang, Minrui Fei, Dajun Du, Qing Sun, T.C. Yang, **Anomaly behavior detection and reliability assessment of control systems based on association rules,** International Journal of Critical Infrastructure Protection (2018), doi: 10.1016/j.ijcip.2018.06.001