

DIOGO SLOVINSKI BOFF

**AVALIAÇÃO DA ASSERTIVIDADE DE
PROJETISTAS EM RELAÇÃO À PREVISÃO
DO IMPACTO DOS DESLIGAMENTOS**

São Leopoldo

2022

DIOGO SLOVINSKI BOFF

**AVALIAÇÃO DA ASSERTIVIDADE DE PROJETISTAS
EM RELAÇÃO À PREVISÃO DO IMPACTO DOS
DESLIGAMENTOS**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica da Universidade do Rio dos Sinos - UNISINOS.

UNIVERSIDADE DO VALE DO RIO DOS SINOS - UNISINOS
UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA
NÍVEL MESTRADO PROFISSIONAL

Orientador: Prof. Dr. Rodrigo Marques de Figueiredo

São Leopoldo

2022

B673a Boff, Diogo Slovinski.
Avaliação da assertividade de projetistas em relação à
previsão do impacto dos desligamentos / Diogo Slovinski
Boff. – 2022.
68 f. : il. ; 30 cm.

Dissertação (mestrado) – Universidade do Vale do Rio dos
Sinos, Programa de Pós-Graduação em Engenharia Elétrica,
2022.
“Orientador: Prof. Dr. Rodrigo Marques de Figueiredo.”

1. Distribuição de energia. 2. Obras. 3. Expansão.
4. Manutenção. 5. Modelagem. 6. Técnico de projetos. I. Título.

CDU 621.3

Dados Internacionais de Catalogação na Publicação (CIP)
(Bibliotecária: Amanda Schuster – CRB 10/2517)

ERRATA

Após a entrega deste trabalho houve atualização da Resolução 414 e o Prodist, passando a vigorar a Resolução 1000 e a Resolução 956 respectivamente.

DIOGO SLOVINSKI BOFF

**AVALIAÇÃO DA ASSERTIVIDADE DE PROJETISTAS
EM RELAÇÃO À PREVISÃO DO IMPACTO DOS
DESLIGAMENTOS**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica da Universidade do Rio dos Sinos - UNISINOS.

Trabalho aprovado. São Leopoldo, 24 de fevereiro de 2022:

**Professora Doutora Alzenira da Rosa
Abaide**

**Professor Doutor Paulo Ricardo da
Silva Pereira**

Professor Doutor Lucio Rene Prade

**Prof. Dr. Rodrigo Marques de
Figueiredo**

São Leopoldo
2022

*Este trabalho é dedicado àqueles que,
não contentes com sua condição, buscaram fazer o que parecia impossível.*

AGRADECIMENTOS

Agradeço primeiramente a Deus, que tem cuidado de minha vida de maneira extraordinária. Mesmo diante de minhas misérias, nunca me abandonou. Agradeço a minha esposa Francine pelo incentivo nos dias ruins, compreensão nos dias difíceis e aos meus filhos, inspiração para seguir em frente e fazer o melhor trabalho que minha circunstância permite.

Ao Professor Doutor Rodrigo Marques de Figueiredo, com sua gratuidade e conhecimento meus sinceros agradecimentos. Você tem sido para mim muito mais que um orientador, um exemplo. Seu zelo e cuidados despertaram em mim o desejo de ensinar, de passar adiante o que sei. De deixar na vida das pessoas um pouco de mim.

Os autores gostariam de agradecer ao grupo CPFL pelo apoio técnico e financeiro, através do projeto de Pesquisa e Desenvolvimento PD-00396-3057/2019 - "PA3057 - Metodologia de Planejamento de Projetos de Redes de Distribuição Otimizando o DEC Programado com a Aplicação de Inteligência Artificial" com recursos do programa de P&D ANEEL.

RESUMO

Os contratos de concessão das distribuidoras de energia exigem entre outras obrigações, a manutenção dos níveis de qualidade do fornecimento bem como da continuidade destes. Para isso, as companhias precisam constantemente manter e expandir o sistema de modo que o mesmo atenda as metas estipuladas. No estado do Rio Grande do Sul a rede de distribuição elétrica possui aproximadamente 40% dos postes de madeira e ainda diversos municípios e centros de carga com apenas uma fonte de alimentação. Há ainda áreas em constante expansão econômica que possuem grande demanda energética.

Diante deste cenário, as distribuidoras precisam anualmente realizar investimentos no sistema através de obras. Estas por sua vez, demandam grande quantidade de material, além de desligamentos programados que ferem os mesmo indicadores que precisam ser melhorados. Esta equação de difícil equilíbrio, passa por um bom projeto e planejamento de obras.

Este trabalho, terá seu escopo no impacto que o Técnico de projetos tem nas obras e nos indicadores de continuidade. Na etapa de projetos será analisado se uma determinada obra é viável do ponto de vista dos indicadores de qualidade e continuidade.

Através da criação de uma ferramenta de modelagem do técnico de projetos será previsto o impacto dos desligamentos programados e o desligamento em si, na execução da obra. como apresentado na revisão sistemática, a ausência de ferramentas computacionais para o planejamento de obras torna o fator humano determinante. Desta forma, esta ferramenta busca modelar o desempenho do técnico de projetos e sua influência no correto planejamento de obras. Pretende-se avaliar se os técnicos estão com nível satisfatório de acordo com sua vivência e dificuldade das atividades.

A modelagem do Técnico de Projetos foi feita utilizando os algoritmos SVM e Redes Neurais. Sendo o SVM *Support Vector Machine* o que trouxe melhor resultado, melhorando 9% a acurácia em relação aos resultados atuais.

Como trata-se de um projeto pioneiro, este trabalho abre caminhos para maior exploração do tema, com incorporação de novas variáveis estima-se que se consiga atingir uma acurácia de até 20%.

Palavras-chave: Distribuição de Energia, Obras, Expansão, Manutenção, Modelagem, Técnico de Projetos.

ABSTRACT

The energy distributors' concession contracts require, among other obligations, the maintenance of supply quality levels as well as their continuity. For this, companies need to constantly maintain and expand the system so that it meets the stipulated goals. In the state of Rio Grande do Sul, the electrical distribution network has approximately 40% of wooden poles and also several municipalities and a load center with only one power supply. There are also areas in constant economic expansion that have high energy demand.

Given this scenario, distributors need to annually invest in the system through works. These, in turn, demand a large amount of material, in addition to scheduled shutdowns that hurt the same indicators that need to be improved. This difficult balance requires a good project and planning of works.

This work will have its scope in the impact that the Project Technician has on the works and on the continuity indicators. In the design stage, it will be analyzed whether a particular work is viable from the point of view of quality and continuity indicators.

Through the creation of a modeling tool for the project technician, the impact of scheduled shutdowns and the shutdown itself, in the execution of the work, will be foreseen. The human factor is currently decisive for the good planning of the works. In this way, this tool seeks to model the performance of the project technician and its influence on the correct planning of works. It is intended to assess whether the technicians are at a satisfactory level according to their experience and difficulty of activities.

The Project Technician's modeling was done using the SVM and Neural Networks algorithms. Being the SVM the one that brought the best result with 54% of accuracy. This value represents a gain of 9% in relation to the current results.

As this is a pioneering project, this work opens the way for further exploration of the topic, with the incorporation of new variables, it is estimated that an accuracy of up to 20%.

Keywords: Power Distribution, Construction, Expansion, Maintenance, Modeling, Project Technician.

LISTA DE ILUSTRAÇÕES

Figura 1 – Ciclo de Melhoria dos Indicadores	24
Figura 2 – Meta e indicador DEC e FEC	31
Figura 3 – Tipos de Variáveis	37
Figura 4 – Processo de filtro	44
Figura 5 – Convergência	46
Figura 6 – Matriz de Calor	47
Figura 7 – Matriz de Calor Objetivo	48
Figura 8 – Field	49
Figura 9 – Estatística	50
Figura 10 – Processo de Modelagem	51
Figura 11 – Pesquisa exploratória	52
Figura 12 – Precision e Recall	54
Figura 13 – Base de Dados	55
Figura 14 – Dados da Companhia	56
Figura 15 – Exemplo de dados	57
Figura 16 – Matriz de Confusão SVM Adaptada	59
Figura 17 – Matriz de Confusão RNA Adaptada	60
Figura 18 – Precisão Por tipo de Atividade de Manutenção	63
Figura 19 – Precisão Por tipo de Atividade de Manutenção	64

LISTA DE TABELAS

Tabela 1 – String de Modelagem de Sistemas	42
Tabela 2 – String de Pesquisa Operacional	43
Tabela 3 – String de Sistema Elétrico de Potência	43
Tabela 4 – Resumo dos Principais Artigos Estudados	45
Tabela 5 – Adaptado de (ZHENG, 2015)	53
Tabela 6 – Hiper Parâmetros Utilizados no Algoritmo SVM	58
Tabela 7 – Hiper Parâmetros Utilizados na RNA	59

SUMÁRIO

1	INTRODUÇÃO	21
1.1	Tema	21
1.2	Delimitação do tema	21
1.3	Problema	22
1.4	Objetivos	22
1.4.1	Objetivo geral	23
1.4.2	Objetivo específico	23
1.5	Justificativa	23
1.6	Estrutura da Dissertação	25
2	O AMBIENTE REGULADO DA DISTRIBUIÇÃO DE ENERGIA	27
2.1	Contrato de concessão	27
2.2	Indicadores da Distribuição	28
2.2.1	Qualidade do Produto	28
2.2.2	Qualidade do serviço	29
2.2.3	Limites de Continuidade	31
2.2.4	Compensações	31
2.3	Expansão do Sistema de distribuição	32
2.3.1	Procedimento para Elaboração dos Estudos de Previsão de Demanda no SDMT	32
2.4	Resolução Normativa 414	33
3	TÉCNICAS DE MODELAGEM	35
3.1	Conceitos	35
3.2	Tipos de Algoritmo de classificação	37
3.3	Avaliação de algoritmos	38
4	ESTADO DA ARTE	39
4.1	Processo de Busca	42
4.2	Filtros de Estudo	44
4.3	Resultados da pesquisa	45
4.4	Técnica de Otimização foi Utilizada	45
4.4.1	Questão Geral 1 (GQ-1): Técnica de otimização foi utilizada	47
4.4.2	Questão Geral 2 (GQ-2): Aspectos (objetivos e restrições) foram otimizados	48
4.4.3	Questão Específica 1 (SQ-1): Campo de trabalho da pesquisa	48
4.4.4	Questão específica 2 (SQ-2): Conclusão do estudo. Solução do problema	49

4.5	Publicações Anuais	49
5	METODOLOGIA	51
5.1	Pré Processamento	51
5.2	Seleção do Algoritmo de Aprendizado	52
5.3	Avaliação de modelos	53
5.4	Previsão	54
6	IMPLEMENTAÇÃO E ANÁLISE DE RESULTADOS	55
6.1	Base de Dados	55
6.2	Escolha das variáveis	56
6.3	Tratamento dos Dados	58
6.4	Implementação do SVM	58
6.5	Implementação de RNA	59
7	DISCUSSÃO E CONCLUSÃO	61
	REFERÊNCIAS	65

1 INTRODUÇÃO

No Brasil, o setor elétrico é administrado por empresas que possuem outorga de explorar este serviço e sua remuneração é provinda da tarifa paga pelo consumidor. Para tal, é necessário que estas empresas cumpram metas estipuladas em seus contratos e são regulados pela Agência Nacional de Energia Elétrica (ANEEL). Segundo a ANEEL (2021a), é obrigação da concessionária adotar serviços e tecnologia adequados para garantir níveis de regularidade, continuidade, eficiência, segurança na prestação dos serviços. Além disto, é obrigação das concessionárias realizar os projetos e obras necessárias ao fornecimento de energia elétrica aos interessados, até o ponto de entrega.

A concessionária obriga-se a manter ou melhorar os níveis de qualidade do fornecimento de energia elétrica, de acordo com os critérios, indicadores, fórmulas e parâmetros definidores da qualidade do serviço, nos termos da legislação em vigor.(ANEEL, 2021a) Caso haja regressão desta obrigação, a concessionária estará sujeita a multas.

É obrigação ainda da concessionária, implantar novas instalações, ampliar e modificar as existentes de modo a garantir o atendimento atual e futuro, observando as normas e recomendações dos órgãos gerenciados do Sistema Elétrico Nacional.(ANEEL, 2021a) As ampliações devem obedecer aos procedimentos legais específicos e às normas do Poder Concedente.

1.1 Tema

Diante do exposto, este trabalho se propõe a estudar uma forma aceitável de atender os requisitos de prover projetos e obras que sejam capazes de concomitantemente atender a obrigação de manter ou melhorar os níveis de qualidade, ampliar e modificar as instalações existentes e garantir os níveis de regularidade, continuidade, eficiência e segurança na prestação do serviço.

1.2 Delimitação do tema

Para manter a qualidade dos serviços, faz-se necessária a manutenção e expansão do Sistema Elétrico de Potência (SEP). Sabe-se que qualquer ativo sofre depreciação ao longo dos anos e além disto, a expansão civilizacional aumenta a densidade demográfica das cidades ou o área urbana e como descrito acima, é obrigação das concessionárias dar suporte a este crescimento e manutenção do sistema.

Para tal, as empresas precisam executar obras de expansão ¹, manutenção e atendimento ao cliente. Neste estudo, a empresa em questão, tem como prática, prover projetos para todas estas obras.

Este trabalho terá seu escopo no impacto que o Técnico de projetos tem nas obras e nos indicadores de continuidade. Esta análise será feita ainda na etapa de projetos, onde é possível verificar se uma determinada obra é viável do ponto de vista dos indicadores de continuidade.

1.3 Problema

Após alguns ciclos de avaliação do resultado dos indicadores de continuidade, qualidade, etc. percebeu-se que um projeto bem executado, é fator determinante no atingimento ou não das metas da companhia. Sejam elas financeiras ou regulatórias. Além disto, um planejamento adequado das obras, permite que os compromissos assumidos perante o regulador possam ser cumpridos.

Embora algumas obras possam ser executadas sem desenergizar a rede elétrica, a grande maioria destas precisa que os clientes sejam desligados para sua execução e desta forma, a continuidade do serviço é afetada. Além disso, um planejamento incorreto faz com que seja necessário desligar mais de uma vez o mesmo cliente para execução de obras.

Diante disto, uma análise do impacto que a obra irá gerar no cliente passa pela área de projeto. Um dos protagonistas do bom desempenho de obras é o técnico de projetos. Acontece que, assim como ocorre em todas as atividades econômicas, fatores como experiência, tempo de serviço, complexidade do trabalho,... são determinantes para a eficiência de um profissional. Desta forma, medir a capacidade do técnico em fazer uma boa análise é importante para prever a sua capacitação. Existem outros fatores que são influenciadores da tomada de decisão como capacidade da empresa executora, condições climáticas, tipo de obra, etc. Evoluir as análises para que haja a garantia de que os projetos possuam um impacto otimizado nos indicadores e assegurem o suporte para os investimentos em expansão e manutenção do sistema torna-se uma questão de suma importância.

1.4 Objetivos

Propõe-se neste trabalho criar uma ferramenta de avaliação da assertividade do técnico de projetos em relação à previsão do impacto dos desligamentos programados e o

¹ As obras de expansão geralmente ocorrem na rede de energia primária convencionalmente chamamos de média tensão. Estas obras são chamadas de Obras Primárias

desligamento em si, na execução da obra. O fator humano, atualmente é determinante para o bom planejamento das obras.

1.4.1 Objetivo geral

O estudo será feito em uma distribuidora de energia do estado do Rio Grande do Sul. Nesta empresa, a área de projetos elabora em torno de duas mil atividades por mês e conta com mais de sessenta técnicos trabalhando. Há uma diversidade de pessoas que atuam na elaboração destes projetos. A área possui três níveis de técnicos: júnior, pleno e sênior. A complexidade do trabalho, é organizada de modo que cada pessoa receba atividades compatíveis com seus conhecimentos porém não há uma divisão por região, sendo que todos fazem projetos para as mais de trinta regiões de abrangência.

Desta forma, buscar-se-á modelar o desempenho do técnico de projetos e sua influência no correto planejamento de obras. Pretende-se com isto, avaliar se os técnicos estão com nível satisfatório de acordo com sua vivência e dificuldade das atividades.

1.4.2 Objetivo específico

Para resolver esta tarefa, serão investigados os dados disponíveis no histórico da distribuidora, comparando características das obras e dos técnicos que elaboraram os projetos. Serão submetidas estas variáveis à modelagem computacional para buscar encontrar qual é o algoritmo que mais se assemelha ao desempenho do técnico buscando:

- Definir variáveis relevantes.
- Avaliar a influência das variáveis no resultado
- Avaliar a influência do nível técnico no resultado
- Avaliar a relação do técnico sobre o resultado
- Avaliar qual é o algoritmo de predição que melhor descreve o comportamento do técnico.
- Avaliar qual é o nível de assertividade deste.

1.5 Justificativa

Diante do exposto, ficam claras as obrigações que empresas de distribuição devem satisfazer. É mandatório para o cumprimentos dos contratos, que obras de manutenção, melhoria, expansão e atendimento ao cliente sejam executadas. Tais obras, precisam ser planejadas de modo que não prejudiquem os indicadores de continuidade e qualidade do

serviço. O primeiro passo ocorre na etapa de projetos onde a estratégia de execução deve ser elaborada no sentido de otimizar o desligamento programado, impactando a menor quantidade possível de clientes.

A área de Obras e Manutenção (O&M), deve garantir que o plano de investimentos da companhia seja atendido, ou seja, que haja a execução física e financeira dos compromissos assumidos nos ciclos anuais. O grande desafio é equilibrar a balança entre investimento e DEC. Pois para melhorar os indicadores é necessário deixar a rede mais robusta com a execução de obras e para fazer estas, é preciso desligar os clientes. A título de simplificarmos os cálculos iremos trabalhar com a dimensão Cliente Hora Interrompido (CHI), que é uma derivação do DEC. A Figura 1 Ilustra este ciclo.

Figura 1 – Ciclo de Melhoria dos indicadores



Fonte: Autor

Fica evidente que é preciso otimizar o impacto das obras para a sustentabilidade do negócio. Não foi mencionada aqui a perda financeira que estes desligamentos provocam no faturamento das empresas. Isto se dá pela complexidade do cálculo e por fugir do escopo deste trabalho. Mas pode-se assegurar que são relevantes e podem inclusive prejudicar o resultado financeiro das empresas.

Após esta breve introdução, busca-se com este trabalho, avaliar se o plano de execução elaborado está de acordo com o executado. Mais especificamente, é foco desta pesquisa entender o nível de assertividade do técnico de projetos na elaboração do planejamento de obras perante o que foi executado e avaliar sua assertividade através da

modelagem deste utilizando técnicas computacionais.

Com a modelagem do técnico de projetos, é possível avaliar o desempenho da equipe e inclusive prever o impacto que obras futuras deverão causar no sistema. Esta análise, na etapa de projetos é fundamental pois possibilita que ações de mitigação sejam tomadas além de auxiliar na melhoria de análise e *feedback* dos técnicos.

1.6 Estrutura da Dissertação

Esta dissertação está dividida em cinco partes. A introdução traz um breve nivelamento para que o leitor entenda o cenário ao qual esta pesquisa está inserida bem como quais objetivos devem ser atingidos. A revisão bibliográfica, traz uma explanação sobre o contexto dos termos utilizados na dissertação e está dividida entre o ambiente regulado da distribuição e as técnicas de modelagem. Nos referenciais teóricos, através de uma revisão sistemática, iremos buscar trabalho correlatos que corroborem no desenvolvimento técnico científico necessário para a solução deste problema. Na metodologia e implementação, como o título já diz, serão explanadas as técnicas de modelagem utilizadas e os resultados colhidos de sua implementação. Por fim, na conclusão, serão discutidos a relevância dos resultados e provocações para trabalhos futuros.

2 O AMBIENTE REGULADO DA DISTRIBUIÇÃO DE ENERGIA

A seguir, será apresentada uma revisão sobre os principais pontos a serem observados sobre o ambiente regulado da distribuição de energia elétrica. O texto busca entregar ao leitor, um subsídio para entender o escopo e o ambiente ao qual esta pesquisa está inserida.

2.1 Contrato de concessão

Segundo (ANEEL, 2021a), o serviço de distribuição de energia elétrica é público. Não obstante, cabe à união, conceder o direito de exploração deste a setor privado. Atualmente, a Rio Grande Energia, é detentora de grande parte da distribuição de energia neste estado. O mesmo contrato, na cláusula segunda, trata das condições de prestação dos serviços e especifica quais são as obrigações da empresa. Dentre estas, destacamos:

- "A CONCESSIONÁRIA obriga-se a adotar, na prestação dos serviços, tecnologia adequada e a empregar equipamentos, instalações e métodos operativos que garantam níveis de regularidade, continuidade, eficiência, segurança, atualidade, generalidade, cortesia na prestação dos serviços e a modicidade das tarifas."
- "A CONCESSIONÁRIA obriga-se a manter ou melhorar os níveis de qualidade do fornecimento de energia elétrica, de acordo com os critérios, indicadores, fórmulas e parâmetros definidores da qualidade do serviço, nos termos da legislação em vigor."
- "Pela inobservância dos índices de continuidade de fornecimento de energia elétrica estabelecidos nos regulamentos específicos para cada conjunto das áreas de concessão, bem como pela violação dos índices de qualidade do serviço relativos à tensão de fornecimento, ou de outros aspectos que afetem a qualidade do serviço de energia elétrica, a CONCESSIONÁRIA estará sujeita a multas pecuniárias, aplicadas pelo PODER CONCEDENTE, em favor dos consumidores afetados."
- "Sem prejuízo do cumprimento do estabelecido na legislação e no presente contrato, a CONCESSIONÁRIA obriga-se a manter e melhorar os atuais níveis de qualidade do fornecimento de energia elétrica globais e por conjunto, tendo os valores indicados nos Quadros 1 e 2 do ANEXO III como referência para os indicadores globais."

Além disso, no mesmo documento, a cláusula quarta cita que é de responsabilidade da concessionária implantar novas instalações e a ampliar e modificar as existentes, de

modo a garantir o atendimento atual e futura demanda de seu mercado de energia elétrica, observadas as normas e recomendações do Sistema Elétrico Nacional.

A cláusula quinta, prevê que a empresa deve realizar as obras necessárias à prestação dos serviços concedidos, reposição de bens, operando as instalações e equipamentos correspondentes, de modo a assegurar a regularidade, continuidade, eficiência, segurança. Além disso, a concessionária precisa informar por escrito ao interessado as condições para a execução das obras e o prazo de sua conclusão.

2.2 Indicadores da Distribuição

Segundo ANEEL (2021c), a qualidade percebida pelo cliente de uma concessionária ou permissionária de serviço público de distribuição de energia elétrica pode ser medido a partir de três grandes aspectos:

- Qualidade do "produto" energia elétrica (relacionada à conformidade da tensão em regime permanente e à ausência de perturbações na forma de onda);
- Qualidade do "serviço" (relacionada à continuidade na prestação do serviço);
- Qualidade do atendimento ao consumidor.

No módulo 8 do Procedimento da Distribuição (PRODIST), encontramos as instruções relativas à qualidade da energia elétrica (QEE). Neste manual, são abordadas a qualidade do produto e a qualidade do serviço prestado, enquanto as Condições Gerais de Fornecimento encontram-se na Resolução Normativa 414/2010.

2.2.1 Qualidade do Produto

Segundo ANEEL (2021d) O PRODIST módulo 8, estabelece os procedimentos relativos à QEE, abordando a qualidade do produto e do serviço prestado e a qualidade do tratamento de reclamações. A tensão em regime permanente, pode ser dividida em três níveis: adequado, crítico e precário. Diante disto, segundo ANEEL (2021d), a tensão em regime permanente deve ser acompanhada em todo o sistema de distribuição, devendo a distribuidora dotar-se de recursos e técnicas modernas para tal acompanhamento. O termo "conformidade de tensão elétrica" refere-se à comparação do valor de tensão obtido por medição apropriada, no ponto de conexão. Os valores de tensão obtidos devem ser comparados à referência que deve ser a tensão nominal ou a contratada, de acordo com o nível de tensão do ponto de conexão. Para fins de registro, após seguir o procedimento de medição descrito em ANEEL (2021d) são calculados o índice de duração relativa da transgressão para tensão precária equivalente (DRP_E) e o índice de duração relativa da

transgressão para tensão crítica (DRC_E) conforme fórmulas abaixo:

$$DRP_E = \sum \frac{DRP_i}{N_L} [\%] \quad (2.1)$$

$$DRC_E = \sum \frac{DRC_i}{N_L} [\%] \quad (2.2)$$

em que:

DRP_i = duração relativa de transgressão de tensão precária individual da unidade consumidora (i);

DRC_i = duração relativa de transgressão de tensão crítica individual da unidade consumidora (i);

DRP_E = duração relativa de transgressão de tensão precária equivalente;

DRC_E = duração relativa de transgressão de tensão crítica equivalente;

N_L = total de unidades consumidoras objeto de medição.

Os limites para estes indicadores são: DRP é de 3% e DRC é de 0,5% e caso haja transgressão, a distribuidora deve pagar uma compensação ao cliente afetado. Este valor deve ser creditado na fatura apresentada no prazo máximo de dois meses subsequentes ao mês civil de referência da última medição que constou a violação.

2.2.2 Qualidade do serviço

A qualidade do serviço segundo (ANEEL, 2021d) é relacionada à continuidade no fornecimento de energia, ou seja, quanto tempo o cliente fica sem energia elétrica em determinado período. Os indicadores de continuidade são medidos de forma individual e em conjuntos. Os conjuntos elétricos são definidos por Subestação de Distribuição (SED). A abrangência destes, deve ser as redes de MT à jusante da SED e de propriedade da distribuidora. Os conjuntos devem possuir entre mil e dez mil consumidores. Os indicadores individuais de continuidade são apurados para todas as unidades consumidoras ou por ponto de conexão e são calculados como segue: A Duração de Interrupção Individual por Unidades Consumidoras ou por Ponto de Conexão (DIC), medido em horas e centésimos de hora é representada pela fórmula:

$$DIC = \sum_{i=1}^n t(i) \quad (2.3)$$

A Frequência de Interrupção Individual por Unidade Consumidora ou por Ponto de Conexão (FIC), expresso em número de interrupções é representada pela equação:

$$FIC = n \quad (2.4)$$

A Duração Máxima de Interrupção Contínua por Unidade Consumidora ou por Ponto de Conexão (DMIC), expresso em horas e centésimos de horas é representada utilizando a

seguinte fórmula:

$$DMIC = t(i)max \quad (2.5)$$

A Duração da interrupção Individual ocorrida em dia Crítico por unidade consumidora ou por ponto de conexão (DICRI) , expresso em horas e centésimos de horas é representada utilizando a seguinte fórmula:

$$DICRI = t_{crítico} \quad (2.6)$$

Onde:

i = índice de interrupções da unidade ocorrida em dia crítico por unidade consumidora ou por ponto de conexão no período de apuração de 1 a n ;

n = número de interrupções da unidade consumidora ou por ponto de conexão considerado no período de apuração;

$t(i)$ = tempo de duração da interrupção (i) da unidade consumidora considerada ou do ponto de conexão, no período de apuração;

$t(i)max$ = valor correspondente ao tempo da máxima duração de interrupção contínua (i), no período de apuração, verificada na unidade consumidora ou no ponto de conexão considerado;

$t_{crítico}$ = duração da interrupção em Dia Crítico.

Segundo ANEEL (2021d), os indicadores de continuidade de conjunto de unidades consumidoras, deverão ser apurados conforme segue: A Duração Equivalente de Interrupção por Unidade Consumidora(DEC), é representada por:

$$DEC = \frac{\sum_{i=1}^{C_c} DIC(i)}{C_c} \quad (2.7)$$

Já a Frequência Equivalente de Interrupção por Unidade Consumidora (FEC), é representada por:

$$FEC = \frac{\sum_{i=1}^{C_c} FIC(i)}{C_c} \quad (2.8)$$

Onde:

DEC = duração equivalente de interrupção por unidade consumidora, expressa em horas e centésimos de hora;

FEC = frequência equivalente de interrupção por unidade consumidora, expressa em número de interrupções e centésimo de número de interrupções;

i = Índice de unidades consumidoras atendida em BT ou MT faturadas do conjunto;

C_c = número total de unidades consumidoras faturadas do conjunto no período de apuração, atendidas em BT ou MT;

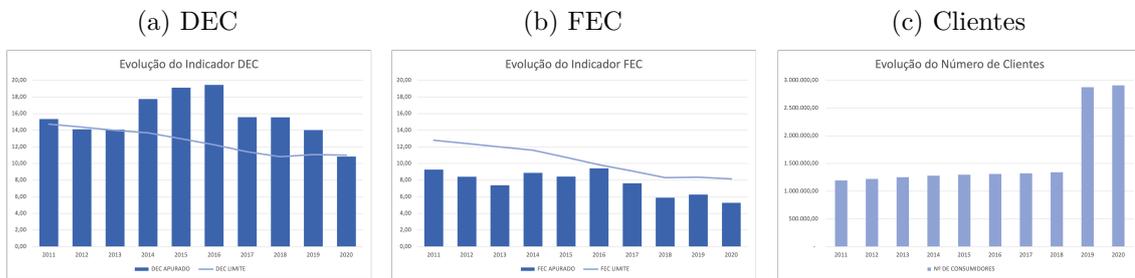
DIC(i) = Duração de interrupção Individual por Unidade Consumidora, excluindo-se as centrais geradores;

FIC(i) = Frequência de Interrupção Individual por Unidade Consumidora, excluindo-se as centrais geradoras.

2.2.3 Limites de Continuidade

A agência de regulação do serviço de energia elétrica edita limites para os indicadores coletivos de continuidade DEC e FEC. As distribuidoras precisam apurar estes indicadores periodicamente e enviá-los à ANEEL. A mesma disponibiliza esta informação em seu site. (ANEEL, 2021c) A Figura 2 Ilustra o acompanhamento da meta e da performance dos indicadores DEC e FEC da distribuidora em questão.

Figura 2 – Gráfico: Evolução da Meta e do Indicador DEC, FEC e Clientes



Fonte: (ANEEL, 2021c)

Podemos observar pelas Figuras 2a e 2b que as metas de DEC e FEC possuem uma curva decrescente ao longo dos anos (representada pela linha). Uma anomalia aconteceu no ano de 2019, onde a meta se manteve igual ao ano anterior. Isto ocorreu pois houve uma junção de duas companhias de energia. A Figura 2c ilustra isto com o expressivo aumento de consumidores. Praticamente o dobro do ano anterior. Outrossim, as metas de continuidade tendem a reduzir novamente nos anos subsequentes, observando a curva dos anos anteriores sendo mandatória a melhoria da confiabilidade do sistema, para trazer o indicador de continuidade para dentro da meta. Apesar de mostrar expressiva melhoria nos últimos quatro anos, o indicador DEC ainda está no limiar da meta.

2.2.4 Compensações

Caso haja comprovada transgressão na qualidade do produto ou serviço, a distribuidora deve compensar financeiramente o cliente.

No caso da qualidade do produto, caso haja transgressão do indicador $DRP > 3\%$ e ou $DRC > 0,5\%$, a mesma é calculada da seguinte forma conforme (ANEEL, 2021d):

$$Valor = \left[\left(\frac{DRP - DRP_{limite}}{100} \right) \cdot K_1 + \left(\frac{DRC - DRP_{limite}}{100} \right) K_1 \right] \cdot EUSD \quad (2.9)$$

sendo:

$K_1 = 0$, se $DRP \leq DRP_{Limite}$;

$K_1 = 3$, se $DRP > DRP_{Limite}$;

$K_2 = 0$, se $DRC \leq DRC_{Limite}$;

$K_2 = 7$, para consumidores atendidos em Baixa Tensão, se $DRC > DRC_{Limite}$;

$K_2 = 5$, para consumidores atendidos em Média Tensão, se $DRC > DRC_{Limite}$;

$K_2 = 3$, para consumidores atendidos em Alta Tensão, se $DRC > DRC_{Limite}$;

DRP = valor do DRP expresso em %, apurado na última medição;

$DRP_{limite} = 3\%$;

DRC = valor do DRC expresso em %, apurado na última medição;

$DRC_{limite} = 0,5\%$;

$EUSD$ = valor do encargo de uso do sistema de distribuição correspondente ao mês de referência da última medição. Cálculo

2.3 Expansão do Sistema de distribuição

Com já mencionado acima, a distribuidora tem por obrigação contratual expandir o sistema elétrico de forma a suportar o crescimento populacional e conseqüentemente de carga. Desta forma ANEEL (2021b) especifica as diretrizes para o planejamento da expansão. Este procedimento, estabelece os requisitos mínimos de informações necessárias para os estudos de planejamento do sistema de distribuição.

2.3.1 Procedimento para Elaboração dos Estudos de Previsão de Demanda no SDMT

A previsão de demanda do Sistema de Distribuição em Média Tensão - SDMT deve fornecer as informações necessárias ao planejamento das redes e linhas de distribuição. Desta forma, devem ser observados alguns requisitos como segue:

- Horizonte deve ser de cinco anos, com revisão anual;
- A previsão deve ter caráter espacial, associada às áreas mais representativas;
- A carga deve ser caracterizada, em patamares, para o período de ponta e fora-de-ponta, de acordo com as curvas de carga verificadas no SDMT em análise.
- Os pontos de interesse são as barras secundárias das subestações de distribuição, e a sua distribuição ao longo dos circuitos SDMT;

- A extensão ou reforço do sistema de distribuição para atendimento a novas áreas com baixa densidade de carga deve ser destacada e classificada como atendimento de mercado emergente; e
- A conexão de geração distribuída, considerada suas características de geração, disponibilidade e sazonalidade.

O planejamento da expansão do SMDT consiste na previsão de novos alimentadores para os próximos cinco anos e reforços nos existentes, de forma detalhada para as obras a serem realizadas nos próximos vinte e quatro meses. Com base na nova topologia da rede, deve-se realizar a avaliação das perdas técnicas, dos perfis de tensão e de carregamento e a estimativa da evolução da confiabilidade, comparando com a situação diagnosticada.

2.4 Resolução Normativa 414

Segundo consta em (ANEEL, 2021e), a resolução 414 de 9 de setembro de 2010 estabelece as condições gerais de fornecimento de energia elétrica de forma atualizada e consolidada. Este documento é escrito e atualizado por meio de consultas públicas a agentes do setor e da sociedade em geral.

No artigo trinta e dois são descritos os parâmetros para as solicitações de ligação do cliente que necessitam de obra para serem viabilizadas. A distribuidora tem um prazo de trinta dias para elaborar os estudos, orçamentos, projetos e informar por escrito ao cliente. O artigo trinta e três estipula os prazos de retorno do cliente com a modalidade escolhida para execução da obra. O mesmo pode optar por executar com recursos próprios em um prazo menor, ou aceitar o prazo informado pela distribuidora. No artigo trinta e quatro, são definidos os prazos de execução de obra conforme a característica das mesmas.

Ao final desta revisão dos principais conceitos relevantes do sistema elétrico pode-se concluir que a área de Obras e Manutenção precisa executar diversas atividades que são necessárias para cumprir os requisitos citados. Desta forma, é mandatório que estas obras "caibam" dentro do DEC e para isto, é preciso garantir que haja uma otimização deste recurso. Como visto acima, caso isto não ocorra, a empresa corre o risco de não cumprir suas obrigações com clientes ou não entregar os indicadores na meta. Uma forma de garantir isto é que a etapa de projetos, com foco na execução, preveja os impactos que este irá trazer ao sistema e, após uma análise, verificar se há otimização na quantidade de clientes desligados. Esta análise, pode ser feita pelo técnico de projetos que elaborou o plano de execução. Sendo uma análise manual, é necessário medir a assertividade desta em relação ao que foi executado realmente em campo. Uma forma de fazer isto, é modelar o projetista. Dada a variedade de obras, nível de conhecimento técnico dos profissionais e a disponibilidade de recursos computacionais. O estudo será conduzido utilizando técnicas

de modelagem baseadas em Aprendizado de Máquina. No capítulo três, teremos uma revisão sobre os principais temas.

3 TÉCNICAS DE MODELAGEM

Esta é a era do *Big Data*. Alpaydin (2014) afirma que atualmente, as pessoas são consumidores de dados e todas as vezes que compra-se algo na internet, aluga-se um filme por *streaming*, visita-se um site, escreve-se em um blog, faz-se um *post* ou simplesmente caminha-se por ai, há a geração de dados. Cada indivíduo não é apenas um gerador, mas um consumidor de dados. As pessoas buscam praticidade e experiências mais personalizadas. Desta forma, os serviços utilizam as informações de consumo, por exemplo para montar a ordem das prateleiras de supermercado. O comportamento das pessoas não é totalmente aleatório, há um padrão que é seguido por determinados grupos de pessoa. Ou seja, existem certo padrões de dados.

Para resolver um problema computacional, é necessário um algoritmo, para a mesma tarefa, podem haver vários algoritmos e porém deve-se encontrar o mais eficiente. Desta forma, pretende-se que a (máquina) extraia automaticamente o algoritmo para esta tarefa. Não há necessidade de aprender a classificar números, pois há algoritmos para isso, mas existem muitas aplicações para as quais não existe um algoritmo, porém tem-se muitos dados. Talvez não seja possível explicar o processo completamente, mas pode-se construir uma aproximação boa e útil. Em outras palavras, embora a identificação do processo completo não seja possível, é plausível detectar certos padrões ou regularidades. Este é o nicho do aprendizado de máquina, em inglês: (*Machine Learning*). Ainda neste sentido, Mohri Afshin Rostamizadeh (2012) afirma que o aprendizado de máquina pode ser amplamente definido como métodos computacionais que utilizam experiência para melhorar o desempenho ou para fazer previsões precisas. Esta precisão, é baseada na complexidade das classes e do tamanho da amostra de treinamento.

3.1 Conceitos

Desta forma, os métodos de aprendizado de máquinas, podem ser classificado de duas formas: Preditivos e Descritivos. Dentro dos métodos Preditivos, temos a Classificação e Regressão. Já nos descritivos temos Associação, Agrupamento, Detecção de desvios, Padrões sequenciais e sumarização. Os algoritmos de classificação, buscam através de características conhecidas, prever determinado resultado. Ex.: baseado em características sócio-culturais, é possível prever a renda de determinada pessoa e se é um bom pagador. Através de indicadores médicos, é possível prever determinadas doenças... Já na regressão, ao invés de trabalhar com rótulos, existem números que são correlatos. Baseado nos investimentos em propaganda, é possível prever qual será o valor de faturamento, baseado em fatores externos, pode-se prever qual será o valor do dólar... Nos métodos descritivos,

há por exemplo as regras de associação que são uma forma de prever quais características em comum existem na situação estudada, um exemplo clássico, são as prateleiras de supermercado, que são montadas de forma que os produtos estão dispostos de acordo com as associações de compra dos clientes. Os algoritmos de agrupamento, buscam características comuns nos dados e desta forma, agrupa-as por similaridade. Ex.: Segmentação de mercado, agrupamento de documentos/notícias, Perfis de Clientes... Detecção de desvios (*outliers*), busca identificar dados fora do padrão. Ex.: Fraude de cartão de crédito, Uso de energia elétrica, monitoramento de máquinas... Descoberta de padrões sequenciais como o próprio nome já diz, procuram no histórico, padrões de comportamento para prever quais serão as possíveis escolhas dos clientes, prevenção de doenças, navegação em sites... Por fim, os algoritmos de sumarização, buscam em uma base de dados os perfis que aderem a certo comportamentos.(ALPAYDIN, 2014)

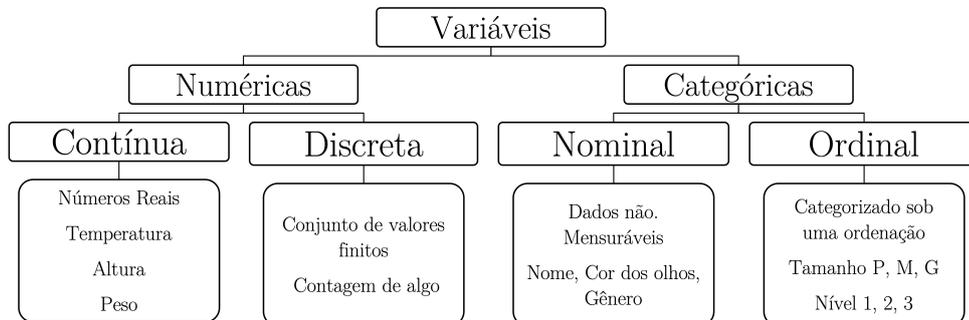
Segundo Mohri Afshin Rostamizadeh (2012) a aprendizagem de máquina pode ser dividida em supervisionada, Não supervisionada e por Reforço. Na aprendizagem supervisionada, o supervisor, entrega ao algoritmo determinadas características e informa o que estas representam. Na sequencia, o algoritmo treinado é capaz de identificar o que certos dados de entrada representam. A não supervisionada, analisa automaticamente os dados (associação, agrupamento) e necessita de análise para determinar o significado dos padrões encontrados. A aprendizagem por reforço é baseada em causa e efeito. Por exemplo, os robôs que aspiram pó.

Após entendidos os conceitos de que se trata o aprendizado de máquina, é possível aprofundar um pouco nos algoritmos de classificação. Segundo Raschka (2017) os dados para serem processados, precisam passar por uma padronização. Esta inicia com o tratamento de valores inconsistentes que são "sujeira" nos dados e também os faltantes. O autor indica que é possível a utilização de algumas técnicas para recuperar os dados ou simplesmente desconsiderar a amostra.

A escolha correta dos algoritmos passa pelo tipo de variável que a base de dados possui. Certos algoritmos não trabalham com dados numéricos por exemplo. Desta forma a figura 3 trás uma classificação dos tipos de variáveis que podemos encontrar.

Até aqui, foram trabalhados conceitos gerais sobre aprendizagem de máquina. Estes são fundamentais e elementares nos estudos destas técnicas. em seguida, será explanado sobre o tratamento e qualidade dos dados. Segundo Raschka (2017) a qualidade dos dados e a quantidade de informações úteis que eles contêm são os principais fatores que determinam o quão bem um algoritmo de aprendizado de máquina pode aprender. Desta forma, analisar e fazer um pré-processamento nestes dados é absolutamente necessário. O autor destaca que não é comum em aplicações reais, faltarem valores ou estes sejam inválidos. A maioria dos aplicativos de tratamento de dados não é capaz de processar tais números ou pode gerar resultados imprevisíveis. Para identificar estes valores, utilizam-se bibliotecas específicas

Figura 3 – Tipos de variáveis utilizadas no aprendizado de máquinas



Fonte: Adaptado de (ALPAYDIN, 2014)

de tratamento de dados Raschka (2017) destaca *Pandas* e *scikit-learn*. A maneira mais simples de eliminar estes dados é desconsiderar a linha com erro da base de dados. O autor destaca que esta conduta, embora pareça ser uma abordagem conveniente, apresenta certas desvantagens pois acaba-se removendo muitas amostras e isso pode gerar a perda de informações valiosas. Uma alternativa para esta tratativa é inserir dados. Embora existam muitas técnicas de interpolação, a mais comumente utilizada é inserir a média dos dados da coluna nos valores faltantes. Outra alternativa é inserir o valor da mediana ou "mais frequente". (ALBON, 2018)

Após esta parte de avaliação dos algoritmos, é possível melhorar os resultados, ajustando os parâmetros do modelo. Uma forma eficiente de se fazer isto, é utilizando a técnica de validação cruzada do inglês *Cross-Validation*. (ZHENG, 2015) No ajuste dos parâmetros, é necessário ficar atento ao *overfitting*. segundo (ALBON, 2018), este efeito ocorre inerentemente de acordo com o algoritmo escolhido e precisa ser tratado com o risco de quando submete-se o modelo à dados reais, este não seja capaz de ter um bom desempenho. O *overfitting* é um ajuste de parâmetros tal que o algoritmo só consegue tratar os dados do qual foi treinado.

3.2 Tipos de Algoritmo de classificação

Segundo Raschka (2017) pode-se expressar o processo de Aprendizado de Máquina em 4 passos: Pré-Processamento, Aprendizado, avaliação e Previsão. O autor destaca que na etapa de aprendizagem, deve-se testar vários algoritmos e verificar qual deles produz o melhor resultado. Este trabalho irá se limitar aos Algoritmos SVM e Redes Neurais

Artificiais. Embora nos experimentos iniciais testou-se outros algoritmos, seus resultados não foram expressivos.

- *Support Vector Machine*(SVM): Segundo (CRISTIANINI, 2000), SVM são um sistema para treinar eficientemente máquinas de aprendizagem linear em espaços de recurso induzidos pelo *kernel* e respeitando o *insight* fornecidos pela teoria de generalização e explorando a teoria de otimização. Por ter uma aprendizagem baseada na teoria da generalização e produzirem representações duais da hipótese, o resultado destes algoritmos é extremamente eficiente.
- Redes Neurais Artificiais: Segundo (L., 1994) Uma rede neural artificial (RNA) é um sistema de processamento de informações que possui certas características de desempenho em comum com redes neurais biológicas. As RNA foram desenvolvidas como generalizações de modelos matemáticos da cognição humana ou biologia neural, com base nas premissas de que: O processamento de informações ocorre em muitos elementos simples chamados neurônios; Os sinais são passados entre os neurônios por meio de *links* de conexão; Cada *link* tem um peso associado, que em uma RNA típica, multiplica o sinal transmitido; Cada neurônio aplica uma função de ativação à sua entrada para determinar seu sinal de saída. Uma RNA é caracterizada por seu padrão de conexões entre neurônios (chamada de arquitetura), seu método de determinar os pesos na conexões (chamado de algoritmo de treinamento ou aprendizagem) e sua função de ativação

3.3 Avaliação de algoritmos

Uma das principais tarefas na construção de um modelo de aprendizado de máquina é avaliar seu desempenho. segundo Zheng (2015), é fundamental e muito difícil. A autora destaca que duas perguntas devem ser respondidas ao final de um projeto de modelagem: "Como posso medir o sucesso deste projeto?" e "Como saberei quando for bem sucedido?" Ela destaca que é importante pensar na avaliação com antecedência. As métricas de avaliação estão vinculadas a tarefa de aprendizado de máquina. Existem métricas diferentes para as tarefas de classificação, regressão, agrupamento, modelagem de tópicos, etc. Terminada a etapa de Revisão Bibliográfica, onde buscou-se entregar ao leitor conhecimentos necessários para interpretar os conceitos técnicos apresentados, o próximo capítulo, tratará sobre o estado da arte, ou seja, quais técnicas e aplicações estão na vanguarda da produção científica.

4 ESTADO DA ARTE

A seguir será apresentada a revisão sistemática onde de acordo com as *strings* de pesquisa utilizadas não foram encontrados trabalhos semelhantes. Este trabalho irá iniciar um caminho de pesquisa para que através da contribuição da comunidade científica, aprofundar o assunto e melhorar a vida das pessoas e o resultado financeiro das empresas de distribuição de energia elétrica. Entende-se que o ponto de partida, é buscar de forma multidisciplinar, trabalhos que nos mostrem os casos e técnicas de solução utilizados atualmente para solucionar situações semelhantes nas diferentes áreas.

O trabalho de Hameed Abdul Raza (2019) busca desenvolver uma ferramenta de apoio a tomada de decisão para a programação da manutenção de uma planta de gás natural. O autor quer encontrar o ponto ótimo entre o custo de manutenção e o risco para a operação. A técnica utilizada é um híbrido entre algoritmo genético e *Pareto-optimal*. O autor consegue encontrar o ponto ótimo, reduzindo substancialmente os custos de manutenção e mantendo a operação em níveis satisfatórios de segurança.

Já o trabalho de Yimer e Demirli (2010) busca resolver um problema de planejamento da produção e materiais de uma empresa com características que a enquadram como *build-to-order*. Esta tem como particularidade a customização em massa. A abordagem do autor, sugere que o *shop floor* seja dividido em duas etapas. Uma que trata da aquisição de matéria prima e componentes e outra que trata da fabricação de peças e da montagem do produto fim. A primeira tem como característica a padronização dos componentes onde as demandas são mais regulares e previsíveis. Segundo Yimer e Demirli (2010) este sistema pode ser modelado com programação linear inteira mista de duas fases (MILP). O objetivo é minimizar os custos agregados a cada subsistema, atendendo os requisitos de serviço ao cliente. Como o espaço de busca para o problema da primeira fase envolve uma paisagem complexa com muitas soluções candidatas, um procedimento de solução baseado em algoritmo genético é proposto.

No estudo de Ghezavati (2011) o autor mostra uma ferramenta capaz de projetar um sistema de manufatura celular integrado com programação de grupo e aspectos de *layout* em um espaço de decisão incerto sob as características de uma cadeia de suprimentos. A técnica de modelagem utilizada é a programação inteira mista que o modelo busca otimizar o custo de manutenção e os custos relativos à rede de fornecedores na cadeia de suprimentos. As decisões de programação em uma estrutura de manufatura celular são segundo Ghezavati (2011) tratadas como um problema de programação de grupo. A técnica utilizada para solucionar o problema é um híbrido entre algoritmo Genético e recozimento simulado. O autor finaliza o estudo comparando esta ferramenta com soluções

globais e também com um algoritmo heurístico, testando este, por meio de problemas de teste.

O Problema de Programação de Fluxo de Permutação *PFSP*, segundo Rahman (2015) é um problema de otimização combinatória complexa. O autor pontua que este tipo de problema, é amplamente estudado como estático e são utilizadas técnicas heurísticas e meta-heurísticas para sua solução. Porém, segundo (RAHMAN, 2015) estes são problemas dinâmicos, pois os pedidos dos clientes são feitos em intervalos de tempo aleatórios. No sistema dinâmico, o autor ressalta que duas questões devem ser respondidas. - Um novo pedido deve ser aceito? e - Se aceito, como essa programação pode ser solicitada, quando alguns pedidos já podem estar em processamento e / ou na fila para processamento? Para responder a estas questões, o artigo sugere duas abordagens distintas. Para a primeira um método heurístico simples e para a segunda uma abordagem baseada no Algoritmo Genético. O autor, realiza diversos testes e juntamente com um modelo de simulação, testa a maximização de receitas. Por fim, é feito um estudo de caso para mostrar a aplicabilidade da abordagem proposta.

O problema de programação Job shop flexível *F.J.S.S.P* segundo Rajan Kumar (2018) é muito importante nos campos de gerenciamento da produção e aparece na categoria de problema de otimização combinatória difícil *NPHard* o autor destaca ainda que *F.J.S.S.P*. é uma extensão do problema de *job shop* tradicional. Ele reforça ainda a ideia de que em muitas organizações, as programações são obrigatórias com a existência de várias interferências repentinas. Esta particularidade torna complicado obter os melhores resultados possíveis dentro de um tempo razoável. Neste artigo, ele cita que o Algoritmo Genético *AG* pode reduzir a complexidade combinatória por decomposição de tarefas métodos de distribuição em tempo real. Através de um estudo de caso, o autor demonstra que é possível otimizar o problema o sistema *F.J.S.S.P* utilizando *AG*.

Muitos autores dedicam seus estudos em encontrar o método mais aderente para a solução de problemas. Os estudos em manufatura, buscam entre outros, resolver problemas do cronograma de produção. Álvarez, Esther Díaz (2007) Em seu trabalho, discute um problema de programação dinâmica em um ambiente fabril com ampla gama de produtos. Segundo o autor, a maioria das metodologias disponíveis em relação à programação da produção, são métodos *off-line* ou propõem todo um processo de reprogramação para responder às perturbações. Ele destaca que na prática, é muito demorado construir novos cronogramas para lidar com esses problemas, então o uso de abordagens aproximadas se justifica. Por um lado, as Regras Heurísticas *HRs* podem fornecer boas soluções para problemas complexos em tempo real. Por outro lado, os Algoritmos Genéticos *AGs* podem se adaptar ao tempo disponível para encontrar uma solução. Seu artigo faz uma comparação entre as duas técnicas e conclui que o *AG* é capaz de reorganizar um cronograma a partir do zero, enquanto o *HRs* só podem reparar um cronograma existente.

Outro problema abordado é o de tempo de atravessamento¹ *makespan*, o trabalho de Tajbakhsh (2014) busca minimizar este tempo, através de um cronograma de produção eficiente. Ele utiliza um modelo matemático inteiro misto. Este é considerado *NP-Hard* e pode ser resolvido por métodos de aproximação, o autor utiliza um modelo híbrido entre AG e enxame de partículas. O trabalho conclui que o modelo consegue encontrar soluções adequadas para o problema apresentado.

Saindo do ambiente de manufatura, atualmente, o *e-commerce* tem chamado a atenção dos estudos de otimização em logística. O trabalho de Mei e Li (2016) busca encontrar uma solução para as farmácias que separam e enviam seus pedidos aos clientes. O autor destaca que é fundamental, encontrar o ponto ótimo entre a quantidade de itens em um pedido e o tempo de separação do mesmo. Estes centros logísticos agrupam vários produtos em um único pacote de envio, isto acaba afetando o tempo de separação do pedido. Para resolver este, o autor propõe um modelo baseado em três algoritmos genéticos multiobjetivo com um projeto cromossômico de três tuplas, incluindo *Vector Evaluated Genetic Algorithm VEGA*, *Multi-Objective Genetic Algorithm MOGA* e algoritmo genético classificado na dominado II *NSGA – II*. O artigo compara estes métodos com o atual que consiste em uma Heurística de Tempo de Processamento mais longo *LPT*. Os estudos concluem que os AGs economizam o tempo de separação em 95,6% enquanto aumentam o *makespan* em apenas 5,62%. Dentre os AGs, o *NSGA – II* obteve melhor desempenho.

No setor elétrico, um assunto atual trata a geração distribuída *GD*. Esta consiste em micro usinas que são geralmente solares e podem ser conectadas ao Sistema elétrico de Potência *SEP*. Um problema comum, é o de minimizar a perda de energia já que com a conexão destas à rede primária, o fluxo passa a ser bi-direcional. O trabalho de Huang Wei Wang (2012), propõe a utilização de um AG aprimorado para otimizar a operação econômica dos micro recursos *GDs*, definindo o sentido de fluxo destas ao *SEP* em determinadas horas do dia baseado no horário de ponta do consumo. O autor conclui que o incentivo à instalação de *GDs*, pode reduzir os custos de energia para o usuário e a necessidade de expansão do sistema por parte das distribuidoras de energia.

Pode-se observar que os trabalhos aqui citados não tratam do problema sobre otimização do *CHI* em projetos, mas mostram que os AGs são amplamente utilizados na solução dos problemas de otimização onde há iteração entre pessoas e sistemas de produção e na busca de soluções no *SEP*. Na distribuidora de energia estudada, a área de projetos funciona como uma fábrica. O produto em questão é o próprio projeto e um dos critério de qualidade é o *CHI* que este irá consumir. Buscar-se-á minimizar este, garantindo que a obra será executada com o mínimo de impacto possível para a sociedade. Será estudado nos artigos acima, um método que seja aderente e contribua para medir se o projetista está conseguindo entregar a qualidade desejada. Nos próximos capítulos, será apresentado,

¹ Tempo decorrido entre o início de operação de um produto em uma manufatura até sua conclusão

com base no trabalho de Petersen (2015), como estes nove artigos foram eleitos.

4.1 Processo de Busca

Para desenvolver uma busca sistêmica, deve-se definir qual será o mecanismo de busca utilizado. Este é o responsável por receber o string de busca como entrada. Existem diversos mecanismos de busca e geralmente cada base de dados possui o seu, este geralmente tem uma abrangência limitada ao próprio banco de dados. Desta forma, para ampliar o campo de busca, foi escolhido o *software Mendeleev*, cuja principal função é a de ser um gestor de referências acadêmicas. Este *software* tem uma particularidade de buscar em vários repositórios digitais o *string* indicado. Isto torna a pesquisa mais profunda e melhora a tendência de serem encontrados resultados significativos.

Definido o método de busca, o próximo passo é o de definir a *string* que será utilizada. Esta consiste em uma série de palavras combinadas a operadores lógicos de acordo com as definições do mecanismo de busca escolhido. A *string* deve ser montada de forma que possa-se fazer uma busca geral e baseado nos resultados, ir refinando a pesquisa de acordo com as questões a serem respondidas. Para este trabalho, optou-se por fazer uma pesquisa multidisciplinar. A *string* de busca será montada com termos utilizados nas áreas de modelagem de Sistemas, Pesquisa Operacional e SEP.

Abaixo, será apresentado esta etapa através de tabelas na Tabela 1 Teremos as strings de Modelagem de Sistemas, na Tabela 2 de Pesquisa Operacional e na Tabela 3 do PDS.

Tabela 1 – String de Modelagem de Sistemas

String de Busca	Número de Artigos
((("Make-to-order") AND ("Make-to-Stock"))	418
((("Make-to-order") AND ("Make-to-Stock") AND ((Manufacturing) OR (Production))))	359
((("Make-to-order") AND ("Make-to-Stock") AND ((Manufacturing) OR (Production)) AND ((scheduling) OR (Programing)) Language (English)))	60

No campo da modelagem de sistemas, foram encontradas com a primeira *sting*, 418 resultados, como o ambiente de estudo contempla uma produção mista de projetos e possui características de uma fábrica, incluímos esta restrição e obtivemos 359 artigos. Por fim, para termos uma melhor aderência, limitamos o tema à programação da produção. Desta forma, nos aproximaremos mais do tema em questão. Não encontrou-se na literatura disponível estudos que tratem da modelagem de projetistas no cálculo do DEC em projetos.

Em pesquisa operacional, buscou-se otimizar um ambiente de manufatura ou projetos, como a modelagem deste tipo de problema caracteriza-se como NPHard, buscou-

Tabela 2 – String de Pesquisa Operacional

String de Busca	Número de Artigos
((("Optimization") AND (("Production") OR ("Manufacturing") OR ("Project")) AND ("Process") AND (("Heuristic") OR ("Artificial Inteligence"))))	1065
((("Optimization") AND (("Production") OR ("Manufacturing"))) AND ("Process") AND (("Heuristic") OR ("Artificial Inteligence"))) AND ("Journal"))	283
((("Optimization") AND (("Production ") OR ("Manufacturing ")) AND ("Scheduling") AND (("Heuristic") OR ("Artificial Inteligence"))) AND ("Journal") AND ("Research"))	122

Tabela 3 – String de Sistema Elétrico de Potência

String de Busca	Número de Artigos
((("Maintenance") AND ("Distribution") AND ("Network") AND ("Expansion") AND ("Planning")))	77
((("Maintenance") OR ("Expansion")) AND ("Distribution")) AND ("Power") AND (("Network") OR ("Grid")) AND ("Planning") AND ("Project"))	72
((("Maintenance") OR ("Expansion")) AND ("Distribution")) AND ("Power") AND (("Network") OR ("Grid")) AND ("Planning") AND ("Project") AND (Engineering))	28
((("Maintenance") OR ("Expansion")) AND ("Power") AND ("Distribution") AND ("Energy") AND ("Grid") OR ("Network")) AND ("Electric") AND ((Journal) OR (Research)))	297
((("Maintenance") OR ("Expansion")) AND (("Power Grid") OR ("Power Network"))) AND ("Electricity Distribution") AND (Journal) AND (Research))	70

se Heurística e IA (Inteligência Artificial) para completar a *string*. A primeira iteração, nos trouxe 1065 resultados. Buscando uma melhor qualidade de estudos, optamos por incluir o termo *journal*, para filtrar trabalhos com validade verificada via publicação, e tivemos 283 resultados. Alinhado com o objetivo, buscou-se trabalhos com temas mais abrangentes e foi inserido o termo *research*. Por fim, obtivemos 122 trabalhos.

No PDS, a maioria dos estudos trata de reconfiguração de sistemas em contingência e ou GD (geração distribuída) de energia assuntos muito relevantes nos dias atuais. O escopo deste trabalho está situado na área de projetos de expansão e manutenção do PDS, desta forma, após testar alguns *string*, optou-se pelo que trouxe 70 resultados, por ter artigos mais relevantes.

4.2 Filtros de Estudo

As questões de pesquisa norteiam o trabalho e auxiliam a definir o que queremos buscar, entretanto, é preciso ser seletivo e ter critérios para definir os artigos que serão considerados. No estudo de Petersen (2015) o autor traz como orientação a utilização de filtros, estes podem ser de inclusão ou exclusão. A seguir, mostra-se a aplicação destes em nosso trabalho, iniciando pelos filtros de inclusão e na sequência os de exclusão.

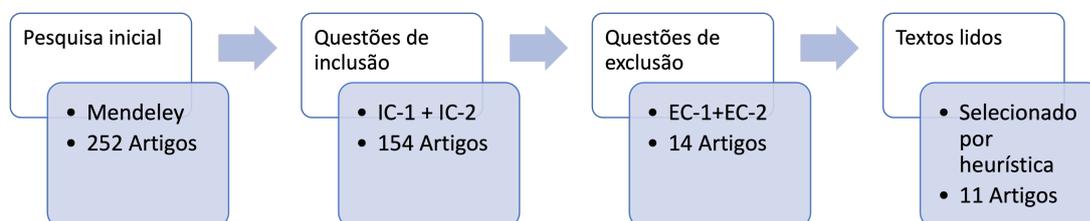
Critério de Inclusão:

- Inclusão 1 (IC-1): Estudos que tratam de modelagem ou otimização.
- Inclusão 2 (IC-2): Estudos publicados entre 2010 e 2020.

Critério de exclusão:

- Exclusão 1 (EC-1): Estudos que utilizam Meta-Heurística
- Exclusão 2 (EC-2): Estudos que utilizam Algoritmo Genético.

Figura 4 – Resultados obtidos pela filtragem



Fonte: Autor

Como na etapa de montagem das String de busca, já foram ajustadas estas para serem mais objetivas, os filtros foram utilizados basicamente para delimitar o período de

busca no tempo e a técnica utilizada. Sabe-se que o tipo de modelagem que buscada é NP-Hard, buscou-se quais técnicas convergem para os três temas e neste caso, encontrou-se a Meta-heurística baseada em Algoritmo Genético. Finalmente, a próxima etapa do fluxo de filtragem é ler os textos completos dos estudos restantes e avaliar sua utilidade para o trabalho.

4.3 Resultados da pesquisa

Seguindo a metodologia descrita na seção III, resultou o fluxo representado na 4. Ela mostra todas as etapas e os resultados do processo de filtragem. Nesta figura, a caixa superior mostra a etapa e a inferior o critério e resultado da filtragem executada. Percebe-se então que de um universo de mais de mil e quinhentos artigos, trabalhando a *string* de busca, e aplicando filtros, conseguiu-se chegar em nove artigos que serão cuidadosamente lidos e a tendência de estudo será melhor observada. Por fim, a Tabela 4 irá apresentar um compilado com as principais informações dos artigos selecionados.

Tabela 4 – Resumo dos Principais Artigos Estudados

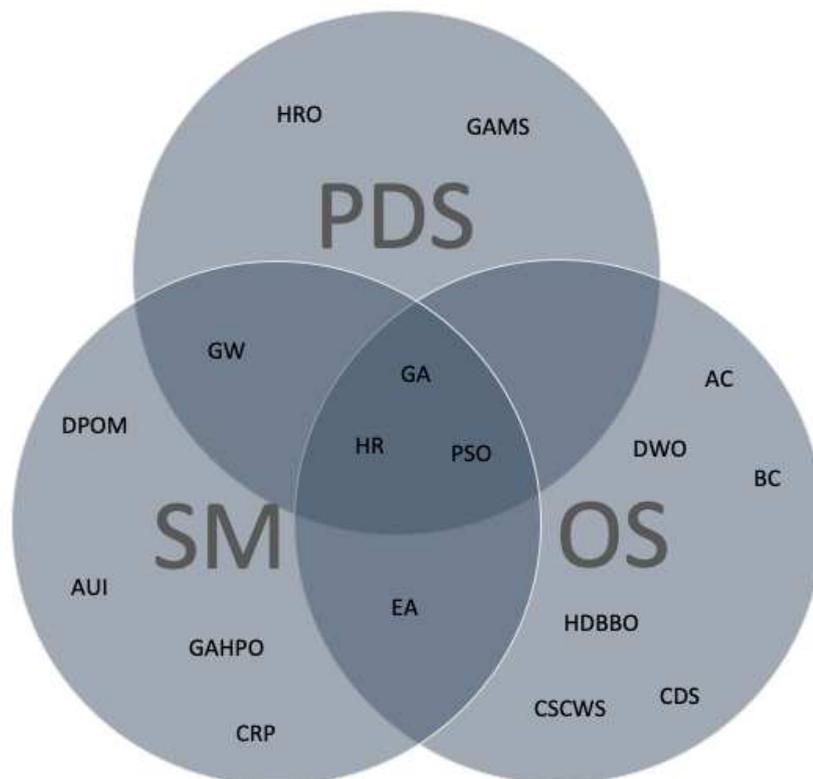
Autor	Título	Relação entre o artigo e o tema estudado
Mei, et al	Multi-objective optimization of collation delay and makespan in mail-order pharmacy automated distribution system	Utilizando Algoritmo Genético, otimiza o sistema de preparação e distribuição de uma farmácia on-line, buscando a melhor combinação entre tempo de separação e tempo de atravessamento (makespan)
Huang, et al	Research on integrated benefit evaluation model of microgrid	Faz uma pesquisa para encontrar a melhor ferramenta de otimização do retorno financeiro em manutenção e expansão do SEP utilizando Micro Grid
Hameed, et al	A decision support tool for bi-objective risk-based maintenance scheduling of an LNG gas sweetening unit	Utilizando Algoritmo Genético, o estudo busca o ponto ótimo entre o custo de manutenção, o risco operacional e as perdas de faturamento em uma planta de gás natural.
Rahman, et al	A real-time order acceptance and scheduling approach for permutation flow shop problems	Utilizando Algoritmos Genético, o estudo busca responder se é possível aceitar uma nova ordem de produção e qual seria o prazo de entrega.
Rajan , et al	A restart scheme based meta heuristic for flexible job shop scheduling	Utilizando Algoritmo Genético, o autor busca criar um modelo que permita ajustar o plano de produção mediante alguma contingência.
Ratnam, et al	Optimization Procedure by Using Genetic Algorithm	Utilizando Algoritmo Genético, o Autor busca um modelo que possa planejar a produção de um sistema de Manufatura Flexível (FMS)
Ghezavati, et al	A new stochastic mixed integer programming to design integrated cellular manufacturing system: A supply chain framework	Utilizando Algoritmo Genético, o Autor busca desenvolver um modelo para dimensionamento de células de manufatura, planejamento da produção e Layout
Tajbakhsh, et al	Multi-objective assembly permutation flow shop scheduling problem: A mathematical model and a meta-heuristic algorithm	O artigo resolve através de Algoritmo Genético, o problema de planejamento da produção de uma manufatura com a fabricação de componentes e montagem do produto final.
Yimer, et al	A genetic approach to two-phase optimization of dynamic supply chain scheduling	Neste estudo, o autor trata de uma manufatura de customização em massa. Ele resolve através de um Algoritmo Genético o problema da fabricação de componentes e da produção de produtos customizados.

4.4 Técnica de Otimização foi Utilizada

Iniciando os estudos conclusivos, iremos responder as questões fundamentais que nortearam a pesquisa bem como identificar as tendências de estudo ao longo dos anos. É importante lembrar que o espaço temporal está entre 2010 e 2020. Para esta análise, agruparam-se

as três áreas. Este agrupamento fará mais sentido pois, como buscou-se trabalhos convergentes, ficará mais fácil esta identificação. A Figura 5 demonstra de maneira gráfica as tendências de pesquisa apontando Algoritmo Genético, Heurística e *Particle Swarm* como estudos comuns nas três áreas. Neste ponto, esta convergência serve como filtro natural de trabalhos. Observa-se também que a comunidade científica empreende muito esforço em inovação. Muitas técnicas e métodos vem sendo testados com os mais diversos objetivos. Observa-se uma preocupação com os tempos de processamento computacional e com os níveis de assertividade. Os estudos no campo da Pesquisa Operacional, focam muito em resolver particularidades das empresas. Isto acaba por trazer um misto de heurística e Meta-heurística. A primeira serve para preparar os dados e ajustar as equações enquanto a segunda realiza as otimizações. Na área de modelagem de sistemas, há uma preocupação em encontrar os modelos que mais aderem ao ambiente operacional e na área de SEP, os estudos focam muito em encontrar o ponto ótimo entre investimento e qualidade do atendimento, nos últimos anos, estudos sobre *Micro Grid* vem ganhando atenção.

Figura 5 – Demonstração da convergência dos estudos



Fonte: Autor

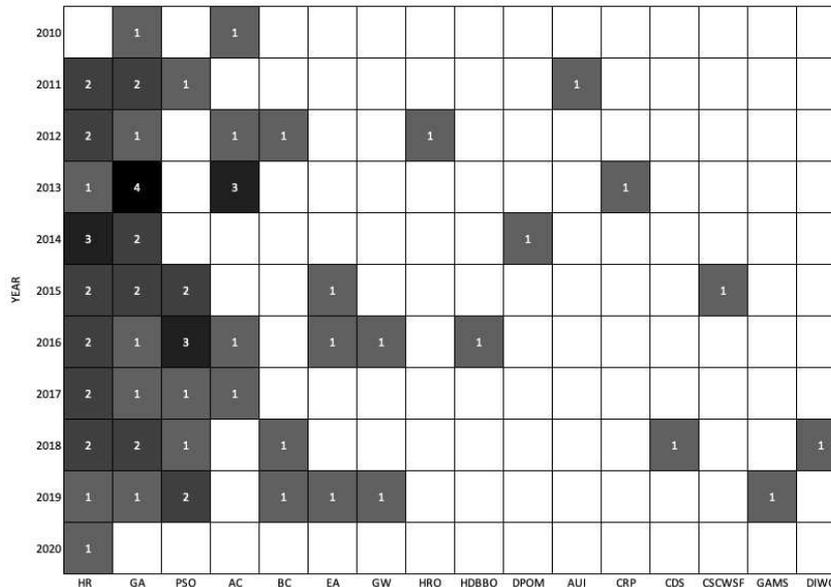
Abaixo, serão demonstrados em forma de tabela, os estudos convergentes e seu comportamento ao longo dos anos. Podemos observar que HR (Heurística) é a técnica mais utilizada, seguido por GA (*Genetic Algorithm*), PSO (*Particle Swarm*), AC (*Ant Colony*), BC (*Bee Colony*), EA (*Evolutionary Algorithm*), GW (*Grey Wolf*), HRO (*High Reliability Organization*) HDBBO (*Biogeography-Based Optimization*), DPOM (*Dynamic Pareto-*

Optimal Method), AUI (*Accumulated Urgency Index*), CRP (*Capacity requirement planning*), CDS (*Calculate Connected Dominating Sets*), CSC (*Cukoo Search Based Community*), GAMS (*Algebraic Modeling System*), DIWO (*Discrete Invasive Weed Optimization*).

4.4.1 Questão Geral 1 (GQ-1): Técnica de otimização foi utilizada

Iniciando com a questão principal, pode-se ter uma visão mais estratégica das principais técnicas que vem sendo utilizadas ao longo do tempo, e qual é a tendência dos estudos apresentados. A Figura 6, representa um mapa de calor das principais técnicas.

Figura 6 – Mapa de calor dos principais estudos ao longo do tempo



Fonte: Autor

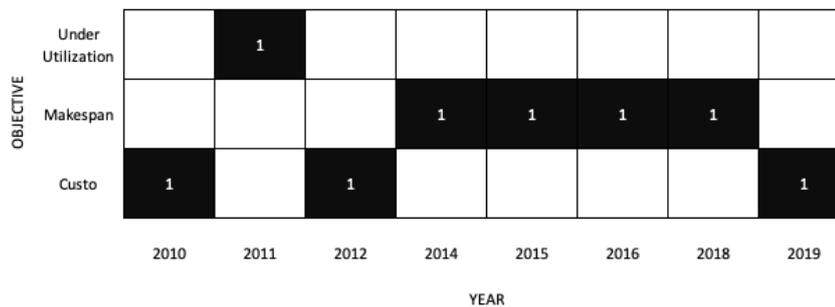
Esta análise mostra que o método baseado em Heurística ganha destaque ao longo dos anos e suas pesquisas mostram que há uma tendência de que o algoritmo é bastante difundido na comunidade acadêmica. Nesta pesquisa, os métodos heurísticos são sequências de decisões baseadas em aprendizado ou seja, os autores, com base no histórico dos acontecimentos, criam um algoritmo pré definido de alternativas pré estabelecidas que devem ser seguidas quando alguma anomalia é encontrada.

O Algoritmo Genético, desponta como solução padronizada. O problema em questão é modelado de forma matemática, e o algoritmo é padronizado de forma que sua solução sempre é válida. Leva vantagem em relação a Heurística pois é adaptável. Já os algoritmos baseados em partículas vem sendo mais estudados nos últimos cinco anos.

4.4.2 Questão Geral 2 (GQ-2): Aspectos (objetivos e restrições) foram otimizados

Ao avaliar os estudos iniciais, percebe-se uma tendência maior entre a GA, HR e PSO entretanto, GA e HR possuem maior relevância por ter mais citações ao longo do tempo. Isto mostra uma tendência de que estes estudos são mais promissores para iniciar a pesquisa. Ao verificar os estudos heurísticos, estes possuem uma serie de regras aplicadas em sua maioria a exclusividade do case em questão, não tendo grande relevância para este estudo. Diante disto, optou-se por aprofundar os estudos nas técnicas que abordam GA. Esta técnica além de ser comum às três áreas, vem se perpetuando ao logo dos anos, o que mostra uma tendência de validade e relevância. A Figura 4 apresenta um mapa de calor com os principais objetivos de estudo. Esta pergunta, aprofunda a pesquisa no sentido de entender qual foi o objetivo e quais foram as restrições utilizadas pelos pesquisadores no emprego do GA. Vale lembrar que apenas os aspectos mais relevantes das pesquisas foram considerados. Em sua maioria estas possuem diversos objetivos secundários. Buscou-se aqui o *core* da pesquisa.

Figura 7 – Mapa de calor com os principais objetivos de otimização



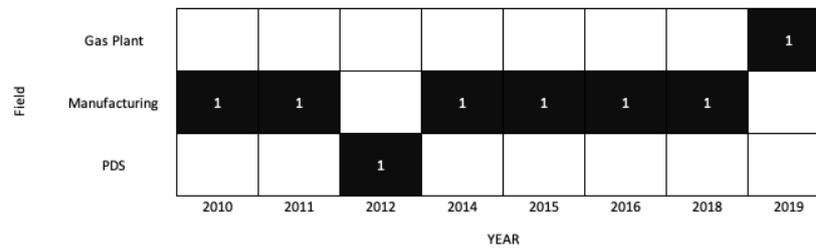
Fonte: Autor

Esta figura nos mostra que a maioria dos estudos conduzidos, buscam melhoria na eficiência operacional. Seja através da redução de custos (diretamente) seja pela redução do *makespan* que é a redução do tempo de atravessamento dos produtos, ou seja a redução do custo de material em processo. Trazendo para o escopo deste trabalho, pode-se encarar o CHI como custo de produção, a eficiência operacional de obras, entre outros indicadores, é a capacidade de realizar um investimento maior utilizando o menor DEC possível.

4.4.3 Questão Específica 1 (SQ-1): Campo de trabalho da pesquisa

Quando o assunto é processo, os estudos de manufatura tornam-se mais frequentes, desta forma, segundo a Figura 5 observa-se que a maioria dos estudos utilizando GA, concentram-se em ambientes produtivos, onde os tempos e custos são fundamentais para o crescimento das companhias.

Figura 8 – Mapa de Calor da aplicação do AG



Fonte: Autor

Esta figura reforça o quão amplo são os estudos no campo da manufatura e o quanto temos oportunidade de aprofundar o tema de estudo referente a otimização de CHI em projetos do sistema elétrico. Um assunto tão relevante para a realidade brasileira que até então, não foi explorado. Esta integração de conhecimento, pode gerar ganhos significativos a população.

4.4.4 Questão específica 2 (SQ-2): Conclusão do estudo. Solução do problema

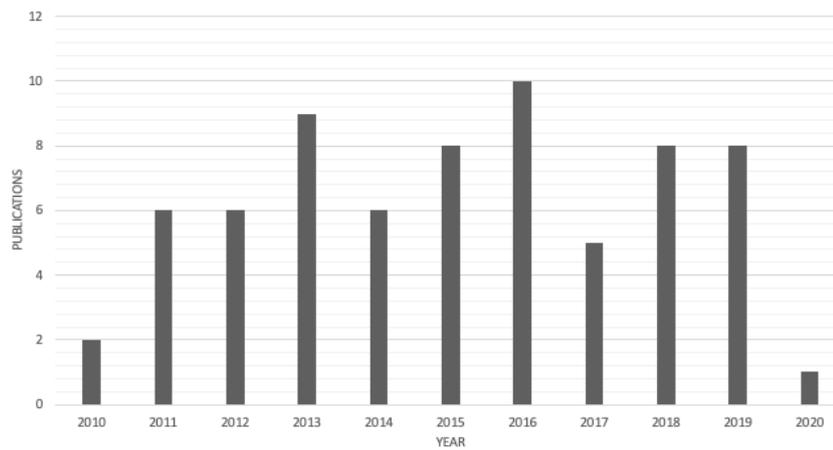
Após ler os artigos, constata-se que todos eles tiveram resultados positivos atingindo os objetivos. No trabalho de Yimer e Demirli (2010) O GA atendeu o objetivo de montar um plano de produção otimizado, já o trabalho de Ghezavati (2011), o modelo de otimização deu conta de otimizar a montagem de células de produção, bem como montar um planejamento e sequenciamento da manufatura. No campo do SEP, o trabalho de Huang Wei Wang (2012) consegue otimizar o modelo de operação de uma *microgrid*, enquanto Tajbakhsh (2014) constrói um cronograma de produção em um ambiente misto, com fabricação e montagem. O trabalho de Rahman (2015), cria um plano de produção adaptável, capaz de se reorganizar de acordo com contingências e o trabalho de Mei e Li (2016) otimiza o tamanho do lote em uma farmácia on-line garantindo o melhor *makespan*. No ambiente FMS, o trabalho de Rajan Kumar (2018) cria um plano de produção otimizado. Por fim, o trabalho de Hameed Abdul Raza (2019) encontra o ponto ótimo entre o custo de manutenção e a segurança de operação de uma planta de GN.

4.5 Publicações Anuais

No universo dos artigos pesquisados, convergindo as três áreas de estudo, observa-se o seguinte comportamento de publicações ao longo dos últimos dez anos.

A revisão sistemática mostra que ao longo da última década, houve uma atividade acadêmica constante. A modelagem e otimização de sistemas vem sendo pesquisada e atualizada anualmente mostrando-se ser um campo de pesquisa atual e relevante. Concluída a revisão sistemática, percebe-se que o assunto desta pesquisa é pioneiro, entretanto a

Figura 9 – Publicações por ano



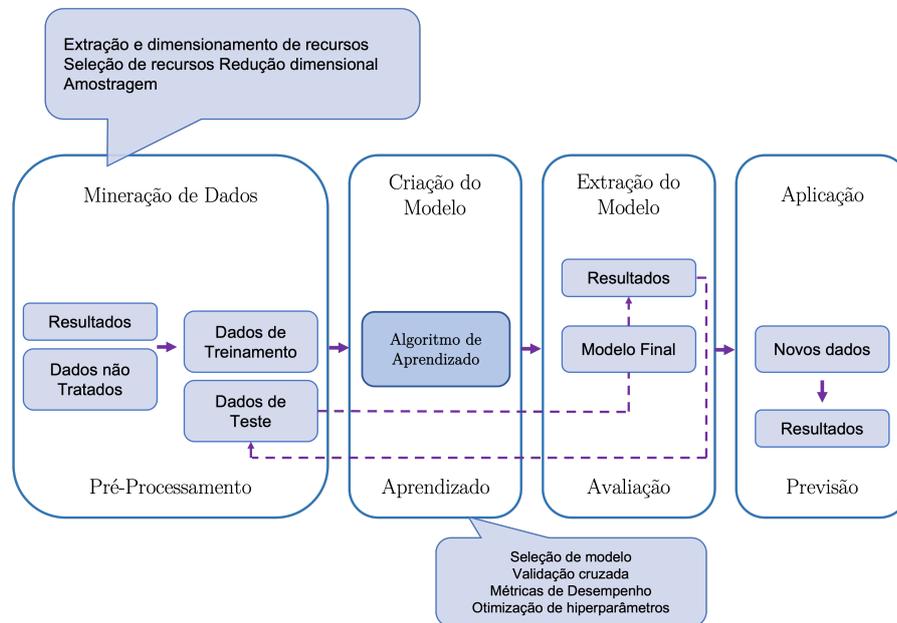
Fonte: Autor

modelagem e otimização são amplamente estudadas. Desta forma, inspirados nas pesquisas existentes, buscar-se-á uma metodologia que seja aderente ao nosso contexto e nos auxilie na construção de um trabalho relevante para a companhia.

5 METODOLOGIA

Neste capítulo, será descrito o processo de modelagem dos dados para montar um modelo de previsão onde, baseado em um histórico de dados, pode-se "ensinar" determinado algoritmo a prever resultados futuros. Serão descritos os passos para tratamento dos dados e a implementação dos algoritmos SVM e RNA. A metodologia utilizada seguirá os passos descritos por Raschka (2017): Pré Processamento, Aprendizado, Avaliação e Previsão como demonstrado na figura 10.

Figura 10 – Processo para modelagem de dados utilizando Aprendizado de Máquina



Fonte: Adaptado de (RASCHKA, 2017)

Nesta pesquisa, observa-se que há uma uniformidade na conduta. Esta metodologia, está consolidada na tratativa de modelos de previsão de classificação. Raschka (2017) expressa isto de forma estruturada. Os demais demonstram na estruturação de seus escritos.

5.1 Pré Processamento

Na construção de sistemas de aprendizagem de máquina, os dados brutos raramente vêm na forma e no formato que são necessários para o desempenho ideal de um algoritmo de aprendizado. desta forma, o pré-processamento dos dados é uma das etapas cruciais em qualquer aplicativo de aprendizagem de máquina. O autor ainda destaca que muitos algoritmos de aprendizado, exigem que as variáveis estejam na mesma escala para um desempenho ideal. Existem algumas formas de se alcançar este resultado, o autor destaca

uma transformação dos recursos em $[0, 1]$ ou uma distribuição normal padrão com média zero e variância unitária.

Após a tratativa dos dados, é necessário dividir a base em dados de treinamento e teste, a primeira irá literalmente treinar o modelo enquanto a segunda simula a entrada de novos dados e destaca que para determinar se o algoritmo de aprendizado possui um bom desempenho esta separação deve ser aleatória. (RASCHKA, 2017)

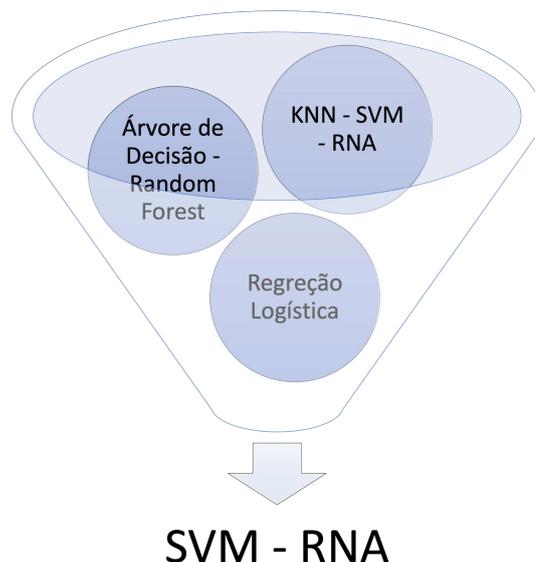
5.2 Seleção do Algoritmo de Aprendizado

Existe uma infinidade de algoritmos de aprendizagem e cada um deles foi desenvolvido para realizar determinada tarefa de modo que cada algoritmo de classificação possui suas tendências inerentes e nenhum dele goza de superioridade se não forem feitas suposições sobre a tarefa. Desta forma, é essencial comparar um punhado de algoritmos diferentes para treinar e selecionar o modelo de melhor desempenho.

Antes de comparar diferentes modelos, primeiro é preciso decidir sobre uma métrica para medir o desempenho. Comumente utilizada a precisão da classificação *classification accuracy* que é definida como a proporção de instâncias classificadas corretamente.

Há uma questão que todo cientista de dados se faz quando trabalha na escolha de algoritmos: Como saber se o resultado encontrado na etapa de modelagem irá se repetir nos dados reais? Para solucionar este problema, as técnicas de validação cruzada (*cross-validation*) são utilizadas de modo que o conjunto de dados de treinamento é dividido em subconjuntos de treinamento e validação para estimar o desempenho de generalização do modelo.

Figura 11 – Processo de pesquisa da aderência dos algoritmos ao tipo de dados



Outra etapa importante trata dos hiper parâmetros. Os algoritmos possuem parâmetros que nas bibliotecas vem configurados como *default*. Não é comum que este padrão atenda a todos os modelos de dados disponíveis. Para isto, técnicas de otimização de Hiper parâmetros são utilizadas e irão ajustar o desempenho do modelo. (RASCHKA, 2017) Neste trabalho, como pesquisa experimental, testou-se os algoritmos: Redes Neurais Artificiais (RNA), K-vizinho mais Próximo do inglês *K-Nearest Neighbour* (KNN), Redes Neurais Artificiais (RNA), Árvores de decisão, *Random Forest* e Regreção Logística. Na 11 demonstra-se que os algoritmos SVM e RNA apresentaram melhor performance.

5.3 Avaliação de modelos

Após selecionado o modelo que foi ajustado ao conjunto de dados de treinamento, pode-se usar o conjunto de dados de teste para estimar o nível de assertividade que ele alcança. Embora Raschka (2017) em seu método utiliza apenas *classification accuracy* para avaliação dos modelos, (ZHENG, 2015) apresenta outros métodos na tomada de decisão. A classificação trata de prever rótulos de classe com base nos dados de entrada, na classificação binária, existem duas classes de saídas possíveis. Há diversas maneiras de medir o desempenho da classificação. Precisão, matriz de confusão, perda de log são os mais populares.

A precisão mede o nível de acertos que o modelo é capaz de gerar ao ser submetido a base de dados porém, não faz distinção entre classes e Isto pode não ser suficiente. Pode ser interessante avaliar quantos exemplos falharam na classe zero e um, pois o custo da falha em determinada classe pode ser maior do que na outra. Por exemplo, quando um médico faz diagnóstico de câncer um falso positivo, tem consequências diferentes do que um falso negativo. Nestes casos, Uma matriz de confusão, mostra uma divisão mais detalhada das classificações corretas e incorretas para cada classe. No quadro 5 é possível verificar um exemplo de matriz de confusão.

	Previsor Positivo	Previsor Negativo
Dados Positivos	80	20
Dados Negativos	5	195

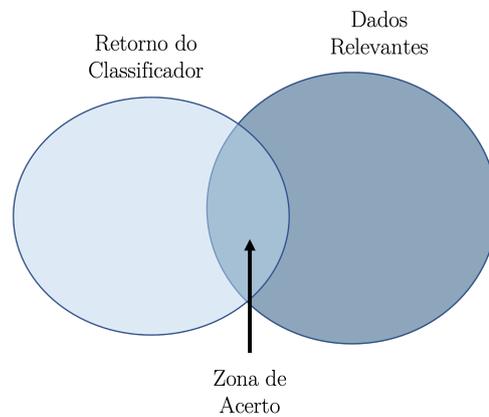
Tabela 5 – Adaptado de (ZHENG, 2015)

As linhas da matriz correspondem aos rótulos de verdade fundamental e as colunas, representam a previsão. Neste caso, podemos ver que a assertividade da previsão positiva é baixa, em torno de oitenta por cento. Já a assertividade da classe negativa chega a noventa e sete por cento. Esta avaliação se perde, quando somente a precisão total que é em torno de noventa por cento é avaliada.

Segundo (ZHENG, 2015), uma outra forma de avaliação dos resultados encontrados pelos

algoritmos são *Precision* e *Recall* e eles buscam responder a duas perguntas: Dos itens que o classificador previu serem relevantes, quantos realmente são relevantes? e De todos os itens que são realmente relevantes, quantos foram encontrados pelo classificador?

Figura 12 – Diagrama explicativo sobre *Precision* e *Recall*



Fonte: Adaptado de (ZHENG, 2015)

A figura acima representa graficamente o significado do *Precision* e *Recall*. Este marcador, representa a capacidade que determinado algoritmo tem de acertar as previsões.

5.4 Previsão

Após realizar o ajuste dos parâmetros e validar o modelo, pode-se "empacotar" e submeter um algoritmo treinado a dados reais e desta forma colher os resultados reais que o modelo é capaz de produzir. Esta etapa basicamente finaliza o ciclo de modelagem e assim, pode-se utilizar o modelo submetendo-o a dados reais. No livro: "*Python Machine Learning Cookbook*", de (ALBON, 2018), embora o autor não traga explicitamente um modelo como citado acima, o livro segue os mesmos passos para chegar a um modelo funcional. Desta forma, o próximo capítulo irá tratar da implementação e avaliação do modelo de dados. Após realizar a mineração dos dados, eles serão tratados e a avaliação do algoritmos citados será feita sob a luz da precisão e capacidade de acerto na Matriz de confusão. Findada a explanação acerca das técnicas de implantação, o próximo capítulo irá mostrar os resultados obtidos com a utilização da técnica apresentada.

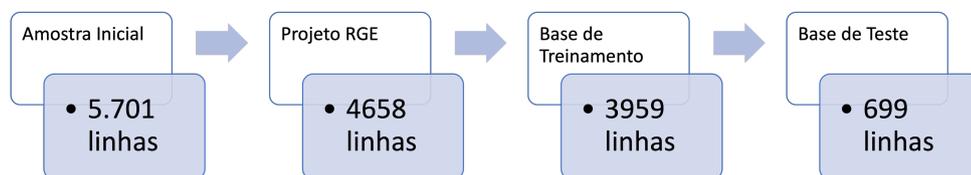
6 IMPLEMENTAÇÃO E ANÁLISE DE RESULTADOS

Segundo (RASCHKA, 2017), Python é uma das linguagens de programação mais populares para ciência de dados e possui um grande número de bibliotecas complementares úteis desenvolvidas. Para tarefas de programação de aprendizagem de máquina, referiu-se principalmente à biblioteca *scikit-learn*, que atualmente é uma das bibliotecas de aprendizado de máquina de código aberto mais populares e acessíveis. Este capítulo irá descrever a implementação da metodologia descrita anteriormente e está dividido em quatro partes: Mineração de dados, Criação do Modelo, Extração e Aplicação.

6.1 Base de Dados

A base de dados utilizada, contou com um histórico de 12 meses onde comparou-se o CHI definido em projetos frente ao CHI executado em obra. Esta base inicial contou com 5.701 linhas ou seja a amostra bruta possui mais de 5.700 registros. Ao analisar estes, percebeu-se inconsistência nos dados provindos de projetos terceirizados. Diante desta característica optou-se por suprimir estes dados e restando 4658 amostras.

Figura 13 – Definição da Base de Dados



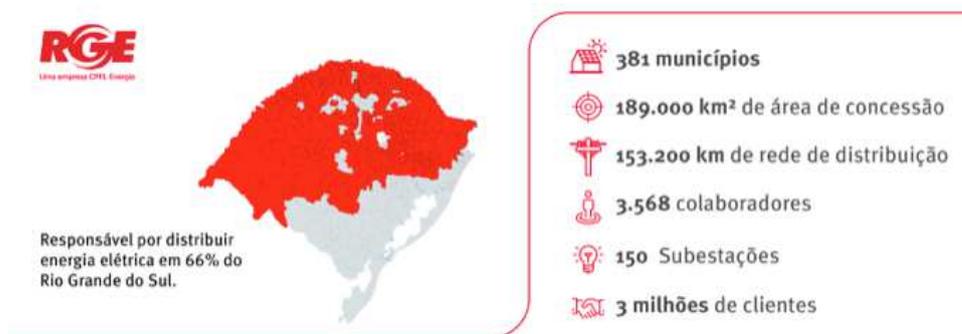
Fonte: Autor

No momento de implantação houve a divisão da base em treinamento e teste. O primeiro com 85% ou seja 3959 linhas e o segundo com 15% e 699 amostras.

6.2 Escolha das variáveis

Como já discutido anteriormente, este trabalho pretende modelar o Técnico de Projetos em sua atividade de prever e dimensionar o CHI que será utilizado para executar determinada obra em redes de distribuição. trabalhou-se com dados reais de uma empresa do estado do Rio Grande do Sul, com mais de três milhões de clientes Figura 14. O desafio a que este trabalho se propõe é o de mitigar o impacto que os desligamentos de energia nas atividades de manutenção e expansão do sistema elétrico causam e desta forma suportar o incremento de investimentos para a melhoria da qualidade do serviço mantendo os limites de continuidade dentro da meta.

Figura 14 – Principais dados sobre a área de concessão RGE



Fonte: (CPFL, 2021)

Atualmente, a área de projetos vem trabalhando para esta redução e possui um histórico de dados que servirão de base para o modelo. As variáveis disponíveis são:

- Cidade. Esta variável diz respeito ao código do município disponível no IBGE. Sua relevância se dá em função das características elétricas. Por exemplo, pequenos municípios com redes radiais, tendem a ter uma manobrabilidade menor do que os que possuem re-alimentação.
- Mês. O mês de execução da obra possui relevância, pois há uma relação entre meses secos e úmidos e a utilização da mão de obra de linha - viva na execução da obra. Importante destacar que estamos tratando do mês de execução da obra e não do projeto. Há uma janela de pelo menos três meses entre a elaboração do projeto e a execução da obra. Em função da reserva de materiais, é preciso que as empresas executoras possuam uma carteira mínima para manter a operação. Isto é monitorado de tal forma que os projetos estejam em uma janela de três a seis meses entre a elaboração do projeto e a execução da obra.
- Técnico. É o responsável pela elaboração do projeto. Baseado nas informações disponíveis no levantamento de campo, topologia e características da rede, este deve

sugerir um plano de execução. Nesta etapa, a experiência do profissional é fator relevante. Existem artifícios para redução de impacto nos desligamentos tais como instalação de chaves para seccionamento do circuito, emprego da atividade de linha - viva, execução de obras em conjunto...

- **Lotação.** Atualmente, a distribuidora trabalha com um modelo misto de elaboração de projetos, tendo em torno de sessenta por cento destes elaborados por técnicos próprios e o restante por empresas contratadas. Desta forma, há um critério maior na elaboração de projetos próprios.
- **Unidade de Serviço (US).** Esta variável trata do tamanho da obra. A quantidade de US que determinada obra utiliza, representa quanto serviço será empregado para a execução da mesma. Há de certa forma, uma relação de horas trabalhadas por US. Quanto maior for esta variável, maior será o fator homem hora empregado para execução da obra.
- **TAM.** Esta variável, representa o tipo de atividade a que cada obra irá se enquadrar. Podemos destacar a manutenção que é representada pela TAM 31, A expansão do sistema é a TAM 14, atendimento a cliente as TAM 24 e 26...
- **Dono.** Esta distribuidora, trabalha com um conceito de Dono de Obra. Este é o técnico responsável pelo levantamento de campo e fiscalização de obra. Sua influência é a de planejar a obra com visão de campo, e entregar informação de qualidade para que o técnico de projetos realizar um orçamento de qualidade. Após esta etapa, o dono deve garantir que a obra seja executada conforme orçamento.

Figura 15 – Tabela com exemplo dos dados utilizados

	CIDADE	MES	TECNICO	LOTACAO	US	TAM	DONO	RESULTADO
0	4316956	6	4000750	DROM-PROJE	176	24		ATENDE
1	4312617	6	2006442	DROM-PROJE	668	35		NATENDE
2	4307807	6	4003934	DROM-PROJE	28	31		NATENDE
3	4320701	6	4000659	DROM-PROJE	145	37		ATENDE
4	4305108	6	4000750	DROM-PROJE	51	24		NATENDE

Fonte: Autor

Além das variáveis previsoras destacadas acima, trabalhou-se com o histórico de assertividade como resultado. Foi medido o nível de acerto que o técnico de projetos tem quando são comparados o CHI previsto com o executado. Para fazer esta medição, considerou-se uma zona de acerto de $\pm 30\%$. Ao admitirmos como alvo 100% , consideramos a janela entre 70% e 130% como acerto e os fora deste, como erro. Dito isto, o modelo atual, baseado na análise do projetista, possui um nível de assertividade de $50,69\%$ ou

seja, há um erro de previsão na ordem de 49,31%. A figura 15, demonstra um exemplo dos dados utilizados.¹

6.3 Tratamento dos Dados

Para tratar os dados deste problema e transformá-los de forma que o modelo possa aprender e fazer previsões, utilizou-se os seguintes passos:

- Divisão dos dados em previsores e classes. A matriz de previsores conta com sete variáveis e 4658 linhas descritas abaixo. o vetor Classe, possui duas dimensões: Atende e Natende.
- Utilizando a biblioteca *Label Encoder*, as variáveis LOTACAO e DONO (do tipo *string*) foram transformadas em numérica.
- Utilizando a biblioteca *One Hot Encoder*, os valores numéricos foram transformados de forma que as variáveis estejam na mesma escala.
- Com a biblioteca *scikit-learn* foi realizada a divisão dos dados entre treinamento e teste. Para este, utilizou-se a proporção de quinze por cento em teste e oitenta e cinco por cento para treinamento. de forma que a base de treinamento ficou com 3959 linhas e a base de teste com 699.

6.4 Implementação do SVM

Seguindo a metodologia descrita, realizou-se a implementação do algoritmo Vetor de Suporte do inglês SVM utilizando a biblioteca *scikit-learn*. Este algoritmo busca classificar os dados encontrando o hiperplano que maximiza a margem entre as classes nos dados de treinamento (ALBON, 2018).

Na implementação do algoritmo SVM, com os dados apresentados na Tabela 6, foram encontrados os resultados citados abaixo.

PARÂMETRO	VALOR
Kernel	sigmoid
Tol	0.001
Fator g	3

Tabela 6 – Hiper Parâmetros Utilizados no Algoritmo SVM

A precisão do algoritmo teve resultado de 54,5%. E a matriz de confusão um resultado conforme segue abaixo.

¹ A variável Dono de obra utiliza os nomes reais dos técnicos. Para preservação de sua identidade, cobrimos os nomes com tarja preta.

Figura 16 – Matriz de Confusão do Algoritmo SVM, Adaptada

NATENDE	154	195
ATENDE	186	164
	ATENDE	NATENDE

Fonte: Autor

Na implementação do algoritmo SVM, encontrou-se uma precisão de 54,5%. Após analisar a Matriz de Confusão, percebe-se que o mesmo é 9% mais eficiente que a média de precisão dos técnicos de projetos.

6.5 Implementação de RNA

Seguindo a metodologia descrita, realizou-se a implementação do algoritmo de Redes Neurais Artificiais utilizando a biblioteca *scikit-learn*. Este algoritmo simula o funcionamento dos neurônios cerebrais. Segundo Raschka (2017) este conceito nasceu em 1943 com *McCulloch e W.Pitts* e foi consolidado mais tarde, em 1957 com *F. Rosenblatt*. O algoritmo basicamente aprende os coeficientes de pesos ideais que são então multiplicados com os recursos de entrada para a tomada de decisão se um neurônio dispara ou não. Assim como no algoritmo SVM, utilizou-se a ferramenta *GridSearch*, buscando os melhores parâmetros de configuração do algoritmo.

Na implementação do algoritmo RNA, foram utilizados os hiper parâmetros apresentados na Tabela 7

PARÂMETRO	VALOR
Função de Ativação	tanh
Learning Rate	constant
Número de Épocas	5000
Número de Neurônios na camada Escondida	100

Tabela 7 – Hiper Parâmetros Utilizados na RNA

A precisão do algoritmo teve resultado de 50,92%. E a matriz de confusão um

resultado conforme segue abaixo. O algoritmo SVM mostra-se mais eficiente do que o algoritmo RNA. Iremos tratar destas diferenças no capítulo que trata das discussões.

Figura 17 – Matriz de Confusão do Algoritmo RNA, Adaptada

NATENDE	174	185
ATENDE	171	169
	ATENDE	NATENDE

Fonte: Autor

Na implementação do algoritmo SVM, encontramos uma precisão de 50,92%. Após analisar a Matriz de Confusão, percebe-se que o mesmo teve uma assertividade de 1,84%. Após a implementação dos algoritmos, percebe-se que o SVM possui uma maior aderência ao objetivo de modelar o projetista. No próximo capítulo, serão explanadas as conclusões acerca dos resultados.

7 DISCUSSÃO E CONCLUSÃO

Este capítulo trata de uma análise mais aprofundada, uma discussão, sobre os resultados e suas interpretações. Portanto o objetivo do capítulo é tangibilizar o conhecimento tácito do sistema sobre as observações do modelo. Os algoritmos de inteligência artificial vem sendo amplamente utilizados para solução de problemas complexos conforme a revisão sistemática nos mostrou. Esta metodologia vem sendo aplicada nas mais variadas áreas e com os mais variados objetivos, sejam otimização, classificação, priorização, etc. No estudo dos trabalhos correlatos, fica evidente que a escolha das variáveis é uma das, senão a mais importante etapa quando buscamos solucionar problemas utilizando estas técnicas. A escolha das variáveis parece ser simples, mas depende daquilo que é medido e está disponível. Nem sempre os dados disponíveis são relevantes ao que se deseja modelar.

Neste caso, não houve uma preparação prévia de medições. Utilizou-se as variáveis disponíveis no processo. Elas basicamente comparam a previsão feita pelo projetista com o resultado real da execução em campo. Não é escopo deste trabalho estes tema, mas faz-se necessário trazer que esta assertividade é da ordem de 50%. Como não encontramos trabalhos com esta abordagem, assumiu-se que 50% de assertividade corresponde a 100% e as análises serão conduzidas sobre esta dimensão. Baseado nisso, buscamos entender o contexto ao qual determinado projeto estava inserido. A cidade, é uma variável que está ligada à empresa executora, ao conjunto elétrico que caracteriza o tipo de rede. O mês de execução, está ligado aos fatores climáticos com chuva sol, umidade. O técnico é objetivo de estudo. A lotação nos traz fatores culturais aos quais os projetos foram elaborados. A Unidade de serviço (US) nos mostra o tamanho da obra, a TAM o tipo de obra e o Dono, traz qual é o responsável por aceitar ou não alterações de projeto.

A melhor modelagem do projetista, nos trouxe um resultado de 9%. O que representa um ganho em relação ao modelo atual. Podemos dizer, em outras palavras, que o modelo utilizando na modelagem é mais eficiente do que a média de acerto dos projetistas. Em primeira análise se o algoritmo for utilizado, teremos uma redução de 9% do DEC Programado (obras) ou que podemos investir 9% a mais anualmente. Para se ter uma ideia de grandeza, os investimentos previstos para 2022 na distribuição possuem um montante da ordem de R\$ 1 bilhão de reais, com a utilização deste modelo, é possível manutencionar em torno de 250 km de rede ou construir 38 km de novos alimentadores e este *saving* pode resolver os problemas de um município inteiro a cada ciclo.

Não obstante, sob a luz do processo atual, restam alguns questionamentos. O principal deles trata de verificar se é possível ter uma acurácia melhor do modelo de previsão. Como visto anteriormente, trabalhamos com variáveis que tratam puramente

do processo estudado. Entretanto, existem algumas variáveis (indicadores) que podem ser incorporados ao estudo e que podem melhorar o resultado. Dentre elas, destacamos a divergência de projetos. Este indicador, trata da diferença do que é projetado e executado. Embora não haja uma correlação direta entre eles ¹, a divergência de projetos, pode ser utilizada como uma constante de cada cidade, ou seja, determinado município é mais ou menos fiel de acordo com a divergência.

Outra variável que pode ser considerada, é o atraso de desligamento. Esta variável, trata da assertividade entre o que foi planejado e executado em relação ao tempo de intervenção na rede e assim como a divergência, seria uma constante vinculada ao município.

Temos também, o retrabalho em obras. Esta trata da qualidade de execução e sua relevância se apresenta pois quando há um retrabalho, é necessário muitas vezes desligar novamente a rede e isto gera um erro.

Como mostrado pela revisão sistemática apresentada nesta pesquisa, estamos desbravando uma área pouco explorada dentro da academia porém de grande relevância no cenário da distribuição de energia no Brasil. Para cumprir as metas estipuladas pelo regulador, as companhias precisam cada vez mais melhorar a infraestrutura aumentando os investimentos e tendo como limitador, os mesmo indicadores que precisam ser melhorados. Esta espiral somente se sustenta, se houver um racionamento do tempo disponível para obras. Em outras palavras, é preciso ter uma inteligência de ao mesmo tempo que se desliga uma rede, haja uma garantia de que a mesma não será desligada em um curto espaço de tempo e que esta intervenção, gere um ganho no indicador de forma que o desligamento de um ciclo, seja compensado no ciclo posterior. Outra questão relevante, é juntar a estas variáveis uma leitura da topologia da rede, tipos de condutor, tipos de equipamento de manobra e demais elementos que podem auxiliar na assertividade da previsão.

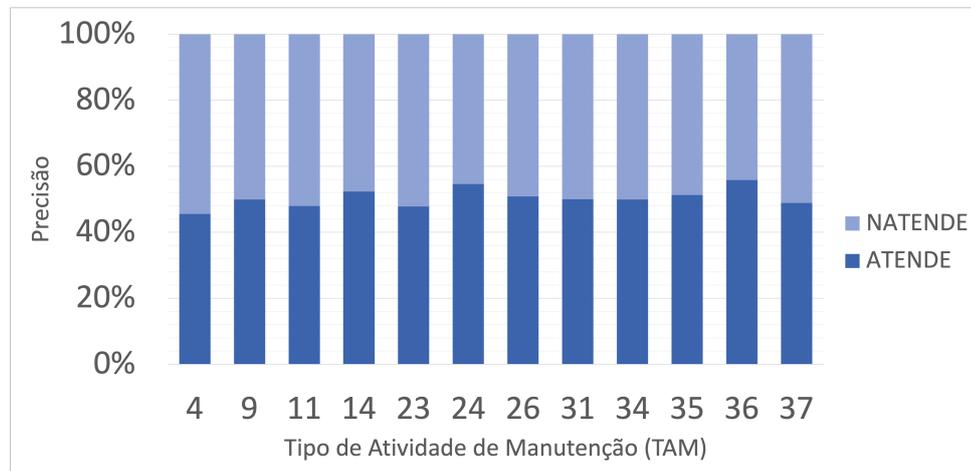
Após este estudo, constatou-se que é possível que um projetista atinja até 109% de asserto em suas previsões, isto mostra que o algoritmo é aderente. Ademais, o estudo pode ser incrementado trazendo outras variáveis que embora indiretamente corroboram para que seja possível melhorar tanto a acurácia do modelo como do técnico real estima-se que acurácias na ordem de 120% podem ser alcançadas, trazendo estas variáveis ao modelo. Este trabalho pode servir de início para novas pesquisas, que aprofundem o tema e possam melhorar os resultados financeiros das empresas de energia elétrica. Para verificar as causas de erro é possível a criação de indicadores para acompanhar os principais causadores de desvio. Em posse destes, pode-se traçar um plano de ação e qualificação dos Técnicos. Técnicas de análise e solução de problemas podem ser utilizadas para estruturar esta análise.

Como contribuição para trabalhos futuros, podemos destacar a influência do Tipo de Atividade de Manutenção (TAM). Percebe-se que algumas atividades possuem maior

¹ As variáveis estudadas tratam do momento da execução, a divergência de projetos só é medida após o pagamento da obra, em torno de setenta e cinco dias após.

influência sobre a precisão.²

Figura 18 – Precisão por Tipo de atividade de Manutenção Análise dos dados coletados. ATENDE, observação dentro do espaço de tolerância, NATENDE observação fora do espaço de tolerância.



Fonte: Autor

Na Figura18 observa-se que embora os dados sejam dispersos, há algumas atividade cujo índice de precisão é maior quando analisamos o acerto (atende), cabe aqui, aprofundar as análises e buscar entender o comportamento desta variável.

Ao observar atentamente, observa-se certa sazonalidade na precisão. As obras executadas no segundo trimestre, possuem uma maior precisão. Percebe-se na Figura19 esta tendência. Quando observamos o cenário anual, temos geralmente o período de férias dos executores no primeiro trimestre. No estado do Rio Grande do Sul, existem episódios de tempo severo em alguns períodos do ano.

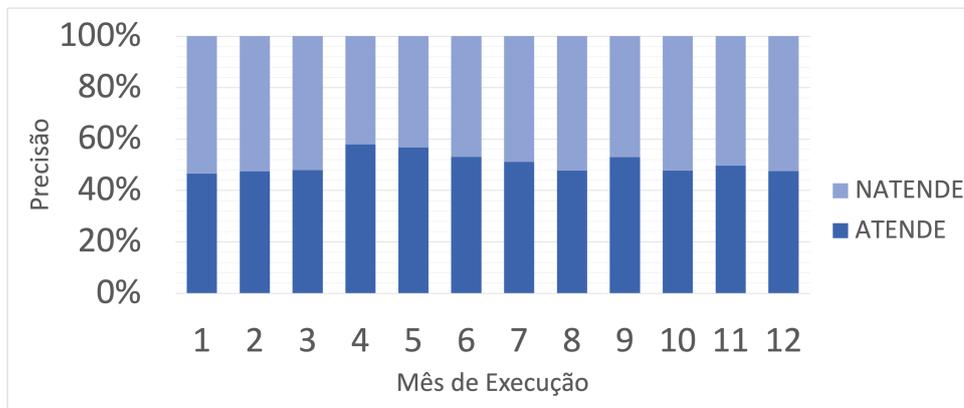
Esta constatação reforça que um bom planejamento e organização são fundamentais para a execução das obras e que estes índices contribuem para a assertividade buscada. As obras executadas com segurança, agilidade e capricho são aquelas cujo planejamento foi bem feito.

Após cumprir os objetivos iniciais, restam algumas provocações que podem servir de inspiração para trabalhos futuros. São elas:

- Dada a relevância do investimento em infraestrutura no Brasil, é preciso avançar na eficiência operacional em Obras e Manutenção de redes elétricas;
- Mapear as empresas para *benchmark* na eficiência de investimento x DEC identificando as melhores práticas;
- Explorar variáveis controladas com balizador na assertividade e acurácia do modelo;

² A fim de tornar as análises qualitativas, foi convencionado que os dados com precisão de $\pm 70\%$ são classificados em ATENDE e os dados fora deste espaço de tolerância como NATENDE.

Figura 19 – Precisão por Tipo de atividade de Manutenção Análise dos dados coletados. ATENDE, observação dentro do espaço de tolerância, NATENDE observação fora do espaço de tolerância.



Fonte: Autor

- identificar o modelo de negócio sustentável para promover o equilíbrio do investimento, DEC e melhoras nos indicadores no médio e longo prazo.

Estas provocações servem de incentivo para que a comunidade acadêmica possa desenvolver melhor a tecnologia do setor elétrico. Diante do desafio de crescimento do país, é preciso evoluir este setor tão importante.

REFERÊNCIAS

- ALBON, C. *Machine Learning with Python Cookbook: Practical Solutions from Preprocessing to Deep Learning*. 1. ed. [S.l.]: O'Reilly Media, 2018. ISBN 1491989386, 9781491989388. Citado 3 vezes nas páginas 37, 54 e 58.
- ALPAYDIN, E. *Introduction to machine learning*. 3ed.. ed. [S.l.]: The MIT Press, 2014. (Adaptive Computation and Machine Learning series). ISBN 0262028182,978-0-262-02818-9. Citado 3 vezes nas páginas 35, 36 e 37.
- Álvarez, Esther Díaz, F. Framework for the dynamic scheduling of complex job shops. *International Journal of Manufacturing Technology and Management*, 2007. ISSN 13682148. Citado na página 40.
- Contrato de Concessão*. Disponível em: <<https://www.aneel.gov.br/documents/10184/15062987//Contrato+de+Concess~ao.pdf>>. Citado 2 vezes nas páginas 21 e 27.
- Prodist Módulo 2*. Disponível em: <https://www.aneel.gov.br/documents/656827/14866914/Módulo2_Revis~ao7.pdf/c1cf5bd8-b2bc-4d57-9b42-285a7fd8c2a5>. Citado na página 32.
- Prodist Módulo 8*. Disponível em: <<https://www.aneel.gov.br/modulo-8>>. Citado 2 vezes nas páginas 28 e 31.
- Prodist Módulo 8*. Disponível em: <<https://www.aneel.gov.br/modulo-8>>. Citado 4 vezes nas páginas 28, 29, 30 e 31.
- Resolução Normativa 414*. Disponível em: <<https://www.aneel.gov.br/documents/656877/14486448/bren2010414.pdf/3bd33297-26f9-4ddf-94c3-f01d76d6f14a?Version=1.0>>. Citado na página 33.
- RGE*. Disponível em: <<https://www.cpfl.com.br/unidades-de-negocios/distribuicao/rge/Paginas/default.aspx>>. Citado na página 56.
- CRISTIANINI, J. S.-T. N. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. 1. ed. [S.l.]: Cambridge University Press, 2000. ISBN 9780521780193,0521780195. Citado na página 38.
- GHEZAVATI, V. R. A new stochastic mixed integer programming to design integrated cellular manufacturing system: A supply chain framework. *International Journal of Industrial Engineering Computations*, 2011. ISSN 19232926. Citado 2 vezes nas páginas 39 e 49.
- HAMEED ABDUL RAZA, t. A decision support tool for bi-objective risk-based maintenance scheduling of an LNG gas sweetening unit. *Journal of Quality in Maintenance Engineering*, 2019. ISSN 13552511. Citado 2 vezes nas páginas 39 e 49.
- HUANG WEI WANG, t. Research on integrated benefit evaluation model of microgrid. 2012. ISSN 21617481. Citado 2 vezes nas páginas 41 e 49.

- L., F. *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*. [S.l.: s.n.], 1994. Citado na página 38.
- MEI, K.; LI, t. Multi-objective optimization of collation delay and makespan in mail-order pharmacy automated distribution system. *International Journal of Advanced Manufacturing Technology*, The International Journal of Advanced Manufacturing Technology, v. 83, n. 1-4, p. 475–488, 2016. ISSN 14333015. Disponível em: <<http://dx.doi.org/10.1007/s00170-015-7555-7>>. Citado 2 vezes nas páginas 41 e 49.
- MOHRI AFSHIN ROSTAMIZADEH, A. T. M. *Foundations of Machine Learning*. [S.l.]: The MIT Press, 2012. (Adaptive Computation and Machine Learning). ISBN 026201825X; 9780262018258. Citado 2 vezes nas páginas 35 e 36.
- PETERSEN, t. Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, v. 64, p. 1–18, aug 2015. ISSN 09505849. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0950584915000646>>. Citado 2 vezes nas páginas 41 e 44.
- RAHMAN, t. A real-time order acceptance and scheduling approach for permutation flow shop problems. *European Journal of Operational Research*, 2015. ISSN 03772217. Citado 2 vezes nas páginas 40 e 49.
- Rajan Kumar, V. A restart scheme based meta heuristic for flexible job shop scheduling. *International Journal of Mechanical and Production Engineering Research and Development*, 2018. ISSN 22498001. Citado 2 vezes nas páginas 40 e 49.
- RASCHKA, V. M. S. *Python Machine Learning*. [S.l.]: Packt Publishing, 2017. v. 2nd ed. Citado 7 vezes nas páginas 36, 37, 51, 52, 53, 55 e 59.
- TAJBAKHSI, t. Multi-objective assembly permutation flow shop scheduling problem: A mathematical model and a meta-heuristic algorithm. *Journal of the Operational Research Society*, 2014. ISSN 14769360. Citado 2 vezes nas páginas 40 e 49.
- YIMER, A. D.; DEMIRLI, K. A genetic approach to two-phase optimization of dynamic supply chain scheduling. *Computers & Industrial Engineering*, v. 58, n. 3, p. 411–422, apr 2010. ISSN 03608352. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0360835209000126>>. Citado 2 vezes nas páginas 39 e 49.
- ZHENG, A. *Evaluating machine learning models : a beginner's guide to key concepts and pitfalls*. First edition. [S.l.]: O'Reilly Media, 2015. Citado 5 vezes nas páginas 17, 37, 38, 53 e 54.