



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CAMPUS ARARANGUÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENERGIA E SUSTENTABILIDADE

Eliei Marcos Rocha Romancini

**Desagregação de cargas residenciais: detecção de eventos e
classificação em ambiente real**

Araranguá

2022

Eliel Marcos Rocha Romancini

**Desagregação de cargas residenciais: detecção de eventos e
classificação em ambiente real**

Dissertação submetida ao Programa de Pós-Graduação em Energia e Sustentabilidade da Universidade Federal de Santa Catarina para a obtenção do título de Mestre em Sistemas de Energia.

Orientador: Prof. Fabrício de Oliveira Ourique, Dr.
Coorientador: Prof. Ricardo Alexandre Reinaldo de Moraes, Dr.

Araranguá

2022

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Romancini, Eliel Marcos Rocha

Desagregação de cargas residenciais : detecção de eventos e classificação em ambiente real / Eliel Marcos Rocha Romancini ; orientador, Fabrício de Oliveira Ourique, coorientador, Ricardo Alexandre Reinaldo de Moraes, 2022. 149 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Campus Araranguá, Programa de Pós-Graduação em Energia e Sustentabilidade, Araranguá, 2022.

Inclui referências.

1. Energia e Sustentabilidade. 2. NILM. 3. Desagregação de Energia. 4. Aprendizado de Máquina. 5. Eficiência Energética. I. Ourique, Fabrício de Oliveira. II. Moraes, Ricardo Alexandre Reinaldo de. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Energia e Sustentabilidade. IV. Título.

Eliei Marcos Rocha Romancini

**Desagregação de cargas residenciais: detecção de eventos e
classificação em ambiente real**

O presente trabalho em nível de Mestrado foi avaliado e aprovado, em 10 de outubro de 2022, pela banca examinadora composta pelos seguintes membros:

Profa. Analúcia Schiaffino Morales, Dra.
Universidade Federal de Santa Catarina

Prof. Iwens Gervásio Sene Junior, Dr.
Universidade Federal de Goiás

Certificamos que esta é a versão original e final do trabalho de conclusão que foi julgado adequado para obtenção do título de Mestre em Energia e Sustentabilidade.

Coordenação do Programa de Pós-Graduação

Prof. Fabrício de Oliveira Ourique, Dr.
Orientador

Prof. Ricardo Alexandre Reinaldo de Moraes, Dr.
Coorientador

Araranguá, 2022.

Dedico os frutos deste estudo aos que acreditaram na minha
capacidade e me incentivaram neste desafio.

RESUMO

A gestão prudente do consumo de eletricidade em residências está se tornando cada vez mais importante. Adotar comportamentos que reduzem o consumo da energia utilizada contribuem na redução de desperdício e favorecem a eficiência energética. O monitoramento de carga não intrusivo (*Non-Intrusive Load Monitoring - NILM*) residencial é uma forma de contribuir na supervisão dos equipamentos elétricos e uma ferramenta de informação para aplicações que envolvem *feedback* aos consumidores, no ajuste de hábitos e uso mais conservador dos equipamentos em suas edificações. Portanto, este trabalho tem como objetivo apresentar um método para desagregação de cargas elétricas em um ambiente real, utilizando sistema de monitoramento não intrusivo com detector probabilístico de eventos e três classificadores distintos com conceitos de aprendizado de máquina, que possuem características desejáveis para aplicações práticas. O método proposto foi avaliado através de um cenário prático, verificando se o monitoramento é eficiente quando a etapa de treinamento é realizada com uma coleção pública de sinais elétricos de outra realidade (país ou região) e, posteriormente, aplicado em ambiente residencial com características distintas. As validações do sistema são realizadas através de métricas definidas na literatura para sistema de monitoramento. Ao aplicar a metodologia verificou-se que um conjunto público de dados, mesmo não correspondente ao conjunto local dos equipamentos monitorados, é suficiente para a base de implementação de sistemas *NILM*. Os resultados mostram através do índice *Kappa*, que o desempenho do classificador por rede neural artificial obteve o maior grau de concordância, entre os dados esperados e apresentados, para a desagregação de cargas em ambiente residencial. Além disso, o trabalho contribui com a disponibilidade de um conjunto de sinais elétricos de acesso público que podem ser utilizados em trabalhos similares. Tal estratégia favorece a comunidade científica nacional, no que diz respeito, à disponibilidade de uma metodologia para o desenvolvimento e avaliação de sistemas *NILM*, visto que, conjuntos de dados são escassos para a realidade brasileira.

Palavras-chave: *NILM*; Desagregação de Energia; Aprendizado de Máquina; Eficiência Energética.

ABSTRACT

The prudent management of electricity consumption in households is becoming increasingly important. The adoption of behaviors that reduce the consumption of energy used contributes to reduce waste and favoring energy efficiency. Non-intrusive load monitoring (NILM) is a way to contribute to the supervision of electrical equipment and an information tool for applications that involve consumer feedback in the adjustment of habits and more conservative use of equipment in their buildings. Therefore, this paper aims to present a method for electrical load disaggregation in a real environment, using non-intrusive monitoring system with probabilistic event detector and three distinct classifiers with machine learning concepts, which have desirable characteristics for practical applications. The proposed method was evaluated through a functional scenario, verifying if the monitoring is efficient when the training stage is performed with a public collection of electrical signals from another reality (country or region) and, later, applied in a residential environment with distinct characteristics. The system validations are performed through metrics defined in the literature for monitoring systems. By using the methodology it was verified that a public data set, even if it does not correspond to the local set of monitored equipment, is sufficient for NILM system implementation. The results show, through the Kappa index, that the performance of the artificial neural network classifier obtained the highest degree of agreement, between the expected and presented data, for load disaggregation in a residential environment. In addition, the work contributes with the availability of a set of publicly accessible electrical signals that can be used in similar work. This strategy favors the national scientific community, regarding the availability of a methodology for developing and evaluating NILM systems, since data sets are scarce for the Brazilian reality.

Keywords: NILM; Energy Disaggregation; Machine Learning; Energy Efficiency.

LISTA DE FIGURAS

Figura 1 – Sistema de monitoramento de cargas intrusivo.	25
Figura 2 – Sistema de monitoramento de cargas semi-intrusivo.....	27
Figura 3 – Sistema de monitoramento de cargas não intrusivo.....	28
Figura 4 – Desagregação por categoria não baseada em eventos.	31
Figura 5 – Desagregação por categoria baseada em eventos.	32
Figura 6 – Diagrama de fluxo <i>NILM</i> , processo para desagregação de cargas.....	34
Figura 7 – Relação da detecção probabilística <i>LLR</i> e as curvas de potência.	39
Figura 8 – Classificação de uma nova instância pelo <i>KNN</i>	44
Figura 9 – Conversão dimensional utilizando <i>Kernel RBF</i> no classificador <i>SVM</i>	46
Figura 10 – Formação do hiperplano em conjuntos linearmente separáveis.	46
Figura 11 – Modelo da rede neural artificial <i>MLP</i> e neurônio artificial <i>Perceptron</i>	48
Figura 12 – Fluxograma da metodologia desenvolvida.	61
Figura 13 – Preenchimento da janela deslizante com amostras da potência.....	65
Figura 14 – Definição dos índices de ativação e desativação de eventos.	69
Figura 15 – Extração das amostras para calcular a diferença de potencial.	73
Figura 16 – Extração da amostra de pico máximo.	74
Figura 17 – Extração das amostras da potência para calcular a distorção do evento.	76
Figura 18 – Sistema de monitoramento sugerido em ambiente real.	81
Figura 19 – Medidor utilizado no monitoramento para o ambiente real.....	82
Figura 20 – Sensores e central de monitoramento instalados no ambiente.....	84
Figura 21 – Processo de verificação do sistema com dados do ambiente real.	85
Figura 22 – Normalização no sinal da potência.....	87
Figura 23 – Aplicação do filtro média móvel no sinal da potência.	89
Figura 24 – Algoritmo <i>LLR</i> detectando eventos no sinal agregado.	91
Figura 25 – Gráfico de eventos do aspirador de pó no conjunto <i>FIREED</i>	92
Figura 26 – Gráfico de eventos da chaleira elétrica no conjunto <i>FIREED</i>	93
Figura 27 – Gráfico de eventos do refrigerador no conjunto <i>FIREED</i>	94
Figura 28 – Gráfico de eventos do secador de cabelo no conjunto <i>FIREED</i>	95
Figura 29 – Gráfico de eventos do televisor no conjunto <i>FIREED</i>	96
Figura 30 – Avaliação de desempenho do <i>LLR</i> com o conjunto <i>FIREED</i>	98

Figura 31 – Avaliação de desempenho do <i>LLR</i> com dados do ambiente real SECRBR.	98
Figura 32 – Eventos de ativação do refrigerador e televisor.	99
Figura 33 – Diferença de potencial das assinaturas no plano PQS – <i>FIRE</i> D.	101
Figura 34 – Diferença de potencial das assinaturas no plano PQS – SECRBR.	101
Figura 35 – Fluxo de trabalho no <i>Orange Data Mining</i>	102
Figura 36 – Avaliação por classe no <i>KNN</i> , <i>SVM</i> e <i>ANN-MLP</i> com a base <i>FIRE</i> D. .	105
Figura 37 – Avaliação por <i>micro-averaging</i> e <i>macro-averaging</i> no <i>KNN</i> , <i>SVM</i> e <i>ANN-MLP</i> com a base <i>FIRE</i> D.	106
Figura 38 – Avaliação por matriz de confusão no <i>KNN</i> , <i>SVM</i> e <i>ANN-MLP</i> com a base <i>FIRE</i> D.	107
Figura 39 – Avaliação por classe no <i>KNN</i> , <i>SVM</i> e <i>ANN-MLP</i> com a base SECRBR.	109
Figura 40 – Avaliação por <i>micro-averaging</i> e <i>macro-averaging</i> no <i>KNN</i> , <i>SVM</i> e <i>ANN-MLP</i> com a base SECRBR.	110
Figura 41 – Avaliação por matriz de confusão no <i>KNN</i> , <i>SVM</i> e <i>ANN-MLP</i> com a base SECRBR.	111
Figura 42 – Melhor ocorrência por matriz de confusão com a base SECRBR.	112
Figura 43 – Boxplot da estatística <i>Kappa</i> com <i>KNN</i> , <i>SVM</i> , <i>ANN-MLP</i>	113

LISTA DE QUADROS

Quadro 1 – Tipos de cargas.....	29
Quadro 2 – Descrição das instâncias utilizadas nas métricas de avaliação.....	52
Quadro 3 – Parâmetros ajustáveis do detector de eventos.....	70
Quadro 4 – Estrutura dos dados da base de assinaturas.	77
Quadro 5 – Hiperparâmetros definidos com auxílio da ferramenta.	79
Quadro 6 – Equipamentos presentes no conjunto <i>FIRE</i> D e ambiente local.....	83

LISTA DE TABELAS

Tabela 1 – Diferentes faixas de <i>Kappa</i>	56
Tabela 2 – Principais conjuntos de dados públicos e suas características.	57
Tabela 3 – Parâmetros de ajuste utilizados na implementação do <i>LLR</i>	90
Tabela 4 – Aspirador de pó, eventos detectados e comparados com o <i>ground truth</i>	92
Tabela 5 – Chaleira elétrica, eventos detectados e comparados com o <i>ground truth</i>	93
Tabela 6 – Refrigerador, eventos detectados e comparados com o <i>ground truth</i>	94
Tabela 7 – Secador de cabelo, eventos detectados e comparados com o <i>ground truth</i>	95
Tabela 8 – Televisor, eventos detectados e comparados com o <i>ground truth</i>	96
Tabela 9 – Contagem dos valores de <i>TP</i> , <i>FP</i> e <i>FN</i> do <i>LLR</i>	97
Tabela 10 – Classes das bases de assinaturas.	100
Tabela 11 – Hiperparâmetros utilizados nos classificadores.	103
Tabela 12 – Índice <i>Kappa</i> nas iterações de avaliação com <i>KNN</i> , <i>SVM</i> e <i>ANN-MLP</i>	113

LISTA DE ABREVIATURAS E SIGLAS

AMPds – Almanac of Minutely Power dataset
ANEEL – Agência Nacional de Energia Elétrica
ANN – Artificial Neural Network
BLUED – Building-Level fUlly-labeled dataset for Electricity Disaggregation
BP – Backpropagation
BSS – Blind Source Separation
CNN – Convolutional Neural Network
COOLL – Controlled On/Off Loads Library
ECO – Electricity Consumption and Occupation
FHMM – Factorial Hidden Markov Model
FIREd – Fully-labeled high-fRequency Electricity Disaggregation
FN – False Negative
FP – False Positive
FSM – Finite State Machine
FT – Fourier Transform
GA – Genetic Algorithms
GLR – Generalized Likelihood Ratio
Hz – Hertz
ILM – Intrusive Load Monitoring
KNN – K-Nearest Neighbors
LLD – Log Likelihood Detector
LLR – Log Likelihood Ratio
MHz – Megahertz
ML – Machine Learning
MLP – Multi-layer Perceptron
MME – Ministério de Minas de Energia
NB – Naïve Bayes
NILM – Non-Intrusive Load Monitoring
ODS – Objetivos de Desenvolvimento Sustentável
ONU – Organização das Nações Unidas
PGM – Probabilistic Graphical Models
PLAID – Plug-Level Appliance Identification Dataset

PROCEL – Programa Nacional de Conservação de Energia Elétrica

RBF – Radial Basis Function

REDD - Reference Energy Disaggregation Data Set

SECRBR – Sinais Elétricos de Cargas Residenciais - Brasil

SILM – Semi Intrusive Load Monitoring

SVM – Support Vector Machine

TC – Transformador de Corrente

TN – True Negative

TP – Transformador de Potencial

TP – True Positive

UK-DALE – UK Domestic Appliance-Level Electricity

W – Watt

WHITE – Worldwide Household and Industry Transient Energy Data Set

ZLSC – Zero Loop Sum Constraint

SUMÁRIO

1	INTRODUÇÃO	17
1.1	MOTIVAÇÃO E JUSTIFICATIVA.....	19
1.2	QUESTÃO DE PESQUISA E HIPÓTESE	21
1.3	OBJETIVOS	21
1.3.1	Objetivo geral.....	22
1.3.2	Objetivos específicos.....	22
1.4	ESTRUTURA DO TRABALHO	23
2	REFERENCIAL TEÓRICO.....	24
2.1	TÉCNICAS DE MONITORAMENTO DE CARGAS ELÉTRICAS	24
2.1.1	Monitoramento intrusivo de cargas.....	24
2.1.2	Monitoramento semi-intrusivo de cargas	26
2.1.3	Monitoramento não intrusivo de cargas	27
2.2	CATEGORIAS DE MONITORAMENTO NÃO INTRUSIVO.....	30
2.2.1	Não baseada em eventos.....	30
2.2.2	Baseada em eventos	32
2.3	ETAPAS PARA DESAGREGAÇÃO	33
2.3.1	Aquisição dos dados e pré-processamento	34
2.3.2	Detecção de eventos.....	36
2.3.3	Extração de características	39
2.3.4	Classificação de eventos.....	41
2.3.4.1	<i>K-Nearest Neighbors.....</i>	44
2.3.4.2	<i>Support Vector Machine</i>	45
2.3.4.3	<i>Artificial Neural Network</i>	47
2.3.5	Predição do consumo de energia	50
2.4	MÉTRICAS DE AVALIAÇÃO.....	52
2.5	CONJUNTO DE DADOS PÚBLICOS.....	56

2.6	CONCLUSÃO	59
3	METODOLOGIA	60
3.1	ETAPA I.....	61
3.1.1	Aquisição dos dados	61
3.1.2	Normalização dos dados e filtragem	62
3.2	ETAPA II.....	63
3.2.1	Identificação dos eventos.....	63
3.2.2	Avaliação do detector de eventos.....	70
3.3	ETAPA III.....	71
3.3.1	Extração das características	71
3.3.2	Seleção dos dados de treinamento	76
3.4	ETAPA IV	78
3.4.1	Treinamento dos classificadores.....	78
3.4.2	Avaliação dos classificadores.....	79
3.5	ETAPA V	80
3.5.1	Avaliação em ambiente real	80
4	RESULTADOS.....	86
4.1	CONJUNTOS DE DADOS UTILIZADOS	86
4.2	TRATAMENTO DOS DADOS	87
4.3	APLICAÇÃO E AVALIAÇÃO DO <i>LLR</i>	90
4.4	FORMAÇÃO E SELEÇÃO DAS ASSINATURAS ELÉTRICAS	100
4.5	TREINO E AVALIAÇÃO DOS CLASSIFICADORES.....	102
4.6	VALIDAÇÃO DO SISTEMA NO AMBIENTE REAL	107
5	CONCLUSÕES	115
5.1	PRINCIPAIS CONTRIBUIÇÕES	117
5.2	ESTUDOS FUTUROS.....	117
	REFERÊNCIAS.....	119

APÊNDICE A – Eventos de ativação e desativação do conjunto público <i>FIRE</i>D	132
APÊNDICE B – Eventos de ativação e desativação do conjunto de sinais do ambiente real <i>SECRBR</i>	133
APÊNDICE C – Recursos no plano PQS do conjunto público <i>FIRE</i>D	134
APÊNDICE D – Recursos no plano PQS do conjunto de sinais do ambiente real <i>SECRBR</i>	135
APÊNDICE E – Resultados de teste e avaliação do conjunto público <i>FIRE</i>D	136
APÊNDICE F – Resultados das predições do conjunto de sinais do ambiente real <i>SECRBR</i>	140
APÊNDICE G – Resultados por matriz de confusão do conjunto público <i>FIRE</i>D	144
APÊNDICE H – Resultados por matriz de confusão do conjunto de sinais do ambiente real <i>SECRBR</i>	147

1 INTRODUÇÃO

A eficiência energética faz parte do sétimo objetivo entre os dezessete Objetivos de Desenvolvimento Sustentável (ODS) definidos pela ONU, apontando que até o ano de 2030 é preciso dobrar a taxa global de melhoria da eficiência energética (ONU, 2015). Logo, para alcançar essa meta são necessárias medidas de incentivo para gestão dos recursos energéticos da geração até o consumo final.

Através das pesquisas e levantamento de dados oficiais da Agência Nacional de Energia Elétrica (ANEEL) e do Ministério de Minas e Energia (MME), os registros entre o período de 2003 até 2021 indicam que, de toda energia elétrica gerada no Brasil: 20% foi consumida pelo setor industrial, 22% pelo setor comercial, 39% pelas residências e o restante pelos demais setores. Os resultados evidenciam que o setor residencial é o maior consumidor de energia elétrica no país, além de apresentarem informações que também é o setor com maior desperdício do recurso (ANEEL, 2021).

A gestão prudente do consumo de eletricidade em residências exige assumir comportamentos que reduzem o consumo e amenizam os impactos da energia utilizada, contribuindo na redução de desperdício e favorecendo a eficiência energética.

Apesar do governo incentivar e desenvolver programas nacionais que contribuem para a redução de desperdício do setor residencial, destacando o selo do Programa Nacional de Conservação de Energia Elétrica (PROCEL), os consumidores não possuem um meio transparente de monitorar o seu uso para modificarem seus comportamentos. Além disso, estudos nacionais (ABRAHÃO; SOUZA, 2021; ELETROBRAS, 2019; EPE, 2018) e internacionais (FUNGSIRIRUT; BENJAPOLAKUL, 2017), sobre o comportamento de usuários residenciais, mostram que muitas vezes eles são falhos em estimar o consumo de toda a casa ou de equipamentos individuais. Os usuários, geralmente, subestimam a energia dos dispositivos com alto consumo e superestimam dispositivos com baixo consumo, levando a um maior gasto total.

Acredita-se que se houvesse um sistema instalado nas residências para fornecer um *feedback* do consumo detalhado, sobre quais dispositivos usam mais energia, os usuários poderiam ajustar seus comportamentos para fazer um uso mais eficiente dos eletrodomésticos. A alteração de hábitos poderia levar a reduções do

consumo e alcançar economia de energia no domínio residencial na ordem de até 20% (EHRHARDT-MARTINEZ et al., 2010; ZHUANG; SHAHIDEHPOUR; LI, 2018).

Medidores inteligentes de eletricidade podem ser um mecanismo para fornecer *feedback* aos consumidores. Um relatório emitido sobre o programa de implementação de medidores inteligentes no Reino Unido, mostra que em uma das companhias envolvidas, *Bristh Gas*, seus consumidores afirmaram ter tomado medidas para reduzir o consumo de energia desde a instalação de um medidor inteligente. Dentre os entrevistados, 62% afirmaram que os medidores inteligentes fizeram diferença na forma como eles usam a energia em casa, além de realizarem melhorias em suas residências em termos de eficiência energética, e 73% alteraram suas atividades de forma mais eficiente e consciente (GOV.UK, 2019).

Entretanto, em cenário nacional as pesquisas relacionadas a implantação de medidores inteligentes ainda precisam maturar. No relatório emitido por *Energy Partnership German-Brazilian* (2021), os projetos pilotos registrados pelas concessionárias junto à ANEEL, iniciados entre 2010 e 2016, não foram completadas etapas de acompanhamento do consumo posteriores aos projetos ou gerados indicadores que representem as mudanças de comportamento esperadas.

Logo, para que o mecanismo de *feedback* seja efetivo no hábito do usuário residencial, é necessário que as informações dos equipamentos sejam coletadas individualmente. Dessa forma, o monitoramento de cargas elétricas é apresentado como uma importante ferramenta para extrair os dados de consumo individual dos eletrodomésticos nas residências, podendo ser utilizado em sistemas de *feedback* ao consumidor final.

Conseqüentemente, o aperfeiçoamento e o desenvolvimento de novas técnicas de monitoramento de consumo, que refletem na mudança de hábitos e redução do uso da energia, têm se mostrado como uma temática cada vez mais relevante. Além disso, o conhecimento detalhado quanto ao tipo de carga, consumo de energia por equipamento na residência e registro dos ciclos de operação, são informações que podem interessar não só aos consumidores, mas também concessionárias, distribuidoras de energia, fabricantes de eletrodomésticos e formuladores de políticas (ALADESANMI; FOLLY, 2015; ZHUANG; SHAHIDEHPOUR; LI, 2018).

Neste contexto, há um esforço significativo de pesquisa dedicado ao desenvolvimento de métodos sobre o monitoramento de carga não intrusivo (*Non-Intrusive Load Monitoring - NILM*). Esta tecnologia é capaz de detectar e desagregar o consumo de energia utilizando um número limitado de medidores instalados na rede de distribuição elétrica, divergindo do monitoramento intrusivo, que necessita da implantação de vários sensores em toda residência.

Apesar da literatura conter vários estudos sobre o monitoramento de carga não intrusivo, a proposta sobre utilizar esta técnica em cenários do mundo real, como uma ferramenta para redução do consumo de energia, ainda é recente e de grande necessidade para adoção nos próximos anos (KELLY; KNOTTENBELT, 2016; SUN et al., 2019).

1.1 MOTIVAÇÃO E JUSTIFICATIVA

Devido à importância do monitoramento dos equipamentos presentes no ambiente residencial para auxiliar o consumidor, neste trabalho é apresentado em maiores detalhes os desafios para o desenvolvimento do monitoramento não intrusivo de cargas elétricas para aplicação em ambiente real.

O monitoramento *NILM* apresenta significativa complexidade no desenvolvimento dos algoritmos que realizam a desagregação da energia, sendo este um campo de pesquisa que tem ganhado relevância na literatura de redes elétricas inteligentes (DONATO et al., 2020). Na busca pelo desenvolvimento dos algoritmos *NILM*, abordagens baseadas em aprendizado de máquina têm sido utilizadas devido à capacidade de processamento e diversidade de aplicações, por exemplo: algoritmos classificadores, redes neurais artificiais, aprendizagem profunda, entre outras (FAUSTINE et al., 2017; HARSHA S. ABEYKOON et al., 2020). E também estratégias interessantes para a detecção de eventos de baixo processamento computacional como modelos heurísticos e probabilísticos (REDDY; GARG; PUDI, 2020).

No entanto, apesar do crescente número de pesquisadores neste campo, ainda existem muitos desafios que devem ser resolvidos antes que seja possível tirar o máximo proveito dos potenciais benefícios da tecnologia *NILM*, entre eles: métodos de monitoramento que suportem a escalabilidade dos diferentes tipos de cargas que contém os distintos ambientes monitorados; sistemas de baixo custo

para aplicações reais com poder de processamento para desagregar as classes de equipamentos com comportamentos versáteis, exigindo técnicas complexas ou combinação de várias técnicas; falta de coleções dos conjuntos de sinais elétricos (assinaturas elétricas) para evolução com as pesquisas no monitoramento de cargas, além de problemas que envolvem privacidade e segurança das informações (CAO et al., 2017; LANGEVIN; GAGNON; CHERIET, 2020; WANG et al., 2021).

É importante mencionar que os conjuntos de dados existentes, geralmente, não são padronizados (formato/estrutura dos dados, sinais/grandezas disponíveis, resolução de dados e tipos de aparelhos) e também apresentam escassez de dados fundamentais para que se possa projetar, estudar, treinar e analisar sistemas e algoritmos de desagregação (SOUZA, 2016).

Outros desafios relacionados aos conjuntos de dados são que o uso da eletricidade diferencia significativamente entre as regiões e países, devido às características de geração elétrica, usabilidade dos equipamentos, padrão de consumo, conjunto de aparelhos e eletrodomésticos monitorados, clima, cultura e economia local (KELLY; KNOTTENBELT, 2015). Portanto, embora necessário, seria inviável implementar para cada localidade uma coleção de dados de acesso público que corresponda com a realidade vigente de monitoramento.

Até a presente data, não foram encontrados na literatura conjuntos de dados brasileiros, de caráter público, que estejam disponíveis e que facilitem as pesquisas com relação às técnicas *NILM*. Logo, há uma necessidade de verificar se conjuntos de dados públicos já disponibilizados em outras realidades atendem ao objetivo de desagregação de cargas em ambientes da realidade local.

A existência de bancos de dados dessa natureza, contribui expressivamente para a comunidade científica em estudos e testes de métodos para desagregação de cargas voltadas para a realidade brasileira (local de pesquisa deste trabalho). Favorecem, também, no avanço de pesquisas relacionadas à gestão do consumo de energia em residências, para que os consumidores finais possam ter responsabilidade sobre o consumo de energia e criar hábitos de incentivo à eficiência energética. Portanto, o conjunto de sinais elétricos gerado neste trabalho foi disponibilizado¹ em um repositório acessível, para contribuir na evolução das pesquisas relacionadas.

¹ Disponível em: <https://github.com/elielmarcos/SECRBR>

1.2 QUESTÃO DE PESQUISA E HIPÓTESE

Tendo em vista a problemática discutida nas seções anteriores, a principal pergunta de pesquisa desta dissertação de mestrado pode ser resumida da seguinte forma: “É possível desagregar cargas residenciais utilizando apenas conjuntos de dados públicos para implementação de sistema *NILM* em ambientes reais?”, deve-se ater aos conceitos de implementação de desagregação de cargas, levando em consideração a realidade do cenário de monitoramento e os padrões de uso.

Para que seja possível desagregar cargas residenciais é necessário que seja utilizado o monitoramento não intrusivo de equipamentos/aparelhos ligados à rede elétrica da edificação. Este monitoramento consiste em avaliar uma coleção de dados públicos disponíveis para pesquisadores da comunidade científica, que abordam as técnicas *NILM*, no intuito de treinar e desenvolver algoritmos de desagregação.

Sabe-se que o desempenho de uma desagregação varia e depende do conjunto de dados em que se está analisando para um país ou região em específico. Isso se justifica porque o uso de eletricidade diverge entre as localidades, por questões de usabilidade (conjunto de aparelhos diferentes), cultura, padrões de consumo, entre outras (KELLY; KNOTTENBELT, 2015). Dessa forma, questiona-se se somente a utilização de conjuntos de dados públicos seria suficiente para desagregar cargas em qualquer ambiente ou realidade de monitoramento, devido às variáveis apresentadas.

Portanto, apresenta-se como hipótese a estruturação de um sistema *NILM*, através de um conjunto de dados público e aplicação do sistema em um ambiente real para a validação da pergunta de pesquisa.

1.3 OBJETIVOS

Aqui nesta seção, são descritos os objetivos gerais e específicos levados em consideração para o desenvolvimento do projeto na desagregação de cargas elétricas residenciais em cenário prático.

1.3.1 Objetivo geral

O objetivo deste trabalho é propor um método que seja capaz de identificar os eventos de ativação e desativação das cargas elétricas residenciais em um ambiente real, utilizando apenas uma base de dados pública. Além disso, desenvolver um sistema de monitoramento *NILM* para validar o método proposto, utilizando métricas de classificação e avaliação sobre os equipamentos monitorados no ambiente.

1.3.2 Objetivos específicos

Como objetivos específicos, apresentam-se:

- realizar uma revisão bibliografia sobre os tipos de monitoramento de cargas elétricas residenciais, bem como, técnicas para a detecção e classificação de eventos de ativação ou desativação de cargas;
- investigar e utilizar banco de dados públicos para a extração da base de assinaturas das medidas agregadas e desagregadas dos equipamentos monitorados;
- desenvolver um método para identificação dos eventos ocorridos nos equipamentos elétricos;
- verificar e validar a eficiência da técnica de detecção de eventos através de métricas de avaliação;
- utilizar abordagens baseadas em aprendizado de máquina para a classificação de eventos e validar as inferências através de métricas de avaliação abordadas na literatura;
- aplicar o método em ambiente real utilizando medidores de sinais elétricos para aquisição dos sinais em medições agregadas e desagregadas;
- produzir um conjunto local de dados das cargas registradas;
- verificar a validade da hipótese em desagregar as cargas locais monitoradas com o uso de um conjunto público de dados.

1.4 ESTRUTURA DO TRABALHO

O restante do trabalho está organizado em cinco capítulos. No capítulo 2 foram abordadas as técnicas para monitoramento de cargas elétricas encontradas na literatura, seguida das categorias baseadas e não baseadas em eventos do monitoramento não intrusivo. Também, são explorados, as etapas para desagregação das cargas, as métricas de avaliação e os principais conjuntos de dados públicos para pesquisadores *NILM*.

No capítulo 3, apresenta-se a metodologia aplicada subdividida em cinco etapas, onde são descritos os processos desenvolvidos para o monitoramento de acordo com os referenciais disponíveis na literatura.

Nos resultados, capítulo 4, são apresentadas as estatísticas relacionadas às desagregações do sistema implementado e a comparação das métricas entre a utilização do conjunto de dados públicos e do ambiente local.

Finalmente, são apresentadas no capítulo 5 as considerações finais do trabalho, com a conclusão, abordando as principais contribuições e estudos futuros. Por último, estão as referências bibliográficas e apêndices.

2 REFERENCIAL TEÓRICO

Através de uma revisão da literatura científica, apresenta-se o estado da arte com as principais etapas e conceitos no monitoramento de carga não intrusivo para desagregação das cargas elétricas residenciais.

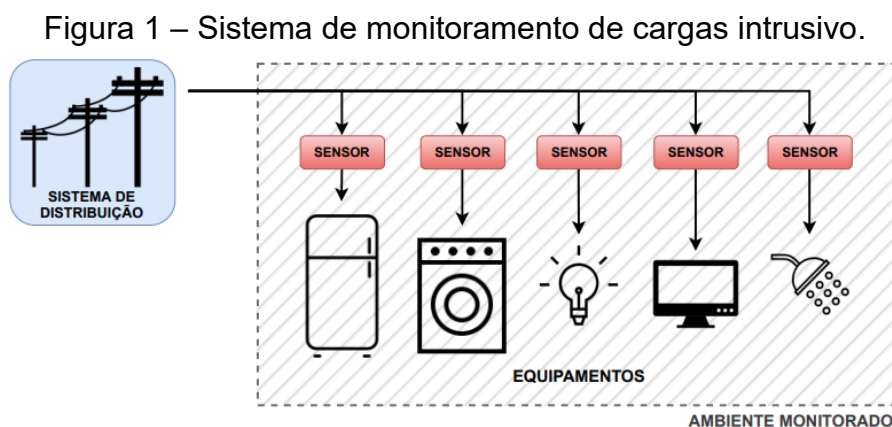
2.1 TÉCNICAS DE MONITORAMENTO DE CARGAS ELÉTRICAS

Atualmente, existem três técnicas para o monitoramento de cargas elétricas residenciais: monitoramento intrusivo, monitoramento semi-intrusivo e monitoramento não intrusivo (DASH; SODHI; SODHI, 2019a). Cada processo tem suas peculiaridades, vantagens e desvantagens para a desagregação de cargas. A seguir são abordadas as características de cada técnica com ênfase no monitoramento não intrusivo.

2.1.1 Monitoramento intrusivo de cargas

O monitoramento intrusivo de cargas (*Intrusive Load Monitoring - ILM*) é realizado por meio de sensores (medidores de consumo elétrico) conectados diretamente às cargas que se deseja medir, monitorando os sinais elétricos do equipamento. Nesta técnica são necessários vários sensores no ambiente, sendo destinados um para cada equipamento que se deseja monitorar, uma vez que as medições são feitas individualmente. Os sensores que realizam a coleta dos dados elétricos podem ser dispositivos como tomadas inteligentes, conhecidos como *Socket-Meters* ou *Smart Plugs*, e geralmente são instalados entre os pontos de acesso da rede elétrica residencial e os equipamentos monitorados. Os dados coletados por cada sensor são enviados para um controlador central, que gerencia as informações, através de técnicas de comunicação sem fio, como por exemplo: *WiFi* ou *ZigBee* (ALADESANMI; FOLLY, 2015).

A Figura 1 representa a distribuição dos sensores em uma residência para o monitoramento das cargas com sistema *ILM*.



A vantagem dessa técnica é o fácil monitoramento das cargas, uma vez que cada *Smart Plug* coleta os dados individualmente de cada equipamento, informando o consumo desagregado sem maiores processos com algoritmos de desagregação de cargas, tornando o sistema mais rápido e permitindo o monitoramento em tempo real.

Portanto, a maioria dos estudos *ILM* concentram-se no desenvolvimento de *hardware*, buscando aprimorar e reduzir os custos de produção dos medidores, também no sistema de detecção e no aplicativo implantado na “nuvem”. Trabalhos como os apresentados por Adiono et al. (2019) e Nguyen et al. (2019), abordam este método e utilizam soquetes inteligentes, operando como um dispositivo ativo ou passivo, para calcular o consumo e transferir os dados das cargas para nuvem ou servidor para posterior análise das informações coletadas.

Outros estudos da literatura, como o publicado por Phan et al. (2021), utilizaram esta técnica de monitoramento para extrair o consumo individual dos aparelhos através de medidores *Smart Plugs*. Os dados de energia coletados são analisados para obter informações úteis sobre o comportamento dos consumidores e contribuíram ao empregar métodos de classificação e agrupamento para informar os ciclos completos de cada equipamento supervisionado. O maior desafio neste caso, segundo os autores, é tratar as informações em nível de dispositivo em tempo real. O trabalho utilizou um classificador compacto no dispositivo de borda, denominado

CompactNet, e apresentou uma redução de 6,6% no tempo da inferência quando comparado com o modelo base.

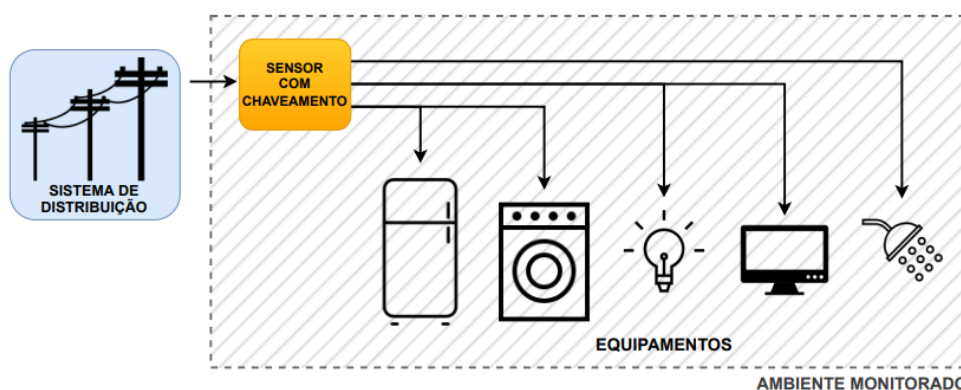
Apesar de não ser necessário desagregar as cargas, a desvantagem dos sistemas *ILM* é seu alto custo com *hardware* de implantação, pois cada equipamento que se deseja monitorar necessita de uma unidade medidora instalada. Dessa forma, o custo individual do monitor é proporcional aos recursos e precisão que oferecem, tornando-se ainda menos atrativo pelo número elevado de manutenções que podem ocorrer no sistema. Além disso, exige a adição de novos sensores à medida que novos equipamentos são adicionados na residência.

2.1.2 Monitoramento semi-intrusivo de cargas

A técnica de monitoramento semi-intrusivo de cargas (*Semi Intrusive Load Monitoring - SILM*) pode ser considerada como a junção do monitoramento intrusivo e não intrusivo. Neste processo as cargas são agrupadas de acordo com a instalação elétrica da residência e o sistema utiliza, geralmente, apenas um sensor central na rede elétrica para obter os dados provenientes das cargas. Porém, é necessário um circuito de chaveamento que desvia a corrente elétrica de cada grupo de equipamentos para serem monitorados pelo sensor em intervalos de tempo periódicos ou em intervalos que se deseja coletar os dados. Assim, os grupos de equipamentos são monitorados por um período de tempo, sendo intercalados com outros grupos de equipamentos existentes no local (DASH; SODHI; SODHI, 2019a).

A Figura 2 representa a instalação do sistema de monitoramento *SILM* em uma residência, contendo um único sensor com circuito de chaveamento que gerencia os vários grupos de cargas.

Figura 2 – Sistema de monitoramento de cargas semi-intrusivo.



Fonte: Autor.

A vantagem dessa técnica, assim como no *ILM*, é a fácil desagregação das cargas, pois a quantidade de dispositivos por grupo é reduzida, além de poder atuar no controle de cada grupo. Porém, apresentam a desvantagem que esses grupos devem ser construídos de forma *offline*, ou seja, na própria instalação elétrica do local, que muitas vezes não possui infraestrutura para receber estas alterações. Além disso, podem ocorrer interrupções de energia para os equipamentos, em caso de falhas do sistema, uma vez que esse controle é realizado por meio de um chaveamento administrado pelo próprio sistema de monitoramento. Seu custo final tende a ser mais elevado que o monitoramento *ILM*, por exigir um *hardware* central mais robusto.

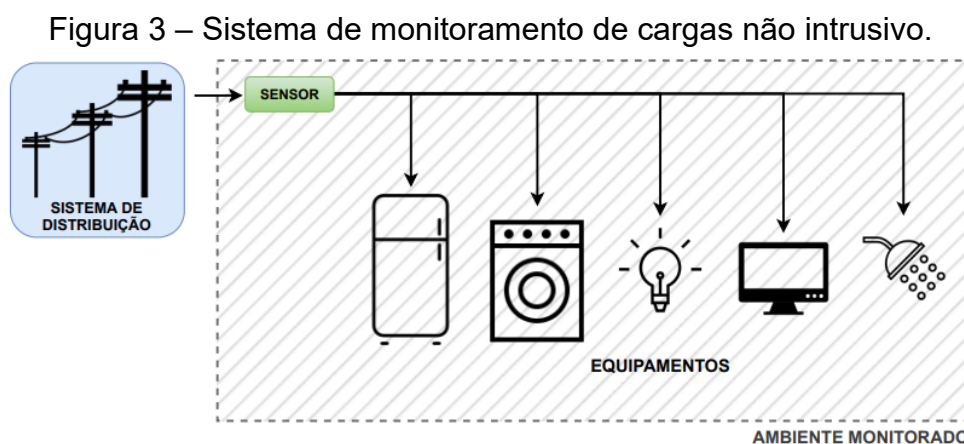
O trabalho de Dash et al. (2019b) utilizou o monitoramento semi-intrusivo, nele uma central administrada por um Raspberry realiza a troca de chaveamentos nos cruzamentos de zero da fase para amenizar os ruídos no sistema elétrico, mas não realizaram a predição do consumo. Tang et al. (2016) propôs um sistema para desagregação semi-intrusiva, utilizando algoritmo de otimização paralela para recuperar os estados dos dispositivos, e aplicou os testes em conjunto de dados públicos e sintéticos, seus resultados mostraram desempenho acima de 90% para até 12 equipamentos monitorados.

2.1.3 Monitoramento não intrusivo de cargas

Em termos simples, o monitoramento não intrusivo de cargas, mais conhecido como *NILM*, é definido como um conjunto de técnicas de processamento de sinais para detecção de eventos e aprendizado de máquina para classificação. A

detecção e desagregação do consumo de energia são realizadas a partir de um número limitado de sensores, geralmente um único sensor central, para estimar o consumo agregado e individual de eletricidade dos aparelhos (HART, 1992).

A Figura 3 representa o monitoramento de cargas em uma residência, por meio de um único sensor, utilizando o sistema *NILM*.



Os primeiros estudos sobre esse tema foram apresentados por Hart (1992) ao desenvolver o primeiro protótipo *NILM*, seu princípio de funcionamento é que cada mudança na carga elétrica total de uma edificação acontece com a mudança de estado de um dispositivo elétrico, por exemplo, um televisor ao ser ligado ou desligado. As primeiras abordagens foram projetadas de forma que fosse possível detectar as mudanças do consumo total de eletricidade da casa e, então, eram extraídas as informações próximas das mudanças de energia. Esses dados foram usados para diferenciar o consumo individual dos aparelhos por meio da aplicação de algoritmos classificadores.

Hart (1992) ainda diferencia e classifica os tipos de cargas, para identificar os diferentes eventos que podem ocorrer durante as operações dos equipamentos no processo de monitoramento das cargas, sendo elas: (I) cargas do tipo LIGADA / DESLIGADA (*ON/OFF*); (II) cargas de estados finitos (*Finite State Machine - FSM*); (III) cargas de potência variável ou sem estados finitos; (IV) cargas de consumo contínuo ou permanente. Uma breve descrição e exemplificação destas cargas são apresentadas no Quadro 1.

Quadro 1 – Tipos de cargas.

Tipos de Cargas	Descrição	Exemplos
LIGADA / DESLIGADA	Dispositivos que possuem apenas dois estados de operação, ligado ou desligado.	Torradeiras e lâmpadas sem controle de brilho.
Estados Finitos	Dispositivos que possuem mais de dois estados de operação ou que vão passar por um número finito de modos de operação.	Máquinas de lavar e refrigeradores com ciclo de descongelamento.
Potência Variável	Dispositivos que não possuem um número limitado de estados fixos, cujo consumo é variável.	Lâmpadas com controle de iluminação e televisores.
Consumo Permanente	Dispositivos que normalmente funcionam com um único estado de operação de maneira contínua.	Alarmes e câmeras de segurança.

Fonte: Hart (1992).

Uma das vantagens do *NILM* é o baixo custo na aquisição e instalação dos sensores de medição. Apresentam também a redução de manutenções em *hardware*, uma vez que, para realizar o monitoramento utilizando esta técnica é necessário apenas um único medidor na demanda central da residência, o que reduz significativamente os custos com *hardware* quando comparado com o monitoramento intrusivo (ABUBAKAR et al., 2015).

De acordo com Ridi, Gisler e Hennebert (2014) a desvantagem é o alto empenho com desenvolvimento do *software* de desagregação. Os algoritmos de monitoramento *NILM* devem realizar várias etapas para serem capazes de desagregar as cargas da energia total medida, que podem implicar no tempo de resposta. Além de necessitar de uma fase de treinamento para identificar as assinaturas elétricas de cada equipamento que serão monitorados.

Mesmo sendo uma técnica desenvolvida nos anos 90, a busca por aprimoração em algoritmos de desagregação ainda é bastante necessária. Novos estudos sobre os trabalhos de Hart (1992) envolvendo a temática de *NILM* utilizam conceitos mais eficientes de aprendizado de máquina, contribuindo como ferramenta para soluções e aplicações referentes ao consumo e economia de energia elétrica (SOUZA, 2016).

Por estas razões essa dissertação é baseada no monitoramento não intrusivo de cargas elétricas residenciais, voltada para a implantação em um ambiente real, a fim de avaliar o método desenvolvido e responder à pergunta de pesquisa empregando algoritmo detector e classificadores já utilizados na literatura *NILM*.

2.2 CATEGORIAS DE MONITORAMENTO NÃO INTRUSIVO

Atualmente o monitoramento não intrusivo está dividido em duas categorias, as abordagens baseadas em eventos e as abordagens não baseadas em eventos. Essas categorias visam determinar o estado de operação dos aparelhos para estimar quais equipamentos estão contribuindo no consumo agregado (ALCALA et al., 2017).

Após as suposições iniciais do *NILM*, que cada mudança no consumo agregado acontece em resposta à mudança de operação de um aparelho, a base para as abordagens mais utilizadas hoje são as baseadas em eventos. Os conceitos de abordagem em eventos e sem eventos foram introduzidos pela primeira vez no *1º Workshop Internacional NILM²* em 2012, e visam fornecer uma categorização clara para as abordagens de desagregação de energia (ANDERSON et al., 2012b; SOUZA, 2016; YASIN; KHAN, 2018).

Nas próximas duas subseções são apresentadas, de forma resumida, algumas pesquisas e suas distinções de acordo com essas categorias.

2.2.1 Não baseada em eventos

Abordagens sem eventos, ou não baseadas em eventos, não dependem da detecção de eventos e classificação. Em vez disso, essas abordagens tentam combinar cada amostra da potência agregada com o consumo de um aparelho específico ou uma combinação de diferentes aparelhos, por meio de métodos estatísticos, otimizando uma estimativa do estado geral do sistema (KOLTER; JOHNSON, 2011).

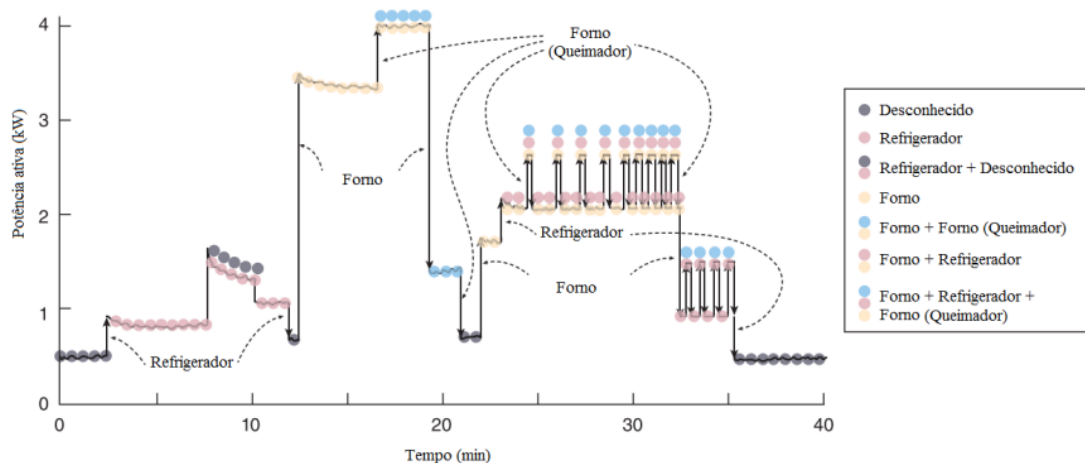
Como este método não se fundamenta pela mudança abrupta de estado dos aparelhos, em geral é utilizado algoritmos que necessitam de uma baixa taxa de amostragem do sinal agregado para aquisição das amostras (na frequência de 1 Hz ou menos). Por meio de métodos estatísticos (métodos de rede Bayesiana), probabilísticos (por exemplo, modelos de Markov ocultos, ou ainda, *Factorial Hidden Markov Model - FHMM*) e métodos de aprendizado de máquina não supervisionada,

² *1st International Workshop on Non-Intrusive Load Monitoring*, nilmworkshop.org/2012/

são investigadas as combinações dos estados dos aparelhos e desagregadas do consumo total agregado (RAIKER et al., 2018).

A Figura 4 representa a curva de potência ativa agregada, apresentando-se em todos os instantes do monitoramento, os possíveis estados de operação para a desagregação das cargas elétricas.

Figura 4 – Desagregação por categoria não baseada em eventos.



Fonte: Adaptada de Souza (2016).

Portanto, os dados de treinamento não requerem nenhuma transição rotulada. Em vez disso, apenas o consumo agregado das cargas de interesse é necessário. Assim, tornando mais simples o processo de coleta de dados de treinamento para abordagens sem eventos.

Além disso, de acordo com Shao, Marwah e Ramakrishnan (2013) as alternativas sem eventos não requerem que os algoritmos de aprendizado de máquina sejam previamente treinados para identificar cada mudança de potência individual no sinal agregado. Essas abordagens baseiam-se, principalmente, no conhecimento existente sobre a operação de aparelhos individuais, através de diferentes técnicas, como Descoberta de Motivos (*Motif Discovery*), Separação Cega de Fontes (*Blind Source Separation - BSS*) e Modelos de Gráficos Probabilísticos (*Probabilistic Graphical Models - PGM*).

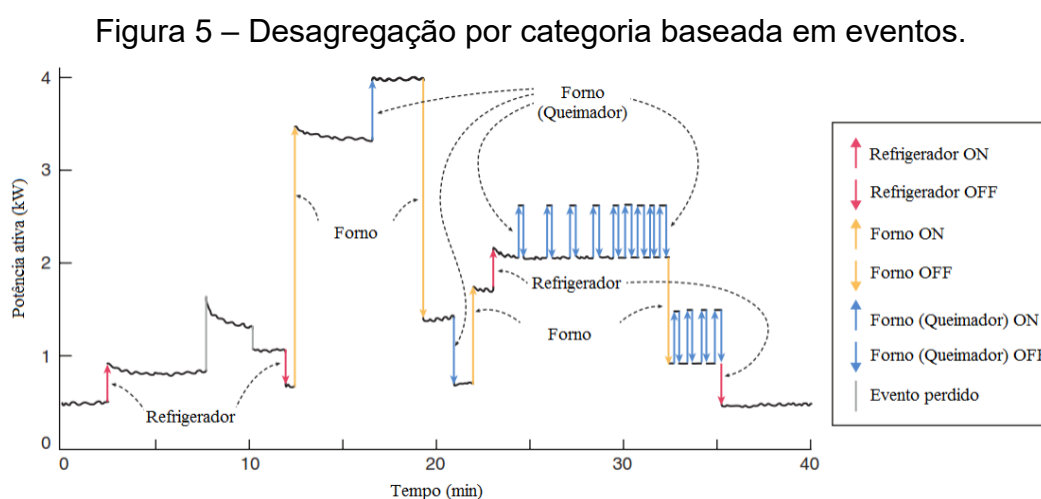
Uma desvantagem desta categoria está relacionada diretamente à etapa de inferência do consumo das cargas, pois o processo de desagregação é realizado para cada uma das amostras coletadas no sinal monitorado, tornando essas abordagens consideravelmente mais intensivas e exigindo mais processamento computacional pelos algoritmos (SOUZA, 2016).

2.2.2 Baseada em eventos

As abordagens baseadas em eventos estão diretamente relacionadas aos primeiros estudos de *NILM*, e buscam desagregar o consumo total por meio de detecção e classificação de cada transição de aparelho no sinal agregado. São utilizados algoritmos de aprendizagem supervisionada ou semi-supervisionada, os quais relacionam as mudanças de estado do sinal com as mudanças de estado do aparelho (ABUBAKAR et al., 2015). Conseqüentemente, as abordagens desta categoria requerem uma etapa de coleta de dados, onde uma série de transições (eventos) dos aparelhos de interesse são coletados, classificados e armazenados, para serem usados posteriormente como dados de treinamento.

Outra característica muito importante neste tipo de abordagem dos sistemas baseados em eventos é a necessidade, em sua maioria, de uma alta taxa de amostragem do sinal de consumo elétrico, variando na faixa de Hz até MHz (ZHUANG; SHAHIDEHPOUR; LI, 2018). Isso se dá pelo fato que nesta categoria os eventos de ativação e desativação das cargas devem ser rastreados continuamente por algoritmos de detecção de borda ou mudanças de nível no consumo agregado. Ou seja, quanto maior a quantidade de informações sobre o sinal elétrico, menor seriam as chances de não detectar os eventos gerados pelos equipamentos monitorados.

A Figura 5 representa a curva de potência ativa agregada com as respectivas sinalizações de ativação e desativação de cada aparelho na desagregação baseada em eventos.



Fonte: Adaptada de Souza (2016).

Segundo Zhuang et al. (2018), altas taxas de amostragem requerem equipamentos mais robustos para extração dos sinais. Contudo, essa característica permite que a análise da forma de onda da tensão e corrente seja realizada, além de poder investigar sinais transitórios e harmônicos que estão relacionados com os eventos de ativação e desativação das cargas elétricas.

Souza (2016) cita em seus estudos que apesar do fato das abordagens baseadas em eventos requererem a execução contínua de algoritmos para detecção, a inferência do aparelho só é realizada quando os eventos são detectados, exigindo menor processamento em cada classificação. Por isso, essa dissertação toma como categoria o método baseado na detecção de eventos, pela vantagem do menor custo computacional dos algoritmos envolvidos em relação às abordagens sem eventos.

2.3 ETAPAS PARA DESAGREGAÇÃO

Um fluxo de trabalho do sistema *NILM* baseado em eventos, contém algumas etapas básicas que devem ser observadas. Yan, Xudong e Zun (2020) e Liu e Chen (2014) citam em seus trabalhos as etapas de aquisição dos dados e pré-processamento, onde os sinais que representam a energia elétrica que flui para a residência são capturados, amostrados e transformados em medições relacionadas à energia (por exemplo, potência ativa e reativa). Uma próxima etapa de detecção de eventos é abordada por Völker et al. (2020a) e Pereira (2017), que é o processo de identificação das mudanças no consumo que ocorrem em resposta à mudança do modo de operação de cada aparelho, incluindo estados de ativação e desativação.

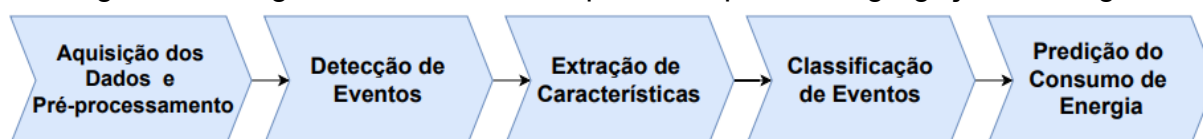
A extração de características, ou formação das assinaturas elétricas dos equipamentos, é uma etapa muito importante citada por Zhang e Zhu (2019) e Zhang, Zhang e Wu (2020), onde diferentes parâmetros são extraídos da vizinhança do evento de potência, formando uma assinatura de evento do equipamento que será utilizada no processo de treinamento e identificação das cargas responsáveis por cada evento identificado.

A etapa seguinte do fluxo é a classificação de eventos, abordada por Garcia et al. (2020) e Bosco et al. (2021), onde algoritmos de aprendizado de máquina, previamente treinados, são aplicados às assinaturas dos eventos de energia detectados anteriormente para fornecer inferências dos aparelhos que provocaram

os eventos. Por último, Souza (2016) descreve a etapa de predição do consumo de energia, onde o consumo das cargas individuais é estimado com base nos eventos de potência rotulados e sua distribuição no tempo.

As cinco etapas mencionadas são representadas por um diagrama de fluxo, conforme mostrado na Figura 6.

Figura 6 – Diagrama de fluxo *NILM*, processo para desagregação de cargas.



Fonte: Autor.

Todas essas etapas também são citadas nos trabalhos de Hart (1992), sendo válido explorá-las nas subseções seguintes para o desenvolvimento da metodologia deste trabalho no monitoramento não intrusivo de cargas.

2.3.1 Aquisição dos dados e pré-processamento

Na etapa de aquisição dos sinais, o sistema deve ser capaz de obter as grandezas elétricas de forma agregada para fase de desagregação e os sinais individuais das cargas para uma fase de treinamento e testes. Os sinais podem ser adquiridos por aparelhos registradores como *Smart Meters* e *Smart Plugs*, para os sinais agregados e desagregados, respectivamente. O registro das grandezas elétricas é o passo inicial importante para serem utilizados em algoritmos de desagregação, sendo eles os sinais de tensão e corrente elétrica que posteriormente são utilizados para calcular as demais grandezas, como potência ativa e reativa por exemplo (VÖLKER; SCHOLL; BECKER, 2019). Portanto, é necessário que os medidores possuam sensores capazes de medir tensão, conhecidos como transformador de potencial (TP), e corrente, conhecidos como transformador de corrente (TC).

De forma resumida, um transformador de potencial trata-se de um dispositivo para instrumentação cujo enrolamento primário é ligado em derivação com um circuito elétrico e o enrolamento secundário fornece uma baixa tensão proporcional à tensão a ser medida. Já um transformador de corrente transforma a

corrente que circula em um enrolamento primário (sendo o próprio condutor do circuito), através de um acoplamento magnético, em uma corrente induzida no secundário de menor grandeza (FILHO, 1997).

Diante dos sinais provenientes dos sensores TP e TC, os dados devem ser amostrados em baixa ou alta taxa de frequência, de acordo com o método de desagregação que se deseja seguir (ZHUANG; SHAHIDEHPOUR; LI, 2018). Visto que esta etapa é relevante para sistemas *NILM*, uma vez que se propaga para os demais estágios do processo, conforme já mencionado nas subseções 2.2.1 e 2.2.2.

Antes das grandezas elétricas, obtidas da rede elétrica, serem utilizadas no próximo passo do fluxo *NILM*, elas podem ou não sofrer um pré-processamento dos dados. Essa manipulação consiste em exclusão de dados duplicados, preenchimento de dados perdidos, filtragem analógica e/ou digital dos sinais, entre outros tipos de intervenção conforme a necessidade (MEDEIROS, 2019).

Para desagregações que utilizam valores de potência como uma grandeza de classificação, Hart (1992) sugere a intervenção de pré-processamento após aquisição dos dados, realizando a normalização das potências (ativa, reativa e aparente), com base na tensão nominal da rede elétrica, conforme Equação (1) por ele utilizada.

$$p_{Norm}(t) = \left(\frac{V_{Nom}}{V(t)} \right)^2 \cdot p(t) \quad (1)$$

Onde, $p_{Norm}(t)$ é a potência normalizada em função do tempo, V_{Nom} a tensão nominal, $V(t)$ e $p(t)$ a tensão e a potência extraídas em função do tempo, respectivamente. Dessa forma as variações da potência elétrica, que são interferidas diretamente pelas variações da tensão elétrica, podem ser ajustadas para minimizar as variações de comportamento dos equipamentos considerados cargas lineares. Por ser um processo simples e reduzir os impactos gerados pelas variações de tensão, este trabalho adota a normalização das potências em função da tensão nominal da rede elétrica.

Após os dados serem convertidos de analógico para digital eles são processados pelos algoritmos das etapas seguintes no processo de desagregação, sendo o próximo passo a detecção de eventos.

Nunes e Pereira (2015) apresentaram um sistema de rotulagem para alguns conjuntos de dados de treinamento e teste de algoritmos *NILM*. Sua abordagem, baseada em eventos, utilizou uma frequência de amostragem de 50 Hz da potência elétrica, obtendo uma probabilidade de até 97% na rotulagem dos eventos. Völker et al. (2020a) utilizaram um sistema de desagregação para avaliar um conjunto de dados de sinais elétricos rotulados a partir dos valores de potência ativa, provenientes da amostragem de cada ciclo da frequência da rede elétrica (50 Hz para região de estudo), atingindo 93% de precisão de reconhecimento entre os 21 equipamentos monitorados.

Após observadas as características sobre a frequência de amostragem e os resultados de desempenho apresentados na literatura, nesta dissertação também será utilizado uma taxa referente ao ciclo de fase elétrico (60 Hz para a frequência nominal da rede elétrica brasileira). Dessa forma, é possível construir uma metodologia padrão que utilize conjuntos de dados com amostragem baseada na frequência da rede elétrica, sendo factível comparar e avaliar o monitoramento não intrusivo em um ambiente real sem maiores preocupações com a taxa de aquisição dos dados.

2.3.2 Detecção de eventos

A segunda etapa de um sistema *NILM* é a detecção de eventos, de acordo com Anderson et al. (2012b) existem diferentes abordagens agrupadas em três categorias: (i) heurística especializada; (ii) filtros de correspondência; (iii) modelos probabilísticos.

Na heurística especializada, os algoritmos são menos complexos e realizam o rastreamento dos sinais da série temporal em busca de mudanças acima de um certo limite (HART, 1992). Biansoongnern e Plungklang (2016) utilizam este método escaneando o sinal de potência previamente filtrado para minimizar a presença de ruído. Os eventos de potência são detectados por meio do cálculo das diferenças entre duas amostras consecutivas, também conhecido como delta ou derivada discreta. O que se nota como maior problema neste tipo de algoritmo é a grande chance de ocorrerem falsos positivos.

Na categoria por filtros de correspondência, os algoritmos detectores de eventos trabalham correlacionando transientes de aparelhos conhecidos (modelos) no sinal de consumo agregado (sinal desconhecido) por meio de técnicas de filtragem. Os autores Wild, Basim e Yang (2015) propõem um detector de eventos que tenta combinar segmentos de transientes (obtidos de treinamento) ao sinal agregado usando dois filtros. O primeiro filtro é usado para encontrar as formas transientes no sinal agregado, e o segundo é usado para garantir as correspondências aos transientes reais e não a algum ruído impertinente. Além disso, o detector utiliza alta frequência de amostragem para extrair e analisar as harmônicas dos sinais.

Outra abordagem para detecção de eventos é a categoria de modelos probabilísticos, onde a detecção ocorre por meio de métodos e algoritmos probabilísticos em duas fases. Na primeira fase é calculado a probabilidade de um evento de ativação ou desativação ocorrer a cada amostra do sinal de potência. Este sinal é considerado como a estatística de detecção após aplicação de testes estatísticos por meio de janelas deslizantes, como por exemplo, razão de probabilidade generalizada (*Generalized Likelihood Ratio - GLR*), introduzido por Luo et al. (2002).

Uma segunda fase é aplicada para extrair os eventos do sinal estatístico resultante. Sempre que a estatística de detecção está acima de um certo limiar a ocorrência de um evento é considerado na amostra de potência, que corresponde ao mesmo índice da amostra do sinal estatístico (JIN et al., 2011).

Mesmo assim, no caso particular de sistemas *NILM*, foram desenvolvidas estratégias mais robustas para implementação de modelos probabilísticos. Pereira (2017) propôs um algoritmo utilizando um detector probabilístico logarítmico (*Log Likelihood Detector - LLD*) na seleção dos eventos, aplicando uma fase adicional para localização de máximos / mínimos ao sinal estatístico de detecção. De acordo com seus resultados, o *LLD* se mostrou mais eficiente que o *GLR*.

Com base nos estudos de Luo et al. (2002) e Pereira (2017), os autores Völker, Scholl e Becker (2019) desenvolveram um algoritmo que realiza a detecção de eventos por meio da razão logarítmica de probabilidade (*Log Likelihood Ratio - LLR*), adotando um limiar variável, de acordo com o sinal médio da potência, e uma janela de votação para decidir onde o evento ocorreu, obtendo o maior sinal de detecção estatístico da janela.

Utilizando a Equação (2), é aplicado o *LLR* nas amostras do sinal da potência elétrica para detectar uma mudança no valor médio do sinal.

$$L[i] = \ln\left(\frac{\sigma_{[h,i]}}{\sigma_{[i,k]}}\right) + \frac{(P[i] - \mu_{[h,i]})^2}{2 \cdot \sigma_{[h,i]}^2} - \frac{(P[i] - \mu_{[i,k]})^2}{2 \cdot \sigma_{[i,k]}^2} \quad (2)$$

Para uma dada amostra do sinal de potência $P[i]$, a probabilidade $L[i]$ de ter ocorrido uma mudança na média dos sinais é calculada. Uma janela deslizante $dw_{[h,k]}$ (janela de detecção) é usada para decidir se uma mudança de estado ocorreu ou não. Esta janela de detecção é dividida em pré-evento $[h, i[$ e janela pós-evento $[i, k]$. Ainda, $\mu_{[h,i]}$ e $\mu_{[i,k]}$ são as médias e $\sigma_{[h,i]}$ e $\sigma_{[i,k]}$ os desvios-padrão das janelas pré-evento e pós-evento, respectivamente.

Em última análise, é calculada pela Equação (3) a probabilidade de um evento ocorrer na amostra i da potência. Se a mudança média estiver abaixo de um limite th_i , $L[i]$ é definido como zero.

$$L[i] = \begin{cases} L[i], & \text{se } |\mu_{[h,i]} - \mu_{[i,k]}| > th_i \\ 0, & \text{se } |\mu_{[h,i]} - \mu_{[i,k]}| \leq th_i \end{cases} \quad (3)$$

Diferente do algoritmo *LLD*, o limiar do método *LLR* aumenta linearmente com a potência média da janela de pré-evento, conforme Equação (4).

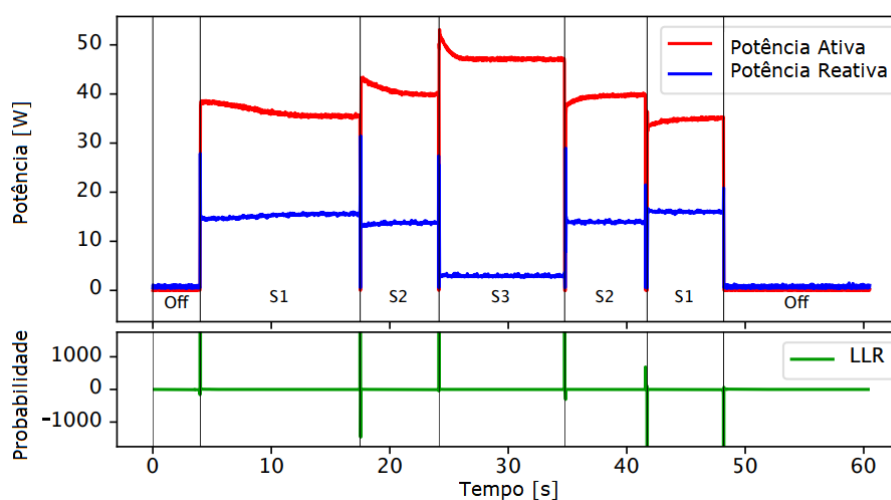
$$th_i = th_{min} + m \cdot \mu_{[h,i]} \quad (4)$$

Onde th_{min} é a mudança mínima de interesse observada e m é um coeficiente que aumenta linearmente o limiar em relação aos valores médios da potência. Esta estratégia adotada é muito interessante, pois equipamentos que consomem alta potência normalmente também mostram uma variação maior no sinal, isso pode causar falsos positivos se um pequeno limite for definido. Entretanto, se um valor limite alto for definido, dispositivos como luzes podem não ser detectados.

Dessa forma, com um limite variável, th_i aumenta linearmente com a potência do dispositivo e se adapta às possíveis flutuações maiores, enquanto reduz a quantidade de detecções falsas.

Por último, o algoritmo localiza o ponto da mudança na média identificando o valor máximo de $L[i]$ na janela (VÖLKER; SCHOLL; BECKER, 2019). A Figura 7 mostra os resultados da relação probabilística entre as mudanças de estados nos sinais das potências e a detecção de eventos pelo algoritmo *LLR*. As mudanças de estados são identificadas pelas barras pretas verticais.

Figura 7 – Relação da detecção probabilística *LLR* e as curvas de potência.



Fonte: Adaptada de Völker, Scholl e Becker (2019).

Os resultados dos testes apresentados pelos autores Völker, Scholl e Becker (2019), aproximaram-se de 97% de desempenho nos eventos detectados. Por ser um método relativamente simples, exigindo baixo poder computacional e apresentando resultados eficientes, essa dissertação utiliza o modelo probabilístico *LLR* como uma abordagem para o problema de detecção de eventos em sistemas *NILM*.

2.3.3 Extração de características

A extração de características é o processo que seleciona parâmetros elétricos no monitoramento das cargas, buscando discriminar características que diferenciam os aparelhos e determinando a formação das assinaturas elétricas dos equipamentos. As características são extraídas e utilizadas no processo de

classificação e treinamento para identificação das cargas responsáveis por cada evento no monitoramento *NILM* (ZHANG; ZHU, 2019). Além disso, as assinaturas devem estar de acordo com a forma que as grandezas elétricas foram adquiridas, baixa ou alta frequência, conforme discutido na subseção 2.3.1.

As características são extraídas aproveitando o conhecimento de domínio em termos da energia elétrica e características do aparelho. Os recursos são normalmente extraídos das amostras em torno do evento de ativação ou desativação de interesse (ZHOU et al., 2021). Os exemplos mais comuns dessas características são as diferenças de potências, formas transientes e estacionárias, componentes harmônicos, formas de onda de tensão e corrente, razões mínimo-máximo, fator de crista, pico máximo, variância, distorção, fator de forma e entre outras (BUCCI et al., 2020).

Características extraídas da forma de onda de tensão e corrente para formação das assinaturas, utilizando alta taxa de amostragem, possibilitam a análise de componentes harmônicos e identificação das cargas elétricas com comportamento não linear. Zhang, Zhang e Wu (2020) propõem um método de extração de características para a identificação de cargas baseado na potência e harmônicas de corrente das cargas, utilizando a transformada de Fourier (*Fourier Transform - FT*). No entanto, a utilização desses recursos não apresenta melhorias significativas para identificação de equipamentos lineares ou resistivos, sendo estes os aparelhos que mais contribuem para o consumo elétrico em uma residência (HOSSEINI et al., 2017).

Assim como Hart (1992) considerou em seu trabalho as potências ativa e reativa, para formação das assinaturas elétricas de ativação e desativação, as assinaturas constituídas a partir da extração das características dos sinais de potência ativa, potência reativa, potência aparente e corrente elétrica, são apontadas como as mais próximas para serem utilizadas em aplicações práticas. Isso se dá devido a sua simplicidade e menor custo com instrumentação de monitoramento (MEDEIROS, 2019).

No trabalho de Biancoongnorn e Plungklang (2016), os autores apresentaram um sistema *NILM* de baixa frequência utilizando as variações das grandezas elétricas de potência ativa e reativa para formação das assinaturas elétricas dos equipamentos. Foram consideradas as diferenças médias nos instantes

pré-evento e pós-evento dos momentos transitórios, mas desconsideradas as características transientes.

Assim como Hart (1992), Biansoongnern e Plungklang (2016), essa dissertação toma os sinais de potência ativa, reativa e aparente para extração das características das assinaturas elétricas de ativação e desativação. Porém, busca explorar mais informações sobre estas grandezas, como dados de regime transitório e permanente dos eventos detectados, permitindo uma metodologia que utilize maior quantidade de recursos para classificação dos eventos detectados.

2.3.4 Classificação de eventos

Na etapa de classificação de eventos, a literatura sobre *NILM* é muito rica em relação aos métodos de aprendizagem empregados, nela os algoritmos devem aprender e reconhecer as assinaturas dos equipamentos.

De acordo com algumas publicações que exploram técnicas de aprendizado de máquina (*Machine Learning - ML*) para classificação no monitoramento não intrusivo, os modelos vão desde os algoritmos de aprendizagem mais tradicionais, como o k-vizinhos mais próximos (*K-Nearest Neighbors - KNN*) (IWAYEMI; ZHOU, 2014), árvores de decisão (*Decision Trees*) (NGUYEN et al., 2015), *Naïve Bayes (NB)* (WEISSHAAR et al., 2018) e máquinas de vetores de suporte (*Support Vector Machine - SVM*) (MORADZADEH et al., 2020), até abordagens mais complexas como redes neurais artificiais (*Artificial Neural Network - ANN*) (BIANSOONGNERN; PLANGKLANG, 2016), redes neurais convolucionais (*Convolutional Neural Network - CNN*) (MEDEIROS et al., 2019) e algoritmos genéticos (*Genetic Algorithms - GA*) (CHANG et al., 2011).

No processo de aprendizado podem ser utilizadas técnicas supervisionada, semi-supervisionada e não supervisionada. Classificadores de aprendizado supervisionado são os mais utilizados em abordagens *NILM* e necessitam de um conjunto de dados elétricos (ou seja, assinaturas elétricas dos equipamentos) devidamente rotulados e identificados, de modo a fornecer uma relação direta entre assinaturas elétricas e sua classificação (ABUBAKAR et al., 2015). Apesar de exigir um maior custo com tempo para se organizar uma base de dados das assinaturas, o treinamento supervisionado tem apresentado melhor desempenho, segundo Zhang et al. (2017).

Classificadores com abordagem semi-supervisionada fazem uso de dados rotulados e não rotulados para treinar algoritmos, ou seja, quando há dados escassos ou são muito custosos de se obter. Dessa forma, são utilizados pequenos conjuntos de exemplos rotulados para inferir os exemplos não rotulados e, com maior confiança, usá-los posteriormente como dados de treinamento (PONRAK; DECHANUPAPRITTHA, 2021).

A categoria de classificadores *NILM* com aprendizado não supervisionado dispensam o conhecimento sobre os dados de treinamento. Como não há necessidade de assinaturas rotuladas, o processo de aquisição das assinaturas é menos trabalhoso, porém, podem apresentar baixa precisão nos resultados, pois inferem o reconhecimento das cargas elétricas a partir de assinaturas análogas existentes em bancos de dados (ZHANG et al., 2017).

Conforme descrito pelos autores Abubakar et al. (2015), Medeiros (2019), Ponrak e Dechanupaprittha (2021), e Zhang et al. (2017), para os sistemas *NILM*, a abordagem supervisionada tende a ser mais eficiente e confiável por utilizar assinaturas totalmente rotuladas. Então, como forma de definir os classificadores utilizados neste trabalho, foi realizada uma pesquisa na base de dados *IEEE Xplore*³ dos últimos cinco anos, buscando identificar publicações que exploraram técnicas de *ML* com aprendizado supervisionado na classificação de equipamentos residenciais.

A pesquisa na base de dados apontou para os modelos de classificadores mais utilizados, e computacionalmente menos complexos (tanto a nível de desenvolvimento, quanto a nível de métrica de comparação para métodos que utilizam outros classificadores), sendo eles os algoritmos:

I) *KNN*, onde o trabalho de Gurbuz et al. (2021) compara o *K-Nearest Neighbors* com outros classificadores (*Decision Trees* e *Random Forest*) na desagregação de cargas, no pior caso *KNN* teve 85% de eficiência, enquanto Khan et al. (2019) testam o desempenho deste classificador com oito equipamentos diferentes de um conjunto de dados público, chegando a 100% de eficiência na inferência de alguns aparelhos.

II) *SVM* é outro algoritmo bem difundido na classificação *NILM*, Chea et al. (2022) utilizam o *Support Vector Machine* na classificação de cargas através das características da trajetória de tensão e corrente, seu resultado mostrou acuracidade

³ *IEEE Xplore Digital Library*, ieeexplore.ieee.org

de 97%. Moradzadeh et al. (2020) aplicaram este algoritmo em dois conjuntos de dados públicos e obtiveram um desempenho maior que 98% em ambos os conjuntos, além de afirmarem que o SVM pode ser utilizado para a proposta de monitoramento no mundo real.

III) *ANN* é uma categoria de classificador que vem ganhando cada vez mais espaço no monitoramento não intrusivo, Ali et al. (2021) implementam em *MATLAB* uma rede neural de três camadas (entrada, oculta e saída) para classificar até 21 combinações de classes simultâneas, sua rede teve erro inferior à 2%. Pirbazari et al. (2020) apresentam em seu estudo a melhoria na previsão em 80% com o uso híbrido de uma *ANN* multicamadas (três camadas) em conjunto de dados públicos.

Modelos mais complexos como *CNN* também são resultados de pesquisas recentes, porém, esta estrutura requer maior poder computacional e matrizes como características de entrada, já que estes modelos são mais empregados para reconhecimento de imagens. Medeiros et al. (2019) utilizam uma rede neural convolucional com assinaturas de equipamentos de dados públicos e sintéticos obtendo até 87% no desempenho. Ciancetta et al. (2021) implementaram uma *CNN* em *Python* para resolver o problema de desagregação, seu modelo precisa de uma entrada bidimensional 101x26 para classificar 11 eventos, os resultados apontaram uma precisão de 99,8% com os equipamentos treinados, sendo necessárias 10.000 amostras de dados na entrada da rede para inferir um equipamento.

Por ser uma abordagem bem explorada e designar resultados naturalmente mais confiáveis em relação às categorias semi-supervisionada e não supervisionada, nesta dissertação foi utilizada a técnica de aprendizagem supervisionada e os três algoritmos de aprendizado de máquina que apresentaram excelentes resultados para o problema *NILM*: *KNN*, *SVM* e *ANN*. Além da pesquisa, outro quesito que contribuiu para a escolha dos classificadores foi a utilização da ferramenta *Orange Data Mining*, que já possui implementado em suas bibliotecas estes algoritmos.

O uso de três algoritmos na metodologia é uma forma de verificar e obter a resposta para a pergunta de pesquisa, mitigando alguma divergência que possa haver na eficiência entre os classificadores. Portanto, durante o processo de classificação, todo e qualquer erro que possa ser observado é efeito das limitações nas técnicas dos classificadores utilizados e não ao método de aprendizado empregado.

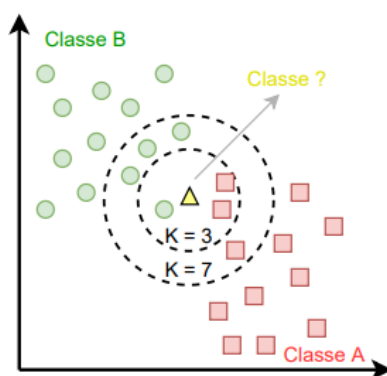
Nas próximas três subseções são apresentados de forma sucinta os modelos de classificação e algumas publicações que implementaram estes algoritmos para resolver o problema de classificação do monitoramento não intrusivo de cargas.

2.3.4.1 K-Nearest Neighbors

O k-vizinhos mais próximo é um algoritmo de aprendizado supervisionado muito simples que classifica novas instâncias com base em uma medida de similaridade (por exemplo, funções de distância: Euclidiana, *Mahalanobis*, *Manhattan*, etc.). Cada nova amostra é classificada pelo voto da maioria de seus vizinhos, com o rótulo sendo atribuído à classe mais comum entre seus “K” vizinhos mais próximos, ou seja, os exemplos do conjunto de dados (BOSCO; GONÇALVES; DE SOUZA, 2021). Essa técnica é possível pois dados semelhantes tendem a estar concentrados na mesma região do espaço de dispersão de dados, formando classes para cada *clusters*, logo este classificador pode resolver problemas lineares e não lineares.

Se for utilizado o valor $K = 1$, então o algoritmo avalia a distância entre todos os pontos do conjunto de dados e a nova amostra, neste caso é simplesmente atribuída à classe de seu vizinho mais próximo. Se $K = 3$ ou 7 , por exemplo, o algoritmo irá verificar as distâncias e realizará uma votação com os 3 ou 7 pontos mais próximos, respectivamente (C; KUMAR; K, 2019). Dessa forma, após calcular as distâncias entre os pontos do conjunto de dados e a nova instância, a votação será concluída e atribuída à qual classe é mais similar, conforme segue a Figura 8.

Figura 8 – Classificação de uma nova instância pelo KNN.



Fonte: Autor.

Neste caso a nova instância será classificada do tipo *A* se o valor 3 para o parâmetro *K* for escolhido, e do tipo *B* se 7 for utilizado. Valores atribuídos a *K* que sejam muito baixos serão influenciados por ruídos nos dados, gerando assim, influência no resultado final obtido. Em contrapartida, valores muito altos serão computacionalmente custosos.

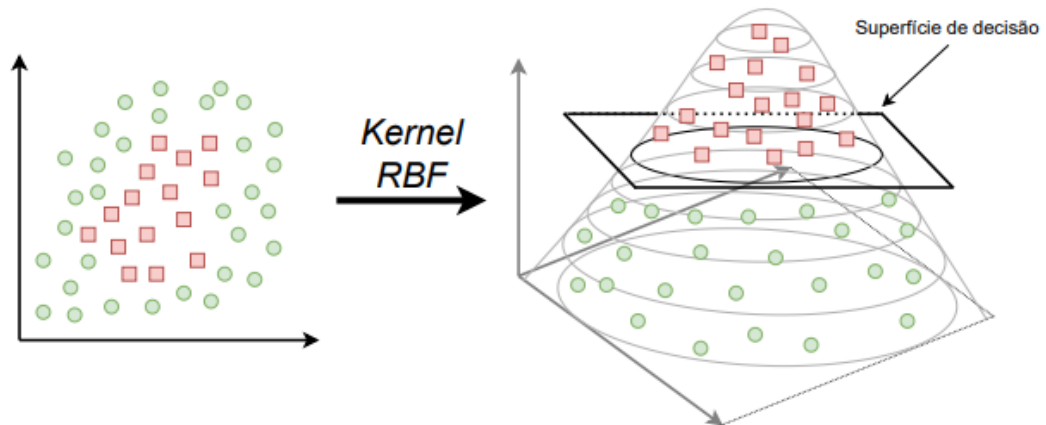
As vantagens deste algoritmo são alta precisão, insensibilidade a *outliers* e nenhuma suposição de dados de entrada. O algoritmo *KNN* pode identificar aparelhos com características não sobrepostas com alta precisão, e seu princípio é simples e fácil de implementar (TAO et al., 2019).

Bosco et al. (2021) mostram em seus estudos uma avaliação entre diferentes valores para *K* na classificação em monitoramento *NILM*, o parâmetro foi variado de 1 até 11, sendo obtido os melhores resultados quando o parâmetro *K* era igual ou menor que 3. Já o estudo de Tao et al. (2019) obtiveram melhores resultados no *KNN*, para classificação de cargas, com *K* igual a 20 dentro de um intervalo de 1 até 27 vizinhos.

2.3.4.2 Support Vector Machine

SVM é uma categoria de algoritmos de aprendizado de máquina supervisionado que pode ser usada para realizar regressões ou classificações lineares e não lineares. Seu foco principal é no treinamento e classificação de um conjunto de dados. Algoritmos nesta categoria funcionam construindo hiperplanos em espaços de alta dimensão, mapeando as entradas (recursos) para espaços de recursos de alta dimensão por meio de funções *Kernel* (linear, polinomial, radial, sigmóide e outras) (CORTES; VAPNIK, 1995). A Figura 9 exemplifica a aplicação da função de núcleo *Radial Basis Function (RBF)* para resolver problemas de classificação não linear convertendo o espaço dimensional.

Figura 9 – Conversão dimensional utilizando *Kernel RBF* no classificador SVM.

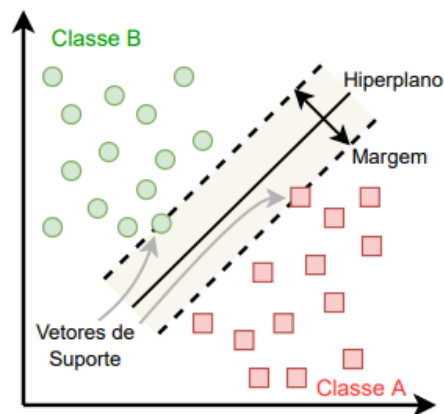


Fonte: Autor.

Uma característica importante desses classificadores é que eles sempre otimizam o hiperplano (ou hiperplanos) com a maior distância em relação aos dados de treinamento de cada classe (vetores de suporte), considerando que quanto maior a margem menor o erro do classificador. Dessa forma, os SVMs permitem que algumas entradas sejam classificadas incorretamente na fase de treinamento, para que a margem possa ser maximizada mais facilmente (CORTES; VAPNIK, 1995).

Um exemplo de conjuntos linearmente separáveis do SVM é apresentado na Figura 10. O hiperplano é representado pela linha divisória dos elementos, os vetores de suporte por linhas tracejadas e a distância mínima entre os vetores de suporte é chamada de margem.

Figura 10 – Formação do hiperplano em conjuntos linearmente separáveis.



Fonte: Autor.

O parâmetro que controla o número de classificações incorretas para maximizar a margem é definido de custo de classificação incorreta (C) e o parâmetro Γ representa a influência de cada exemplo de treinamento, para definir a curvatura em um limite de decisão. Valores menores de C significam que uma pequena penalidade é aplicada às classificações incorretas, resultando em margens maiores. Valores maiores de C resultarão em margens menores, uma vez que o algoritmo tentará cometer o mínimo de erros de classificação possível para evitar as altas penalizações (ABHILASH; KUMARI; PALLE, 2020).

Nos estudos de Rao et al. (2016), o classificador SVM foi testado e comparado com outros classificadores (*Naïve Bayes*, *Decision Trees* e *Artificial Neural Network*) para resolver o problema de desagregação. O trabalho mostra que usando *Kernel* linear, o *Support Vector Machine* teve melhor desempenho que os outros classificadores testados. Além disso, o SVM prova que é mais preciso quando os dados do conjunto não são tão grandes.

Ao comparar dois tipos de *Kernel* do algoritmo SVM para realizar a classificação de cargas, Abhilash et al. (2020) obtiveram 63% de acuracidade utilizando *Kernel* linear, parâmetro C igual a 10 e Γ igual a 0.00001, para identificar 3 aparelhos. E 100% de acuracidade com *Kernel RBF*, C igual a 10 e Γ igual a 0.000001, para até 7 cargas elétricas, mostrando um excelente desempenho na desagregação de equipamentos quando aplicada a conversão dimensional.

2.3.4.3 Artificial Neural Network

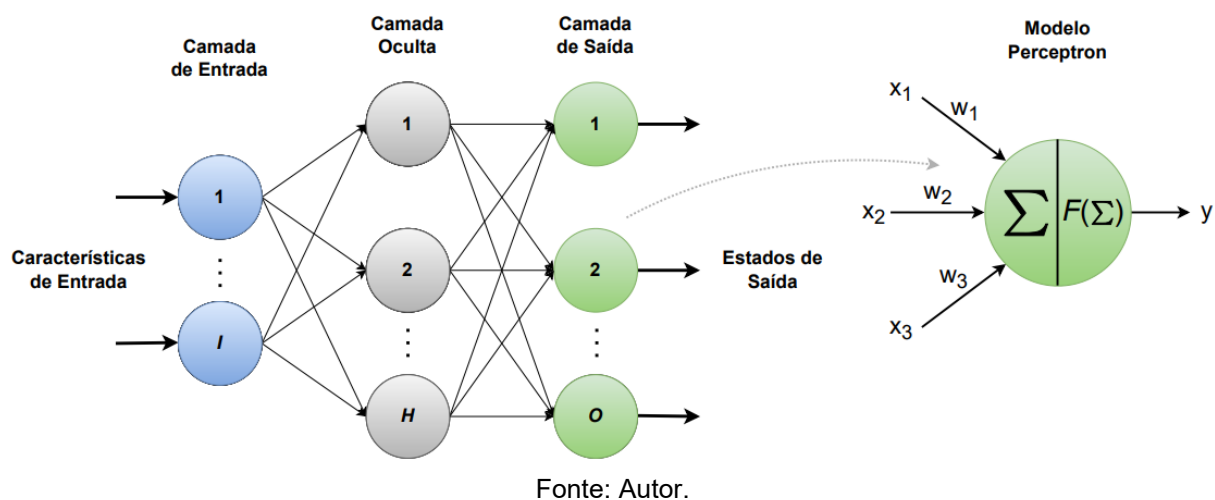
Redes neurais artificiais são outro método de aprendizado para construir modelos preditivos a partir de dados de treinamento rotulados. Elas imitam uma rede neural biológica interligando vários modelos de neurônios artificiais. Existem diversos tipos de redes neurais e sua configuração depende do tipo de problema e sua complexidade (VEEN, 2016).

O modelo mais comum de ANN é obtido pela implementação do neurônio artificial *Perceptron*. O princípio de funcionamento do modelo *Perceptron* é realizar a soma ponderada das entradas com seus respectivos pesos e acionar a saída através de uma função de ativação (por exemplo, linear, sigmóide, tangente hiperbólica e entre outras) (SHARMA, 2017).

Uma característica chave das redes neurais é o processo de aprendizado iterativo no qual os dados de treinamento são apresentados à rede um por vez, e os pesos associados aos valores de entrada são ajustados a cada iteração. Durante esta fase de aprendizagem, a rede aprende ajustando os pesos de forma que seja capaz de prever o rótulo de classe correto das amostras de treinamento (ELMSLEY, 2019b).

Uma rede neural construída com o modelo *Perceptron* pode conter uma ou mais camadas de neurônios, também conhecida como *Multi-layer Perceptron (MLP)*, e pode aprender tanto modelos lineares quanto não lineares (JAIN, 2021; JUNG, 2021; YAEMPRAYOON; SRINONCHAT, 2022). Um exemplo da *ANN-MLP* e do modelo *Perceptron* são mostrados na Figura 11. Os valores de entrada do neurônio artificial são representados por x , os pesos por w e y a saída.

Figura 11 – Modelo da rede neural artificial *MLP* e neurônio artificial *Perceptron*.



No caso das *ANNs*, há um número considerável de parâmetros que devem ser definidos com antecedência, podendo resultar em diferenças consideráveis na estrutura da rede e no processo de aprendizagem. Os parâmetros mais comuns incluem o número de camadas (*Layers*), o número (I) de neurônios na camada de entrada (*Input Layers*), o número (O) de neurônios na camada de saída (*Output Layers*), o número (H) de neurônios ocultos (*Hidden Neurons*), a taxa de aprendizado (*Learning Rate*), o otimizador para ajuste de pesos (*Solver*) e a função de ativação (*Activation Function*). Os quatro primeiros afetam principalmente a

estrutura da rede, enquanto os três últimos influenciam diretamente no processo de aprendizagem e inferência (ELMSLEY, 2019a; SINGH, 2020).

O método de aprendizagem mais utilizado e conhecido para ajustar os pesos nas redes neurais é o algoritmo de retropropagação (*Backpropagation - BP*) (TIAN et al., 2017). O objetivo da retropropagação é otimizar os pesos para que a rede neural possa aprender como mapear corretamente entradas arbitrárias em saídas, de modo que a saída real fique mais próxima da saída desejada. O ajuste é realizado verificando o erro da saída da rede e corrigindo os pesos de cada neurônio da saída para entrada, minimizando assim o erro para cada neurônio e a rede como um todo (BIANSOONGNERN; PLANGKLANG, 2016; ELMSLEY, 2019b).

A capacidade das redes neurais com ajuste por retropropagação, de reconhecer padrões e identificação, pode resolver problemas no monitoramento e identificação do uso de energia elétrica com resultados precisos. A vantagem da ANN é que a função utilizada pode ser não linear, possui alta precisão e não possui modelo definido, de modo que com este método de rede neural artificial não há necessidade de suposições a partir de dados multivariados que são normalmente distribuídos (ALI et al., 2021).

Ainda, de acordo com Tian et al. (2017), as características de múltiplas entradas e múltiplas saídas da rede neural *MLP* podem ser usadas no reconhecimento de estado de vários aparelhos elétricos, uma vez que as entradas são as características de amostragem dos dispositivos e as saídas são os rótulos correspondentes aos estados de execução dos equipamentos.

Conforme apresentado nos estudos de Ali et al. (2021), Chang et al. (2012), Tian et al. (2017) e Zhao et al. (2021), para o problema de desagregação de cargas não intrusivo, uma rede *ANN-MLP* pode ser modelada da seguinte forma:

I) Camada de entrada: informações de assinatura do evento, sendo as características selecionadas e extraídas, por exemplo, potência ativa e reativa. Neste caso, o número de neurônios de entrada é igual ao número de características das assinaturas.

II) Camada de saída: o número de neurônios de saída é igual ao número de estados das cargas individuais que se deseja identificar. Cada saída binária serve como um indicador de status para cada carga, por exemplo *ON/OFF*.

III) Camada oculta: em geral apenas uma camada oculta é usada para este tipo de problema. Algumas heurísticas foram desenvolvidas para determinar o

número de neurônios em uma camada oculta. O número comum de neurônios para a camada oculta pode ser definido como sendo a média aritmética ou a média geométrica da soma entre o número de neurônios da camada de entrada e de saída (FURTADO, 2019; TIAN et al., 2017).

IV) Função de ativação: deve ser uma função diferenciável e não linear (quando os dados não são lineares) para aprender automaticamente a detectar os recursos de entrada e saída no sistema e aproximar o mapeamento não linear com precisão arbitrária.

No modelo sugerido por Tian et al. (2017), a *ANN-MLP* foi implementada com a camada de entrada contendo 2 neurônios, pois as características de entrada das cargas eram apenas duas, uma camada oculta com 7 neurônios e 11 neurônios na camada de saída para identificar os estados de nove aparelhos. A função de ativação dos neurônios foi do tipo sigmóide e a retropropagação para ajuste dos pesos da rede foi utilizada.

Zhao et al. (2021) utilizam uma *MLP* mais complexa, contendo duas camadas ocultas, pois seu modelo foi estruturado de forma *online* com *multi-thread*. A entrada correspondeu com as características dos aparelhos, sendo 13 neurônios na camada de entrada da rede, 27 neurônios na primeira e 55 na segunda camada oculta que foram definidos pela equação $2n + 1$, onde n é o número neurônios correspondente a camada antecessora. A função de ativação utilizada em todos os nós da rede foi a função *ReLU*. Além disso, na publicação os autores compararam a *ANN-MLP* com o *SVM* e *KNN*, os resultados mostraram um desempenho de 99%, 98% e 96%, respectivamente, entre os classificadores testados na desagregação de equipamentos.

2.3.5 Predição do consumo de energia

A última etapa do fluxo *NILM* é a predição do consumo, ou estimativa de energia, aqui os eventos de energia classificados são usados para inferir o consumo dos aparelhos individuais. Este tópico foi explorado brevemente no trabalho de Hart (1992), onde foi proposto modelar o consumo individual dos aparelhos por meio de heurísticas como a *Zero Loop Sum Constraint (ZLSC)*, que afirma que a soma das mudanças de potência em qualquer ciclo de transições de estado é igual a zero. No

entanto, esta abordagem é muito simples e pode apresentar diversas falhas, pois assume que as transições de potência de um determinado aparelho são simétricas e que não há eventos simultâneos (SOUZA, 2016).

Outras abordagens para estimativa de energia também foram exploradas, He et al. (2019) apresentam um método não supervisionado de estimar o comportamento do aparelho com base nas curvas de potência usando Algoritmos Genéticos.

Azaza e Wallin (2017) propõem a modelagem do comportamento do aparelho usando formulações de *FSMs*, separando as curvas de energia dos aparelhos em modos transitórios da potência ativa e reativa. Porém, ambas abordagens ainda precisam ser validadas e testadas em ambiente real e algumas das suposições feitas pelos autores mostraram algumas desvantagens em relação ao tempo e acuracidade na predição.

Em resumo, a natureza sequencial das abordagens baseadas em eventos no fluxo *NILM* implica que cada etapa no processo propaga os resultados para seus sucessores. Logo, o objetivo dessas soluções é encontrar a melhor combinação de algoritmos e recursos entre as diferentes etapas, de modo que a energia desagregada seja adequadamente maximizada.

No entanto, essa tarefa é particularmente desafiadora por uma série de razões. De acordo com Souza (2016) os algoritmos de estimativa de energia requerem conjuntos de dados totalmente rotulados com dados suficientes para servir como conjuntos de treinamento e teste. Ainda, algumas das abordagens propostas também precisam de dados submedidos em nível de equipamento para aprender automaticamente os modelos dos aparelhos, o que por si só é outro problema diferente para os pesquisadores de técnicas *NILM* resolverem.

Embora pesquisada e exposta no referencial teórico, esta dissertação não implementa a etapa da predição do consumo dos equipamentos no fluxo do processo *NILM* para a metodologia aplicada. O foco e objetivo deste trabalho concentra-se na detecção e classificação de eventos de ativação e desativação, apesar do amplo campo de pesquisa necessário devido à sua complexidade e alguns desafios existentes apresentados para etapa de previsão do consumo elétrico.

2.4 MÉTRICAS DE AVALIAÇÃO

O termo precisão de desagregação tem sido utilizado por pesquisadores *NILM* quando se referem ao desempenho e avaliação de seus algoritmos. Muitas métricas foram apresentadas de acordo com a finalidade da avaliação que se destinam para definir a proximidade entre os resultados *NILM* e o valor real.

Estudos envolvendo o monitoramento não intrusivo podem avaliar diferentes etapas, desde identificação dos estados dos aparelhos, para o caso do monitoramento baseado em eventos, até a avaliação do consumo desagregado para análises mais complexas. Por exemplo, os autores Zoha et al. (2012) realizaram um estudo para definir métricas de precisão para detecção de eventos, classificação por aparelho individual, precisão de classificação global e cálculo de potência para verificar a diferença entre o consumo real e o previsto.

Métricas na detecção de eventos para quantificar algoritmos detectores também foram apresentadas por Houidi et al. (2018), Liu et al. (2019) e Luan et al. (2022), três métricas que permitem avaliar o processo são exploradas: (I) *Precision*; (II) *Recall*; e (III) *F1-Score*.

As técnicas de desempenho relacionam valores referentes às instâncias de verdadeiros positivos (*True Positive - TP*), verdadeiros negativos (*True Negative - TN*), falsos positivos (*False Positive - FP*) e falsos negativos (*False Negative - FN*). O Quadro 2 descreve as instâncias citadas.

Quadro 2 – Descrição das instâncias utilizadas nas métricas de avaliação.

Instâncias	Descrição
Verdadeiros Positivos (TP)	Número de vezes que o sistema reconhece algo como sendo verdadeiro e a saída real é verdadeira.
Verdadeiros Negativos (TN)	Número de vezes que o sistema reconhece algo como sendo falso e a saída real é falsa.
Falsos Positivos (FP)	Número de vezes que o sistema reconhece algo como sendo verdadeiro e a saída real é falsa.
Falsos Negativos (FN)	Número de vezes que o sistema reconhece algo como sendo falso e a saída real é verdadeira.

Fonte: Adaptada de Houidi et al. (2018).

A avaliação executada por *Precision* relata a fração de todas as instâncias recuperadas que são relevantes. Em outras palavras, é a proporção de instâncias verdadeiras que foram relatadas como verdadeiras (TP) em relação a todas as instâncias que foram relatadas ($TP + FP$). A Equação (5) define esta métrica.

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

A avaliação executada por *Recall* relata a fração de instâncias relevantes que foram realmente recuperadas. Ou seja, é a proporção de instâncias verdadeiras que foram relatadas como verdadeiras (TP) em relação a todas as instâncias verdadeiras ($TP + FN$). A Equação (6) apresenta a métrica *Recall*.

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

Se avaliadas distintamente *Precision* e *Recall*, as duas métricas relatam instâncias recuperadas corretamente, mas a custos diferentes. Isso pode trazer um problema no momento de analisar os resultados. Para o caso de algoritmos de detecção de eventos, em um detector mais liberal (dispara muitos eventos ao custo de um grande número de detecções incorretas), *Recall* aumentará com a redução de FN , mas *Precision* irá diminuir devido ao elevado número de FP . Por outro lado, um detector mais conservador (dispara poucos eventos com uma probabilidade alta de estar correto) terá *Precision* maior às custas de um *Recall* muito menor (HOUIDI et al., 2018; LIU et al., 2019).

Dessa forma, uma alternativa de compensação é a avaliação executada por *F1-Score* que realiza uma média harmônica entre as métricas *Precision* e *Recall*. A Equação (7) representa matematicamente a métrica *F1-Score*.

$$F1_Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (7)$$

Nos trabalhos de Medeiros et al. (2019), Pereira (2017), Liu et al. (2019) e Völker et al. (2020a) foram utilizados a métrica *F1-Score* para avaliar o desempenho dos detectores de eventos implementados no monitoramento não intrusivo. Assim

como nestes estudos, essa dissertação também utiliza o avaliador *F1-Score*, além de *Precision* e *Recall*, para quantificar o desempenho do fluxo *NILM* na etapa de detecção de eventos.

Quanto à identificação dos estados das cargas elétricas, a tarefa de classificação é um problema de várias classes, onde cada evento de energia pode ser classificado em mais de dois aparelhos diferentes. Portanto, uma métrica interessante para avaliar o desempenho dos algoritmos classificadores é através da matriz de confusão. Nela são expressas as instâncias classificadas corretamente e possíveis erros de classificação, conhecidos como as confusões na classificação das instâncias (LIU; WANG; YOU, 2019).

Uma forma de quantificar o desempenho na classificação, além das métricas *Precision*, *Recall* e *F1_Score*, é por meio da métrica de avaliação *Accuracy*. Ela relata a fração de instâncias classificadas corretamente em relação às classes corretas. Ou seja, é a proporção de instâncias classificadas de acordo com sua classe ($TP + TN$) em relação a todas as instâncias classificadas ($TP + TN + FP + FN$). A Equação (8) define a métrica *Accuracy*.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

As métricas de classificação podem ser calculadas sobre toda a coleção de classes ou sobre a média do desempenho de cada classe individual, com o intuito de avaliar de forma global o desempenho obtido na etapa de classificação. Essa diferença é particularmente importante quando as classes não são simétricas, sendo este o caso do *NILM*, pois em uma residência é normal que alguns aparelhos provoquem muito mais eventos do que outros. Portanto, Medeiros et al. (2019) e Souza (2016) sugerem a utilização das abordagens micro-média (*micro-averaging*) e macro-média (*macro-averaging*) para o emprego das métricas de avaliação.

Na análise micro-média é realizado um processo interno nas métricas, executando o somatório das instâncias (TP , TN , FP e FN) de todas as classes de interesse para obter as métricas de avaliação. Já na análise macro-média é realizada em um processo externo às métricas, executando a média simples para cada métrica de avaliação considerando todas as classes de interesse.

Assim sendo, como realizados por Ciancetta et al. (2021), Liu, Wang e You (2019), Medeiros et al. (2019) e Saha et al. (2020), nesta dissertação as avaliações dos classificadores empregados na metodologia são expressas através da matriz de confusão. Sendo possível analisar as classificações que apresentam eventuais problemas e entre quais classes esses erros ocorrem, além da quantificação dos testes de classificação de eventos pela métrica *Accuracy*.

Por fim, uma métrica pouco explorada em avaliações de desempenho para sistemas *NILM* é conhecida como a estatística *Kappa*, ou coeficiente de *Cohen's Kappa* (LANDIS; KOCH, 1977). É uma métrica interessante e suas origens estão no campo da psicologia, usada para medir a concordância entre dois avaliadores humanos, que mais tarde foi apropriado pela comunidade de aprendizado de máquina para medir o desempenho da classificação (SHMUELI, 2019).

O índice *Kappa* (κ) é o grau de concordância entre dois ou mais avaliadores, por isso é uma medida mais robusta para avaliar o desempenho de modelos de *ML*. É usada não apenas para quantificar um único classificador, mas também comparar classificadores entre si. A métrica compara a precisão observada (*Accuracy*) com a precisão esperada (ϵ), ou seja, os resultados do classificador utilizado para a classificação e os valores reais de cada instância a ser classificada (SAKIYAMA et al., 2008).

Matematicamente a estatística *Kappa* é calculada conforme a Equação (9):

$$\kappa = \frac{Accuracy - \epsilon}{1 - \epsilon} \quad (9)$$

A precisão esperada ϵ é definida como a precisão que qualquer classificador aleatório espera obter com base na matriz de confusão (REHMAN et al., 2021). Este avaliador é obtido pela Equação (10).

$$\epsilon = \frac{(TP + FN) \cdot (TP + FP) + (TN + FN) \cdot (TN + FP)}{(TP + TN + FP + FN)^2} \quad (10)$$

Os resultados de *Kappa* são valores tipicamente entre 0 e 1, onde resultados próximos de 1 indicam alta concordância entre os avaliadores, enquanto resultados próximos de 0 indicam baixa concordância. Lands e Koch (1977) atribuíram rótulos

em termos de força de concordância a diferentes faixas de κ , conforme apresentados na Tabela 1.

Tabela 1 – Diferentes faixas de *Kappa*.

κ	Rótulo
Menor que 0	Pobre
0 – 0.20	Leve
0.21 – 0.40	Razoável
0.41 – 0.60	Moderado
0.61 – 0.80	Substancial
0.81 – 1.00	Quase perfeito

Fonte: Adaptada de Rehman et al. (2021).

Geralmente, $\kappa > 0.40$ é desejável para problemas de classificação com aprendizado de máquina (REHMAN et al., 2021; SAKIYAMA et al., 2008). Portanto, este trabalho utiliza o índice *Kappa* para verificar o desempenho do sistema na desagregação, quanto a validação em ambiente real, descrito na metodologia e assumindo hipótese para rótulo “Moderado” ou superior.

2.5 CONJUNTO DE DADOS PÚBLICOS

Um conjunto de dados para desagregação de cargas elétricas residenciais é uma coleção de medições elétricas extraídas de casas em cenários do mundo real, tentando manter os dados o mais próximo possível da realidade. Estas coleções fornecem dados para pesquisadores *NILM* desenvolverem, treinarem, testarem e avaliarem seus algoritmos, semelhante às bases de dados visuais voltadas para pesquisas em reconhecimento de imagens e identificação de padrões através de *ML*.

Cada conjunto de dados possui características próprias, que variam desde o local em que foram gerados (localização geográfica), frequência de amostragem, grandezas elétricas disponíveis, dados agregados e desagregados, período de registro, até a quantidade e variedade de equipamentos domésticos registrados (KELLY; KNOTTENBELT, 2015).

Além disso, os dados podem ser fornecidos através de medições do consumo de toda a edificação (painel elétrico), medições das cargas individuais ou por circuito individual compreendendo um conjunto de cargas conectadas.

Entretanto, este último método nem sempre garante que o consumo individual de todas as diferentes cargas esteja realmente disponível (ANDERSON et al., 2012a).

Os conjuntos de dados também podem ser categorizados como baseados em eventos ou sem eventos. Neste caso, os baseados em eventos requerem a identificação de todas as mudanças de consumo junto aos dados de medição, enquanto a segunda categoria não necessita. Logo, é maior a disponibilidade de conjuntos de dados públicos sem eventos, pois a coleta é mais simples e rápida (VÖLKER et al., 2020b).

Na literatura *NILM* existem diversos trabalhos que enfatizam pesquisas sobre os diferentes conjuntos de dados públicos. Os autores Nunes e Pereira (2015), Renaux et al. (2018) e Laouali et al. (2020) resumem alguns bancos de dados disponíveis e, de acordo com estudos, já é possível encontrar mais de 20 conjuntos públicos de categoria residencial. Porém, até a presente data, não foram encontrados conjuntos de dados públicos de origem brasileira. Logo, os registros dos sinais elétricos coletados do ambiente real para este trabalho ficarão disponíveis publicamente, uma vez que coleções de dados nacionais são de grande importância para pesquisadores *NILM* no Brasil.

A Tabela 2 apresenta os principais conjuntos de dados públicos para desagregação de cargas elétricas residenciais com suas principais características.

Tabela 2 – Principais conjuntos de dados públicos e suas características.

Conjunto (Ano)	País	Duração	Baseado Eventos	Agregado / Desagregado	Grandezas	Resolução	Aparelhos (Casas)
REDD ¹ (2011)	USA	2-4 Semanas	Não	Agr.	V, I, P	15 kHz, 1Hz, 0.5Hz	N.D. (6)
BLUED ² (2012)	USA	1 Semana	Sim	Agr.	V, I, P	12 kHz, 60 Hz	43 (1)
AMPds ³ (2013)	Canada	2 Anos	Não	Agr. / Desagr.	V, I, P, Q, S, F	1/60 Hz	19 (1)
Dataport ⁴ (2013)	USA	4 Anos	Não	Agr. / Desagr.	P, S	1/60 Hz	N.D. (1400)
ECO ⁵ (2014)	Suíça	8 Meses	Não	Agr. / Desagr.	V, I, P	1 Hz	47 (6)
PLAID ⁶ (2014)	USA	1094 Eventos	Sim	Desagr.	V, I	30 kHz	235 (55)
REFIT ⁷ (2014)	UK	2 Anos	Não	Agr. / Desagr.	P	1/8 Hz	11 (20)
UK-DALE ⁸ (2014)	UK	2 Anos	Não	Agr. / Desagr.	V, I, P, Q	16 kHz, 1/6 Hz	53 (5)
COOLL ⁹ (2016)	França	840 Eventos	Sim	Desagr.	V, I	100 kHz	42 (N.D.)

Tabela 2 – Principais conjuntos de dados públicos e suas características.

Conjunto (Ano)	País	Duração	Baseado Eventos	Agregado / Desagregado	Grandezas	Resolução	Aparelhos (Casas)
<i>WHITED</i> ¹⁰ (2016)	Alemanha	5123 Eventos	Sim	Desagr.	V, I	44 kHz	110 (N.D.)
<i>FIRE</i> ¹¹ (2020)	Alemanha	8 Semanas	Sim	Agr. / Desagr.	V, I, P, Q, S	8 kHz, 2kHz, 50Hz	35 (1)

V - Tensão; I - Corrente; P - Potência Ativa; Q - Potência Reativa; S - Potência Aparente; F - Frequência; N.D. - Não Definido.

Fonte: [1] Kolter e Johnson (2011), [2] Anderson et al. (2012a), [3] Makonin et al. (2013), [4] Holcomb (2014), [5] Beckel et al. (2014), [6] Gao et al. (2014), [7] Murray et al. (2015), [8] Kelly e Knottenbelt (2015), [9] Picon et al. (2016), [10] Kahl et al. (2016), [11] Völker et al. (2020b).

Dos conjuntos de dados listados na Tabela 2, é possível ver que 6 deles não são baseados em eventos (*REDD*, *AMPds*, *Dataport*, *ECO*, *REFIT* e *UK-DALE*), ou seja, não possuem rotulagem dos eventos, logo não satisfazem a condição utilizada nesta dissertação para abordagem baseada em eventos no monitoramento não intrusivo (BECKEL et al., 2014; HOLCOMB, 2014; KELLY; KNOTTENBELT, 2015; KOLTER; JOHNSON, 2011; MAKONIN et al., 2013; MURRAY et al., 2015).

PLAID, *COOLL* e *WHITED* são coleções que contêm apenas dados dos transientes de inicialização e traços espectrais de vários dispositivos individuais em alta frequência de amostragem (GAO et al., 2014; KAHL et al., 2016; PICON et al., 2016). Por último, *BLUED* e *FIRE* são os únicos conjuntos de dados baseados em eventos que foram listados, porém o primeiro não possui dados desagregados individuais, apenas dados por circuito individual, o que não garante o consumo individual dos equipamentos (ANDERSON et al., 2012a). Por outro lado, *FIRE* possui os dois tipos de medição, agregado e por aparelho individual, além de possuir arquivo de rotulagem que garante o gabarito em aplicações de testes e avaliações para uma maior gama de algoritmos *NILM*, sendo possível utilizá-lo nos dois modelos de abordagem, as baseadas e não baseadas em eventos (VÖLKER et al., 2020b).

Dessa forma, para esta dissertação foi utilizado o conjunto de dados *FIRE* por ser baseado em eventos, fornecer dados agregados e de equipamentos individuais, além de possuir rotulagem dos dados. Também, apresenta como requisito a frequência de amostragem com base no ciclo da rede elétrica (50 Hz). Fatores esses, condicionantes para o treinamento do sistema de detecção e classificação de eventos, considerando as métricas de avaliação.

2.6 CONCLUSÃO

Foram apresentadas, neste capítulo, as técnicas de monitoramento de cargas elétricas no ambiente residencial, tratando os principais conceitos envolvidos no monitoramento não intrusivo de cargas. Nas duas categorias de monitoramento, não baseada em eventos e baseada em eventos, foram destacados os aspectos envolvidos, representando ser um divisor das técnicas utilizadas para o desenvolvimento de sistemas *NILM*, visto que a segunda categoria é uma abordagem mais próxima de ser empregada na prática.

Além disso, foi explorado o fluxo *NILM* baseado na literatura, considerando desde os primeiros trabalhos sobre o tema, onde cada processo (Figura 6) é apresentado sobre a perspectiva de alguns autores no estudo do monitoramento não intrusivo de cargas. Também foi descrito um método utilizado para detectar a mudança do estado de operação dos aparelhos, baseado em eventos, através da razão logarítmica de probabilidade, que apesar de simples apresenta bons resultados.

Para realizar a classificação de eventos no monitoramento *NILM* foram explanados três diferentes algoritmos de aprendizado de máquina bem difundidos na literatura. Os classificadores *KNN*, *SVM* e *ANN* são métodos menos complexos que empregam aprendizado supervisionado para realizar a desagregação de cargas, porém muito utilizados em estudos *NILM* por terem resultados consideráveis e possuírem a vantagem de serem computacionalmente mais leves para aplicações reais.

Como forma de avaliar o desempenho dos sistemas *NILM*, métricas de avaliação são utilizadas para quantificar os métodos de detecção e classificação. Por fim, foram apresentados os principais conjuntos de dados públicos disponíveis para aplicação de testes, treinamento e avaliação *NILM*, onde cada coleção é constituída de características distintas, o que possibilita uma variedade de práticas para o desenvolvimento de sistema de monitoramento de cargas.

3 METODOLOGIA

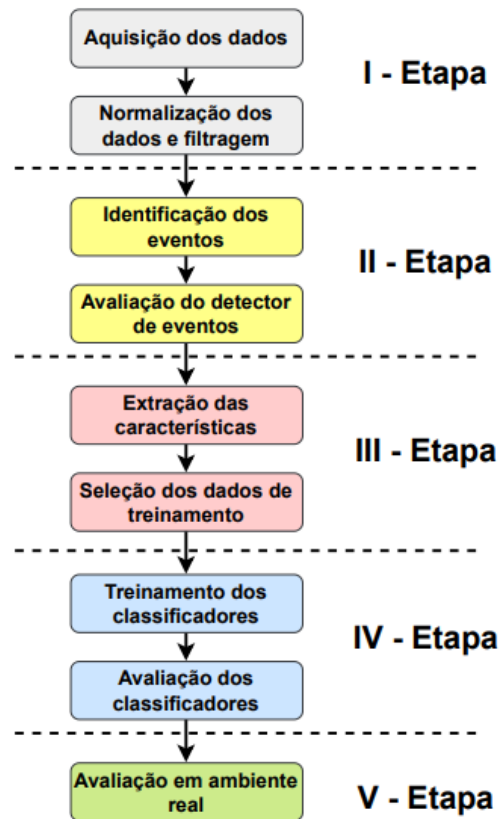
Neste capítulo, apresenta-se a metodologia proposta para adquirir os dados de entrada, identificar e avaliar a detecção dos eventos, realizar a extração das características para gerar as assinaturas das cargas, treinar e avaliar os classificadores. Por fim, o sistema devidamente treinado com um conjunto de dados público é validado com base em testes utilizando sinais elétricos originados em um cenário de ambiente real local.

Vale salientar que nesta metodologia, a aquisição dos dados foi realizada na frequência do ciclo elétrico da rede de distribuição, sendo apresentado o modelo de assinaturas que extrai dados transitórios e estacionários das cargas de ativação e desativação.

Outro ponto importante a ressaltar, trata-se da avaliação por três diferentes classificadores, ideais para sistemas *NILM* de caráter prático e com baixo poder computacional. Dessa forma é possível fornecer o estado de operação das cargas elétricas residenciais, identificando qual equipamento foi ativado ou desativado poucos segundos após sua detecção, tornando o método adequado para ser empregado em aplicações de gerenciamento de energia.

Os processos e etapas da metodologia desenvolvida para o monitoramento *NILM* são apresentados na Figura 12.

Figura 12 – Fluxograma da metodologia desenvolvida.



Fonte: Autor.

3.1 ETAPA I

A primeira etapa da metodologia é composta por dois processos para realizar a aquisição, normalização e filtragem dos dados de entrada. Nesta etapa, as grandezas elétricas são obtidas e pré-processadas antes de serem aplicadas ao detector de eventos.

3.1.1 Aquisição dos dados

Neste processo é considerado que os dados de entrada estejam disponíveis através de monitoramento local (ambiente real), conjunto de dados públicos ou até mesmo dados sintéticos. A frequência de amostragem dos sinais compreende o ciclo de fase da rede elétrica (60 Hz para o caso da rede elétrica brasileira) ou a frequência da rede elétrica do conjunto de dados públicos.

Como a metodologia desenvolvida utiliza o método de aprendizado supervisionado e abordagem baseada em eventos, os sinais do monitoramento

agregado e desagregado são necessários para identificação dos equipamentos com rotulagem de ativação e desativação. Dessa forma é possível extrair as assinaturas elétricas das cargas para as fases de treinamento e avaliação do sistema.

As grandezas elétricas utilizadas na aquisição dos dados são a tensão elétrica, potência ativa, potência reativa e potência. Estes sinais são a base da metodologia para realizar a normalização, detecção dos eventos, extração das características e classificação das cargas elétricas.

3.1.2 Normalização dos dados e filtragem

A normalização dos dados tende reduzir os impactos das variações dos sinais de potência ativa, potência reativa e potência aparente em relação às variações da tensão elétrica. Com esta abordagem pretende-se diminuir as variações das potências causadas pelas alterações da tensão, gerando um conjunto de assinaturas mais robusto e menos influenciado pelas oscilações da rede elétrica que supre uma residência.

Assim, como sugerido por Hart (1992), a normalização dos valores das potências são adotadas nesta metodologia. A Equação (11) realiza a normalização para os três sinais de potência P , Q e S .

$$\{P, Q, S\}_{Norm}[i] = \left(\frac{V_{Nom}}{V[i]}\right)^2 \cdot \{P, Q, S\}[i] \quad (11)$$

Sendo:

$P_{Norm}[i]$ a potência ativa normalizada na amostra i ;

$Q_{Norm}[i]$ a potência reativa normalizada na amostra i ;

$S_{Norm}[i]$ a potência aparente normalizada na amostra i ;

V_{Nom} a tensão nominal;

$V[i]$ a tensão elétrica na amostra i ;

$P[i]$ a potência ativa na amostra i ;

$Q[i]$ a potência reativa na amostra i ;

$S[i]$ a potência aparente na amostra i ;

i a i -ésima amostra.

A aplicação de um filtro média móvel para a potência ativa normalizada é necessária para eliminar variações rápidas e curtas dos sinais, muitas vezes ocasionadas por ruídos no sinal, que podem dificultar a identificação dos eventos pelo algoritmo detector.

A Equação (12) representa o filtro média móvel que suavizam as curvas da potência.

$$P_{Filt}[i] = \frac{1}{N+1} \sum_{n=0}^N P_{Norm}[i+n] \quad (12)$$

Sendo:

$P_{Filt}[i]$ a potência ativa filtrada na amostra i ;

$P_{Norm}[i+n]$ a potência ativa normalizada na amostra $i+n$;

N a quantidade de amostras para realizar a média.

3.2 ETAPA II

Nesta etapa é realizada a identificação dos eventos, aplicando algoritmo de abordagem probabilística aos sinais anteriormente amostrados e pré-processados. A segunda atividade desta etapa é a avaliação do detector de eventos desenvolvido na metodologia, utilizando as métricas de avaliação para detecção de eventos.

3.2.1 Identificação dos eventos

Para aplicação desta metodologia, a identificação dos eventos provocados pelas cargas elétricas monitoradas é necessária. Somente a partir deles é possível localizar onde um equipamento teve seu estado de operação alterado, utilizando o sinal da potência elétrica para análise. Com a identificação dos eventos, as assinaturas elétricas podem ser extraídas na vizinhança do local detectado.

Como esta metodologia adota o modelo probabilístico para detecção de eventos, é utilizado o algoritmo da razão logarítmica de probabilidade (*LLR*), conforme proposto por Völker, Scholl e Becker (2019). As grandezas analisadas pelo

algoritmo são as amostras da potência ativa, aplicadas ao *LLR* após executada a primeira etapa de aquisição, normalização e filtragem dos dados.

A função de identificação de eventos desenvolvida utiliza uma janela de detecção deslizante (dw) sobre as amostras do sinal da potência ativa. Esta janela é dividida em duas partes, janela de pré-evento (w_{pre}), antes da amostra analisada, e janela de pós-eventos (w_{pos}), depois da amostra analisada. As janelas são preenchidas conforme Equações (13), (14) e (15).

$$dw = [P_{Filt}[i - h], \dots, P_{Filt}[i - 1], P_{Filt}[i], P_{Filt}[i + 1], \dots, P_{Filt}[i + j]]_{1 \times (h+j+1)} \quad (13)$$

$$w_{pre} = [dw[1], dw[2], \dots, dw[h]]_{1 \times h} \quad (14)$$

$$w_{pos} = [dw[h + 1], dw[h + 2], \dots, dw[h + j + 1]]_{1 \times (j+1)} \quad (15)$$

Sendo:

dw a janela de detecção deslizante, preenchida com as amostras entre $i - h$ até $i + j$ da potência ativa, contendo $h + j + 1$ amostras;

w_{pre} a janela de pré-evento, preenchida com as primeiras h amostras da janela de detecção, contendo h amostras;

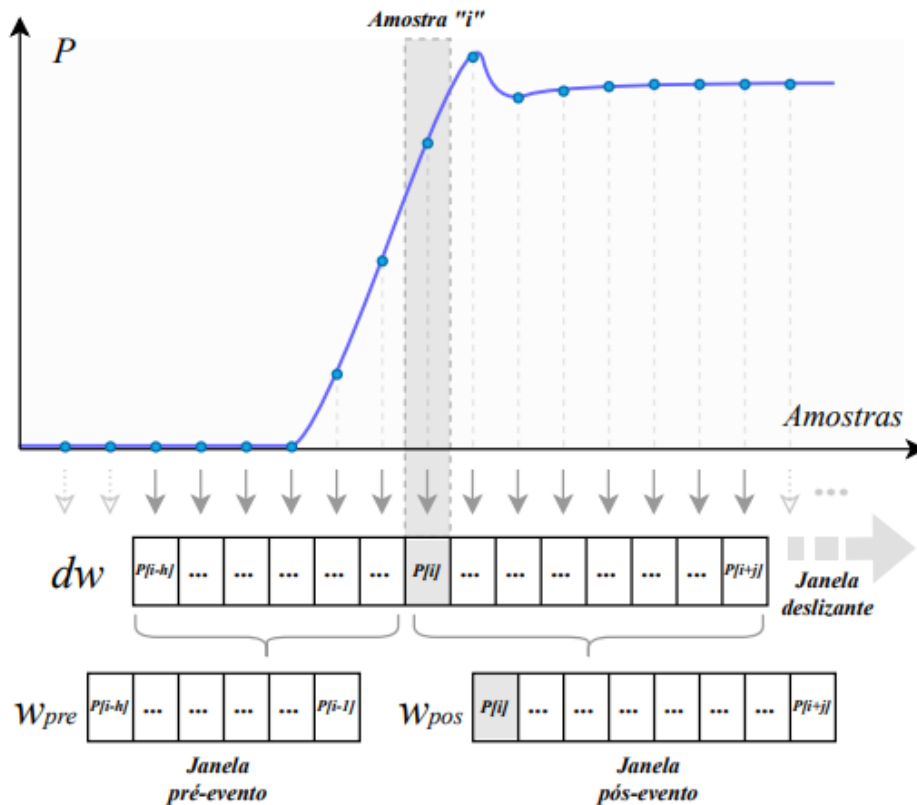
w_{pos} a janela de pós-evento, preenchida com as amostras entre $h + 1$ e $h + j + 1$ da janela de detecção, contendo $j + 1$ amostras;

h a quantidade de amostras antes do evento;

j a quantidade de amostras depois do evento.

Uma abstração do preenchimento das janelas deslizantes em um possível evento de ativação é ilustrada na Figura 13:

Figura 13 – Preenchimento da janela deslizante com amostras da potência.



Fonte: Autor.

Através das janelas pré-evento e pós-evento são calculadas as médias dos conjuntos, conforme Equações (16) e (17).

$$\mu_{pre} = \frac{1}{h} \sum_{n=1}^h w_{pre}[n] \quad (16)$$

$$\mu_{pos} = \frac{1}{j+1} \sum_{n=1}^{j+1} w_{pos}[n] \quad (17)$$

Sendo:

μ_{pre} a média aritmética das amostras da janela de pré-evento;

μ_{pos} a média aritmética das amostras da janela de pós-evento;

$w_{pre}[n]$ a janela de pré-evento processada no índice n ;

$w_{pos}[n]$ a janela de pós-evento processada no índice n ;

h a quantidade de amostras antes do evento;

j a quantidade de amostras depois do evento.

Também são calculados os desvios-padrão das amostras das janelas, de acordo com as Equações (18) e (19).

$$\sigma_{pre} = \sqrt{\frac{1}{h} \sum_{n=1}^h (\mu_{pre} - w_{pre}[n])^2} \quad (18)$$

$$\sigma_{pos} = \sqrt{\frac{1}{j+1} \sum_{n=1}^{j+1} (\mu_{pos} - w_{pos}[n])^2} \quad (19)$$

Sendo:

σ_{pre} o desvio-padrão das amostras da janela de pré-evento;

σ_{pos} o desvio-padrão das amostras da janela de pós-evento;

μ_{pre} a média aritmética da janela de pré-evento;

μ_{pos} a média aritmética da janela de pós-evento;

$w_{pre}[n]$ a janela de pré-evento processada no índice n ;

$w_{pos}[n]$ a janela de pós-evento processada no índice n ;

h a quantidade de amostras antes do evento;

j a quantidade de amostras depois do evento.

Antes de calcular a probabilidade da ocorrência de um evento, é verificada a diferença entre as janelas pré-evento e pós-evento pela Equação (20). Se esta diferença absoluta for menor que um limiar, então é atribuído o valor zero de probabilidade. Se a diferença absoluta foi maior que o limiar, então é calculada a probabilidade de ocorrência do evento, conforme Equação (21).

$$\delta w = \mu_{pos} - \mu_{pre} \quad (20)$$

$$LLR[i] = \begin{cases} L[i], & \text{se } |\delta w| > threshold_i \\ 0, & \text{se } |\delta w| \leq threshold_i \end{cases} \quad (21)$$

Sendo:

δw a diferença entre as médias das janelas pré-evento e pós-evento;

$LLR[i]$ o sinal estatístico de detecção de eventos correspondente ao sinal de potência ativa $P_{Filt}[i]$;

$L[i]$ o analisador da razão logarítmica de probabilidade LLR , correspondente a amostra i da potência ativa $P_{Filt}[i]$;

μ_{pre} a média aritmética da janela de pré-evento;

μ_{pos} a média aritmética da janela de pós-evento;

$threshold_i$ o limiar que define quando realizar o cálculo da probabilidade.

Normalmente as cargas que consomem alta potência apresentam uma variação maior no sinal, de acordo com Völker, Scholl e Becker (2019). Dessa forma, esta metodologia adota um limiar variável que aumenta linearmente com a média da potência de pré-evento do dispositivo e se adapta às possíveis flutuações maiores. O limiar variável, $threshold_i$, é definido pela Equação (22).

$$threshold_i = thres_{min} + m \cdot \mu_{pre} \quad (22)$$

Sendo:

$threshold_i$ o limiar para realizar o cálculo da probabilidade;

$thres_{min}$ o limite mínimo de interesse nas variações da potência ativa observada;

m o coeficiente linear que aumenta o limiar em relação à média da janela de pré-evento;

μ_{pre} a média aritmética da janela de pré-evento.

Finalmente, se a diferença absoluta entre as médias pré e pós-evento for maior que o limiar variável, então é calculada a probabilidade do evento correspondente à amostra da potência, conforme Equação (21). O analisador da razão logarítmica de probabilidade LLR é definido pela Equação (23).

$$L[i] = \ln\left(\frac{\sigma_{pre}}{\sigma_{pos}}\right) + \frac{(dw[h+1] - \mu_{pre})^2}{2 \cdot \sigma_{pre}^2} - \frac{(dw[h+1] - \mu_{pos})^2}{2 \cdot \sigma_{pos}^2} \quad (23)$$

Sendo:

$L[i]$ o analisador da razão logarítmica de probabilidade LLR ;

σ_{pre} o desvio-padrão da janela de pré-evento;

σ_{pos} o desvio-padrão da janela de pós-evento;

μ_{pre} a média aritmética da janela de pré-evento;

μ_{pos} a média aritmética da janela de pós-evento;

$dw[h + 1]$ a janela de detecção, processada na amostra $h + 1$, que corresponde a amostra da potência ativa $P_{Filt}[i]$.

Se um evento for detectado na amostra i do sinal da potência, $LLR[i]$ também é diferente de zero em torno da amostra i , visto que a mudança na média ainda é observada nas proximidades do evento.

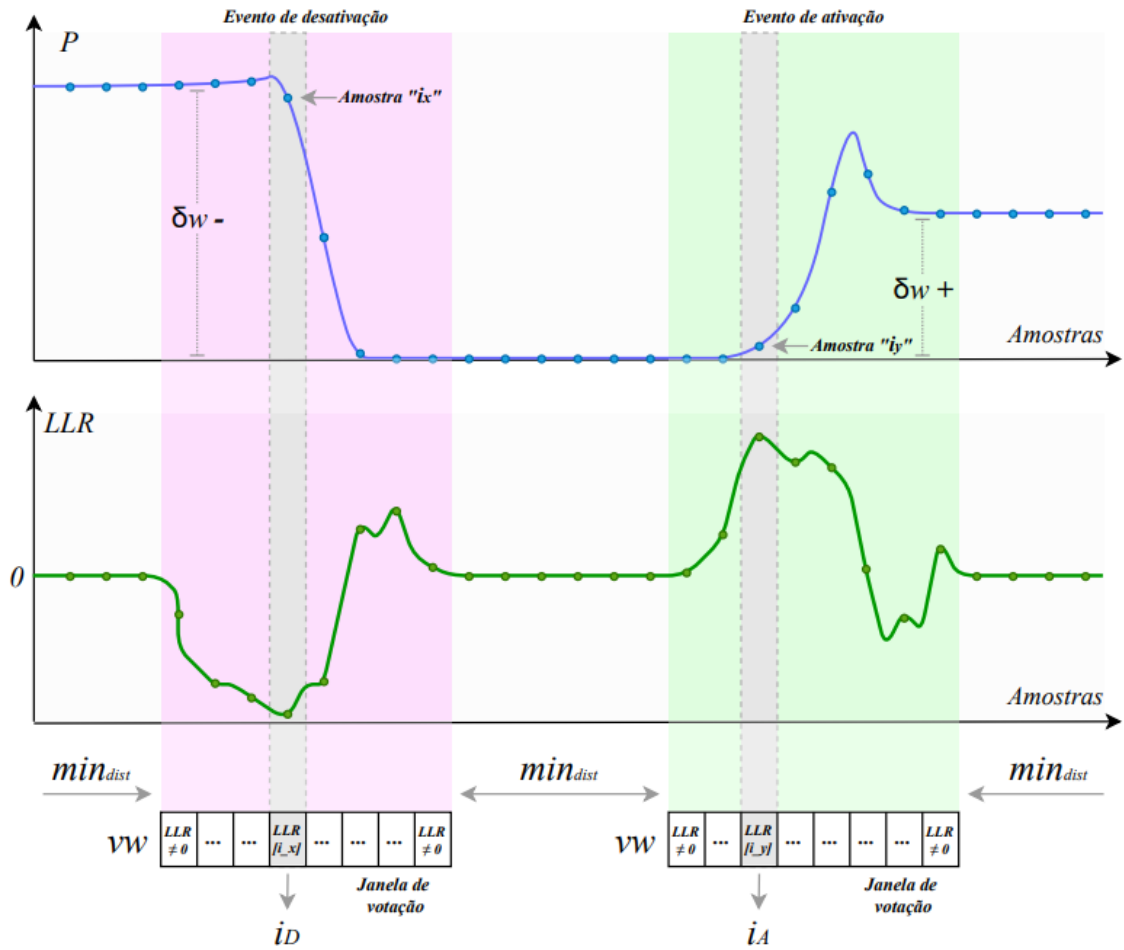
Para definir o índice da amostra onde ocorreu um evento, uma janela de votação (vw) é aplicada ao sinal estatístico de detecção para verificar, entre os intervalos não nulos, qual índice possui o maior valor. Após a votação, dentro da janela apenas o índice do máximo valor absoluto de $LLR[i]$ é mantido.

Além disso, uma distância mínima (min_{Dist}) entre os eventos é definida para reduzir a detecção de falsos positivos, sendo considerada uma próxima votação de vw após min_{Dist} amostras da última detecção em $LLR[i]$.

Quando o índice do valor máximo de $LLR[i]$ é definido pela janela de votação, a amostra i é considerada o instante em que ocorreu um evento, correspondendo a mesma amostra da potência ativa. Se neste instante a diferença entre as janelas pré e pós-evento (δw) for positiva, então um evento de ativação foi detectado, atribuindo o instante i ao instante i_A (índice da amostra de ativação). Se a diferença δw for negativa, então um evento de desativação foi detectado, atribuindo o instante i ao instante i_D (índice da amostra de desativação).

A Figura 14 abstrai um suposto evento de desativação e ativação de cargas, extraíndo as amostras da potência e aplicando a janela de votação sobre o cálculo do LLR para definir a amostra que representa o momento do evento.

Figura 14 – Definição dos índices de ativação e desativação de eventos.



Fonte: Autor.

Os índices i_A de ativação e i_D de desativação são utilizados, posteriormente, para extrair as características na vizinhança do local de detecção dos eventos. Também é utilizado na formação das assinaturas elétricas para treinamento, testes dos classificadores ou classificação efetiva de equipamentos. A cada nova amostra do sinal da potência a janela de detecção é deslocada e novos valores são calculados.

O algoritmo proposto é implementado em linguagem *Python* e como este detector possibilita e necessita de ajustes, para um melhor desempenho sobre os dados analisados, os parâmetros mostrados no Quadro 3 são definidos na fase de testes para detectar os eventos dos equipamentos monitorados. De maneira geral, estes parâmetros são estabelecidos de forma empírica até se atingir um desempenho satisfatório do detector.

Quadro 3 – Parâmetros ajustáveis do detector de eventos.

Parâmetro	Descrição
dw	Janela de detecção.
w_{pre}	Janela de pré-evento.
w_{pos}	Janela de pós-evento.
$thres_{min}$	Limite mínimo das variações do sinal.
m	Coefficiente linear de ajuste do limiar.
min_{Dist}	Mínima distância entre eventos.

Fonte: Autor.

3.2.2 Avaliação do detector de eventos

O detector de eventos recomendado nesta metodologia pode ser avaliado utilizando um conjunto de dados público que contenha os sinais agregados e individuais dos equipamentos inspecionados. A rotulagem dos eventos é uma informação que indica as mudanças de estado provocadas pelas cargas, contendo o índice ou grandeza temporal do exato momento de ativação ou desativação de um equipamento. O conjunto de dados públicos *FIRE*D é contemplado com estes requisitos, o que possibilita realizar as métricas com maior facilidade (VÖLKER et al., 2020b).

A ausência de informação rotulada não impede a avaliação do detector, mas exige uma inspeção “visual” nos sinais da potência desagregada para identificar os momentos onde os eventos em análise foram provocados. Dessa forma é possível obter o “gabarito”, ou verdadeiro fundamental (*ground truth*), dos eventos ocasionados por cada equipamento da coleção de sinais sem rotulagem. Portanto, um conjunto de dados local (extraído do ambiente real onde é executada esta metodologia) também pode ser utilizado, desde que contenha os sinais agregados e desagregados dos equipamentos monitorados para inspeção dos eventos gerados.

Com base no conhecimento dos eventos, rotulagem ou inspeção visual, é selecionado um período do conjunto de dados para ser verificado pelo detector e avaliado através das métricas *Precision*, *Recall*, e *F1-Score*, conforme expostas na seção 2.4.

Estes avaliadores quantificam o algoritmo detector de eventos através das instâncias *TP* (eventos detectados que correspondem ao *ground truth*), *FP* (eventos detectados sem correspondência ao *ground truth*) e *FN* (eventos do *ground truth* que não são detectados). Após a contabilidade das instâncias, a avaliação do detector de eventos é executada através das Equações (5) *Precision*, (6) *Recall* e (7) *F1-Score*.

3.3 ETAPA III

Assim como no diagrama de fluxo do sistema *NILM* (Figura 6, seção 2.3), onde cada etapa do processo propaga seus resultados para as etapas seguintes, nesta metodologia não é diferente. Os resultados obtidos na Etapa II do detector de eventos foram utilizados na Etapa III.

Os eventos dos equipamentos identificados são utilizados como ponto de partida para extrair as características das assinaturas elétricas, conforme descrito nas duas subseções seguintes.

3.3.1 Extração das características

A formação das assinaturas, ou extração das características das cargas, foram realizadas com base nos sinais da potência ativa, potência reativa e potência aparente normalizadas.

As assinaturas elétricas geradas são utilizadas para a identificação e treinamento dos equipamentos no processo de classificação que correspondem aos aparelhos inspecionados do ambiente monitorado. Portanto, os eventos detectados no monitoramento servem de indicação para obter os recursos dos sinais elétricos nos instantes listados.

A partir dos índices i_A e i_D são analisados os sinais das potências, que correspondem a estes instantes, e extraídas informações para compor as características de cada equipamento. Nesta etapa, foram extraídas a diferença de potencial e as características transitórias ocorridas no evento, tais como: pico máximo do transiente e a distorção.

Um dos recursos mais importantes que compõem a assinatura é a diferença de potencial. Esta diferença é realizada através da subtração de amostras pré e pós-

evento de potência, que equivalem aos instantes anterior e posterior da detecção no sinal $LLR[i]$. Ou seja, o índice com amostra de valor zero antes do início da detecção (i_{pre}) e o índice com amostra de valor zero após a detecção (i_{pos}) do sinal $LLR[i]$, onde é instanciado uma detecção i_A ou i_D .

Os instantes escolhidos para extração das amostras correspondem ao sinal estabilizado da potência, uma vez que $LLR[i]$ é zero quando não há variações significativas nas médias das potências. A Equação (24) representa a diferença calculada em P , Q e S para eventos de ativação e desativação.

$$\delta\{P, Q, S\} = \{P, Q, S\}_{pos} - \{P, Q, S\}_{pre} \quad (24)$$

Sendo:

δP a diferença da potência ativa do evento;

δQ a diferença da potência reativa do evento;

δS a diferença da potência aparente do evento;

P_{pos} a amostra da potência ativa $P_{Norm}[i_{pos}]$ depois do evento;

Q_{pos} a amostra da potência reativa $Q_{Norm}[i_{pos}]$ depois do evento;

S_{pos} a amostra da potência aparente $S_{Norm}[i_{pos}]$ depois do evento;

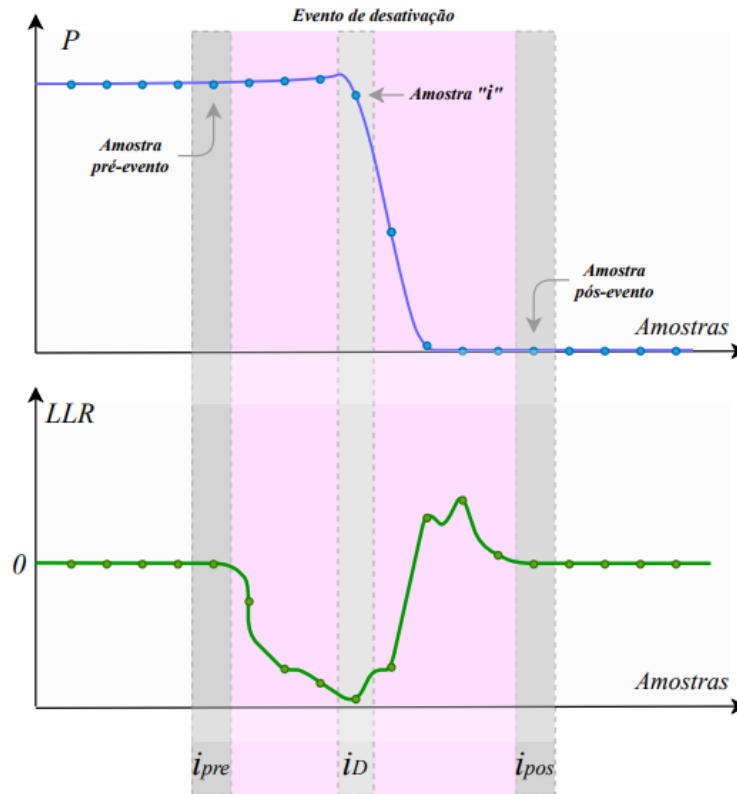
P_{pre} a amostra da potência ativa $P_{Norm}[i_{pre}]$ antes do evento;

Q_{pre} a amostra da potência reativa $Q_{Norm}[i_{pre}]$ antes do evento;

S_{pre} a amostra da potência aparente $S_{Norm}[i_{pre}]$ antes do evento.

Um exemplo é ilustrado na Figura 15 para extração das amostras pré e pós-evento da potência nos índices i_{pre} e i_{pos} . Após obtidos os valores é calculada a diferença de potência para um possível evento de desativação.

Figura 15 – Extração das amostras para calcular a diferença de potencial.



Fonte: Autor.

Outro recurso extraído é o transiente de pico máximo da potência que geralmente é oriunda de cargas não lineares. Esta propriedade é obtida a partir das amostras contidas dentro do evento com o maior valor da potência (índice i_{pmax}), ou seja, entre as amostras de índices i_{pre} e i_{pos} . Essa característica de assinatura é extraída da potência ativa, reativa e aparente para os eventos de ativação e desativação, conforme as Equações (25) e (26), respectivamente.

$$p_{\{P,Q,S\}} = \{P, Q, S\}_{Norm}[i_{pmax}] - \{P, Q, S\}_{Norm}[i_{pre}] \quad (25)$$

$$p_{\{P,Q,S\}} = \{P, Q, S\}_{Norm}[i_{pmax}] - \{P, Q, S\}_{Norm}[i_{pos}] \quad (26)$$

Sendo:

- p_P o pico máximo da potência ativa no evento;
- p_Q o pico máximo da potência reativa no evento;
- p_S o pico máximo da potência aparente no evento;

$P_{Norm}[i_{pmax}]$, $Q_{Norm}[i_{pmax}]$ e $S_{Norm}[i_{pmax}]$ amostras de pico máximo das potências ativa, reativa e aparente, respectivamente;

$P_{Norm}[i_{pre}]$, $Q_{Norm}[i_{pre}]$ e $S_{Norm}[i_{pre}]$ amostras de pré-evento das potências ativa, reativa e aparente, respectivamente;

$P_{Norm}[i_{pos}]$, $Q_{Norm}[i_{pos}]$ e $S_{Norm}[i_{pos}]$ amostras de pós-evento das potências ativa, reativa e aparente, respectivamente;

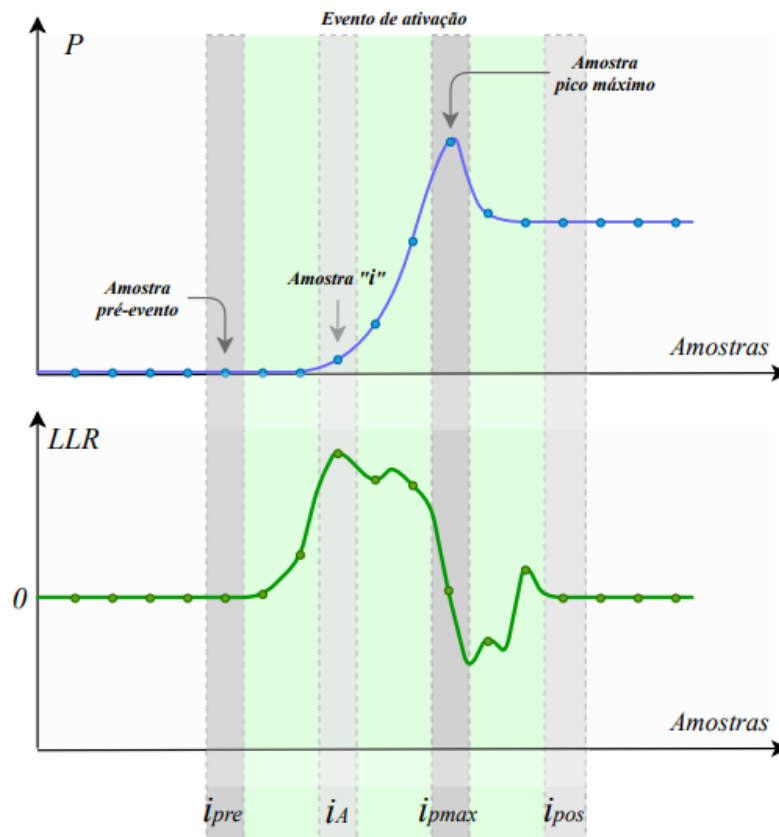
i_{pre} o índice da amostra pré-evento;

i_{pos} o índice da amostra pós-evento;

i_{pmax} o índice da amostra com pico máximo de potência no evento.

Uma abstração para extrair as amostras de pico máximo é mostrado na Figura 16 em um possível evento de ativação detectado.

Figura 16 – Extração da amostra de pico máximo.



Fonte: Autor.

A última característica que compõe a assinatura elétrica assumida nesta metodologia é a distorção do sinal de potência, onde representa o grau de

assimetria dos dados de transição em relação à média. Esta característica é calculada com as amostras das potências que compreendem toda a vizinhança de um evento, identificado por i_A ou i_D , onde as amostras do sinal $LLR[i]$ são diferentes de zero.

Um ponto importante no resultado da distorção é sua capacidade de escalabilidade, portanto eventos semelhantes com intensidades diferentes resultam em distorções semelhantes.

A partir do conjunto de amostras da potência, contidas entre os índices de pré e pós-evento, é realizado o cálculo da distorção provocado pelo evento detectado, como segue as Equações (27), (28) e (29).

$$\mu_{\{P,Q,S\}} = \frac{1}{i_{pos} - i_{pre} - 1} \sum_{n=i_{pre}+1}^{i_{pos}-1} \{P, Q, S\}_{Norm}[n] \quad (27)$$

$$\sigma_{\{P,Q,S\}} = \sqrt{\frac{1}{i_{pos} - i_{pre} - 2} \sum_{n=i_{pre}+1}^{i_{pos}-1} (\mu_{\{P,Q,S\}} - \{P, Q, S\}_{Norm}[n])^2} \quad (28)$$

$$g_{\{P,Q,S\}} = \frac{i_{pos} - i_{pre} - 1}{(i_{pos} - i_{pre} - 2) \cdot (i_{pos} - i_{pre} - 3)} \sum_{n=i_{pre}+1}^{i_{pos}-1} \left(\frac{\{P, Q, S\}_{Norm}[n] - \mu_{\{P,Q,S\}}}{\sigma_{\{P,Q,S\}}} \right)^3 \quad (29)$$

Sendo:

g_P a distorção da potência ativa, entre os índices i_{pre} e i_{pos} da vizinhança do evento;

g_Q a distorção da potência reativa, entre os índices i_{pre} e i_{pos} da vizinhança do evento;

g_S a distorção da potência aparente, entre os índices i_{pre} e i_{pos} da vizinhança do evento;

μ_P , μ_Q e μ_S a média aritmética das amostras da potência ativa, reativa e aparente, respectivamente, entre os índices i_{pre} e i_{pos} da vizinhança do evento;

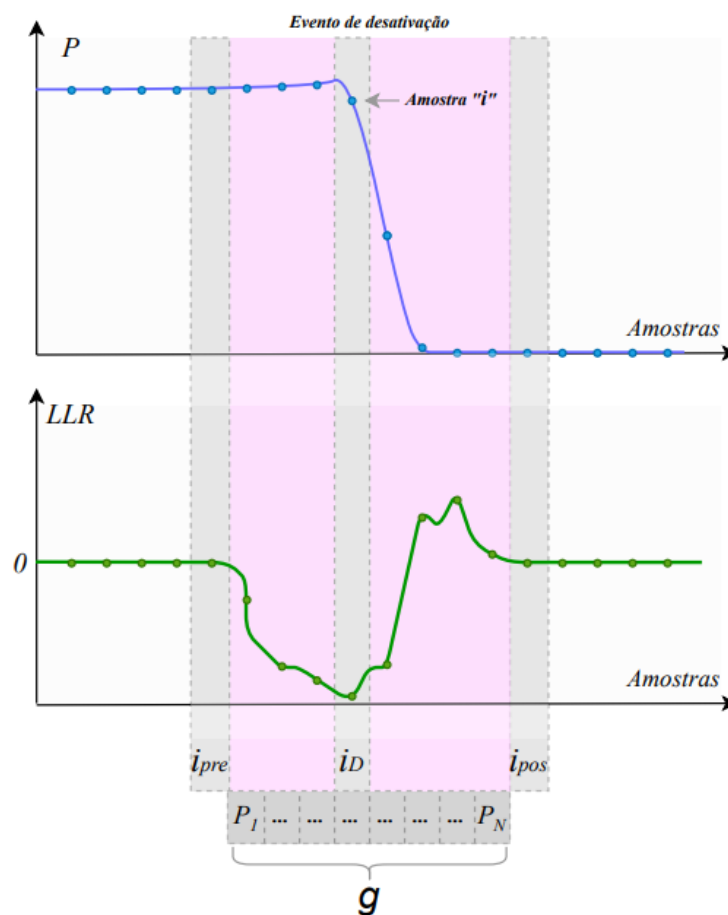
σ_P , σ_Q e σ_S o desvio padrão das amostras da potência ativa, reativa e aparente, respectivamente, entre os índices i_{pre} e i_{pos} da vizinhança do evento;

i_{pre} o índice da amostra pré-evento;

i_{pos} o índice da amostra pós-evento.

A partir da Figura 17 é mostrado a extração do conjunto de amostras do sinal da potência para calcular a distorção de um possível evento de desativação de carga. As amostras estão contidas entre o primeiro índice após o pré-evento e o último índice antes do pós-evento.

Figura 17 – Extração das amostras da potência para calcular a distorção do evento.



Fonte: Autor.

3.3.2 Seleção dos dados de treinamento

A classificação de dados está intrinsecamente relacionada com as classes que se deseja classificar e os dados que são utilizados para o treinamento dos classificadores, independentemente de sua natureza. Logo, para sistemas *NILM* de categoria baseada em eventos e aprendizado supervisionado, é necessário a pré-

seleção de dados das classes que se pretende desagregar, ou seja, os equipamentos monitorados.

Os dados para realizar o treinamento dos classificadores nada mais são que as assinaturas elétricas das cargas obtidas no processo anterior, formando uma base que contém várias assinaturas de cada equipamento para futura classificação.

A criação da base de assinaturas, pode ser realizada utilizando os sinais desagregados (individual de cada equipamento) do conjunto de dados utilizado. Como já mencionado anteriormente, o conjunto de dados pode ser uma coleção pública, extraído do ambiente real ou até mesmo dados sintéticos, desde que possuam os sinais desagregados e rotulagem.

Primeiramente, as classes que serão monitoradas do conjunto de dados são selecionadas (por exemplo: refrigerador, micro-ondas e outras) e, a partir dos sinais individuais destas cargas, é realizada a detecção dos eventos.

Após a detecção dos eventos, sobre os sinais desagregados de cada equipamento, são extraídas somente as características dos eventos de ativação ou desativação que correspondem ao *ground truth* (verdadeiros positivos - *TP*). Dessa forma é possível obter uma base de assinaturas confiáveis de cada classe de equipamento para realizar o processo de treinamento, evitando a introdução de erros nos algoritmos de classificação por adição de assinaturas desconhecidas.

A base é estruturada pelas nove variáveis das características que compõem as assinaturas e suas classes correspondentes através de uma lista ou tabela, conforme o cabeçalho (*header*) do Quadro 4, que dita a formação da base.

Quadro 4 – Estrutura dos dados da base de assinaturas.

<i>Classe</i>	δP	δQ	δS	p_P	p_Q	p_S	g_P	g_Q	g_S
---------------	------------	------------	------------	-------	-------	-------	-------	-------	-------

Fonte: Autor.

Com uma base de assinaturas contendo várias características de cada classe dos equipamentos, os dados são selecionados e processados para o treinamento dos classificadores, sendo a próxima etapa na metodologia.

3.4 ETAPA IV

Nesta etapa são realizados os treinamentos dos classificadores, utilizando ferramentas para otimizar a variação dos principais parâmetros dos algoritmos. Em seguida, são aplicadas as métricas de avaliação em cada classificador para verificar os resultados com dados de testes.

3.4.1 Treinamento dos classificadores

O treinamento dos classificadores é um procedimento essencial para determinar o ajuste dos hiperparâmetros (variáveis ajustáveis nos modelos de predições) e obter o melhor desempenho no aprendizado de cada método. Como forma de acelerar a otimização das variáveis de cada classificador, esta metodologia utiliza uma ferramenta que auxilia na definição dos hiperparâmetros.

A ferramenta utilizada, *Orange Data Mining*, é uma plataforma de código aberto (*open source*) que permite criar e visualizar todo fluxo de trabalho de projetos voltados ao aprendizado de máquina (*ML*) e mineração de dados, sem a necessidade de código ou implementação dos algoritmos, possuindo uma interface totalmente gráfica. Com ela é possível treinar e avaliar várias técnicas de inferências com os modelos já disponíveis em sua biblioteca, tornando a análise exploratória rápida e interativa (DEMSAR et al., 2013). Portanto, esta ferramenta auxilia e acelera o processo de ajuste dos hiperparâmetros de cada classificador, além de servir para treinar e avaliar o sistema com os dados do conjunto público e também a validação da etapa com dados do ambiente real.

Como já mencionado, nesta metodologia foram utilizados três diferentes classificadores, sendo eles: *KNN*, *SVM* e *ANN*. Esta seleção segue a premissa de serem os algoritmos mais explorados na literatura *NILM* que requerem menor processamento em sistemas ou dispositivos de monitoramento para ambiente real (embora retornem bons resultados de acordo com as publicações exploradas no referencial). Além disso, a utilização de mais de um classificador possibilita verificar a ocorrência de possíveis desconformidades e garante a robustez dos algoritmos de inferência.

Neste trabalho cada classificador foi treinado com 70% dos dados da base de assinaturas, criada na Etapa III, para estabelecer as principais variáveis de ajuste dos métodos de predição com o *Orange Data Mining*. O restante dos dados (30%) foi utilizado para o processo seguinte em testes e avaliações através das métricas.

A definição dos principais hiperparâmetros para os classificadores utilizados são especificados no Quadro 5:

Quadro 5 – Hiperparâmetros definidos com auxílio da ferramenta.

Classificador	Hiperparâmetros avaliados
KNN	O número de vizinhos (K) e a métrica de distância.
SVM	O custo (C), o tipo de <i>Kernel</i> e <i>Gamma</i> , se aplicável.
ANN	Número de neurônios na camada de entrada (I), oculta (H) e saída (O), a função de ativação (<i>Activation Function</i>) e o otimizador (<i>Solver</i>).

Fonte: Autor.

Com os hiperparâmetros ajustados, considerando o melhor desempenho, os classificadores são submetidos ao treinamento na plataforma *Orange Data Mining* com os dados da base de assinaturas disponíveis e, posteriormente, para avaliação dos modelos.

3.4.2 Avaliação dos classificadores

Após a definição dos parâmetros e processo de treinamento, para verificar se os classificadores são capazes de realizar a identificação dos eventos de ativação e desativação das cargas, foi utilizada no treinamento 30% da base de assinaturas para avaliação dos classificadores. Como teste alternativo, poderia ser construído um novo conjunto de assinaturas, contendo as mesmas classes utilizadas no treinamento.

Os resultados de desempenho dos classificadores são quantificados por meio da métrica de avaliação *Accuracy* nas versões *micro-averaging* e *macro-averaging*. Esta métrica informa a fração de instâncias classificadas corretamente em relação a todas as classes inferidas, calculada de acordo com a Equação (8).

Outra métrica utilizada nesta metodologia foi a exposição das inferências através da matriz de confusão. Nela são colocadas as instâncias classificadas, para

que sejam vistos os acertos e os erros, ou ainda confusões, ocorridas nas classificações das amostras de avaliação, conforme descrito na seção 2.4.

A ferramenta *Orange Data Mining* é capaz de fornecer as métricas definidas para análise dos resultados (*Accuracy* e matriz de confusão), além das métricas *Precision*, *Recall* e *F1-Score* no desempenho de classificadores. Portanto, estes valores podem ser extraídos e visualizados diretamente na plataforma como meio de avaliação para cada classificador utilizado, sendo possível verificar e comparar, caso ocorra, alguma divergência nos desempenhos.

As métricas levam em consideração cada classe informada na etapa de treinamento (os estados de ativação e desativação dos equipamentos monitorados). Já os resultados são baseados nos dados do conjunto de teste devidamente rotulados para contabilizar as instâncias.

3.5 ETAPA V

A última etapa da metodologia desenvolvida consiste na aplicação do sistema *NILM* em um ambiente real para verificar e validar a desagregação das cargas. Após a implementação dos processos e treinamento dos classificadores com um conjunto de dados, o sistema é colocado à prova e apresenta-se os resultados de um cenário prático residencial.

3.5.1 Avaliação em ambiente real

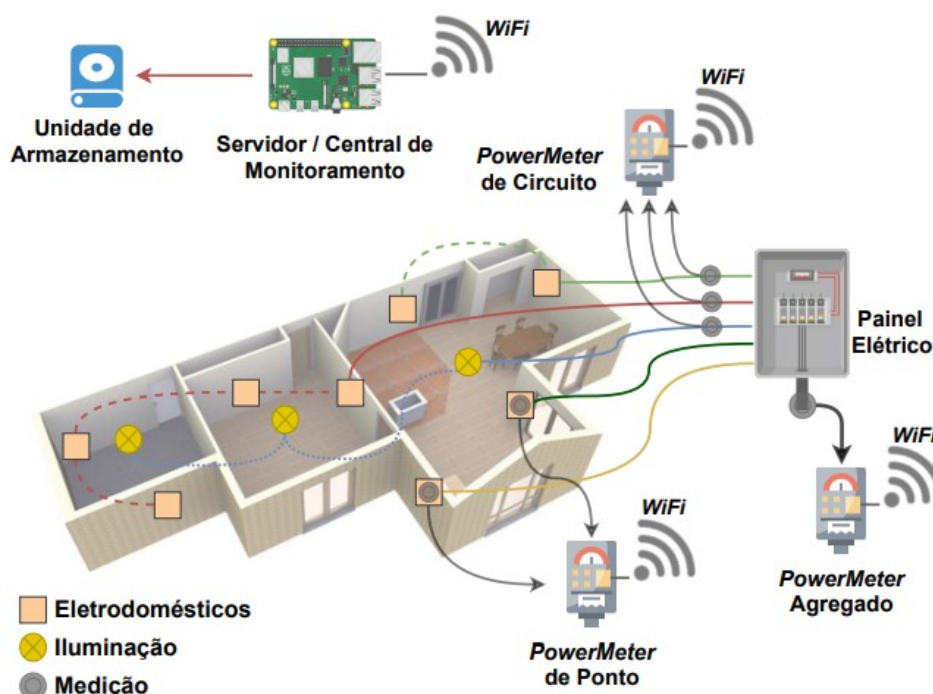
O teste do monitoramento *NILM* com dados de ambiente real é válido, uma vez que a utilização de um conjunto de dados público é empregado para o desenvolvimento, treinamento e avaliação do fluxo de monitoramento não intrusivo de cargas. Os dados de uma coleção de sinais nem sempre correspondem à realidade do cenário residencial, principalmente quando o conjunto de dados utilizado no projeto não são extraídos do mesmo local da implementação do monitoramento. Vale salientar que gerar conjuntos de sinais dos ambientes que se deseja monitorar é uma tarefa dispendiosa e muitas vezes inviável.

Portanto, nesta etapa da metodologia, o sistema é instalado em ambiente real para registrar os dados e verificar o seu desempenho. Além disso, os sinais

armazenados podem vir a ser utilizados para gerar um conjunto de dados local e serem disponibilizados publicamente.

A Figura 18 é uma representação abstrata do sistema de monitoramento distribuído na residência, que pode ser visto como uma composição de sensores eletrônicos de aquisição dos sinais (aqui nomeados de *PowerMeters*) e um servidor de gerenciamento para monitorar e armazenar os dados. Desta forma é possível desagregar as cargas e obter os sinais de cada equipamento que se deseja monitorar, para posterior validação dos resultados apresentados.

Figura 18 – Sistema de monitoramento sugerido em ambiente real.



Fonte: Autor.

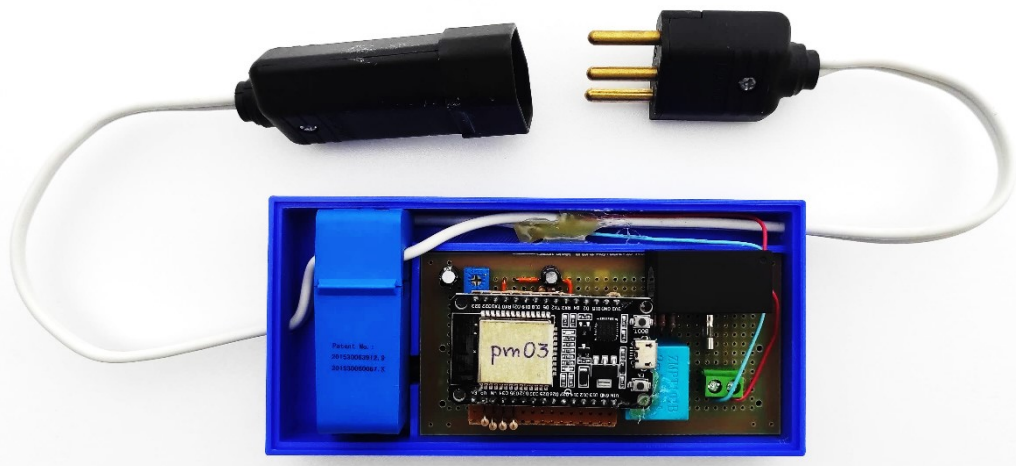
A taxonomia dos medidores, geralmente, é dividida em três grupos: (I) *PowerMeter* Agregado - usado para coletar o sinal agregado no painel elétrico de distribuição da residência e a partir dele desagregar as cargas; (II) *PowerMeter* de Circuito - usado para coletar os sinais de um grupo específico de eletrodomésticos, como o sistema de iluminação, por exemplo; e (III) *PowerMeter* de Ponto - que coleta os sinais desagregados por equipamento individualmente para gerar o *ground truth* do cenário (ANDERSON et al., 2012a). Neste trabalho são dispostos 1 sensor de medição agregada (*PowerMeter* Agregado) e 5 sensores distribuídos para eletrodomésticos individuais (*PowerMeter* de Ponto).

Os sensores utilizados são medidores desenvolvidos em trabalhos no âmbito deste projeto, sendo que alguns resultados estão publicados em Romancini (2020) e Romancini et al. (2022). Os dispositivos anteriormente desenvolvidos foram adaptados para realizar a leitura da tensão e corrente elétrica, potência ativa, reativa e aparente das cargas na frequência do ciclo de fase (60 Hz) com carimbo de hora. Estes dados são enviados para a central de monitoramento através de uma rede de comunicação sem fio (*WiFi*).

A central de monitoramento é implementada em um pequeno computador *Raspberry Pi 4*. Este servidor tem a função de administrar a infraestrutura de rede de comunicação dos medidores distribuídos no local e armazenar os dados enviados por cada *PowerMeter* em uma unidade de memória não volátil (para posterior recuperação dos dados registrados). Esses dados podem ser usados para validar as classificações dos equipamentos, gerar um conjunto de dados locais, uma base de assinaturas e treinar o sistema.

A Figura 19 traz uma imagem real do modelo utilizado para monitorar os dados dos equipamentos da residência.

Figura 19 – Medidor utilizado no monitoramento para o ambiente real.



Fonte: Autor.

Para os testes com dados do cenário residencial, os *PowerMeters* de Ponto são instalados apenas nas classes de equipamentos que foram treinadas no desenvolvimento do sistema. Uma vez que a desagregação não seria reconhecida em aparelhos que não participaram do processo de aprendizado supervisionado.

Além disso, o *PowerMeter* de Circuito também não é utilizado, pois não agrega informação adicional quando já são extraídos os sinais desagregados e agregados.

Dessa forma, os eletrodomésticos monitorados pelos sensores individualmente no ambiente real devem ser de classe semelhante aos disponíveis no conjunto público de dados. O conjunto utilizado *FIRE*D é composto por 35 tipos de aparelhos medidos individualmente, dentre estes são selecionados equipamentos que correspondem aos contidos no cenário prático. Através do Quadro 6 foram listados os eletrodomésticos existentes no conjunto de dados *FIRE*D e no ambiente local de análise.

Quadro 6 – Equipamentos presentes no conjunto *FIRE*D e ambiente local.

Conjunto <i>FIRE</i> D ⁴	Ambiente local monitorado
Lâmpada de calor para bebês	Lâmpada com reator
Lâmpada halógena de cozinha	Lâmpada resistiva
Refrigerador	Lâmpada de LED
Chaleira elétrica	Refrigerador
Torradeira	Chaleira elétrica
Máquina de café expresso	Torradeira
Moedor de café	Forno elétrico
Forno elétrico	Micro-ondas
Máquina de waffles	Batedeira
Batedeira	Ferro de passar roupa
Mixer de cozinha	Aspirador de pó
Exaustor de fumaça	Máquina de lavar roupa
Ferro de passar roupa	Secador de cabelo
Aspirador de pó	Chuveiro elétrico
Window Cleaner	Televisor
Máquina de costura	Notebook
Secador de cabelo	Roteador
Escova elétrica	Carregador de Smartphone
Barbeador elétrico	
Televisor	
Sistema Hifi	
Subwoofer	
Smart Speaker	
Notebook	
Computador de escritório	
Tela de PC	
Mídia de PC	
Alto-falante de PC	
PC de gravação	
Impressora	
Roteador	
Access point	
Telefone	
Carregador de Smartphone	
Carregador de bateria	

Fonte: Völker et al. (2020b) e autor.

⁴ Conjunto de dados *FIRE*D extraído de: *Fully-labeled high-frequency Electricity Disaggregation*, https://github.com/voelkerb/FIREdataset_helper

Com a disposição e limitação de 5 medidores de ponto, os aparelhos escolhidos para detectar, treinar, avaliar o sistema e monitorar no ambiente local foram: refrigerador, chaleira elétrica, aspirador de pó, secador de cabelo e televisor. Estes aparelhos foram selecionados por corresponderem em ambos os conjuntos e possuírem representatividade no ambiente residencial.

Uma imagem contendo os sensores instalados no ambiente real e a central de monitoramento é apresentada na Figura 20.

Figura 20 – Sensores e central de monitoramento instalados no ambiente



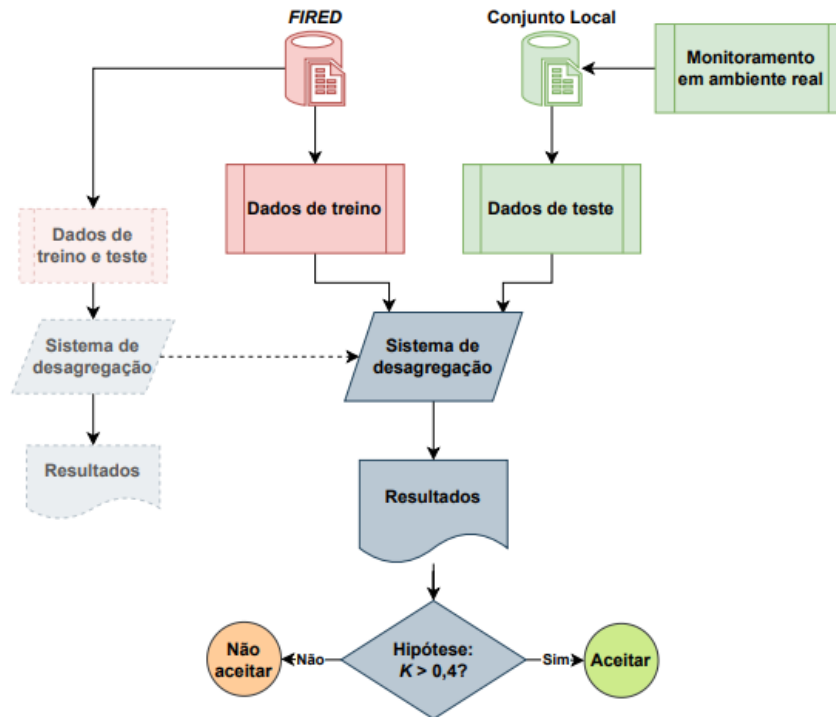
Fonte: Autor.

Legenda: **(1)** Refrigerador; **(2)** Chaleira; **(3)** Secador de cabelo; **(4)** Televisor; **(5)** Aspirador de pó; **(6)** PowerMeter Agregado; **(7)** Central de monitoramento.

Por fim, após o sistema implementado, treinado e extraídos os dados na residência, são verificados os resultados de desagregação utilizando as métricas de avaliação *Accuracy*, matriz de confusão e o índice *Kappa*, Equação (9) da seção 2.4.

As inferências do sistema são comparadas com o *ground truth* produzido e armazenado no intervalo de tempo em que o monitoramento esteve ativo no ambiente, conforme o processo mostrado na Figura 21.

Figura 21 – Processo de verificação do sistema com dados do ambiente real.



Fonte: Autor.

Os dados do conjunto *FIRED* foram utilizados para treinar e avaliar o sistema proposto nesta metodologia. Em seguida, com os dados de monitoramento do ambiente real local, o sistema foi colocado à prova para verificar se conjuntos de dados públicos são suficientes para desagregar cargas de um cenário residencial distinto. Considerando a hipótese para o valor de *Kappa* maior que 0.40 (rótulo Moderado - Tabela 1), conforme sugerido por Rehman et al. (2021) e Sakiyama et al. (2008), é possível aceitar ou não aceitar a premissa e responder a pergunta de pesquisa.

4 RESULTADOS

Neste capítulo são apresentados os resultados obtidos com a implementação do sistema de desagregação, utilizando conjunto público de sinais elétricos, e a validação com os sinais elétricos extraídos do ambiente real local.

4.1 CONJUNTOS DE DADOS UTILIZADOS

Com o objetivo de avaliar a metodologia proposta foi realizada a aquisição dos sinais para a implementação, treinamento e testes do sistema de desagregação, utilizando os dados de monitoramento disponibilizados pelo banco de dados público *FIRE*D da Alemanha. Sendo os principais motivos para a escolha desse conjunto: a disponibilidade das grandezas necessárias na frequência de ciclo da rede elétrica, o registro dos sinais elétricos individuais dos eletrodomésticos e a presença de rotulagem no *ground truth* dos eventos monitorados.

O período de análise no monitoramento do conjunto *FIRE*D, utilizado no levantamento para extração dos eventos, foi compreendido entre 00:00 h do dia 01/07/2020 até às 23:59 h do dia 31/08/2020, cerca de 62 dias. Foram realizadas as aquisições dos sinais de medição agregada (*smartmeter001*) e de cinco medidores individuais que correspondem aos aparelhos já propostos: aspirador de pó (*powermeter11*), chaleira elétrica (*powermeter16*), refrigerador (*powermeter09*), secador de cabelo (*powermeter17*) e televisor (*powermeter23*). Os dados dos medidores individuais contêm as grandezas da potência ativa, reativa e aparente, na frequência de 50 Hz, enquanto a medição agregada contêm os sinais das potências e da tensão elétrica necessárias para essa metodologia.

Para realizar a validação do sistema de desagregação com os dados do ambiente real local, foram utilizados os sinais elétricos registrados de um apartamento na região de Criciúma/SC. A residência monitorada possui, aproximadamente, 45 metros quadrados (sala, cozinha, sacada, banheiro e 2 quartos) e habitada por dois moradores. Os dados registrados pelos medidores locais compreendem o período entre 00:00 h do dia 05/07/2022 até às 23:59 h do dia 05/08/2022, totalizando 31 dias. Os equipamentos monitorados individualmente corresponderam com as mesmas cinco classes de aparelhos utilizadas para

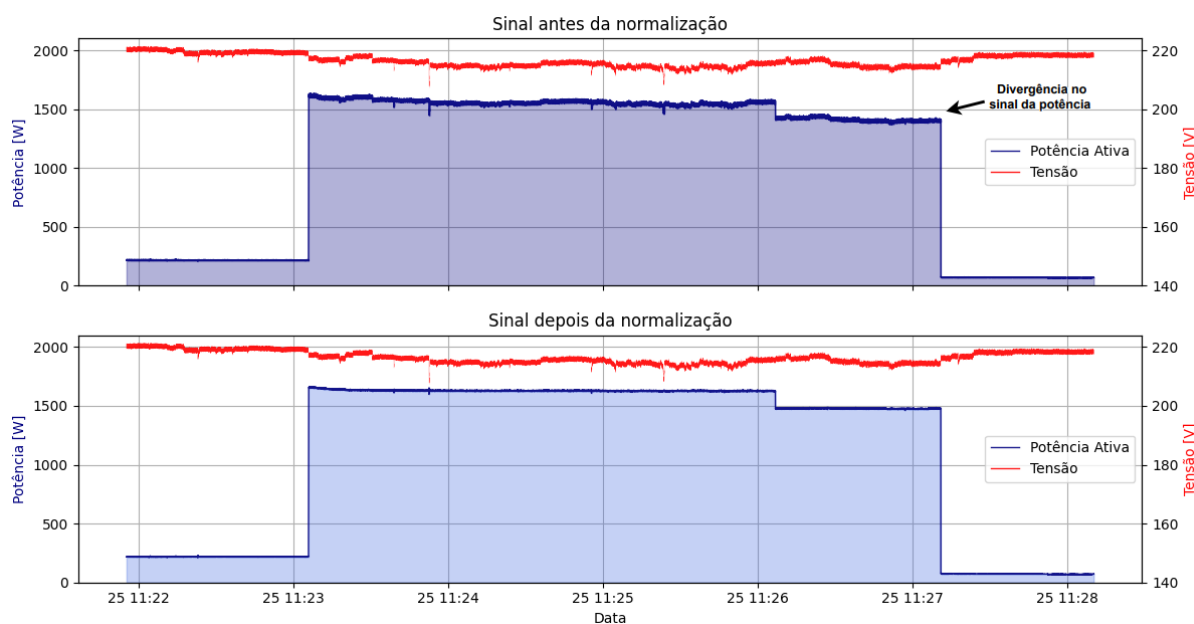
implementação do sistema com o conjunto público *FIRE*D, conforme exposto na subseção 3.5.1 da metodologia.

As grandezas monitoradas pelos *PowerMeters* de Ponto e Agregado do ambiente real foram as potências ativa, reativa e aparente, além dos valores de tensão e corrente elétrica na frequência de 60 Hz. Devido à escassez de conjuntos de sinais elétricos com caráter público, para pesquisadores *NILM* do Brasil, os registros obtidos do ambiente real local para este trabalho foram disponibilizados em plataforma de acesso livre e chamado de Sinais Elétricos de Cargas Residenciais – Brasil (SECRBR). O acesso ao conjunto de sinais SECRBR pode ser realizado através do endereço eletrônico <https://github.com/elielmarcos/SECRBR>.

4.2 TRATAMENTO DOS DADOS

Utilizando a Equação (11), subseção 3.1.2, foi realizada a normalização dos sinais das potências para todo o intervalo de estudo considerando a tensão nominal de 220 V, conforme demonstrado nos gráficos da Figura 22. Eles representam a ação de normalização em um espaço de tempo no dia 25/07/2020 do conjunto *FIRE*D.

Figura 22 – Normalização no sinal da potência.



Fonte: Autor.

Os gráficos apresentam a série temporal do sinal da tensão elétrica (sinais na cor vermelha) e os sinais da potência ativa antes e depois do processo de normalização (sinal azul do gráfico superior e inferior). Os sinais das potências reativa e aparente também sofreram tratamento de normalização em ambos os conjuntos *FIRE*D e SECRBR.

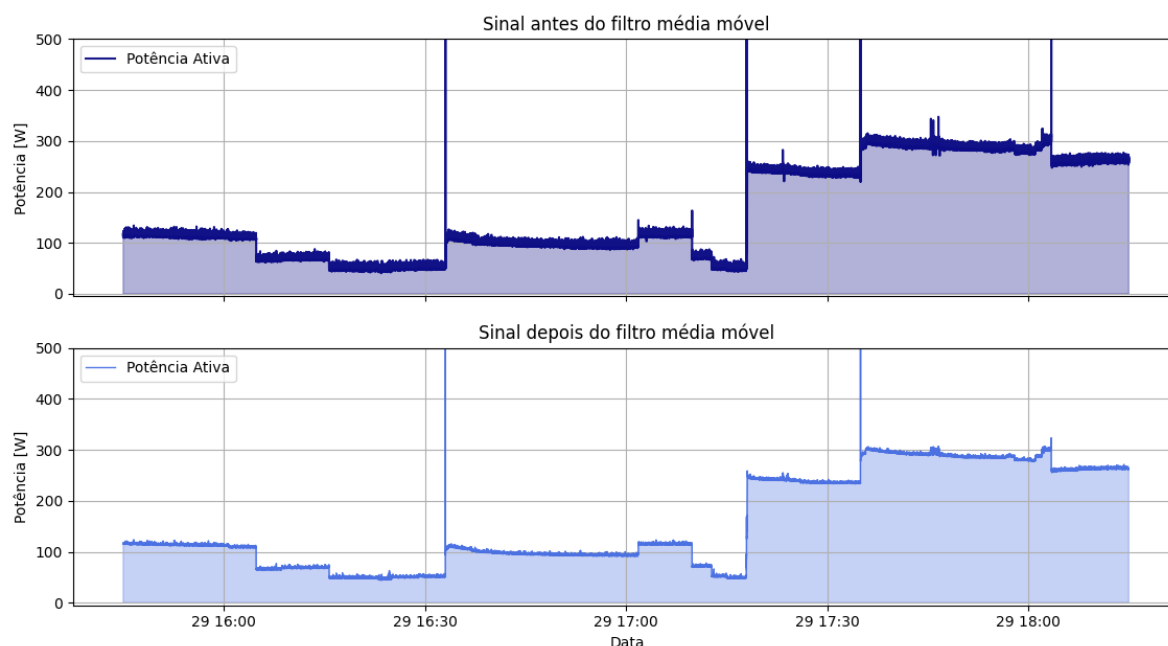
Ao analisar os gráficos, é possível verificar que as oscilações da tensão elétrica no monitoramento interferem significativamente nos sinais elétricos da potência, alterando o valor nominal de consumo dos aparelhos ao longo da série temporal. E inclusive nos momentos de transições, onde são as regiões de extração das características para as assinaturas dos eletrodomésticos.

Nota-se que próximo da região de um evento (horário 11:27 h) a potência média no sinal depois de normalizada está em torno de 1500 W, enquanto no mesmo horário do sinal não normalizado a potência média é relativamente inferior devido às oscilações da tensão. Dessa forma, após a normalização dos sinais, é possível extrair dados mais fidedignos dos aparelhos e com menores flutuações das grandezas elétricas.

As oscilações da tensão elétrica no ambiente residencial podem ser ocasionadas pela própria geração, distribuição e principalmente pelos eletrodomésticos que apresentam maior demanda de consumo no local, como cargas resistivas de alta potência.

Após o processo de normalização dos dados, foi realizada a filtragem dos sinais da potência ativa, de acordo com a Equação (12) da subseção 3.1.2. A Figura 23 apresenta um intervalo de tempo da medição agregada que teve o sinal filtrado para reduzir a presença de ruídos.

Figura 23 – Aplicação do filtro média móvel no sinal da potência.



Fonte: Autor.

O processo de filtragem do sinal da potência é realizado através de um filtro média móvel no conjunto *FIRE*D e no conjunto do ambiente real - SECRBR. Foram utilizadas 9 amostras para cada índice da série temporal na implementação do filtro digital, parâmetro N da Equação (12). Este parâmetro foi utilizado por reduzir significativamente os ruídos do sinal e apresentar um atraso, ou deslocamento, praticamente desprezível para implementações de tempo real.

O tratamento do sinal elétrico pelo filtro foi empregado apenas na potência ativa, pois o detector de eventos desta metodologia investiga os possíveis eventos com base na potência real. Assim, a filtragem do sinal antes de ser aplicado no algoritmo detector probabilístico reduz a quantidade de detecções falsas, ou seja, falsos positivos (*FP*).

Através dos gráficos da Figura 23, nota-se a atenuação de ruídos no sinal da potência ativa, sendo o gráfico superior o sinal antes do processo de filtragem e o gráfico inferior o sinal depois do filtro média móvel. Após a aquisição e tratamento dos sinais elétricos realizados em ambos conjuntos de dados, condizente com a primeira etapa da metodologia, foi executada a segunda etapa com a aplicação do algoritmo detector de eventos.

4.3 APLICAÇÃO E AVALIAÇÃO DO LLR

O algoritmo para detecção probabilística de eventos *LLR*, proposto na subseção 3.2.1 deste trabalho para identificar os eventos de ativação e desativação dos aparelhos, foi implementado em linguagem *Python*. O detector *LLR* foi testado por uma série de vezes nos vários sinais da potência de cada aparelho utilizados nas análises, até se obter um desempenho na detecção que correspondesse à maior quantidade de acertos e a menor quantidade de falsos eventos.

Podendo haver uma variedade infinita de combinações dos parâmetros do detector, os valores utilizados na implementação obtiveram um bom desempenho na detecção para os cinco aparelhos monitorados, sendo eles listados na Tabela 3:

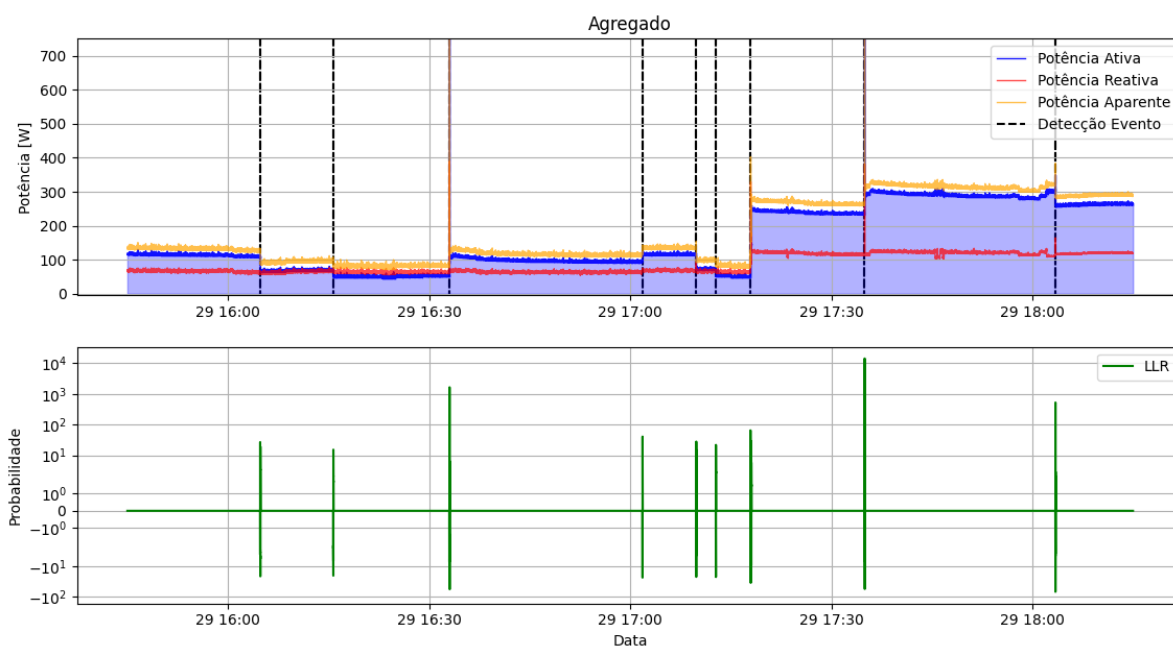
Tabela 3 – Parâmetros de ajuste utilizados na implementação do *LLR*.

Parâmetro	Valor
dw	4,5 s
w_{pre}	2,0 s
w_{pos}	2,5 s
$thres_{min}$	5 W
m	0,05
min_{Dist}	3,0 s

Fonte: Autor.

Os valores dos parâmetros em unidades de tempo (s) são convertidos em quantidades de amostras pelo algoritmo *LLR*, realizando a divisão do valor do parâmetro pela frequência de amostragem do conjunto de dados. Dessa forma são preenchidas as janelas deslizantes, de acordo com o número de amostras necessárias da potência.

Após o ajuste dos parâmetros do detector, pode ser visto através dos gráficos da Figura 24, o algoritmo probabilístico em funcionamento sobre o sinal agregado da potência ativa. Realizando uma inspeção visual é possível observar uma ótima eficiência do *LLR* na identificação de degraus significativos das potências. Ao mesmo tempo, o detector ignorou pequenas oscilações e ruídos não mitigados no processo de filtragem que poderiam ser identificados como falsos eventos.

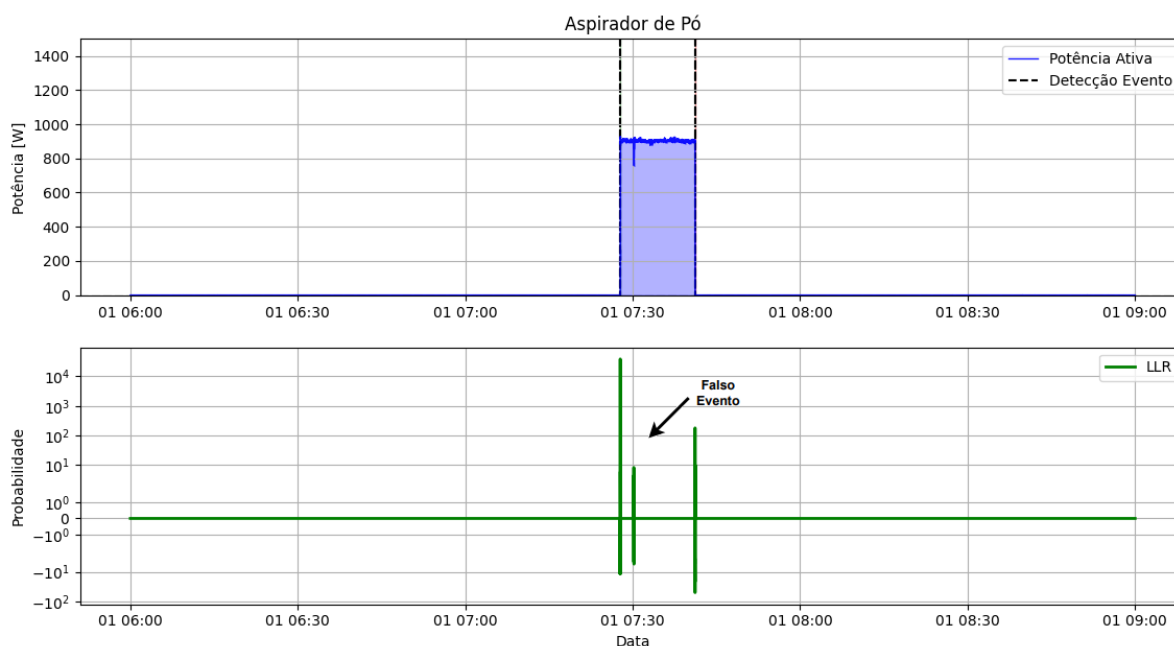
Figura 24 – Algoritmo *LLR* detectando eventos no sinal agregado.

Fonte: Autor.

O gráfico superior mostra os sinais das potências ativa, reativa e aparente, bem como os locais de identificação dos eventos por linhas tracejadas verticais. Enquanto o gráfico inferior ilustra o resultado do sinal probabilístico para detecção de eventos em uma escala logarítmica de probabilidade, onde os valores diferentes de zero indicam as detecções.

Para executar a avaliação do detector de eventos, foram analisados os sinais individuais de cada aparelho e comparados com o *ground truth* disponível no conjunto público *FIRE*D. Dessa forma foi obtido a quantidade exata dos eventos detectados que eram eventos (*TP*), que não eram eventos (*FP*) e eventos que não foram detectados (*FN*). Para isso, foi verificado o *timestamp* do índice da detecção com o *timestamp* do “gabarito”.

A Figura 25 mostra as detecções do aparelho aspirador de pó para contabilizar os eventos no dia 01/08/2020.

Figura 25 – Gráfico de eventos do aspirador de pó no conjunto *FIRE*D.

Fonte: Autor.

Analisando a figura acima é possível perceber um falso evento mensurado pelo resultado da probabilidade no gráfico inferior (sinal verde) próximo às 07:30 h, logo após o evento de ativação do aparelho. Esta falsa detecção foi extraída pelo sinal da potência ativa devido uma variação brusca, mas que não representa um evento real ou mudança de estado do aparelho. Este tipo de problema não teve sua fonte causadora identificada, uma vez que a potência antes e depois do falso evento se manteve praticamente a mesma. Logo, foi assumida uma estratégia para reduzir a detecção de falsos positivos neste tipo de situação. Antes de considerar a ocorrência de um evento, foi adicionado ao algoritmo *LLR* a condição para descartar a detecção caso as diferenças entre as médias das janelas pré e pós-evento (δw) forem inferiores a 10% da maior média.

Com esta condição, observa-se na Figura 25 que o *LLR* mensurou o evento, mas foi descartado e não registrado para ser avaliado. A Tabela 4 apresenta os eventos detectados no gráfico do aspirador de pó com o *ground truth*:

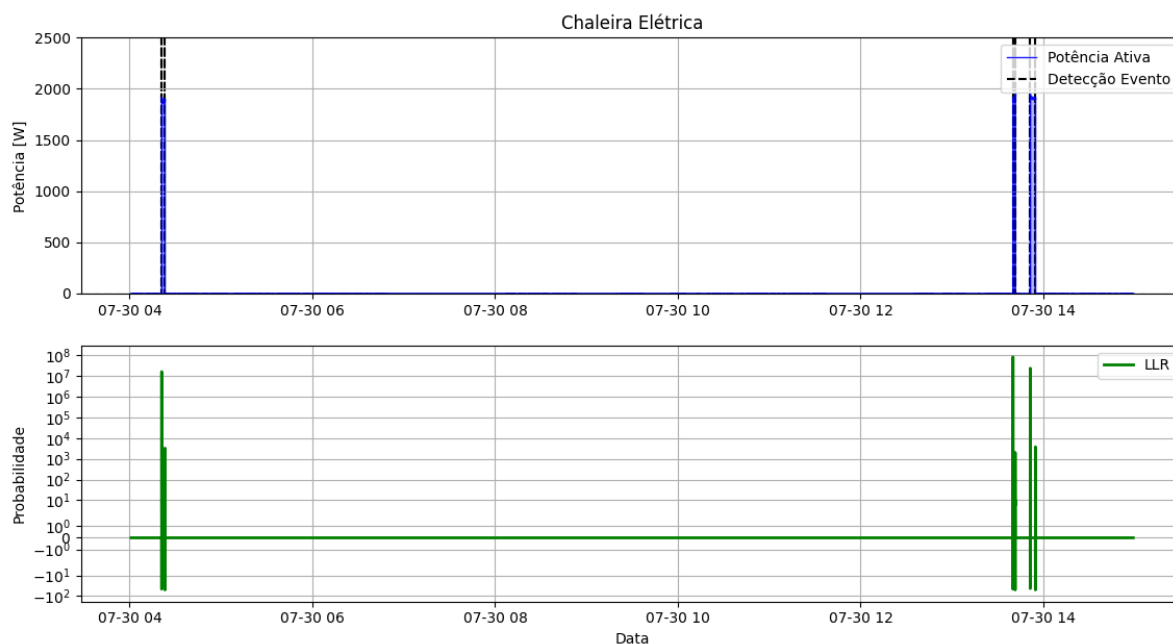
Tabela 4 – Aspirador de pó, eventos detectados e comparados com o *ground truth*.

Aspirador de Pó – 01/08/2020 (<i>FIRE</i> D)			
Eventos <i>LLR</i>		<i>Ground Truth</i>	
<i>Timestamp</i>	Tipo	<i>Timestamp</i>	Tipo
1596277666	Ativação	1596277666.025	vacuum cleaner - on
1596278470	Desativação	1596278470.371	vacuum cleaner - off

Fonte: Autor.

Ao comparar o registro de rotulagem com os eventos de ativação e desativação identificados, é possível inferir quais eventos o detector acertou e contabilizar os resultados para calcular o desempenho. A Figura 26 e a Tabela 5 mostram os eventos da chaleira elétrica detectados na data 30/07/2020 do conjunto *FIRE*D.

Figura 26 – Gráfico de eventos da chaleira elétrica no conjunto *FIRE*D.



Fonte: Autor.

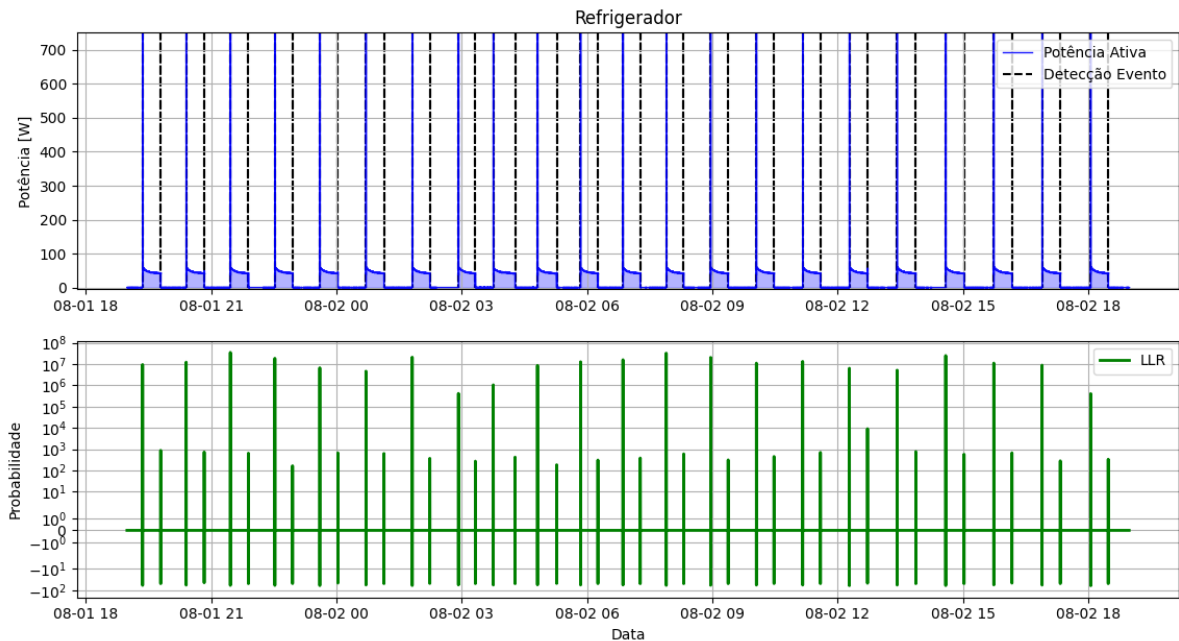
Tabela 5 – Chaleira elétrica, eventos detectados e comparados com o *ground truth*.

Chaleira Elétrica – 30/07/2020 (<i>FIRE</i> D)			
Eventos LLR		<i>Ground Truth</i>	
<i>Timestamp</i>	Tipo	<i>Timestamp</i>	Tipo
1596093666	Ativação	1596093665.68	on
1596093793	Desativação	1596093793.28	off
1596127201	Ativação	1596127201.16	on
1596127289	Desativação	1596127289.28	off
1596127884	Ativação	1596127883.92	on
1596128087	Desativação	1596128086.72	off

Fonte: Autor.

Os eventos identificados corresponderam todos com o *ground truth*. No dia em análise, não houve a ocorrência de falsos eventos e nem eventos não identificados no eletrodoméstico chaleira elétrica. Uma inspeção para mensurar as detecções do refrigerador também são mostradas através da Figura 27 e Tabela 6 abaixo:

Figura 27 – Gráfico de eventos do refrigerador no conjunto *FIRE*D.



Fonte: Autor.

Tabela 6 – Refrigerador, eventos detectados e comparados com o *ground truth*.

(continua)

Refrigerador – 02/08/2020 (<i>FIRE</i> D)			
Eventos LLR		<i>Ground Truth</i>	
Timestamp	Tipo	Timestamp	Tipo
1596320481	Ativação	1596320480.72	Cooling
1596322041	Desativação	1596322040.88	Off
1596324238	Ativação	1596324237.64	Cooling
1596325795	Desativação	1596325795.12	Off
1596328049	Ativação	1596328048.62	Cooling
1596329604	Desativação	1596329603.92	Off
1596331887	Ativação	1596331886.56	Cooling
1596333410	Desativação	1596333409.64	Off
1596335760	Ativação	1596335759.7	Cooling
1596337322	Desativação	1596337322.16	Off
1596339743	Ativação	1596339742.82	Cooling
1596341293	Desativação	1596341292.76	Off
1596343729	Ativação	1596343728.74	Cooling
1596345246	Desativação	1596345245.72	Off
1596347715	Ativação	1596347715.0	Cooling
1596349173	Desativação	1596349172.78	Off
1596350717	Ativação	1596350717.34	Cooling
1596352597	Desativação	1596352596.58	Off
1596354529	Ativação	1596354529.18	Cooling
1596356185	Desativação	1596356184.96	Off
1596358246	Ativação	1596358245.96	Cooling
1596359737	Desativação	1596359736.74	Off
1596361899	Ativação	1596361898.6	Cooling
1596363386	Desativação	1596363386.02	Off
1596365630	Ativação	1596365629.88	Cooling
1596367136	Desativação	1596367135.92	Off
1596369459	Ativação	1596369459.22	Cooling
1596370968	Desativação	1596370968.16	Off
1596373403	Ativação	1596373402.7	Cooling
1596374922	Desativação	1596374921.64	Off
1596377391	Ativação	1596377391.06	Cooling
1596378921	Desativação	1596378921.14	Off
1596381419	Ativação	1596381418.7	Cooling
1596382982	Desativação	1596382981.58	Off
1596385537	Ativação	1596385536.82	Cooling
1596387148	Desativação	1596387147.66	Off
1596389723	Ativação	1596389722.52	Cooling
1596391305	Desativação	1596391304.66	Off

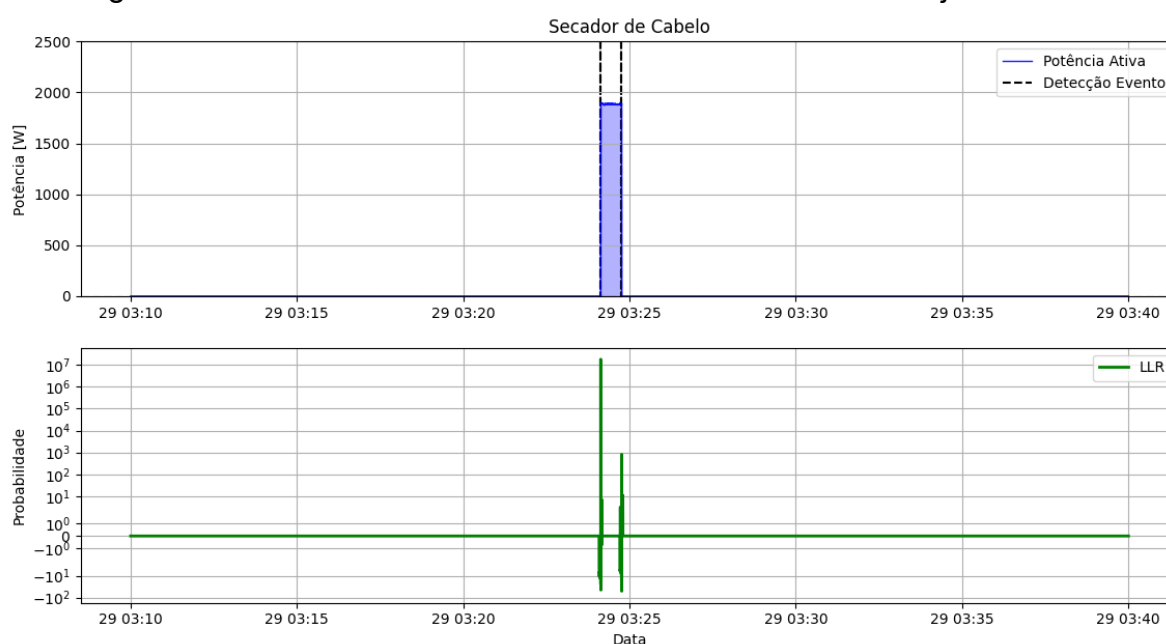
Tabela 6 – Refrigerador, eventos detectados e comparados com o *ground truth*.
(conclusão)

Refrigerador – 02/08/2020 (FIRED)			
Eventos LLR		Ground Truth	
Timestamp	Tipo	Timestamp	Tipo
1596393872	Ativação	1596393871.6	Cooling
1596395434	Desativação	1596395433.7	Off
1596398031	Ativação	1596398030.96	Cooling
1596399602	Desativação	1596399602.4	Off
1596402227	Ativação	1596402227.42	Cooling

Fonte: Autor.

O refrigerador foi o aparelho que teve o maior número de eventos registrados durante o período de monitoramento e, também, o que apresentou a maior quantidade de detecções corretas em relação ao total de eventos. Ao verificar os eventos do *LLR* com a rotulagem, observa-se que todos os eventos foram listados perfeitamente, apesar da grande quantidade de ocorrências em um único dia. A Figura 28 e a Tabela 7 apresentam o gráfico de detecções e a relação de eventos para o aparelho secador de cabelo do conjunto *FIRED*, no dia 29/07/2020.

Figura 28 – Gráfico de eventos do secador de cabelo no conjunto *FIRED*.



Fonte: Autor.

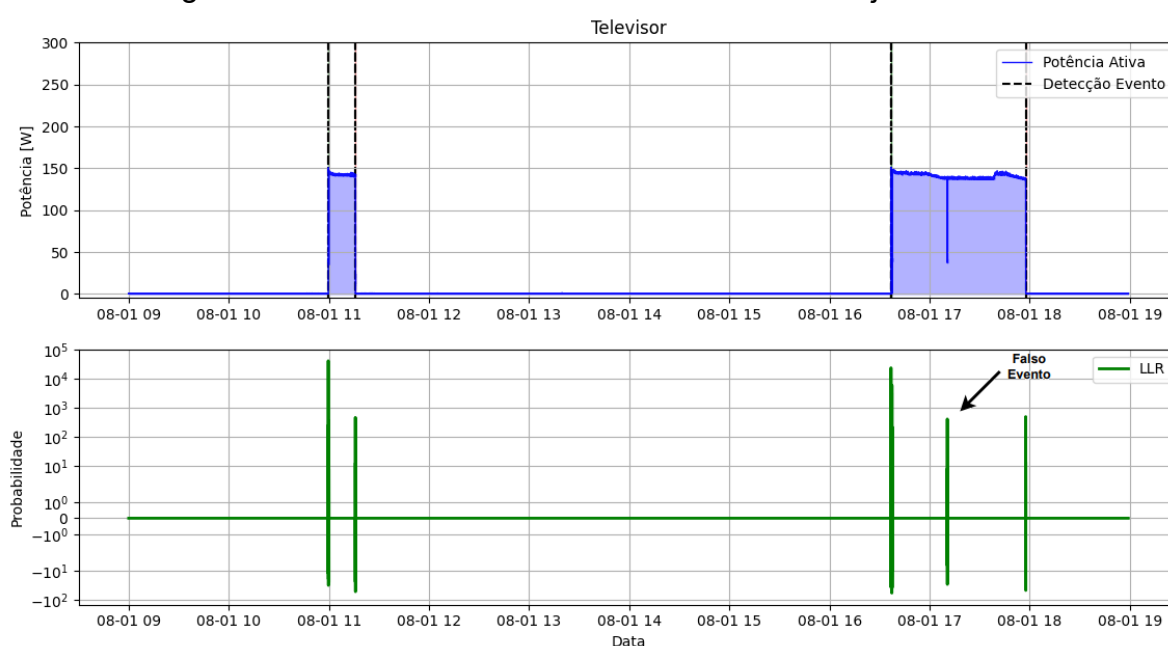
Tabela 7 – Secador de cabelo, eventos detectados e comparados com o *ground truth*.

Secador de Cabelo – 29/07/2020 (FIRED)			
Eventos LLR		Ground Truth	
Timestamp	Tipo	Timestamp	Tipo
1596003848	Ativação	1596003848.04	on
1596003886	Desativação	1596003885.76	off

Fonte: Autor.

Este aparelho foi um dos equipamentos com o menor número de eventos registrados em todo o período de análise, uma vez que seu acionamento é realizado de forma manual, diferente do refrigerador que tem seu funcionamento automático. Os dois eventos listados corresponderam com o arquivo de rotulagem, tanto no horário do evento, quanto no tipo de evento. Por fim, o último eletrodoméstico verificado no conjunto público foi o aparelho televisor. Os dados do dia 01/08/2020 são mostrados na Figura 29 e Tabela 8.

Figura 29 – Gráfico de eventos do televisor no conjunto *FIRE*D.



Fonte: Autor.

Tabela 8 – Televisor, eventos detectados e comparados com o *ground truth*.

Televisor – 01/08/2020 (<i>FIRE</i> D)			
Eventos LLR		<i>Ground Truth</i>	
<i>Timestamp</i>	Tipo	<i>Timestamp</i>	Tipo
1596290376	Ativação	1596290376.32	turning on
1596290383	Ativação	1596290383.1	on
1596291351	Desativação	1596291350.983	turning off
1596291355	Desativação	1596291355.12	off
1596310610	Ativação	1596310609.349	turning on
-	-	1596310611.08	on
1596310620	Desativação	1596310619.82	off
1596310628	Ativação	1596310628.34	on
1596310639	Desativação	1596310638.659	turning off
-	-	1596310640.52	off
1596310645	Ativação	1596310645.18	turning on
1596310652	Ativação	1596310651.7	on
1596315454	Desativação	1596315454.34	turning off

Fonte: Autor.

O televisor foi o equipamento com o maior número de falso eventos (*FP*) e não detecções (*FN*), em relação ao seu total de eventos. Isso se deu pelas próprias características do aparelho, onde tiveram vários períodos ruidosos em seu funcionamento não sendo eliminados na etapa de tratamento dos sinais e, também, por apresentar dois níveis de eventos para ativação e desativação. Essa característica ocasionou, em alguns momentos, a não identificação de eventos consecutivos com curto espaço de tempo, conforme apresentados na tabela de comparação com o *ground truth*.

Depois de listados os eventos do conjunto *FIRE*D, para registrar as detecções sobre os dados do SECRBR foram analisados os mesmos cinco tipos de aparelhos durante os 31 dias. As anotações dos eventos foram realizadas por inspeção visual, através dos sinais adquiridos pelos medidores no ambiente real. Visto que o conjunto SECRBR não possui arquivo de rotulagem, diferente do conjunto *FIRE*D.

Ao final do processo de detecção de eventos dos eletrodomésticos, foram contabilizadas as instâncias de cada aparelho e o valor total de eventos em todo o período de monitoramento. As contagens foram realizadas tanto para o conjunto de dados *FIRE*D, quanto para SECRBR. A Tabela 9 informa os valores de *TP*, *FP* e *FN* do detector *LLR*.

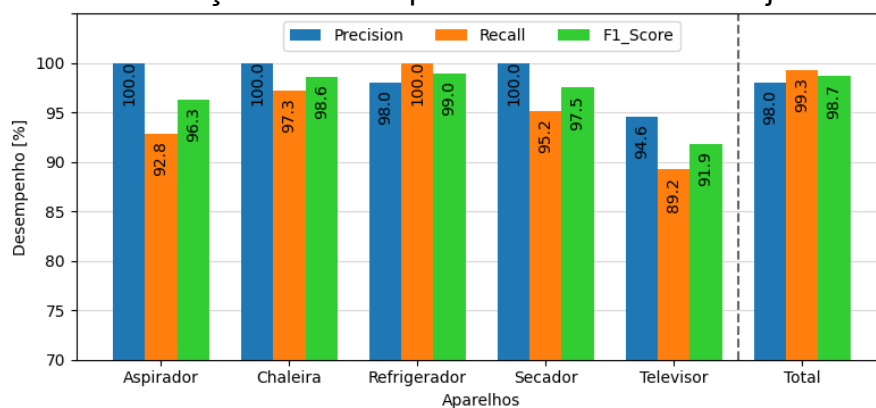
Tabela 9 – Contagem dos valores de *TP*, *FP* e *FN* do *LLR*.

Aparelhos	<i>FIRE</i> D (62 dias)			SECRBR (31 dias)		
	<i>TP</i>	<i>FP</i>	<i>FN</i>	<i>TP</i>	<i>FP</i>	<i>FN</i>
Aspirador	39	0	3	84	0	2
Chaleira	142	0	4	323	6	5
Refrigerador	3585	72	0	5248	9	2
Secador	79	0	4	100	0	2
Televisor	141	8	17	107	29	0
Total	3986	80	28	5862	44	11

Fonte: Autor.

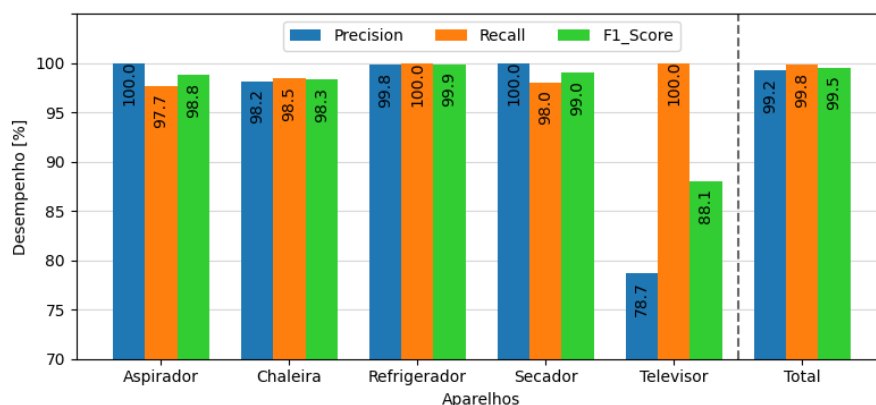
Com os valores contabilizados, foram realizadas as avaliações através das métricas *Precision*, *Recall* e *F1-Score*, de acordo com as Equações (5), (6) e (7) da seção 2.4. Os gráficos apresentam os resultados por equipamento e o desempenho total do detector de eventos *LLR* sobre o conjunto *FIRE*D (Figura 30) e o conjunto SECRBR (Figura 31).

Figura 30 – Avaliação de desempenho do *LLR* com o conjunto *FIRE*D.



Fonte: Autor.

Figura 31 – Avaliação de desempenho do *LLR* com dados do ambiente real *SECR*BR.



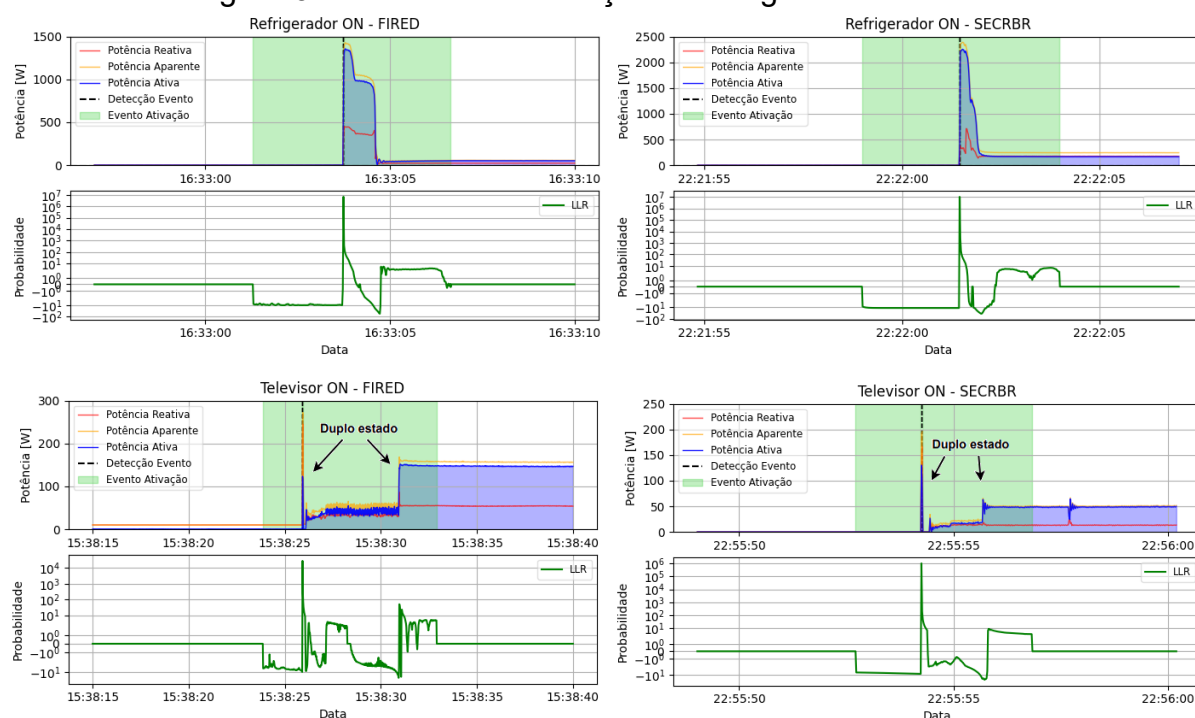
Fonte: Autor.

Ao analisar os resultados individuais dos aparelhos, observa-se que o detector obteve o melhor desempenho da métrica *F1-Score* para o refrigerador nos dois conjuntos de dados. Sendo o aparelho com a maior quantidade de eventos registrados, também foi o que teve o menor erro nas identificações. O *LLR* atingiu 99.0% no conjunto *FIRE*D e 99.9% no conjunto *SECR*BR para o refrigerador.

Em contrapartida, o televisor foi o eletrodoméstico com a maior quantidade de erros de detecção e, conseqüentemente, sendo o aparelho com o pior desempenho do algoritmo probabilístico, através da métrica *F1-Score*. Sua baixa atuação para este equipamento está relacionada com as próprias características do eletrodoméstico, devido às várias oscilações em estado estacionário de funcionamento e por apresentar estado duplo nos eventos de ativação e desativação. O resultado para este aparelho foi 91.9% e 88.1%, no conjunto *FIRE*D e *SECR*BR, respectivamente.

Ao comparar ambos os gráficos de barras, nota-se que a distribuição de desempenho de detecção nos eletrodomésticos, possui uma correlação positiva entre os dois conjuntos de dados, se comportando de forma análoga mesmo em coleções de sinais elétricos diferentes. Isso também pode ser analisado pela Figura 32, quando comparadas as curvas de potência dos eventos de ativação detectados dos aparelhos refrigerador e televisor:

Figura 32 – Eventos de ativação do refrigerador e televisor.



Fonte: Autor.

As séries temporais dos gráficos à esquerda mostram os eventos do conjunto *FIRED* e, a direita, os eventos do conjunto de sinais do ambiente real. Quando observados os eventos dos equipamentos, percebe-se as semelhanças nas curvas da potência entre ambos conjuntos. Inclusive o duplo estado na ativação do aparelho televisor. As demais séries temporais de potência e probabilidade, para os eventos de ativação e desativação detectados dos cinco eletrodomésticos, podem ser visualizadas no APÊNDICE A (para o *FIRED*) e APÊNDICE B (para o *SECRBR*).

Por fim, a avaliação geral do *LLR*, considerando todos os eventos detectados, foi 98.7% para o conjunto público e 99.5% para o conjunto do ambiente real. Estes valores de *F1-Score* foram superiores aos resultados apresentados por Völker, Scholl e Becker (2019), que atingiram 97% na eficiência das detecções.

4.4 FORMAÇÃO E SELEÇÃO DAS ASSINATURAS ELÉTRICAS

Contemplando a segunda etapa da metodologia após avaliar o *LLR*, a terceira etapa representa a extração das características dos eventos de ativação e desativação para a formação e seleção das assinaturas elétricas dos eletrodomésticos.

Através dos índices de identificação de eventos (i_A e i_D), foram coletadas as características dos aparelhos pelos sinais das potências normalizadas, sendo elas: a diferença de potencial, o pico máximo registrado dentro da região de evento e a distorção do conjunto de amostras de cada evento. Todos os recursos foram extraídos dos três sinais de potência ativa, reativa e aparente, conforme características sugeridas na subseção 3.3.1 da metodologia aplicada.

As assinaturas elétricas foram definidas como classe *ON*, para eventos de ativação, e classe *OFF*, para eventos de desativação, diferenciadas em cada tipo de aparelho. Além disso, duas bases de assinaturas foram preparadas, uma oriunda do conjunto público e a outra do ambiente real. Dessa forma, cada base de assinaturas foi constituída por um total de 10 classes e 9 recursos, estruturadas de acordo com o Quadro 4 da subseção 3.3.2.

Apenas as assinaturas de eventos verdadeiros foram mantidas, ou seja, que corresponderam ao *ground truth*, tornando as bases mais confiáveis para o processo de treinamento e testes do sistema, com dados *FIRE*D, e a validação com os dados do ambiente real. O Tabela 10 mostra as classes definidas para o sistema de desagregação e a quantidade de assinaturas contidas em cada base.

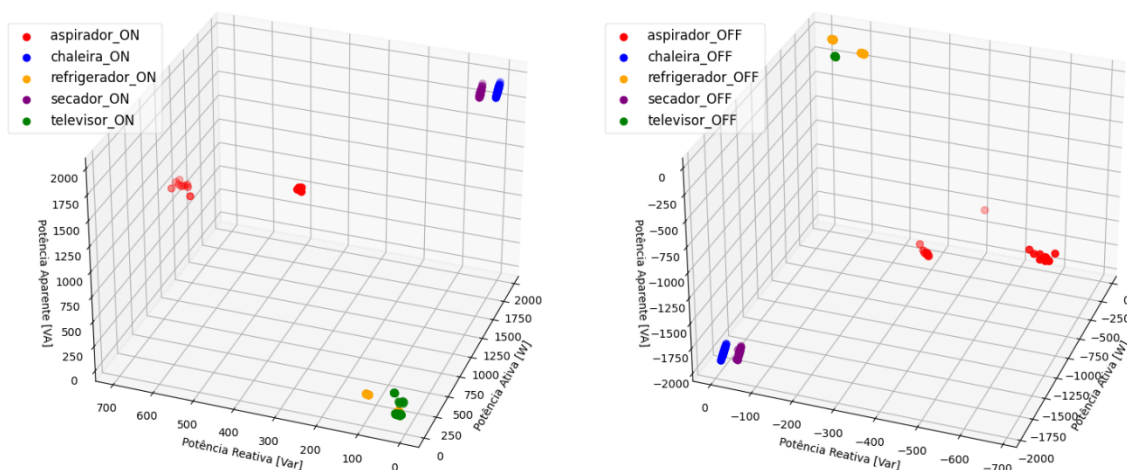
Tabela 10 – Classes das bases de assinaturas.

Classes	Base <i>FIRE</i> D	Base <i>SECR</i> BR
	Qtd. de Assinaturas	Qtd. de Assinaturas
<i>aspirador_OFF</i>	18	42
<i>aspirador_ON</i>	16	42
<i>chaleira_OFF</i>	70	161
<i>chaleira_ON</i>	69	162
<i>refrigerador_OFF</i>	1758	2608
<i>refrigerador_ON</i>	1777	2640
<i>secador_OFF</i>	36	50
<i>secador_ON</i>	37	50
<i>televisor_OFF</i>	63	53
<i>televisor_ON</i>	71	54

Fonte: Autor.

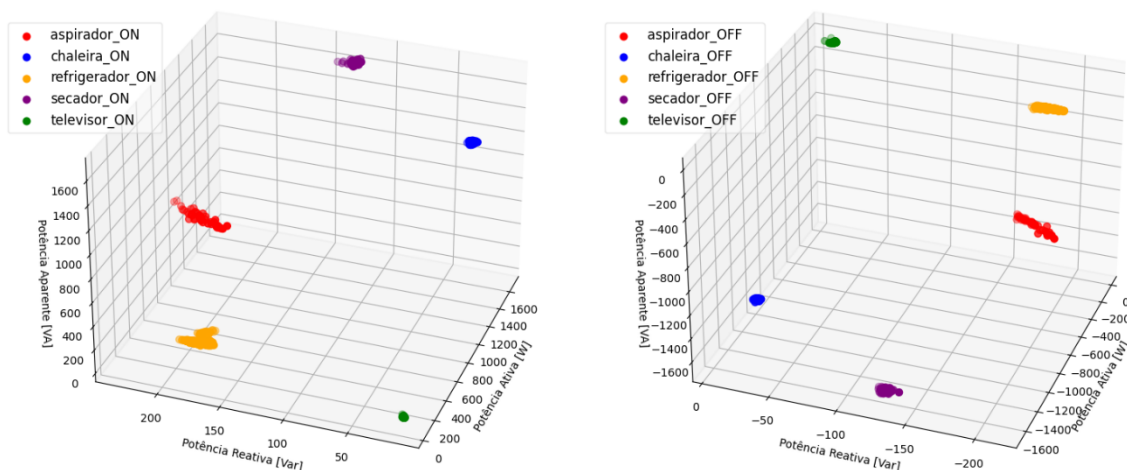
Dispondo os recursos das classes em representações de três dimensões, é possível verificar os agrupamentos das assinaturas através dos planos PQS. Os planos na Figura 33 e Figura 34 mostram as aglomerações da diferença de potencial para as assinaturas de ativação e desativação extraídas de *FIRE*D e SECRBR.

Figura 33 – Diferença de potencial das assinaturas no plano PQS – *FIRE*D.



Fonte: Autor.

Figura 34 – Diferença de potencial das assinaturas no plano PQS – SECRBR.



Fonte: Autor.

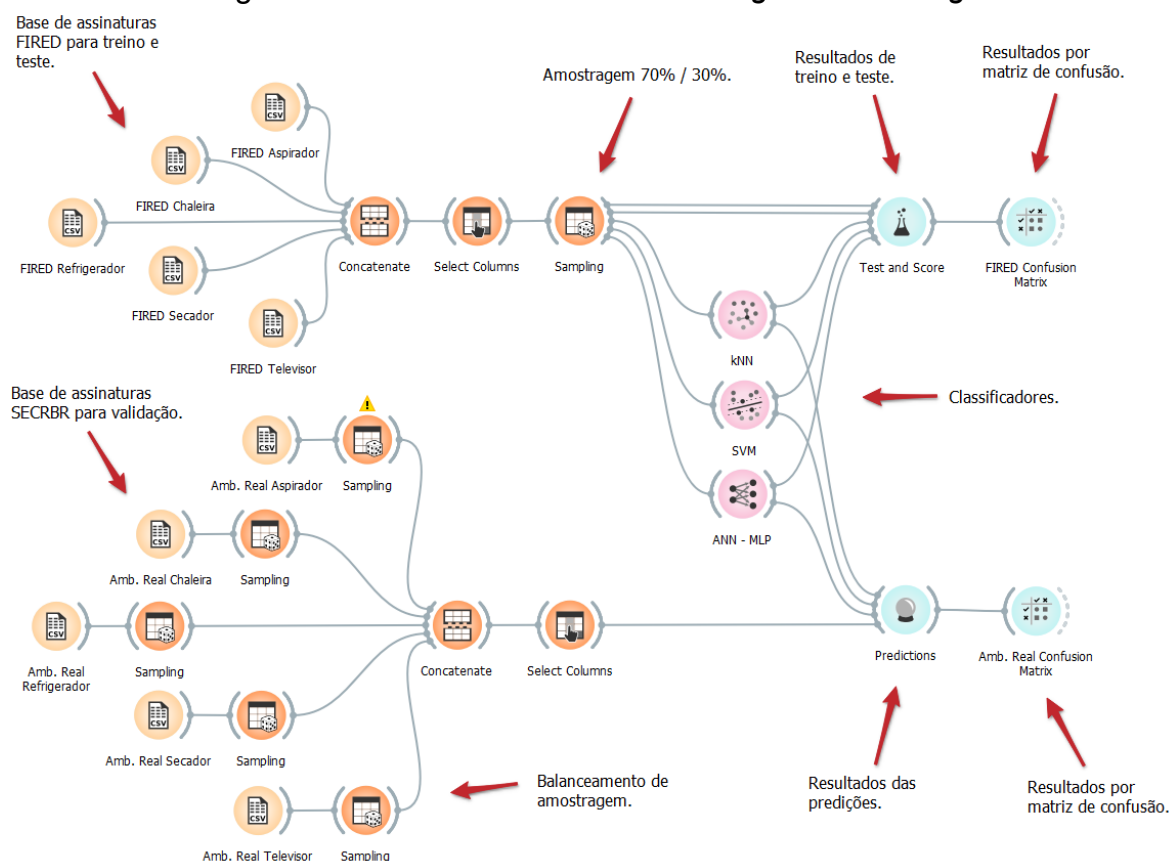
É possível verificar as semelhanças e correlações entre os planos PQS, quando relacionados os dois conjuntos de dados. Percebe-se que os agrupamentos das classes de ativação e desativação dos aparelhos possuem similaridades, mesmo sendo recursos extraídos de eletrodomésticos em locais distintos (Alemanha e Brasil).

Por exemplo, as características das assinaturas *refrigerador_ON* e *televisor_ON* estão próximas umas das outras e localizadas em regiões equivalentes do espaço tridimensional em ambos os conjuntos de dados. O mesmo ocorre com as demais classes. Isso evidencia a possibilidade de classificar aparelhos monitorados em um ambiente utilizando sistema de desagregação implementado com dados de outra realidade. Os outros recursos extraídos dos dois conjuntos para formação das assinaturas estão representados nos planos PQS, através do APÊNDICE C e APÊNDICE D.

4.5 TREINO E AVALIAÇÃO DOS CLASSIFICADORES

A quarta etapa da metodologia consiste em aplicar a base de assinaturas para treinamento e testes do sistema, utilizando os três classificadores de aprendizado de máquina: *KNN*, *SVM* e *ANN-MLP*. Para realizar o processo completo da definição dos hiperparâmetros, treinamento e avaliação, foi utilizado a ferramenta *Orange Data Mining* e construído o *workflow* mostrado na Figura 35:

Figura 35 – Fluxo de trabalho no *Orange Data Mining*.



Fonte: Autor.

O fluxo de trabalho consiste na base de assinaturas *FIRE*D, como entrada de informações para treino e teste, seguido de um objeto para amostragem das assinaturas. Nele foi definido a separação dos dados de entrada em 70% para o treino do sistema e 30% para testar e avaliar. Além disso, a amostragem é configurada para amostras estratificadas e não determinísticas da base *FIRE*D.

O objeto de teste e avaliação da ferramenta mostra os resultados individuais dos classificadores pelas métricas *Precision*, *Recall*, *F1-Score*, *Accuracy* e a área sob a curva *ROC (Receiver Operating Characteristic)*.

Já o objeto matriz de confusão apresenta os resultados de acertos e confusões pelos classificadores, de acordo com o objeto que ele está conectado (teste ou predição).

Os classificadores são os objetos que constituem os algoritmos de inferências por aprendizado de máquina. Os demais objetos (base de assinaturas *SECRBR* e predição) são utilizados nos resultados da quinta e última etapa da metodologia, abordada na próxima seção.

Após realizar diversas iterações no fluxo de trabalho, variando os parâmetros dos classificadores e analisando os resultados de treino e teste, foram alterados os valores dos hiperparâmetros até obter os melhores resultados. Os valores definidos são mostrados na Tabela 11.

Tabela 11 – Hiperparâmetros utilizados nos classificadores.

Classificadores	Hiperparâmetros	Métrica
KNN	Métrica de distância	<i>Mahalanobis</i>
	Número de vizinhos (<i>K</i>)	3
SVM	Custo (<i>C</i>)	20,00
	<i>Kernel</i>	<i>RBF</i>
	<i>Gamma</i>	0,10
ANN-MLP	Neurônios na entrada (<i>I</i>)	9
	Camadas ocultas	1
	Neurônios na oculta (<i>H</i>)	19
	Neurônios na saída (<i>O</i>)	10
	Função de ativação	<i>ReLU</i>
	<i>Solver</i>	<i>L-BFGS-B</i>

Fonte: Autor.

No classificador *KNN* o número de vizinhos que melhor se adaptou ao problema foram 3 vizinhos, semelhante ao estudo de Bosco et al. (2021), que também utilizaram este valor para a classificação de cargas. O hiperparâmetro *Kernel* do classificador *SVM* foi definido com a função de base radial, assim como o

trabalho de Abhilash et al. (2020). Este tipo de *Kernel* é melhor empregado para problemas multidimensionais não-lineares, sendo este o caso para o sistema de desagregação.

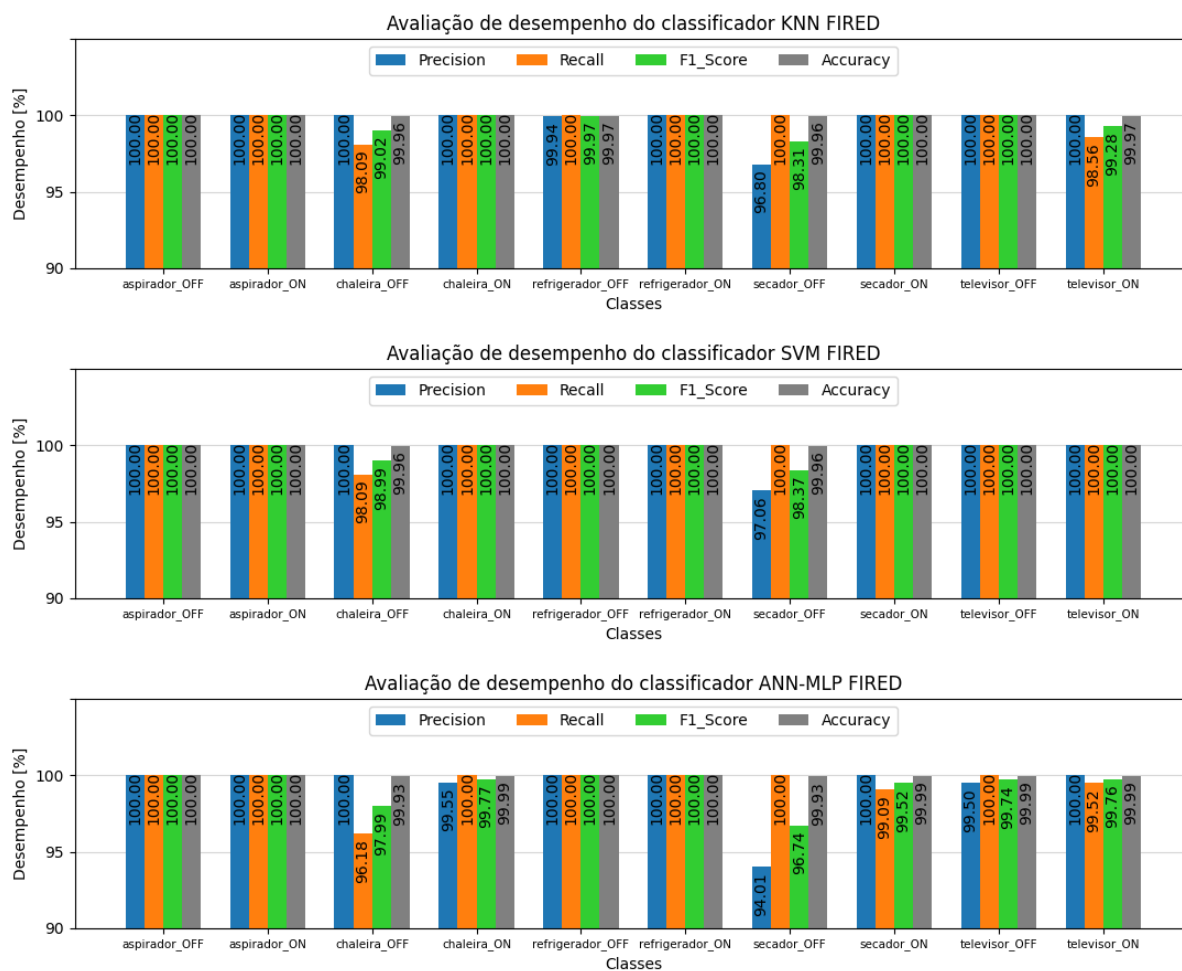
O último classificador utilizado, *ANN-MLP*, teve seus hiperparâmetros ajustados similarmente ao estudo dos autores Zhao et al. (2021), porém apenas uma camada oculta foi implementada. A quantidade de neurônios nesta camada, com melhor resultado, foi definida pela função $2n + 1$, onde n é o número de neurônios da camada de entrada. Além disso, a função de ativação *ReLU* também mostrou o melhor desempenho. Por fim, o *Solver* escolhido para a rede neural foi o *L-BFGS-B* (um otimizador da família de métodos *Quasi-Newton*). Além de apresentar resultado semelhante/superior, em relação aos outros otimizadores disponíveis na ferramenta, ainda obteve o menor tempo para treinamento do modelo (menos de 0.2 s).

Com os hiperparâmetros definidos em cada classificador, foram realizadas 10 repetições seguidas de treino e teste, com o objetivo de obter resultados mais confiáveis através da média sobre as métricas.

Em cada iteração foram alteradas as amostras da entrada de maneira randômica (não determinística) e estratificada, mas mantendo a proporção 70/30. Os resultados dos classificadores foram registrados, conforme as métricas definidas na subseção 3.4.2 da metodologia (*Accuracy* e matriz de confusão). Além de *Precision*, *Recall* e *F1-Score*, que são fornecidas nos relatórios do *Orange Data Mining*.

Os relatórios mostram os resultados das métricas individualmente para cada classe, além da média geral interna (*micro-averaging*) e a média geral externa (*macro-averaging*) dos classificadores. Os registros das 10 repetições de treino e teste, com dados *FIRE*D, estão disponíveis no APÊNDICE E.

A partir das iterações foram realizadas as médias das métricas para as classes de aparelhos e gerado os gráficos de barras com as avaliações de desempenho dos classificadores *KNN*, *SVM* e *ANN-MLP*, como segue na Figura 36.

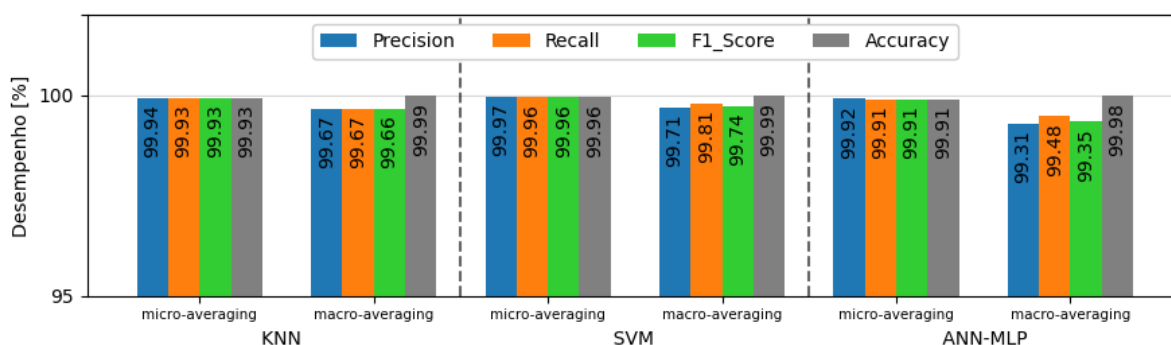
Figura 36 – Avaliação por classe no *KNN*, *SVM* e *ANN-MLP* com a base *FIRE*D.

Fonte: Autor.

Pelas avaliações de desempenho podem ser visualizados que *chaleira_OFF* e *secador_OFF* são as classificações com o menor desempenho em relação às demais classes pelos classificadores. Sendo o desempenho mais baixo apresentado na rede neural artificial, com 99.93% de *Accuracy* nas duas classes. Os gráficos representam a eficiência das inferências dos eventos de ativação e desativação de maneira muito similar entre os três classificadores, com os dados de teste da base de assinaturas *FIRE*D. Isso significa que os resultados estão coerentes, além de um ótimo desempenho dos classificadores.

Também foram dispostos em gráficos de barras as avaliações nas versões *micro-averaging* e *macro-averaging*, ou seja, o desempenho total dos classificadores com os dados de teste. Os resultados são mostrados na Figura 37.

Figura 37 – Avaliação por *micro-averaging* e *macro-averaging* no *KNN*, *SVM* e *ANN-MLP* com a base *FIRE*D.



Fonte: Autor.

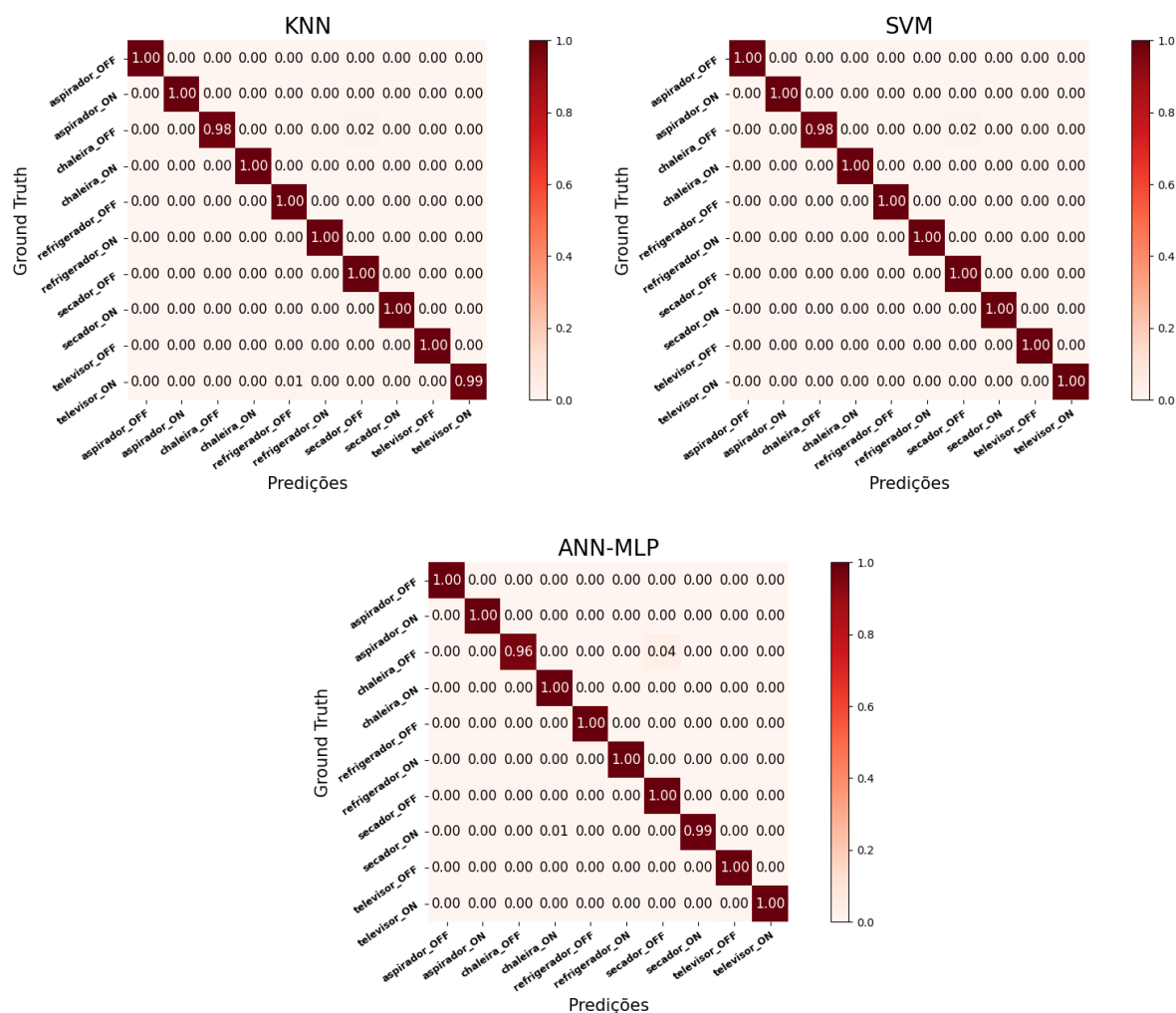
Ao inspecionar o gráfico de barras acima, percebe-se que a métrica *Accuracy* tem seu desempenho praticamente idêntico entre as versões *micro-averaging* e *macro-averaging* para cada classificador. Isso pode ser explicado, pois os algoritmos de aprendizado de máquina identificaram corretamente quase todas as amostras de teste. Portanto, o não balanceamento de carga amostral das classes não interferiu nos resultados de desempenho final.

Analisando a *micro-averaging*, os classificadores apresentaram na fase de treino e teste 99.93%, 99.96% e 99.91% de desempenho na *Accuracy* com o *KNN*, *SVM* e *ANN-MLP*, respectivamente. Estes valores superaram os resultados de Gurbuz et al. (2021) que avaliaram o *KNN* com 85%, Moradzadeh et al. (2020) que avaliaram o *SVM* com 98% e Zhao et al. (2021) que avaliaram e compararam o *KNN*, *SVM* e *ANN-MLP*, com desempenho de 96%, 98% e 99%, respectivamente.

Dessa forma, conclui-se que os classificadores avaliados nesta metodologia, atingiram um excelente desempenho para a implementação do sistema de desagregação de cargas proposto.

A última métrica aplicada para verificar os classificadores com os dados de treino e teste, é a matriz de confusão que mostra os erros e acertos em cada classe. Para esta análise foram acumuladas todas as matrizes das 10 repetições e normalizada a matriz de confusão resultante. A Figura 38 apresenta as três matrizes de confusão dos classificadores *KNN*, *SVM* e *ANN-MLP*, sendo cruzado o *ground truth* com as predições das classes.

Figura 38 – Avaliação por matriz de confusão no *KNN*, *SVM* e *ANN-MLP* com a base *FIRED*.



Fonte: Autor.

A maior quantidade de confusões das classes entre os classificadores é apresentada pela matriz de confusão do *ANN-MLP*. Como já era de se esperar, alguns eventos da *chaleira_OFF* foram cruzados com *secador_OFF*, onde apenas 4% das amostras da classe *chaleira_OFF* são classificadas incorretamente como sendo da classe *secador_OFF*. Todas as matrizes de confusão capturadas nas 10 repetições, para os resultados de treino e teste com a base *FIRED*, estão disponíveis no APÊNDICE G.

4.6 VALIDAÇÃO DO SISTEMA NO AMBIENTE REAL

A última etapa da metodologia é a validação do sistema de desagregação com os dados coletados do ambiente real. Através da estatística *Kappa* foi verificado

se, utilizando dados de um conjunto público na implementação do sistema, é possível classificar cargas de um cenário real local.

Após coletados os sinais elétricos dos eletrodomésticos monitorados na residência, tratados, identificados os eventos, extraídas as características e formada a base de assinaturas, foi utilizado o *workflow* (Figura 35) para validar o sistema com a ferramenta *Orange Data Mining*.

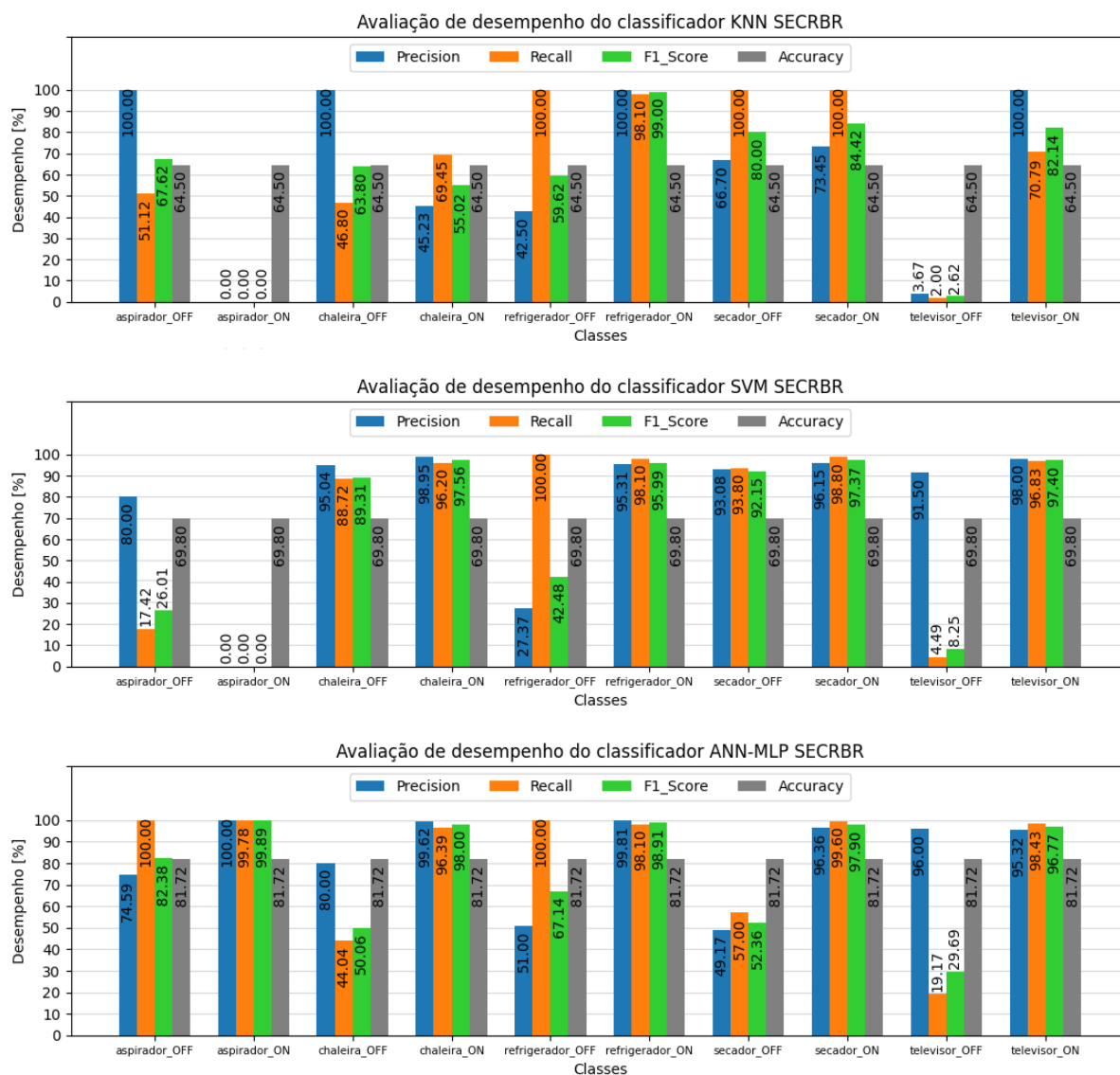
A base SECRBR foi utilizada como dados de entrada para a validação, seguido do objeto de amostragem para balanceamento das amostras. As assinaturas foram amostradas igualmente para cada tipo de aparelho, sendo obtidas 500 amostras de maneira determinística, estratificada e com reposição. Dessa forma, cada classe ficou com, aproximadamente, 50 assinaturas do conjunto para classificação.

Depois de treinados os três classificadores de aprendizado de máquina com os dados *FIRE*D, o objeto de predição do fluxo de trabalho é utilizado para inferir os dados de SECRBR e apresentar os resultados pelas métricas *Precision*, *Recall*, *F1-Score* e *Accuracy*.

Assumindo a mesma estratégia de treino e teste, foram realizadas 10 repetições seguidas na avaliação, obtendo resultados mais confiáveis para validar o sistema. Os registros das predições, com os dados do ambiente real, estão disponíveis no APÊNDICE F.

A partir das 10 iterações foram realizadas as médias dos resultados das classes de aparelhos. Através da Figura 39 são apresentados os gráficos de barras com as médias das avaliações de desempenho dos classificadores *KNN*, *SVM* e *ANN-MLP*.

Figura 39 – Avaliação por classe no KNN, SVM e ANN-MLP com a base SECRBR.

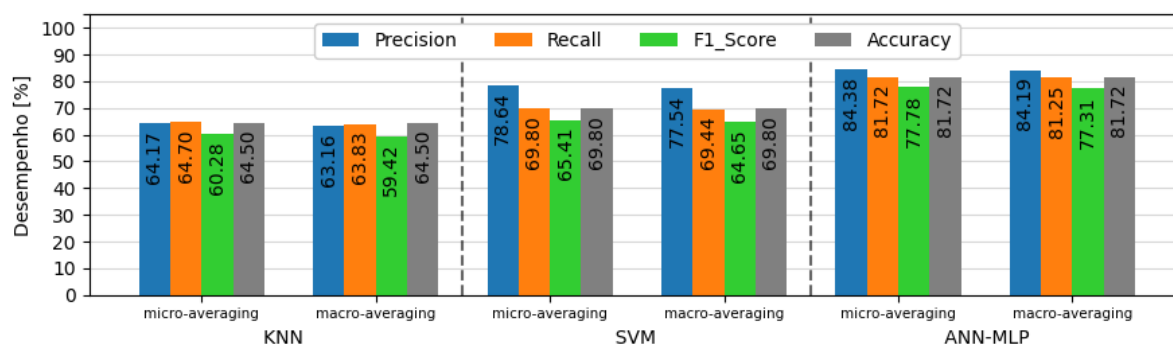


Fonte: Autor.

Pelas avaliações individuais dos eletrodomésticos, os gráficos mostram que a classe *aspirador_ON* não foi identificada pelos modelos *KNN* e *SVM*. Porém, teve um ótimo desempenho no *ANN-MLP*, atingindo quase 100% de identificação pela métrica *F1-Score*. Além disso, a classe *televisor_OFF* também teve um baixo desempenho de classificação por *KNN* e *SVM*, mas alcançou uma média de, aproximadamente, 30% com o *ANN-MLP*, sendo a classe com o menor desempenho neste classificador. Portanto, é possível afirmar que a rede neural artificial se destacou em relação à classificação individual dos aparelhos quando comparada com os outros dois modelos de *ML*.

As métricas nas versões *micro-averaging* e *macro-averaging*, para avaliar o desempenho total dos classificadores com os dados do ambiente real, também são apresentadas na Figura 40.

Figura 40 – Avaliação por *micro-averaging* e *macro-averaging* no KNN, SVM e ANN-MLP com a base SECRBR.



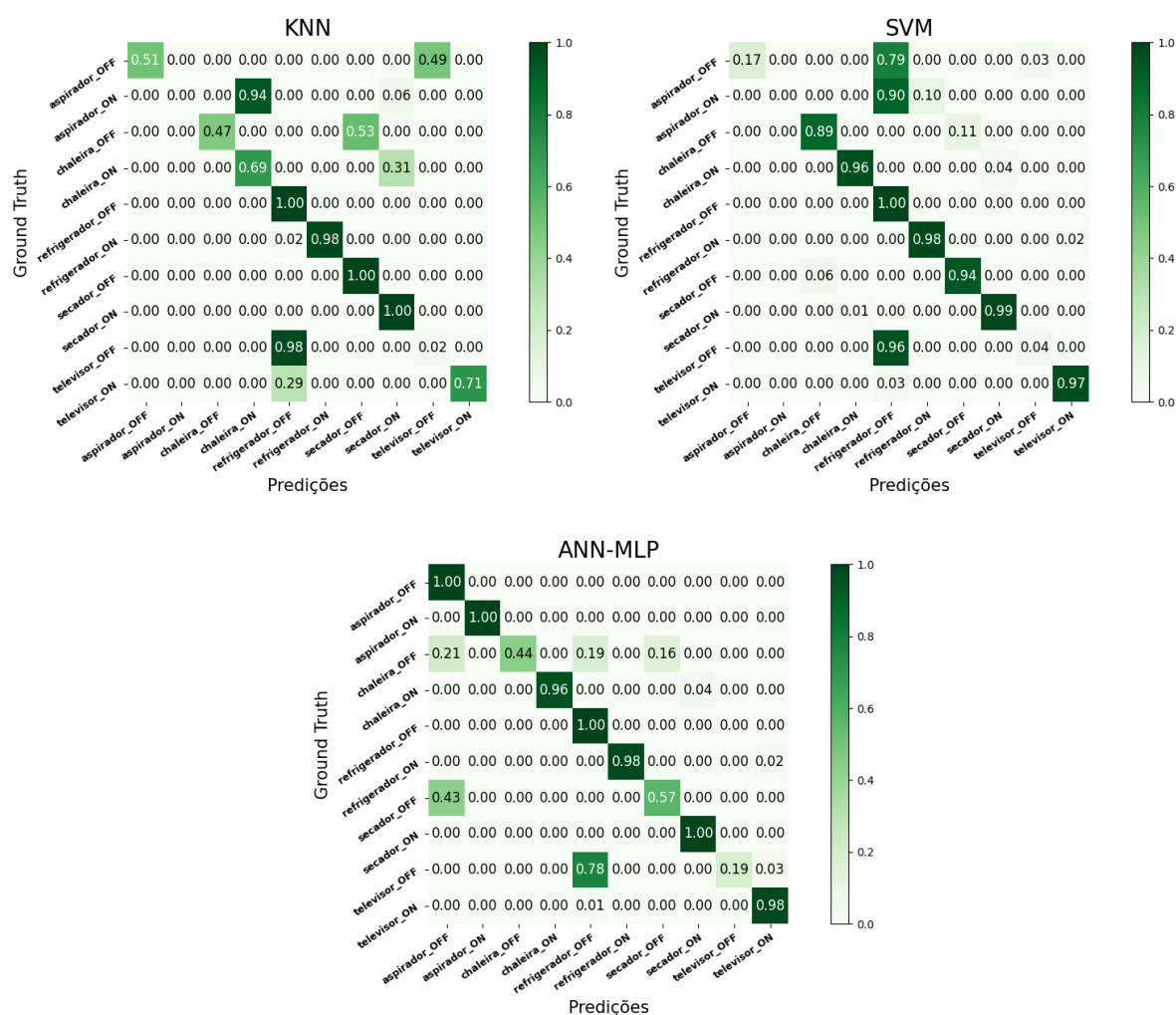
Fonte: Autor.

Em cada modelo classificador as métricas *Precision*, *Recall*, *F1-score* e *Accuracy* tiveram desempenho praticamente idênticos entre *micro-averaging* e *macro-averaging*. Essa característica é esclarecida uma vez que os dados de entrada foram balanceados entre as classes da base de assinaturas. O equilíbrio dos dados é desejável para a validação desta metodologia, pois o principal objetivo é verificar se o sistema é capaz de classificar corretamente todos os aparelhos, sem desviar os resultados dos modelos por classes com maior ou menor quantidade de amostras.

Apurando os valores das médias das 10 repetições por *micro-averaging*, os classificadores *KNN*, *SVM* e *ANN-MLP* apresentaram na fase de validação 64.50%, 69.80% e 81.72% de desempenho por *Accuracy*, respectivamente. De acordo com os resultados da avaliação individual dos aparelhos, já era previsto a capacidade superior do modelo *ANN-MLP* sobre o *KNN* e *SVM*.

Também foram expostos os erros e acertos dos classificadores por matriz de confusão com os dados SECRBR. A Figura 41 mostra as três matrizes acumuladas e normalizadas.

Figura 41 – Avaliação por matriz de confusão no *KNN*, *SVM* e *ANN-MLP* com a base *SECRBR*.



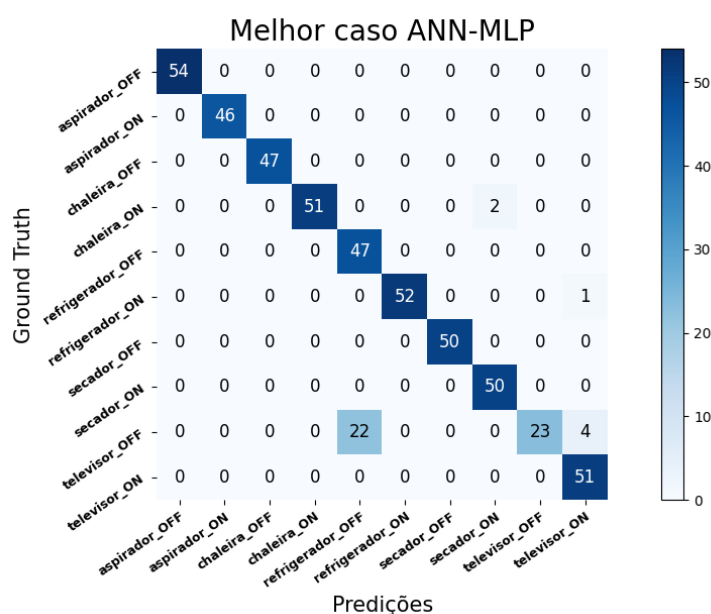
Fonte: Autor.

Ao verificar as confusões dos três classificadores, percebe-se que a classe *televisor_OFF* foi prevista como *refrigerador_OFF* em grande parte das ocorrências por todos os modelos. Logo, é possível deduzir que o baixo desempenho na classe *televisor_OFF* está relacionado com as semelhanças das características nos eventos de desativação entre os dois eletrodomésticos. Além disso, as classes *aspirador_OFF* e *aspirador_ON* também tiveram trocas nos classificadores *KNN* e *SVM*, mas as confusões ocorreram em classes distintas entre os dois modelos. Portanto, é possível inferir que estes resultados possuem maior relação com as limitações dos modelos *KNN* e *SVM*, uma vez que o classificador *ANN-MLP* não apresentou essa divergência.

Explorando todos os resultados por matriz de confusão, a Figura 42 traz o melhor caso de todas as avaliações nas 10 iterações. Nota-se que a maior

ocorrência de confusão foi na classe *televisor_OFF*, onde metade das amostras foram classificadas corretamente. A rede neural artificial apresentou um ótimo desempenho nesta avaliação, atingindo 94.20% de *Accuracy*. Todas as matrizes de confusão para avaliação do sistema com os dados do ambiente real, estão no APÊNDICE H.

Figura 42 – Melhor ocorrência por matriz de confusão com a base SECRBR.



Fonte: Autor.

Por fim, a última avaliação do sistema e da metodologia, é o índice *Kappa*, que mede o grau de concordância em modelos de *machine learning*. Conforme exposto na subseção 3.5.1, a estatística *Kappa* é utilizada neste trabalho como forma de validar o sistema e responder à pergunta de pesquisa. Uma vez que não se têm referências para avaliar este tipo de problema, foi assumido o rótulo “Moderado” (Tabela 1) como limite inferior do valor *Kappa*. E, assim sendo, aceitar ou recusar a hipótese de utilizar um conjunto público de dados elétricos para implementar sistemas de desagregação e classificar cargas em ambiente real local.

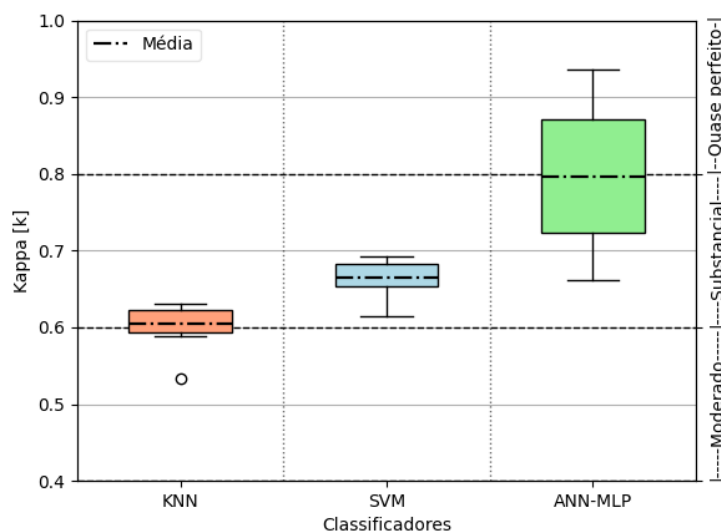
Os valores da estatística *Kappa* são exibidos na Tabela 12 para as 10 iterações com os três classificadores utilizados:

Tabela 12 – Índice *Kappa* nas iterações de avaliação com *KNN*, *SVM* e *ANN-MLP*.

Iterações	<i>KNN</i>	<i>SVM</i>	<i>ANN-MLP</i>
<i>Kappa</i> #1	0.622	0.645	0.893
<i>Kappa</i> #2	0.626	0.687	0.713
<i>Kappa</i> #3	0.593	0.676	0.662
<i>Kappa</i> #4	0.624	0.669	0.795
<i>Kappa</i> #5	0.62	0.614	0.757
<i>Kappa</i> #6	0.533	0.665	0.807
<i>Kappa</i> #7	0.631	0.649	0.907
<i>Kappa</i> #8	0.589	0.685	0.793
<i>Kappa</i> #9	0.622	0.665	0.936
<i>Kappa</i> #10	0.593	0.693	0.704
Média	0.61	0.66	0.80

Fonte: Autor.

A partir da tabela com os índices *Kappa* foi construído o diagrama de caixa (boxplot), que é mostrado na Figura 43, abaixo:

Figura 43 – Boxplot da estatística *Kappa* com *KNN*, *SVM*, *ANN-MLP*.

Fonte: Autor.

Diante dos resultados de índice *Kappa*, o classificador *KNN* obteve valores restritos, o que indica que este modelo é limitado quando aplicado ao problema em questão. Mesmo assim, o *K-Nearest Neighbors* teve um grau médio de κ igual a 0.61, apresentando resultados entre os rótulos “Moderado” e “Substancial”.

Assim como o *KNN*, o *SVM* também foi mais conservador sobre a faixa de resultados, porém alcançou um maior índice. O modelo *Support Vector Machine* obteve valor médio de κ igual a 0.66 e permaneceu dentro do rótulo “Substancial” em todas as iterações.

Já o classificador *ANN-MLP* apresentou um grau de concordância mais amplo, o que indica que as características deste modelo são mais mutáveis e se adapta melhor ao problema com um grande poder de generalização. O *Artificial Neural Network* teve o valor médio de κ igual a 0.80, atingindo o valor máximo de 0.94 e apresentando resultados entre os rótulos “Substancial” e “Quase perfeito”.

Portanto, ao utilizar a hipótese de $\kappa > 0.40$, pode ser aceita a premissa que: implementar sistemas de desagregação de cargas residências, com conjunto público de dados, é possível classificar cargas de um ambiente residencial de outra realidade. Além disso, a implementação do sistema com três classificadores, responde à pergunta de pesquisa com mais convicção, pois todos os classificadores utilizados neste trabalho mantiveram-se acima da condição mínima avaliada de *Kappa*.

5 CONCLUSÕES

O monitoramento de cargas elétricas residenciais vem se apresentando como uma alternativa para que os consumidores tomem ciência de seus gastos e modifiquem seus comportamentos com a reeducação de hábitos. Favorecendo na redução do consumo, minimizando o desperdício de energia elétrica e contribuindo para a eficiência energética. Se assim seguir, será possível alcançar o sétimo objetivo de desenvolvimento sustentável estabelecido pela ONU.

Visto que o conhecimento detalhado do perfil de consumo dos eletrodomésticos em uma residência auxilia na gestão do recurso energético, técnicas de monitoramento não intrusivo de cargas são exploradas cada vez mais. Esta categoria apresenta vantagens sobre os demais métodos, com o principal destaque por ser uma opção de baixo custo de implantação no ambiente residencial. Além de apresentar um grande potencial para ser incluída em medidores de energia do tipo *Smart Meter* e servir de ferramenta para aplicações de *feedback*.

Neste contexto, é necessário o desenvolvimento de novos métodos e algoritmos para que seja possível implementar e extrair os benefícios do monitoramento não intrusivo de cargas no mundo real. Pois, é importante mencionar, que ainda existem vários desafios a serem solucionados até que este tipo de técnica se torne viável em um ambiente residencial. Um dos principais problemas está relacionado com as bases de dados utilizadas para a construção e o treinamento dos sistemas de desagregação, visto que o monitoramento necessita de referências para inferir o consumo dos aparelhos.

Apesar de existirem alguns conjuntos de dados públicos, pouco se sabe se essas bases de dados podem servir para implementação de sistemas de monitoramento *NILM* que viabilizem o uso em uma realidade diferente. Isso ocorre devido ao fato que as características de consumo, tipos de aparelhos, região de extração e outras particularidades são diferentes de um local para outro, ou seja, o local de extração da base de dados e o local do monitoramento.

Portanto, essa dissertação investigou este aspecto, uma vez que não foi encontrado na literatura estudos que abordassem esse tema, também por não estar disponível um conjunto de sinais elétricos residencial público de âmbito regional e pela complexidade na criação de conjunto de dados. Visto que a criação de uma

base de assinaturas para treinamento, em cada local onde for instalado o sistema de monitoramento, pode ser uma tarefa desafiadora e até inviável.

Logo, este estudo apresentou um método de desenvolvimento e implementação de sistemas *NILM* para o ambiente real utilizando um conjunto público de dados de outra realidade. A metodologia sugerida empregou a categoria de monitoramento baseada em eventos, com a frequência de aquisição no ciclo da rede elétrica, utilizando o detector de eventos probabilístico *Log Likelihood Ratio*. Além disso, foram aplicados três modelos de aprendizado de máquina para classificação de eventos de ativação e desativação, sendo eles, o algoritmo *K-Nearest Neighbors*, *Support Vector Machine* e *Artificial Neural Network - Multi-layer Perceptron*. O estudo ainda propôs um método para extrair os sinais elétricos do ambiente real e gerar um conjunto de dados local.

A avaliação do sistema foi realizada através de métricas disponíveis na literatura *NILM*, sendo avaliado o desempenho na fase de treino e teste com dados do conjunto público *FIREED*. O sistema apresentou desempenho superior aos estudos encontrados na literatura e descritos no referencial teórico, tanto no quesito detecção de eventos, quanto na classificação dos eventos.

Na avaliação com dados do ambiente real foi verificado, através da estatística *Kappa*, que os três classificadores de aprendizado de máquina demonstraram eficiência suficiente para validar o método e responder à pergunta de pesquisa investigada nesta dissertação. Sendo o modelo *ANN-MLP* mais indicado para solucionar este tipo de problema, por apresentar o melhor desempenho (chegando a 94.2% de *Accuracy*) e atingindo o maior grau de concordância entre os modelos testados (“Quase perfeito”), demonstrando excelente poder de generalização. Entretanto, para aplicações em cenários onde há a inclusão de novas classes monitoradas, é necessária uma reconfiguração e estruturação da rede neural artificial para que elas possam ser identificadas pelo sistema.

Dessa forma, conclui-se neste estudo que é possível desagregar cargas residenciais utilizando apenas conjuntos de dados públicos para implementação de sistema *NILM* em ambientes reais.

5.1 PRINCIPAIS CONTRIBUIÇÕES

As principais contribuições deste estudo se referem ao:

- Desenvolvimento de um método de monitoramento que utiliza apenas conjuntos de dados públicos para desagregação de cargas em qualquer ambiente ou realidade de monitoramento sem a necessidade de extrair uma base de assinaturas local para implementação de sistemas *NILM*;
- A criação de uma estratégia de validação do método desenvolvido pela estatística *Kappa* para julgar a hipótese da pergunta de pesquisa;
- A possibilidade de incluir em sistemas reais a desagregação de classes de equipamentos utilizando tecnologias de aprendizado de máquina e algoritmos que exigem baixo poder de processamento;
- A avaliação de modelos de *machine learning* (*KNN*, *SVM* e *ANN-MLP*) que melhor se adequa para a classificação de cargas. Sendo a rede neural artificial a que se destacou para a solução do problema;
- A disponibilidade de coleções de conjuntos de sinais elétricos para a evolução de pesquisas no monitoramento de cargas para a realidade brasileira.

5.2 ESTUDOS FUTUROS

Como trabalho futuros, destacam-se:

- Avaliar mais equipamentos na metodologia a fim de consolidar a generalização do resultado;
- Mesclar as características dos equipamentos com outros bancos de dados públicos para contemplar uma maior variedade de assinaturas elétricas sobre o sistema de inferência;
- Utilizar outros recursos na extração de assinaturas, visto que, as características dos equipamentos são os maiores influenciadores no desempenho dos classificadores;

- Investigar outras técnicas de classificações que não exijam a reestruturação com a inclusão de novas classes;
- Implementar o sistema em um dispositivo embarcado para testes em tempo real;
- Ampliar a quantidade de equipamentos monitorados individualmente para aumentar a quantidade de informações da base de dados. E adicionar um arquivo de rotulagem no SECRBR.

REFERÊNCIAS

ABHILASH, D.; KUMARI, L. V. R.; PALLE, V. R. **Design and development of the secondary energy meter for non-intrusive load monitoring using machine learning method**. 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC). **Anais...** Em: 2020 FOURTH INTERNATIONAL CONFERENCE ON I-SMAC (IOT IN SOCIAL, MOBILE, ANALYTICS AND CLOUD) (I-SMAC). Palladam, India: IEEE, 7 out. 2020. Disponível em: <<https://ieeexplore.ieee.org/document/9243346/>>. Acesso em: 17 jul. 2022

ABRAHÃO, K. C. DE F. J.; SOUZA, R. G. V. DE. Estimativa da evolução do uso final de energia elétrica no setor residencial do Brasil por região geográfica. **Ambiente Construído**, v. 21, n. 2, p. 383–408, abr. 2021.

ABUBAKAR, I. et al. **An overview of Non-intrusive load monitoring methodologies**. 2015 IEEE Conference on Energy Conversion (CENCON). **Anais...** Em: 2015 IEEE CONFERENCE ON ENERGY CONVERSION (CENCON). Johor Bahru: IEEE, out. 2015. Disponível em: <<http://ieeexplore.ieee.org/document/7409513/>>. Acesso em: 18 nov. 2021

ADIONO, T. et al. **Using A Smart Plug based on Consumer Electronics to Support Low Power Smart Home**. 2019 4th International Conference on Intelligent Green Building and Smart Grid (IGBSG). **Anais...** Em: 2019 4TH INTERNATIONAL CONFERENCE ON INTELLIGENT GREEN BUILDING AND SMART GRID (IGBSG). Hubei, Yi-chang, China: IEEE, set. 2019. Disponível em: <<https://ieeexplore.ieee.org/document/8886272/>>. Acesso em: 18 set. 2022

ALADESANMI, E. J.; FOLLY, K. A. Overview of non-intrusive load monitoring and identification techniques. **IFAC-PapersOnLine**, v. 48, n. 30, p. 415–420, 2015.

ALCALA, J. et al. Event-Based Energy Disaggregation Algorithm for Activity Monitoring From a Single-Point Sensor. **IEEE Transactions on Instrumentation and Measurement**, v. 66, n. 10, p. 2615–2626, out. 2017.

ALI, M. et al. **Monitoring and Identification Electricity Load Using Artificial Neural Network**. 2021 7th International Conference on Electrical, Electronics and Information Engineering (ICEEIE). **Anais...** Em: 2021 7TH INTERNATIONAL CONFERENCE ON ELECTRICAL, ELECTRONICS AND INFORMATION ENGINEERING (ICEEIE). Malang, Indonesia: IEEE, 2 out. 2021. Disponível em: <<https://ieeexplore.ieee.org/document/9616922/>>. Acesso em: 15 jul. 2022

ANDERSON, K. et al. **BLUED: a fully labeled public dataset for event-based non-intrusive load monitoring research**. Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability. **Anais...** Em: ACM SUSTKDD'12. Beijing, China: dez. 2012a.

ANDERSON, K. D. et al. **Event detection for Non Intrusive load monitoring**. IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society. **Anais...** Em: IECON 2012 - 38TH ANNUAL CONFERENCE OF IEEE INDUSTRIAL

ELECTRONICS. Montreal, QC, Canada: IEEE, out. 2012b. Disponível em: <<http://ieeexplore.ieee.org/document/6389367/>>. Acesso em: 18 nov. 2021

ANEEL. **Relatórios de Consumo e Receita de Distribuição**. Agência Nacional de Energia Elétrica, , 2021. Disponível em: <<https://www.aneel.gov.br/relatorios-de-consumo-e-receita>>. Acesso em: 25 nov. 2021

AZAZA, M.; WALLIN, F. Finite State Machine Household's Appliances Models for Non-intrusive Energy Estimation. **Energy Procedia**, v. 105, p. 2157–2162, maio 2017.

BECKEL, C. et al. **The ECO data set and the performance of non-intrusive load monitoring algorithms**. Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings. **Anais...** Em: SENSYS '14: THE 12TH ACM CONFERENCE ON EMBEDDED NETWORK SENSOR SYSTEMS. Memphis Tennessee: ACM, 3 nov. 2014. Disponível em: <<https://dl.acm.org/doi/10.1145/2674061.2674064>>. Acesso em: 28 nov. 2021

BIANSOONGNERN, S.; PLANGKLANG, B. **Nonintrusive load monitoring (NILM) using an Artificial Neural Network in embedded system with low sampling rate**. 2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON). **Anais...** Em: 2016 13TH INTERNATIONAL CONFERENCE ON ELECTRICAL ENGINEERING/ELECTRONICS, COMPUTER, TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY (ECTI-CON). Chiang Mai, Thailand: IEEE, jun. 2016. Disponível em: <<http://ieeexplore.ieee.org/document/7561398/>>. Acesso em: 18 nov. 2021

BIANSOONGNERN, S.; PLUNGKLANG, B. Non-Intrusive Appliances Load Monitoring (NILM) for Energy Conservation in Household with Low Sampling Rate. **Procedia Computer Science**, v. 86, p. 172–175, 2016.

BOSCO, T. B.; GONÇALVES, F. A. S.; DE SOUZA, W. A. **A Comparative Study of Machine Learning Classifiers for Electric Load Disaggregation based on an extended NILM dataset**. 2021 14th IEEE International Conference on Industry Applications (INDUSCON). **Anais...** Em: 2021 14TH IEEE INTERNATIONAL CONFERENCE ON INDUSTRY APPLICATIONS (INDUSCON). São Paulo, Brazil: IEEE, 15 ago. 2021. Disponível em: <<https://ieeexplore.ieee.org/document/9529824/>>. Acesso em: 18 nov. 2021

BUCCI, G. et al. **Load Identification System for Residential Applications Based on the NILM Technique**. 2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC). **Anais...** Em: 2020 IEEE INTERNATIONAL INSTRUMENTATION AND MEASUREMENT TECHNOLOGY CONFERENCE (I2MTC). Dubrovnik, Croatia: IEEE, maio 2020. Disponível em: <<https://ieeexplore.ieee.org/document/9128599/>>. Acesso em: 18 nov. 2021

C, S.; KUMAR, M.; K, I. **Design and Implementation of Non-Intrusive Load Monitoring using Machine Learning Algorithm for Appliance Monitoring**. 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and

Signal Processing (INCOS). **Anais...** Em: 2019 IEEE INTERNATIONAL CONFERENCE ON INTELLIGENT TECHNIQUES IN CONTROL, OPTIMIZATION AND SIGNAL PROCESSING (INCOS). Tamilnadu, India: IEEE, abr. 2019. Disponível em: <<https://ieeexplore.ieee.org/document/8951312/>>. Acesso em: 17 jul. 2022

CAO, H. et al. **A Privacy Preserving Model for Energy Internet Base on Differential Privacy**. 2017 IEEE International Conference on Energy Internet (ICEI). **Anais...** Em: 2017 IEEE INTERNATIONAL CONFERENCE ON ENERGY INTERNET (ICEI). Beijing: IEEE, abr. 2017. Disponível em: <<https://ieeexplore.ieee.org/document/7926875/>>. Acesso em: 14 jul. 2022

CHANG, H.-H. et al. **Feature Extraction of Non-intrusive Load-Monitoring System Using Genetic Algorithm in Smart Meters**. 2011 IEEE 8th International Conference on e-Business Engineering. **Anais...** Em: 2011 IEEE 8TH INTERNATIONAL CONFERENCE ON E-BUSINESS ENGINEERING (ICEBE). Beijing, China: IEEE, out. 2011. Disponível em: <<http://ieeexplore.ieee.org/document/6104632/>>. Acesso em: 18 nov. 2021

CHANG, H.-H. et al. **Particle Swarm Optimization based non-intrusive demand monitoring and load identification in smart meters**. 2012 IEEE Industry Applications Society Annual Meeting. **Anais...** Em: 2012 IEEE INDUSTRY APPLICATIONS SOCIETY ANNUAL MEETING. Las Vegas, NV, USA: IEEE, out. 2012. Disponível em: <<http://ieeexplore.ieee.org/document/6373990/>>. Acesso em: 18 nov. 2021

CHEA, R.; THOURN, K.; CHHORN, S. **Improving V-I Trajectory Load Signature in NILM Approach**. 2022 International Electrical Engineering Congress (iEECON). **Anais...** Em: 2022 INTERNATIONAL ELECTRICAL ENGINEERING CONGRESS (iEECON). Khon Kaen, Thailand: IEEE, 9 mar. 2022. Disponível em: <<https://ieeexplore.ieee.org/document/9741688/>>. Acesso em: 15 jul. 2022

CIANCETTA, F. et al. A New Convolutional Neural Network-Based System for NILM Applications. **IEEE Transactions on Instrumentation and Measurement**, v. 70, p. 1–12, 2021.

CORTES, C.; VAPNIK, V. Support-vector networks. **Machine Learning**, v. 20, n. 3, p. 273–297, set. 1995.

DASH, S.; SODHI, R.; SODHI, B. **A Semi-Intrusive Load Monitoring Approach for Demand Response Applications**. 2019 8th International Conference on Power Systems (ICPS). **Anais...** Em: 2019 8TH INTERNATIONAL CONFERENCE ON POWER SYSTEMS (ICPS). Jaipur, India: IEEE, dez. 2019a. Disponível em: <<https://ieeexplore.ieee.org/document/9067320/>>. Acesso em: 18 nov. 2021

DASH, S.; SODHI, R.; SODHI, B. A Novel Instrumentation Approach for Clustered Appliance Load Monitoring. **IEEE Transactions on Power Delivery**, v. 34, n. 6, p. 2257–2259, dez. 2019b.

DEMSˆAR, J. et al. Orange: Data Mining Toolbox in Python. **Journal of Machine Learning Research**, v. 14, p. 2349–2353, 2013.

DONATO, P. G. et al. **Review of NILM applications in smart grids: power quality assessment and assisted independent living**. 2020 Argentine Conference on Automatic Control (AADECA). **Anais...** Em: 2020 ARGENTINE CONFERENCE ON AUTOMATIC CONTROL (AADECA). Buenos Aires, Argentina: IEEE, 28 out. 2020. Disponível em: <<https://ieeexplore.ieee.org/document/9301641/>>. Acesso em: 28 nov. 2021

EHRHARDT-MARTINEZ, K. et al. **Advanced Metering Initiatives and Residential Feedback Programs: A Meta-Review for Household Electricity-Saving Opportunities**. American Council for an Energy-Efficient Economy. p. 140, 2010.

ELETROBRAS, C. E. B. S. A. **Pesquisa de posse e hábitos de uso de equipamentos elétricos na classe residencial**. [s.l.] Eletrobras - Procel, 2019. Disponível em: <https://eletrobras.com/pt/SiteAssets/Paginas/PPH-2019/RESUMO_EXECUTIVO_BRASIL_BR.pdf>. Acesso em: 9 jun. 2022.

ELMSLEY, A. **Thinking artificially: Artificial Neural Networks, part I**. Disponível em: <<https://medium.com/the-sound-of-ai/thinking-artificially-artificial-neural-networks-part-i-fb0ea78a9f6>>. Acesso em: 17 jul. 2022a.

ELMSLEY, A. **Training to train: Artificial Neural Networks, part II**. Disponível em: <<https://medium.com/the-sound-of-ai/training-to-train-artificial-neural-networks-part-ii-b1b4efd944be>>. Acesso em: 17 jul. 2022b.

ENERGY PARTNERSHIP, G.-B. **Uso de novas tecnologias digitais para medição de consumo de energia e níveis de eficiência energética no Brasil - Baseado nas experiências da Alemanha: Bilateral Energy Partnerships in Developing and Emerging Countries**. [s.l: s.n.]. Disponível em: <www.energypartnership.com.br>. Acesso em: 15 out. 2022.

EPE. **Estudos de Longo Prazo - Considerações sobre o Comportamento do Consumidor**. Empresa de Pesquisa Energética, , dez. 2018. Disponível em: <<https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-227/topico-457/Comportamento%20do%20Consumidor.pdf>>. Acesso em: 9 jun. 2022

FAUSTINE, A. et al. **A Survey on Non-Intrusive Load Monitoring Methodies and Techniques for Energy Disaggregation Problem**. arXiv, , 10 mar. 2017. Disponível em: <<http://arxiv.org/abs/1703.00785>>. Acesso em: 10 jun. 2022

FILHO, S. D. M. **Medição de energia elétrica**. 4. ed. Rio de Janeiro: Editora LTC, 1997.

FUNGSIRIRUT, W.; BENJAPOLAKUL, W. **A study on energy reduction based on responsive behavior of users**. 2017 14th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON). **Anais...** Em: 2017 14TH INTERNATIONAL CONFERENCE ON ELECTRICAL ENGINEERING/ELECTRONICS, COMPUTER, TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY (ECTI-CON). Phuket: IEEE, jun. 2017. Disponível em: <<http://ieeexplore.ieee.org/document/8096304/>>. Acesso em: 28 nov. 2021

FURTADO, M. I. V. **Redes Neurais Artificiais: Uma Abordagem Para Sala de Aula**. 1. ed. Ponta Grossa, PR: Atena Editora, 2019.

GAO, J. et al. **PLAID: a public dataset of high-resolution electrical appliance measurements for load identification research: demo abstract**. Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings. **Anais...** Em: SENSYS '14: THE 12TH ACM CONFERENCE ON EMBEDDED NETWORK SENSOR SYSTEMS. Memphis Tennessee: ACM, 3 nov. 2014. Disponível em: <<https://dl.acm.org/doi/10.1145/2674061.2675032>>. Acesso em: 28 nov. 2021

GARCIA, F. D. et al. NILM-based approach for energy efficiency assessment of household appliances. **Energy Informatics**, v. 3, n. 1, p. 10, dez. 2020.

GOV.UK. **Smart Metering Implementation Programme: a report on progress of the realisation of smart meter consumer benefits**. [s.l.] Department for Business, Energy & Industrial Strategy, 2019. Disponível em: <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/830668/smart-meters-benefits-realisation.pdf>. Acesso em: 25 nov. 2021.

GURBUZ, F. B.; BAYINDIR, R.; VADI, S. **Comprehensive Non-Intrusive Load Monitoring Process: Device Event Detection, Device Feature Extraction and Device Identification Using KNN, Random Forest and Decision Tree**. 2021 10th International Conference on Renewable Energy Research and Application (ICRERA). **Anais...** Em: 2021 10TH INTERNATIONAL CONFERENCE ON RENEWABLE ENERGY RESEARCH AND APPLICATION (ICRERA). Istanbul, Turkey: IEEE, 26 set. 2021. Disponível em: <<https://ieeexplore.ieee.org/document/9598679/>>. Acesso em: 15 jul. 2022

HARSHA S. ABEYKOON, A. M. et al. **A Machine Learning Approach for NILM based on Superimposed Current Profiles**. 2020 Moratuwa Engineering Research Conference (MERCon). **Anais...** Em: 2020 MORATUWA ENGINEERING RESEARCH CONFERENCE (MERCON). Moratuwa, Sri Lanka: IEEE, jul. 2020. Disponível em: <<https://ieeexplore.ieee.org/document/9185203/>>. Acesso em: 10 jun. 2022

HART, G. W. Nonintrusive appliance load monitoring. **Proceedings of the IEEE**, v. 80, n. 12, p. 1870–1891, dez. 1992.

HE, J. et al. An Efficient and Accurate Nonintrusive Load Monitoring Scheme for Power Consumption. **IEEE Internet of Things Journal**, v. 6, n. 5, p. 9054–9063, out. 2019.

HOLCOMB, C. **Pecan Street Inc. : A Test-bed for NILM**. . Em: PRESENTED AT THE INTERNATIONAL WORKSHOP ON NON-INTRUSIVE LOAD MONITORING. Pittsburgh, PA: 2014.

HOSSEINI, S. S. et al. Non-intrusive load monitoring through home energy management systems: A comprehensive review. **Renewable and Sustainable Energy Reviews**, v. 79, p. 1266–1274, nov. 2017.

HOUIDI, S. et al. **Statistical assessment of abrupt change detectors for non-intrusive load monitoring**. 2018 IEEE International Conference on Industrial Technology (ICIT). **Anais...** Em: 2018 IEEE INTERNATIONAL CONFERENCE ON INDUSTRIAL TECHNOLOGY (ICIT). Lyon: IEEE, fev. 2018. Disponível em: <<https://ieeexplore.ieee.org/document/8352368/>>. Acesso em: 22 jul. 2022

IWAYEMI, A.; ZHOU, C. **Leveraging smart meters for residential energy disaggregation**. 2014 IEEE PES General Meeting | Conference & Exposition. **Anais...** Em: 2014 IEEE PES GENERAL MEETING: CONFERENCE & EXPOSITION. National Harbor, MD: IEEE, jul. 2014. Disponível em: <<https://ieeexplore.ieee.org/document/6939461/>>. Acesso em: 18 nov. 2021

JAIN, V. **Multi-layer perceptrons as non-linear classifiers**. Disponível em: <<https://medium.com/the-feynman-journal/the-linear-and-nonlinear-nature-of-feedforward-84199eb3edea>>. Acesso em: 17 jul. 2022.

JIN, Y. et al. **Robust adaptive event detection in non-intrusive load monitoring for energy aware smart facilities**. 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). **Anais...** Em: ICASSP 2011 - 2011 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP). Prague, Czech Republic: IEEE, maio 2011. Disponível em: <<http://ieeexplore.ieee.org/document/5947314/>>. Acesso em: 18 nov. 2021

JUNG, T. **The Linear and Nonlinear Nature of Feedforward**. Disponível em: <<https://medium.com/the-feynman-journal/the-linear-and-nonlinear-nature-of-feedforward-84199eb3edea>>. Acesso em: 17 jul. 2022.

KAHL, M. et al. **WHITED - A Worldwide Household and Industry Transient Energy Data Set**. . Em: PRESENTED AT THE INTERNATIONAL WORKSHOP ON NON-INTRUSIVE LOAD MONITORING. Vancouver, BC, Canada: 2016.

KELLY, J.; KNOTTENBELT, W. The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. **Scientific Data**, v. 2, n. 1, p. 150007, dez. 2015.

KELLY, J.; KNOTTENBELT, W. **Does disaggregated electricity feedback reduce domestic electricity consumption? A systematic review of the literature**. 3rd International NILM Workshop. **Anais...** Em: 3RD INTERNATIONAL NILM WORKSHOP. 2016.

KHAN, M. M. R.; SIDDIQUE, MD. A. B.; SAKIB, S. **Non-Intrusive Electrical Appliances Monitoring and Classification using K-Nearest Neighbors**. 2019 2nd International Conference on Innovation in Engineering and Technology (ICIET). **Anais...** Em: 2019 2ND INTERNATIONAL CONFERENCE ON INNOVATION IN ENGINEERING AND TECHNOLOGY (ICIET). Dhaka, Bangladesh: IEEE, 23 dez. 2019. Disponível em: <<https://ieeexplore.ieee.org/document/9290671/>>. Acesso em: 15 jul. 2022

KOLTER, J. Z.; JOHNSON, M. J. **REDD: A Public Data Set for Energy Disaggregation Research**. Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability (SustKDD). **Anais...**2011.

LANDIS, J. R.; KOCH, G. G. The Measurement of Observer Agreement for Categorical Data. **Biometrics**, v. 33, n. 1, p. 159, mar. 1977.

LANGEVIN, A.; GAGNON, G.; CHERIET, M. Crosstalk Suppression in Semi-Intrusive Load Monitoring Systems Using Hall Effect Sensors. **IEEE Transactions on Smart Grid**, v. 11, n. 6, p. 5019–5027, nov. 2020.

LAOUALI, I. H. et al. **A Survey On Computational Intelligence Techniques For Non Intrusive Load Monitoring**. 2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS). **Anais...** Em: 2020 IEEE 2ND INTERNATIONAL CONFERENCE ON ELECTRONICS, CONTROL, OPTIMIZATION AND COMPUTER SCIENCE (ICECOCS). Kenitra, Morocco: IEEE, 2 dez. 2020. Disponível em: <<https://ieeexplore.ieee.org/document/9314383/>>. Acesso em: 18 nov. 2021

LIU, Q. et al. Low-Complexity Non-Intrusive Load Monitoring Using Unsupervised Learning and Generalized Appliance Models. **IEEE Transactions on Consumer Electronics**, v. 65, n. 1, p. 28–37, fev. 2019.

LIU, Y.; CHEN, M. **A review of nonintrusive load monitoring and its application in commercial building**. The 4th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent. **Anais...** Em: 2014 IEEE 4TH ANNUAL INTERNATIONAL CONFERENCE ON CYBER TECHNOLOGY IN AUTOMATION, CONTROL, AND INTELLIGENT SYSTEMS (CYBER). Hong Kong, Hong Kong: IEEE, jun. 2014. Disponível em: <<https://ieeexplore.ieee.org/document/6917536>>. Acesso em: 18 nov. 2021

LIU, Y.; WANG, X.; YOU, W. Non-Intrusive Load Monitoring by Voltage–Current Trajectory Enabled Transfer Learning. **IEEE Transactions on Smart Grid**, v. 10, n. 5, p. 5609–5619, set. 2019.

LUAN, W. et al. An Adaptive Two-Stage Load Event Detection Method for Nonintrusive Load Monitoring. **IEEE Transactions on Instrumentation and Measurement**, v. 71, p. 1–14, 2022.

LUO, D. et al. **Monitoring HVAC Equipment Electrical Loads from a Centralized Location—Methods and Field Test Results**. ASHRAE Transactions. **Anais...**: 1. Em: ASHRAE TRANSACTIONS: SYMPOSIA. 2002.

MAKONIN, S. et al. **AMPds: A public dataset for load disaggregation and eco-feedback research**. 2013 IEEE Electrical Power & Energy Conference. **Anais...** Em: 2013 IEEE ELECTRICAL POWER & ENERGY CONFERENCE (EPEC). Halifax, NS, Canada: IEEE, ago. 2013. Disponível em: <<http://ieeexplore.ieee.org/document/6802949/>>. Acesso em: 28 nov. 2021

MEDEIROS, A. P. **Classificação de eventos em monitoramento NILM de cargas elétricas residenciais utilizando rede neural convolucional**. Pós-Graduação em Engenharia Elétrica—Santa Maria, RS: Universidade Federal de Santa Maria, 2019.

MEDEIROS, A. P. et al. **Event Classification in Non-Intrusive Load Monitoring Using Convolutional Neural Network**. 2019 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America). **Anais...** Em: 2019 IEEE PES INNOVATIVE SMART GRID TECHNOLOGIES CONFERENCE - LATIN AMERICA (ISGT LATIN AMERICA). Gramado, Brazil: IEEE, set. 2019. Disponível em: <<https://ieeexplore.ieee.org/document/8895291/>>. Acesso em: 18 nov. 2021

MORADZADEH, A. et al. **Support Vector Machine-Assisted Improvement Residential Load Disaggregation**. 2020 28th Iranian Conference on Electrical Engineering (ICEE). **Anais...** Em: 2020 28TH IRANIAN CONFERENCE ON ELECTRICAL ENGINEERING (ICEE). Tabriz, Iran: IEEE, 4 ago. 2020. Disponível em: <<https://ieeexplore.ieee.org/document/9260869/>>. Acesso em: 18 nov. 2021

MURRAY, D. et al. A data management platform for personalised real-time energy feedback. p. 15, 2015.

NGUYEN, M. et al. **A novel feature extraction and classification algorithm based on power components using single-point monitoring for NILM**. 2015 IEEE 28th Canadian Conference on Electrical and Computer Engineering (CCECE). **Anais...** Em: 2015 IEEE 28TH CANADIAN CONFERENCE ON ELECTRICAL AND COMPUTER ENGINEERING (CCECE). Halifax, NS, Canada: IEEE, maio 2015. Disponível em: <<http://ieeexplore.ieee.org/document/7129156/>>. Acesso em: 18 nov. 2021

NGUYEN, V. K. et al. **Demo Abstract: An End-to-End Real-Time Efficient System for Smart Energy Monitoring**. IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). **Anais...** Em: IEEE INFOCOM 2019 - IEEE CONFERENCE ON COMPUTER COMMUNICATIONS WORKSHOPS (INFOCOM WKSHPS). Paris, France: IEEE, abr. 2019. Disponível em: <<https://ieeexplore.ieee.org/document/8845260/>>. Acesso em: 18 set. 2022

NUNES, N. J.; PEREIRA, L. **Semi-automatic labeling for public non-intrusive load monitoring datasets**. 2015 Sustainable Internet and ICT for Sustainability (SustainIT). **Anais...** Em: 2015 SUSTAINABLE INTERNET AND ICT FOR SUSTAINABILITY (SUSTAINIT). Madrid, Spain: IEEE, abr. 2015. Disponível em: <<http://ieeexplore.ieee.org/document/7101378/>>. Acesso em: 18 nov. 2021

ONU. **Transformando Nosso Mundo: A Agenda 2030 para o Desenvolvimento Sustentável**. Organização das Nações Unidas, , 2015. Disponível em: <<https://www.undp.org/content/dam/brazil/docs/agenda2030/undp-br-Agenda2030-completo-pt-br-2016.pdf>>. Acesso em: 25 nov. 2021

PEREIRA, L. **Developing and evaluating a probabilistic event detector for non-intrusive load monitoring**. 2017 Sustainable Internet and ICT for Sustainability (SustainIT). **Anais...** Em: 2017 SUSTAINABLE INTERNET AND ICT FOR SUSTAINABILITY (SUSTAINIT). Funchal: IEEE, dez. 2017. Disponível em: <<https://ieeexplore.ieee.org/document/8379796/>>. Acesso em: 18 nov. 2021

PHAN, M. H. et al. CompactNet: A Light-Weight Deep Learning Framework for Smart Intrusive Load Monitoring. **IEEE Sensors Journal**, v. 21, n. 22, p. 25181–25189, 15 nov. 2021.

PICON, T. et al. COOLL: Controlled On/Off Loads Library, a Public Dataset of High-Sampled Electrical Signals for Appliance Identification. **arXiv:1611.05803 [cs]**, 17 nov. 2016.

PIRBAZARI, A. M. et al. **Improving Load Forecast Accuracy of Households Using Load Disaggregation Techniques**. 2020 International Conferences on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics). **Anais...** Em: 2020 INTERNATIONAL CONFERENCES ON INTERNET OF THINGS (ITHINGS) AND IEEE GREEN COMPUTING AND COMMUNICATIONS (GREENCOM) AND IEEE CYBER, PHYSICAL AND SOCIAL COMPUTING (CPSCOM) AND IEEE SMART DATA (SMARTDATA) AND IEEE CONGRESS ON CYBERMATICS (CYBERMATICS). Rhodes, Greece: IEEE, nov. 2020. Disponível em: <<https://ieeexplore.ieee.org/document/9291555/>>. Acesso em: 15 jul. 2022

PONRAK, T.; DECHANUPAPRITTHA, S. **A Semi-supervised Method of Non-intrusive Load Monitoring with Answer-Question Technique**. 2021 9th International Electrical Engineering Congress (iEECON). **Anais...** Em: 2021 9TH INTERNATIONAL ELECTRICAL ENGINEERING CONGRESS (IEECON). Pattaya, Thailand: IEEE, 10 mar. 2021. Disponível em: <<https://ieeexplore.ieee.org/document/9440251/>>. Acesso em: 18 nov. 2021

RAIKER, G. A. et al. **Approach to Non-Intrusive Load Monitoring using Factorial Hidden Markov Model**. 2018 IEEE 13th International Conference on Industrial and Information Systems (ICIIS). **Anais...** Em: 2018 IEEE 13TH INTERNATIONAL CONFERENCE ON INDUSTRIAL AND INFORMATION SYSTEMS (ICIIS). Rupnagar, India: IEEE, dez. 2018. Disponível em: <<https://ieeexplore.ieee.org/document/8721436/>>. Acesso em: 18 nov. 2021

RAO, K. M.; RAVICHANDRAN, D.; MAHESH, K. Non-Intrusive Load Monitoring and Analytics for Device Prediction. **Hong Kong**, p. 6, 2016.

REDDY, R.; GARG, V.; PUDI, V. A feature fusion technique for improved non-intrusive load monitoring. **Energy Informatics**, v. 3, n. 1, p. 9, dez. 2020.

REHMAN, A. U. et al. Comparative Evaluation of Machine Learning Models and Input Feature Space for Non-intrusive Load Monitoring. **Journal of Modern Power Systems and Clean Energy**, v. 9, n. 5, p. 1161–1171, 2021.

RENAUX, D. et al. **Designing a Novel Dataset for Non-intrusive Load Monitoring**. 2018 VIII Brazilian Symposium on Computing Systems Engineering (SBESC). **Anais...** Em: 2018 VIII BRAZILIAN SYMPOSIUM ON COMPUTING SYSTEMS ENGINEERING (SBESC). Salvador, Brazil: IEEE, nov. 2018. Disponível em: <<https://ieeexplore.ieee.org/document/8691975/>>. Acesso em: 18 nov. 2021

RIDI, A.; GISLER, C.; HENNEBERT, J. **A Survey on Intrusive Load Monitoring for Appliance Recognition**. 2014 22nd International Conference on Pattern Recognition. **Anais...** Em: 2014 22ND INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION (ICPR). Stockholm, Sweden: IEEE, ago. 2014. Disponível em: <<http://ieeexplore.ieee.org/document/6977348/>>. Acesso em: 18 nov. 2021

ROMANCINI, E. M. R. **Desenvolvimento de Smart Meter integrado com IoT para monitoramento da energia elétrica em residências**. Graduação em Engenharia de Computação—Araranguá, SC: Universidade Federal de Santa Catarina, 2020.

ROMANCINI, E. M. R. et al. **Monitoramento inteligente do consumo de energia elétrica em residências utilizando recursos de IoT**. Anais do XIII Computer on the Beach - COTB'22. **Anais...** Em: COMPUTER ON THE BEACH. Itajaí - Santa Catarina - Brasil: Universidade do Vale do Itajaí, 13 jul. 2022. Disponível em: <<https://periodicos.univali.br/index.php/acotb/article/view/18743>>. Acesso em: 22 jul. 2022

SAHA, D. et al. Comprehensive NILM Framework: Device Type Classification and Device Activity Status Monitoring Using Capsule Network. **IEEE Access**, v. 8, p. 179995–180009, 2020.

SAKIYAMA, Y. et al. Predicting human liver microsomal stability with machine learning techniques. **Journal of Molecular Graphics and Modelling**, v. 26, n. 6, p. 907–915, fev. 2008.

SHAO, H.; MARWAH, M.; RAMAKRISHNAN, N. **A Temporal Motif Mining Approach to Unsupervised Energy Disaggregation: Applications to Residential and Commercial Buildings**. AAAI Conference on Artificial Intelligence. **Anais...** Em: AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE. 2013.

SHARMA, S. **Activation Functions in Neural Networks**. Disponível em: <<https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6>>. Acesso em: 17 jul. 2022.

SHMUELI, B. **Multi-Class Metrics Made Simple, Part III: the Kappa Score (aka Cohen's Kappa Coefficient)**. Disponível em: <<https://towardsdatascience.com/multi-class-metrics-made-simple-the-kappa-score-aka-cohens-kappa-coefficient-bdea137af09c>>. Acesso em: 28 jul. 2022.

SINGH, A. P. **Steps You Should Follow To Successfully Train MLP**. Disponível em: <<https://medium.com/analytics-vidhya/steps-you-should-follow-to-successfully-train-mlp-40a98c3b5bb3>>. Acesso em: 17 jul. 2022.

SOUZA, A. L. P. DE S. **Hardware and software platforms to deploy and evaluate non-intrusive load monitoring systems**. Doctorate in Informatics Engineering specialty: Software Engineering—Portugal: Universidade da Madeira, 2016.

SUN, M. et al. **Non-Intrusive Load Monitoring System Framework and Load Disaggregation Algorithms: A Survey**. 2019 International Conference on Advanced Mechatronic Systems (ICAMechS). **Anais...** Em: 2019 INTERNATIONAL CONFERENCE ON ADVANCED MECHATRONIC SYSTEMS (ICAMECHS).

Kusatsu, Shiga, Japan: IEEE, ago. 2019. Disponível em: <<https://ieeexplore.ieee.org/document/8861646/>>. Acesso em: 28 nov. 2021

TANG, G.; WU, K.; LEI, J. A Distributed and Scalable Approach to Semi-Intrusive Load Monitoring. **IEEE Transactions on Parallel and Distributed Systems**, v. 27, n. 6, p. 1553–1565, 1 jun. 2016.

TAO, P. et al. **Multi-level non-intrusive load identification based on k-NN**. 2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2). **Anais...** Em: 2019 IEEE 3RD CONFERENCE ON ENERGY INTERNET AND ENERGY SYSTEM INTEGRATION (EI2). Changsha, China: IEEE, nov. 2019. Disponível em: <<https://ieeexplore.ieee.org/document/9061896/>>. Acesso em: 17 jul. 2022

TIAN, J. et al. **Residential load disaggregation based on resident behavior learning and neural networks**. 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2). **Anais...** Em: 2017 IEEE CONFERENCE ON ENERGY INTERNET AND ENERGY SYSTEM INTEGRATION (EI2). Beijing: IEEE, nov. 2017. Disponível em: <<http://ieeexplore.ieee.org/document/8245665/>>. Acesso em: 18 nov. 2021

VEEN, F. V. **The Neural Network Zoo**. Disponível em: <<https://www.asimovinstitute.org/neural-network-zoo/>>. Acesso em: 10 jul. 2022.

VÖLKER, B. et al. **Annoticity: A Smart Annotation Tool and Data Browser for Electricity Datasets**. Proceedings of the 5th International Workshop on Non-Intrusive Load Monitoring. **Anais...** Em: BUILDSYS '20: THE 7TH ACM INTERNATIONAL CONFERENCE ON SYSTEMS FOR ENERGY-EFFICIENT BUILDINGS, CITIES, AND TRANSPORTATION. Virtual Event Japan: ACM, 18 nov. 2020a. Disponível em: <<https://dl.acm.org/doi/10.1145/3427771.3427844>>. Acesso em: 18 nov. 2021

VÖLKER, B. et al. **FIREd: A Fully-labeled high-fRequency Electricity Disaggregation Dataset**. Proceedings of the 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation. **Anais...** Em: BUILDSYS '20: THE 7TH ACM INTERNATIONAL CONFERENCE ON SYSTEMS FOR ENERGY-EFFICIENT BUILDINGS, CITIES, AND TRANSPORTATION. Virtual Event Japan: ACM, 18 nov. 2020b. Disponível em: <<https://dl.acm.org/doi/10.1145/3408308.3427623>>. Acesso em: 28 nov. 2021

VÖLKER, B.; SCHOLL, P. M.; BECKER, B. **Semi-Automatic Generation and Labeling of Training Data for Non-intrusive Load Monitoring**. Proceedings of the Tenth ACM International Conference on Future Energy Systems. **Anais...** Em: E-ENERGY '19: THE TENTH ACM INTERNATIONAL CONFERENCE ON FUTURE ENERGY SYSTEMS. Phoenix AZ USA: ACM, 15 jun. 2019. Disponível em: <<https://dl.acm.org/doi/10.1145/3307772.3328295>>. Acesso em: 18 nov. 2021

WANG, H. et al. **Privacy Preserving in Non-Intrusive Load Monitoring: A Differential Privacy Perspective**. 2021 IEEE Power & Energy Society General Meeting (PESGM). **Anais...** Em: 2021 IEEE POWER & ENERGY SOCIETY GENERAL MEETING (PESGM). Washington, DC, USA: IEEE, 26 jul. 2021.

Disponível em: <<https://ieeexplore.ieee.org/document/9638107/>>. Acesso em: 14 jul. 2022

WEISSHAAR, D. et al. **Device Classification for NILM using FIT-PS compared with Standard Signal Forms**. 2018 International IEEE Conference and Workshop in Óbuda on Electrical and Power Engineering (CANDO-EPE). **Anais...** Em: 2018 INTERNATIONAL IEEE CONFERENCE AND WORKSHOP IN ÓBUDA ON ELECTRICAL AND POWER ENGINEERING (CANDO-EPE). Budapest, Hungary: IEEE, nov. 2018. Disponível em: <<https://ieeexplore.ieee.org/document/8601150/>>. Acesso em: 18 nov. 2021

WILD, B.; BARSIM, K. S.; YANG, B. **A new unsupervised event detector for non-intrusive load monitoring**. 2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP). **Anais...** Em: 2015 IEEE GLOBAL CONFERENCE ON SIGNAL AND INFORMATION PROCESSING (GLOBALSIP). Orlando, FL, USA: IEEE, dez. 2015. Disponível em: <<http://ieeexplore.ieee.org/document/7418159/>>. Acesso em: 18 nov. 2021

YAEMPRAYOON, S.; SRINONCHAT, J. **Non-Intrusive Load Monitoring using Multi-Layer Perceptron for Appliances Classification**. 2022 19th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON). **Anais...** Em: 2022 19TH INTERNATIONAL CONFERENCE ON ELECTRICAL ENGINEERING/ELECTRONICS, COMPUTER, TELECOMMUNICATIONS AND INFORMATION TECHNOLOGY (ECTI-CON). Prachuap Khiri Khan, Thailand: IEEE, 24 maio 2022. Disponível em: <<https://ieeexplore.ieee.org/document/9795518/>>. Acesso em: 18 jul. 2022

YAN, Q.; XUDONG, W.; ZUN, W. **Applications of NILM in the Optimization Management of Intelligent Home Energy Management System**. 2020 IEEE Sustainable Power and Energy Conference (iSPEC). **Anais...** Em: 2020 IEEE SUSTAINABLE POWER AND ENERGY CONFERENCE (ISPEC). Chengdu, China: IEEE, 23 nov. 2020. Disponível em: <<https://ieeexplore.ieee.org/document/9351198/>>. Acesso em: 18 nov. 2021

YASIN, A.; KHAN, S. A. **Unsupervised Event Detection and On-Off Pairing Approach Applied to NILM**. 2018 International Conference on Frontiers of Information Technology (FIT). **Anais...** Em: 2018 INTERNATIONAL CONFERENCE ON FRONTIERS OF INFORMATION TECHNOLOGY (FIT). Islamabad, Pakistan: IEEE, dez. 2018. Disponível em: <<https://ieeexplore.ieee.org/document/8616978/>>. Acesso em: 22 jul. 2022

ZHANG, L.; ZHU, L. **Basic Summary of Non-intrusive Load Monitoring**. 2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS). **Anais...** Em: 2019 IEEE 10TH INTERNATIONAL CONFERENCE ON SOFTWARE ENGINEERING AND SERVICE SCIENCE (ICSESS). Beijing, China: IEEE, out. 2019. Disponível em: <<https://ieeexplore.ieee.org/document/9040726/>>. Acesso em: 18 nov. 2021

ZHANG, X. et al. **IEHouse: A non-intrusive household appliance state recognition system**. 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI). **Anais...** Em: 2017 IEEE SMARTWORLD, UBIQUITOUS INTELLIGENCE & COMPUTING, ADVANCED & TRUSTED COMPUTED, SCALABLE COMPUTING & COMMUNICATIONS, CLOUD & BIG DATA COMPUTING, INTERNET OF PEOPLE AND SMART CITY INNOVATION (SMARTWORLD/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI). San Francisco, CA: IEEE, ago. 2017. Disponível em: <<https://ieeexplore.ieee.org/document/8397510/>>. Acesso em: 18 nov. 2021

ZHANG, X.; ZHANG, R.; WU, F. **A Method of Conforming Feature Extraction in Non - Invasive Load Monitoring**. 2020 IEEE 9th International Power Electronics and Motion Control Conference (IPEMC2020-ECCE Asia). **Anais...** Em: 2020 IEEE 9TH INTERNATIONAL POWER ELECTRONICS AND MOTION CONTROL CONFERENCE (IPEMC2020-ECCE ASIA). Nanjing, China: IEEE, 29 nov. 2020. Disponível em: <<https://ieeexplore.ieee.org/document/9367739/>>. Acesso em: 18 nov. 2021

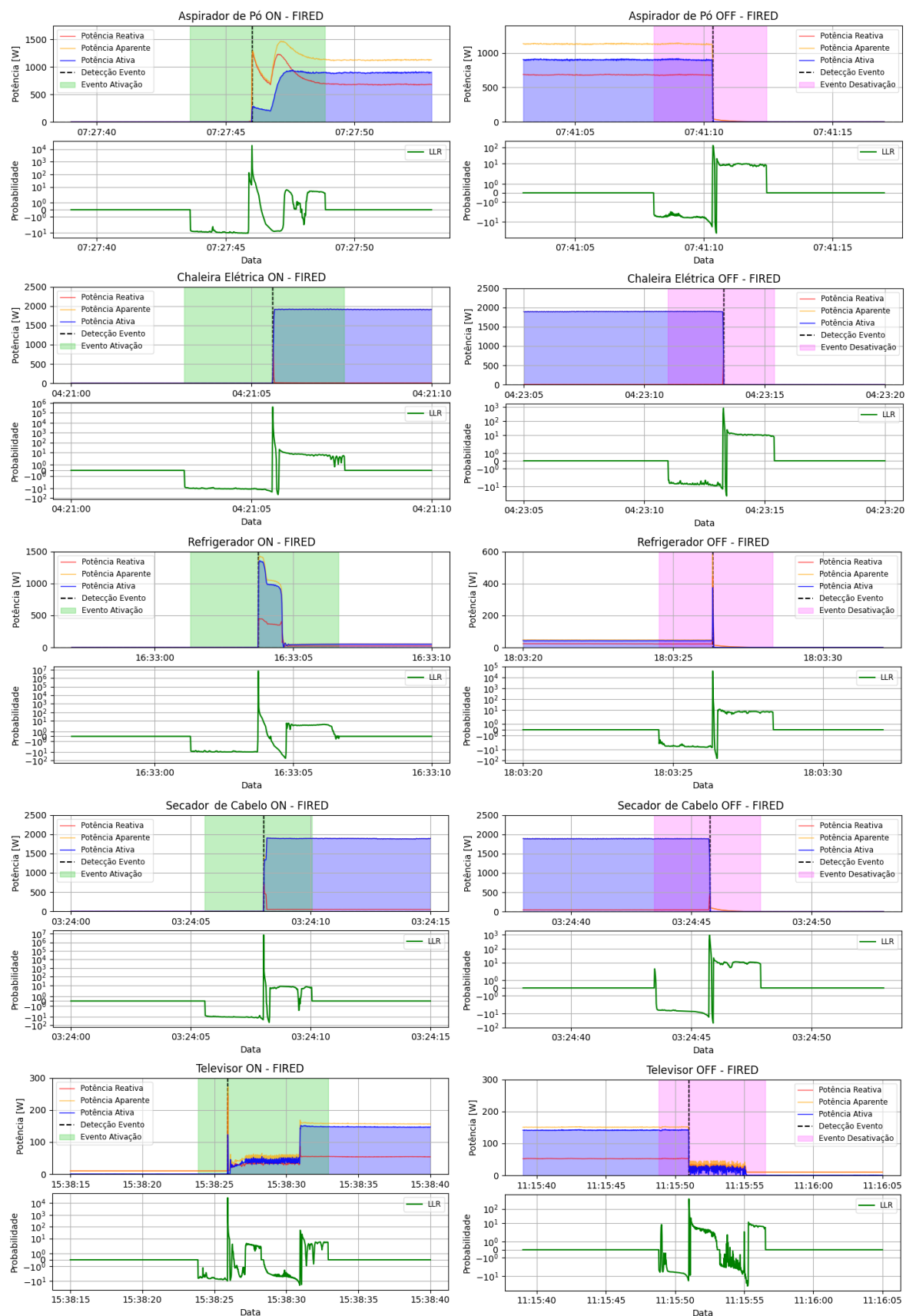
ZHAO, H. et al. **Research on online non-intrusive load identification system based on multi-threaded CUSUM-MLP algorithm**. 2021 IEEE Sensors. **Anais...** Em: 2021 IEEE SENSORS. Sydney, Australia: IEEE, 31 out. 2021. Disponível em: <<https://ieeexplore.ieee.org/document/9639843/>>. Acesso em: 18 jul. 2022

ZHOU, Q. et al. **Feature Extraction for Non-intrusive Load Monitoring System**. 2021 6th Asia Conference on Power and Electrical Engineering (ACPEE). **Anais...** Em: 2021 6TH ASIA CONFERENCE ON POWER AND ELECTRICAL ENGINEERING (ACPEE). Chongqing, China: IEEE, abr. 2021. Disponível em: <<https://ieeexplore.ieee.org/document/9436971/>>. Acesso em: 18 nov. 2021

ZHUANG, M.; SHAHIDEHPOUR, M.; LI, Z. **An Overview of Non-Intrusive Load Monitoring: Approaches, Business Applications, and Challenges**. 2018 International Conference on Power System Technology (POWERCON). **Anais...** Em: 2018 INTERNATIONAL CONFERENCE ON POWER SYSTEM TECHNOLOGY (POWERCON). Guangzhou: IEEE, nov. 2018. Disponível em: <<https://ieeexplore.ieee.org/document/8601534/>>. Acesso em: 18 nov. 2021

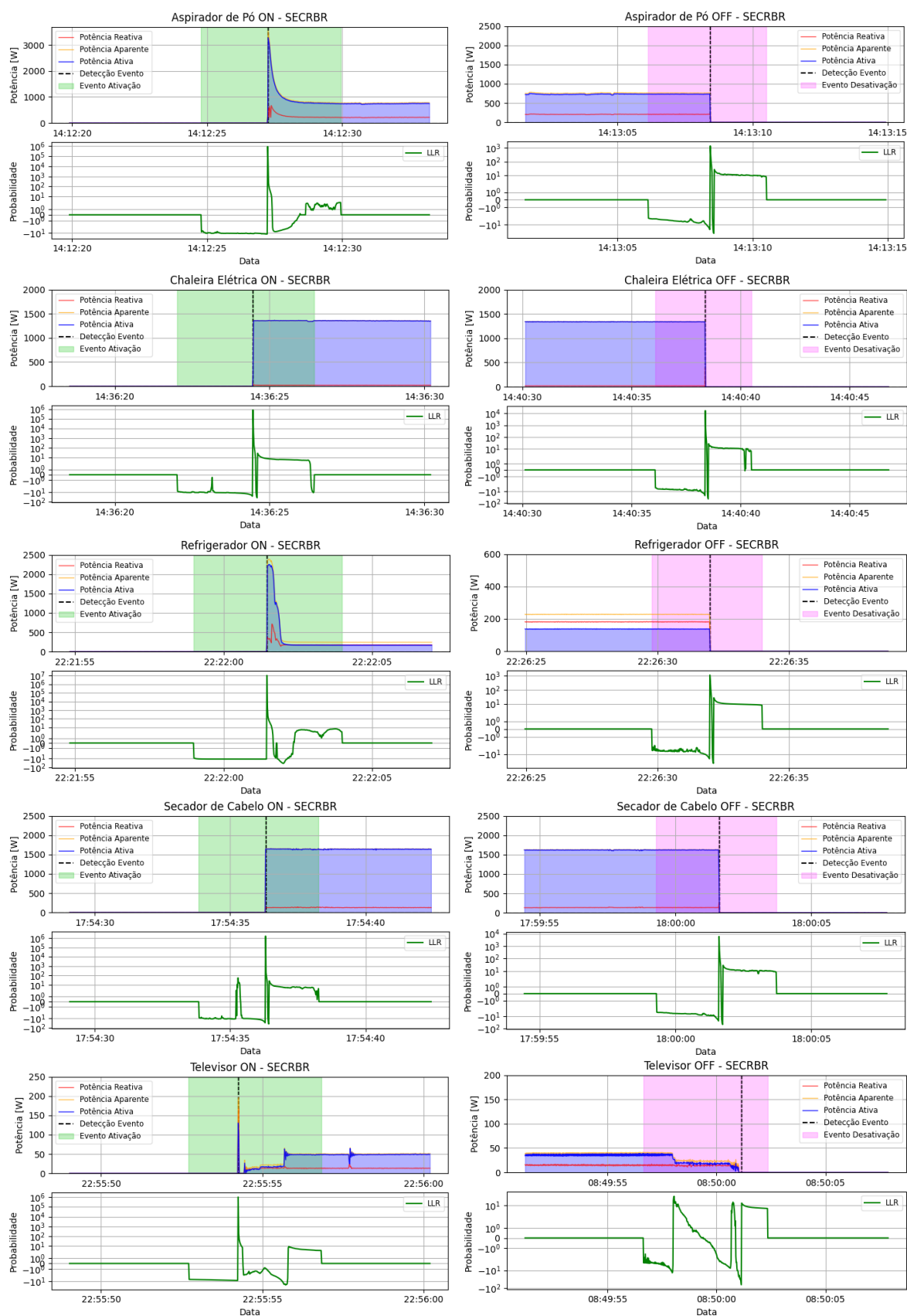
ZOHA, A. et al. Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey. **Sensors**, v. 12, n. 12, p. 16838–16866, 6 dez. 2012.

APÊNDICE A – Eventos de ativação e desativação do conjunto público *FIRE*



Fonte: Autor.

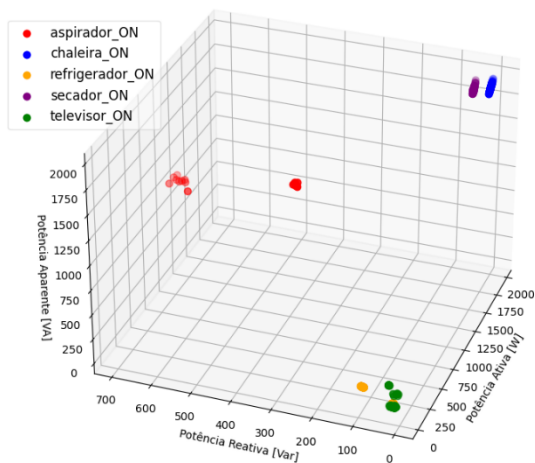
APÊNDICE B – Eventos de ativação e desativação do conjunto de sinais do ambiente real SECRBR



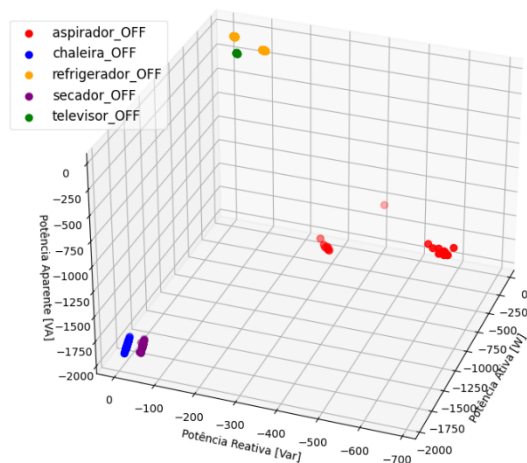
Fonte: Autor.

APÊNDICE C – Recursos no plano PQS do conjunto público *FIRED*

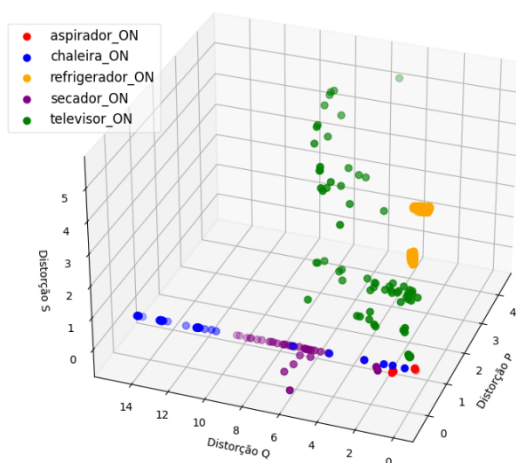
Diferença de potencial nos eventos de ativação - FIRED



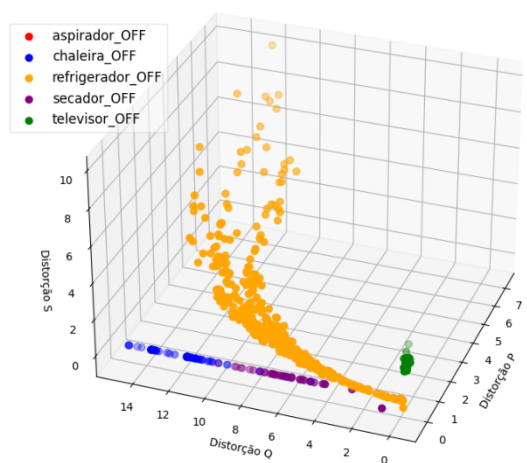
Diferença de potencial nos eventos de desativação - FIRED



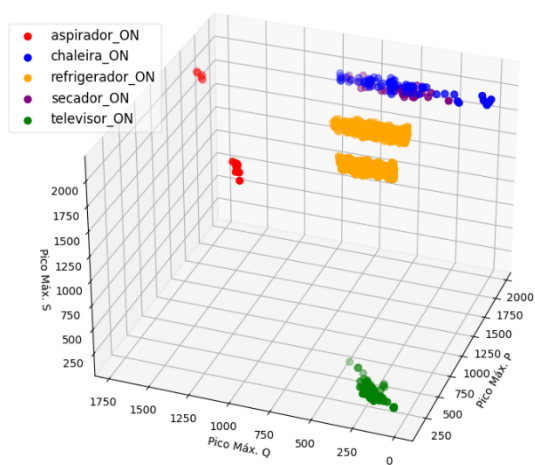
Distorção nos eventos de ativação - FIRED



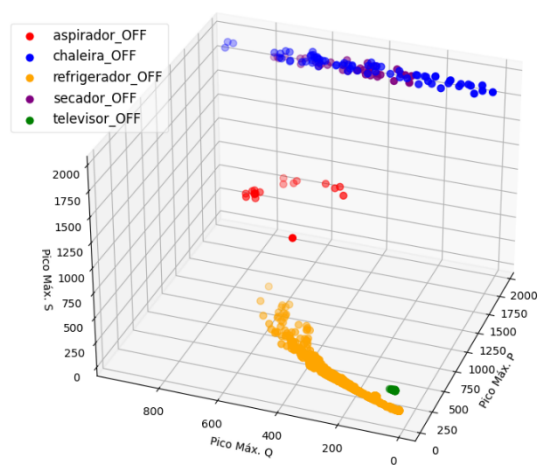
Distorção nos eventos de desativação - FIRED



Pico máximo nos eventos de ativação - FIRED



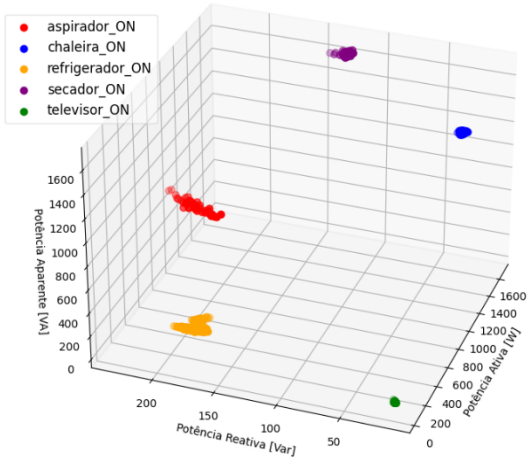
Pico máximo nos eventos de desativação - FIRED



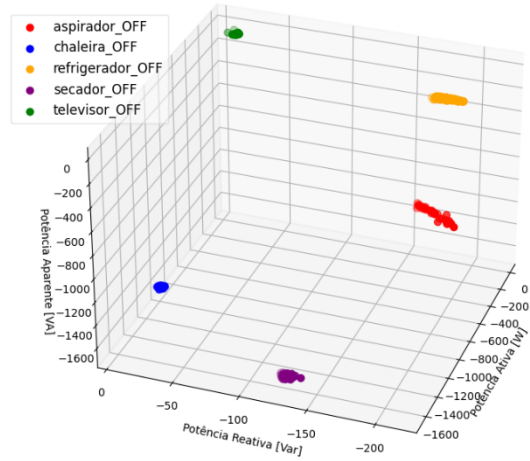
Fonte: Autor.

APÊNDICE D – Recursos no plano PQS do conjunto de sinais do ambiente real SECRBR

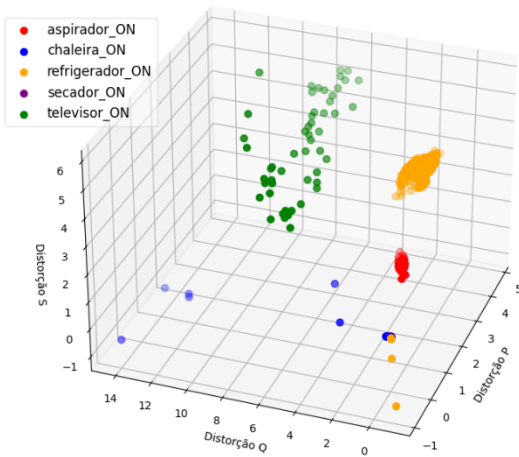
Diferença de potencial nos eventos de ativação - SECRBR



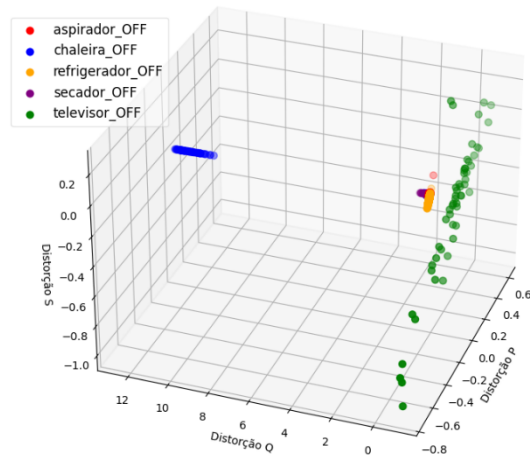
Diferença de potencial nos eventos de desativação - SECRBR



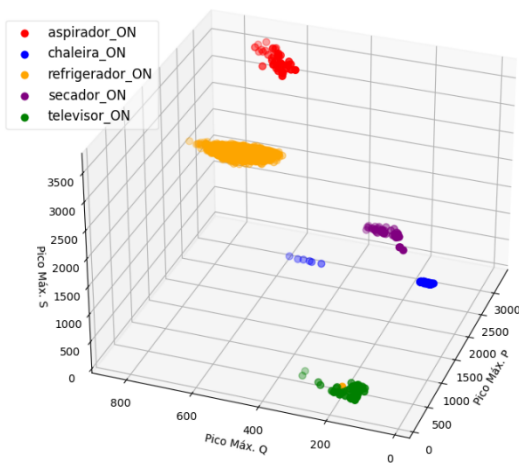
Distorção nos eventos de ativação - SECRBR



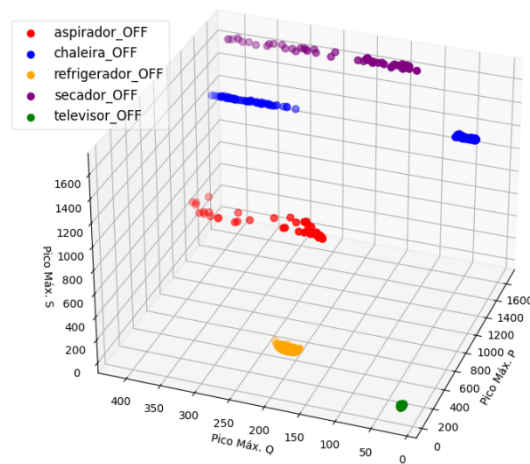
Distorção nos eventos de desativação - SECRBR



Pico máximo nos eventos de ativação - SECRBR



Pico máximo nos eventos de desativação - SECRBR



Fonte: Autor.

APÊNDICE E – Resultados de teste e avaliação do conjunto público *FIRE*D

Test and Score #1	KNN					SVM					ANN-MLP				
	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
Eventos FIRE															
<i>aspirador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>aspirador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>chaleira_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>chaleira_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>refrigerador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>refrigerador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>secador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>secador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>televisor_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>televisor_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>Micro-averaging</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>Macro-averaging</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Test and Score #2	KNN					SVM					ANN-MLP				
	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
Eventos FIRE															
<i>aspirador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>aspirador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>chaleira_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>chaleira_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>refrigerador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>refrigerador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>secador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>secador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>televisor_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>televisor_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>Micro-averaging</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>Macro-averaging</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Test and Score #3	KNN					SVM					ANN-MLP				
	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
Eventos FIRE															
<i>aspirador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>aspirador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>chaleira_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>chaleira_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.977	.955	1.000
<i>refrigerador_OFF</i>	.999	.999	.999	.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>refrigerador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>secador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>secador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.952	1.000	0.909
<i>televisor_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.974	.950	1.000
<i>televisor_ON</i>	.976	.999	.976	1.000	.952	1.000	1.000	1.000	1.000	1.000	1.000	.999	.976	1.000	.952
<i>Micro-averaging</i>	.999	.999	.999	.999	.999	1.000	1.000	1.000	1.000	1.000	1.000	.998	.998	.998	.998
<i>Macro-averaging</i>	.9975	.9998	.9975	.9998	.9952	1.000	1.000	1.000	1.000	1.000	1.000	.9996	.9879	.9905	.9861

Test and Score #4	KNN					SVM					ANN-MLP				
Eventos FIRED	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
<i>aspirador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>aspirador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>chaleira_OFF</i>	1.000	.999	.976	1.000	.952	1.000	1.000	1.000	1.000	1.000	1.000	.999	.976	1.000	.952
<i>chaleira_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>refrigerador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>refrigerador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>secador_OFF</i>	1.000	.999	.957	.917	1.000	1.000	1.000	1.000	1.000	1.000	.999	.999	.957	.917	1.000
<i>secador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>televisor_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>televisor_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>Micro-averaging</i>	1.000	.999	.999	.999	.999	1.000	1.000	1.000	1.000	1.000	1.000	.999	.999	.999	.999
<i>Macro-averaging</i>	1.000	.9998	.9933	.9917	.9952	1.000	1.000	1.000	1.000	1.000	.9999	.9998	.9933	.9917	.9952

Test and Score #5	KNN					SVM					ANN-MLP				
Eventos FIRED	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
<i>aspirador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>aspirador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>chaleira_OFF</i>	.976	.998	.950	1.000	.905	1.000	.997	.923	1.000	.857	.992	.997	.895	1.000	.810
<i>chaleira_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>refrigerador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>refrigerador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>secador_OFF</i>	1.000	.998	.917	.846	1.000	1.000	.997	.880	.789	1.000	.999	.997	.846	.733	1.000
<i>secador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>televisor_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>televisor_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>Micro-averaging</i>	.999	.998	.998	.999	.998	1.000	.997	.997	.998	.997	1.000	.997	.997	.998	.997
<i>Macro-averaging</i>	.9976	.9996	.9867	.9846	.9905	1.000	.9994	.9803	.9789	.9857	.9991	.9994	.9741	.9733	.9810

Test and Score #6	KNN					SVM					ANN-MLP				
Eventos FIRED	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
<i>aspirador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>aspirador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>chaleira_OFF</i>	1.000	.999	.976	1.000	.952	1.000	1.000	1.000	1.000	1.000	1.000	.999	.976	1.000	.952
<i>chaleira_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>refrigerador_OFF</i>	.999	.999	.999	.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>refrigerador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>secador_OFF</i>	1.000	.999	.957	.917	1.000	1.000	1.000	1.000	1.000	1.000	.999	.999	.957	.917	1.000
<i>secador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>televisor_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>televisor_ON</i>	.976	.999	.976	1.000	.952	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>Micro-averaging</i>	.999	.998	.998	.998	.998	1.000	1.000	1.000	1.000	1.000	1.000	.999	.999	.999	.999
<i>Macro-averaging</i>	.9975	.9996	.9908	.9915	.9904	1.000	1.000	1.000	1.000	1.000	.9999	.9998	.9933	.9917	.9952

Test and Score #10	KNN					SVM					ANN-MLP				
Eventos FIRED	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
<i>aspirador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>aspirador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>chaleira_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	.999	.976	1.000	.952	1.000	.999	.976	1.000	.952
<i>chaleira_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>refrigerador_OFF</i>	.999	.999	.999	.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>refrigerador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>secador_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	.999	.957	.917	1.000	.999	.999	.957	.917	1.000
<i>secador_ON</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>televisor_OFF</i>	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>televisor_ON</i>	.976	.999	.976	1.000	.952	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
<i>Micro-averaging</i>	.999	.999	.999	.999	.999	1.000	.999	.999	.999	.999	1.000	.999	.999	.999	.999
<i>Macro-averaging</i>	.9975	.9998	.9975	.9998	.9952	1.000	.9998	.9933	.9917	.9952	.9999	.9998	.9933	.9917	.9952

Fonte: Autor.

Legenda: **(Auc)** área sob a curva ROC; **(Acc)** Accuracy; **(F1)** F1-score; **(Pre)** Precision; **(Rec)** Recall.

APÊNDICE F – Resultados das predições do conjunto de sinais do ambiente real SECRBR

Predictions #1	KNN					SVM					ANN-MLP				
	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
Eventos SECRBR															
<i>aspirador_OFF</i>	.787	.660	.667	1.000	.500	.844	.680	.200	1.000	.111	1.000	.904	1.000	1.000	1.000
<i>aspirador_ON</i>	.500	.660	.000	.000	.000	.676	.680	.000	.000	.000	1.000	.904	1.000	1.000	1.000
<i>chaleira_OFF</i>	.976	.660	.638	1.000	.468	1.000	.680	.862	.758	1.000	1.000	.904	1.000	1.000	1.000
<i>chaleira_ON</i>	.865	.660	.617	.512	.774	.999	.680	.981	1.000	.962	.999	.904	.981	1.000	.962
<i>refrigerador_OFF</i>	.945	.660	.610	.439	1.000	.971	.680	.398	.249	1.000	.987	.904	.676	.511	1.000
<i>refrigerador_ON</i>	.991	.660	.990	1.000	.981	.990	.680	.990	1.000	.981	.991	.904	.990	1.000	.981
<i>secador_OFF</i>	.999	.660	.800	.667	1.000	1.000	.680	.824	1.000	.700	1.000	.904	1.000	1.000	1.000
<i>secador_ON</i>	.993	.660	.840	.725	1.000	1.000	.680	.980	.962	1.000	1.000	.904	.980	.962	1.000
<i>televisor_OFF</i>	.470	.660	.026	.036	.020	.328	.680	.078	1.000	.041	.946	.904	.151	1.000	.082
<i>televisor_ON</i>	.990	.660	.879	1.000	.784	.994	.680	.980	.980	.980	.999	.904	.990	.981	1.000
<i>Micro-averaging</i>	.854	.660	.616	.648	.660	.883	.680	.637	.809	.680	.992	.904	.880	.948	.904
<i>Macro-averaging</i>	.8516	.6600	.6067	.6379	.6527	.8802	.6800	.6293	.7949	.6775	.9922	.9040	.8768	.9454	.9025

Predictions #2	KNN					SVM					ANN-MLP				
	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
Eventos SECRBR															
<i>aspirador_OFF</i>	.870	.664	.699	1.000	.537	.910	.718	.338	1.000	.204	1.000	.742	.590	.419	1.000
<i>aspirador_ON</i>	.500	.664	.000	.000	.000	.558	.718	.000	.000	.000	1.000	.742	1.000	1.000	1.000
<i>chaleira_OFF</i>	.822	.664	.638	1.000	.468	1.000	.718	1.000	1.000	1.000	1.000	.742	.531	1.000	.362
<i>chaleira_ON</i>	.865	.664	.617	.512	.774	.999	.718	.981	1.000	.962	.999	.742	.981	1.000	.962
<i>refrigerador_OFF</i>	.946	.664	.610	.439	1.000	.975	.718	.405	.254	1.000	.978	.742	.648	.480	1.000
<i>refrigerador_ON</i>	.991	.664	.990	1.000	.981	.990	.718	.990	1.000	.981	.990	.742	.990	1.000	.981
<i>secador_OFF</i>	.983	.664	.800	.667	1.000	1.000	.718	1.000	1.000	1.000	.944	.742	.000	.000	.000
<i>secador_ON</i>	.993	.664	.840	.725	1.000	1.000	.718	.980	.962	1.000	1.000	.742	.980	.962	1.000
<i>televisor_OFF</i>	.470	.664	.027	.038	.020	.358	.718	.115	1.000	.061	.969	.742	.115	1.000	.061
<i>televisor_ON</i>	1.000	.664	.879	1.000	.784	.991	.718	.960	.980	.941	.998	.742	.990	.981	1.000
<i>Micro-averaging</i>	.848	.664	.619	.649	.664	.882	.718	.685	.832	.718	.988	.742	.686	.782	.742
<i>Macro-averaging</i>	.8440	.6640	.6100	.6381	.6564	.8781	.7180	.6769	.8196	.7149	.9878	.7420	.6825	.7842	.7366

Predictions #3	KNN					SVM					ANN-MLP				
	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
Eventos SECRBR															
<i>aspirador_OFF</i>	.898	.634	.667	1.000	.500	.914	.708	.338	1.000	.204	.927	.696	.590	0.419	1.000
<i>aspirador_ON</i>	.500	.634	.000	.000	.000	.557	.708	.000	.000	.000	1.000	.696	1.000	1.000	1.000
<i>chaleira_OFF</i>	.948	.634	.638	1.000	.468	1.000	.708	1.000	1.000	1.000	1.000	.696	.000	.000	.000
<i>chaleira_ON</i>	.961	.634	.596	.477	.792	.996	.708	.927	.895	.962	1.000	.696	.990	1.000	.981
<i>refrigerador_OFF</i>	.921	.634	.560	.388	1.000	.971	.708	.407	.255	1.000	.992	.696	.566	.395	1.000
<i>refrigerador_ON</i>	.991	.634	.990	1.000	.981	.990	.708	.990	1.000	.981	.989	.696	.990	1.000	.981
<i>secador_OFF</i>	1.000	.634	.800	.667	1.000	1.000	.708	1.000	1.000	1.000	.944	.696	.000	.000	.000
<i>secador_ON</i>	.998	.634	.901	.820	1.000	.997	.708	.917	.957	.880	1.000	.696	.990	.980	1.000
<i>televisor_OFF</i>	.476	.634	.026	.036	.020	.328	.708	.078	1.000	.041	.930	.696	.111	.600	.061
<i>televisor_ON</i>	.765	.634	.675	1.000	.510	.997	.708	.980	.980	.980	.994	.696	.907	.957	.863
<i>Micro-averaging</i>	.849	.634	.595	.649	.634	.879	.708	.671	.821	.708	.977	.696	.621	.641	.696
<i>Macro-averaging</i>	.8458	.6340	.5853	.6388	.6271	.8750	.7080	.6637	.8087	.7048	.9776	.6960	.6144	.6351	.6886

Predictions #4	KNN					SVM					ANN-MLP				
Eventos SECRBR	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
<i>aspirador_OFF</i>	.778	.662	.667	1.000	.500	.809	.702	.036	1.000	.019	1.000	.816	1.000	1.000	1.000
<i>aspirador_ON</i>	.500	.662	.000	.000	.000	.557	.702	.000	.000	.000	1.000	.816	1.000	1.000	1.000
<i>chaleira_OFF</i>	.787	.662	.638	1.000	.468	1.000	.702	1.000	1.000	1.000	1.000	.816	.042	1.000	.021
<i>chaleira_ON</i>	.882	.662	.586	.471	.774	.999	.702	.981	1.000	.962	.997	.816	.981	1.000	.962
<i>refrigerador_OFF</i>	.946	.662	.614	.443	1.000	.975	.702	.392	.244	1.000	.987	.816	.606	.435	1.000
<i>refrigerador_ON</i>	.991	.662	.990	1.000	.981	.990	.702	.990	1.000	.981	.988	.816	.990	1.000	.981
<i>secador_OFF</i>	.978	.662	.800	.667	1.000	1.000	.702	1.000	1.000	1.000	.962	.816	.800	.667	1.000
<i>secador_ON</i>	.998	.662	.893	.806	1.000	1.000	.702	.980	.962	1.000	1.000	.816	.980	.962	1.000
<i>televisor_OFF</i>	.480	.662	.026	.036	.020	.349	.702	.115	1.000	.061	.979	.816	.218	1.000	.122
<i>televisor_ON</i>	.990	.662	.891	1.000	.804	.997	.702	.980	.980	.980	.998	.816	.962	.927	1.000
<i>Micro-averaging</i>	.837	.662	.620	.652	.662	.871	.702	.653	.831	.702	.991	.816	.767	.902	.816
<i>Macro-averaging</i>	.8330	.6620	.6105	.6423	.6547	.8676	.7020	.6474	.8186	.7003	.9911	.8160	.7579	.8991	.8086

Predictions #5	KNN					SVM					ANN-MLP				
Eventos SECRBR	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
<i>aspirador_OFF</i>	.787	.658	.667	1.000	0.500	.815	.652	.200	1.000	.111	1.000	.782	.590	.419	1.000
<i>aspirador_ON</i>	.500	.658	.000	.000	.000	.562	.652	.000	.000	.000	1.000	.782	1.000	1.000	1.000
<i>chaleira_OFF</i>	.734	.658	.638	1.000	.468	.969	.652	.576	1.000	.404	.956	.782	.638	1.000	.468
<i>chaleira_ON</i>	.848	.658	.586	.471	.774	.999	.652	.981	1.000	.962	1.000	.782	.981	1.000	.962
<i>refrigerador_OFF</i>	.945	.658	.606	.435	1.000	.971	.652	.397	.247	1.000	.987	.782	.764	.618	1.000
<i>refrigerador_ON</i>	.991	.658	.990	1.000	.981	.990	.652	.990	1.000	.981	.990	.782	.990	1.000	.981
<i>secador_OFF</i>	.972	.658	.800	.667	1.000	1.000	.652	.781	.641	1.000	.944	.782	.000	.000	.000
<i>secador_ON</i>	.993	.658	.893	.806	1.000	1.000	.652	.980	.962	1.000	1.000	.782	.980	.962	1.000
<i>televisor_OFF</i>	.480	.658	.026	.036	.020	.330	.652	.040	1.000	.020	1.000	.782	.537	1.000	.367
<i>televisor_ON</i>	.990	.658	.867	1.000	.765	.997	.652	.980	.980	.980	.998	.782	.971	.944	1.000
<i>Micro-averaging</i>	.828	.658	.616	.652	.658	.866	.652	.602	.796	.652	.988	.782	.746	.792	.782
<i>Macro-averaging</i>	.8240	.6580	.6073	.6415	.6508	.8633	.6520	.5925	.7830	.6458	.9875	.7820	.7451	.7943	.7778

Predictions #6	KNN					SVM					ANN-MLP				
Eventos SECRBR	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
<i>aspirador_OFF</i>	.787	.580	.667	1.000	.500	.805	.698	.000	.000	.000	1.000	.826	.878	.783	1.000
<i>aspirador_ON</i>	.500	.580	.000	.000	.000	.557	.698	.000	.000	.000	1.000	.826	1.000	1.000	1.000
<i>chaleira_OFF</i>	.734	.580	.638	1.000	.468	.988	.698	1.000	1.000	1.000	1.000	.826	.638	1.000	.468
<i>chaleira_ON</i>	.848	.580	.293	.246	.283	1.000	.698	.981	1.000	.962	1.000	.826	.981	1.000	.962
<i>refrigerador_OFF</i>	.926	.580	.560	.388	1.000	.971	.698	.388	.241	1.000	.981	.826	.686	.522	1.000
<i>refrigerador_ON</i>	.991	.580	.990	1.000	.981	.990	.698	.990	1.000	.981	.988	.826	.990	1.000	.981
<i>secador_OFF</i>	.972	.580	.800	.667	1.000	1.000	.698	1.000	1.000	1.000	.945	.826	.636	.583	.700
<i>secador_ON</i>	.993	.580	.725	.568	1.000	1.000	.698	.980	.962	1.000	1.000	.826	.980	.962	1.000
<i>televisor_OFF</i>	.471	.580	.026	.036	.020	.329	.698	.078	1.000	.041	.907	.826	.185	1.000	.102
<i>televisor_ON</i>	.824	.580	.675	1.000	.510	.997	.698	.980	.980	.980	.998	.826	.981	.962	1.000
<i>Micro-averaging</i>	.808	.580	.541	.600	.580	.867	.698	.645	.723	.698	.982	.826	.800	.882	.826
<i>Macro-averaging</i>	.8046	.5800	.5374	.5905	.5762	.8637	.6980	.6397	.7183	.6964	.9819	.8260	.7955	.8812	.8213

Predictions #7	KNN					SVM					ANN-MLP				
Eventos SECRBR	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
<i>aspirador_OFF</i>	.898	.668	.714	1.000	.556	.933	.684	.338	1.000	.204	1.000	.916	1.000	1.000	1.000
<i>aspirador_ON</i>	.500	.668	.000	.000	.000	.557	.684	.000	.000	.000	1.000	.916	1.000	1.000	1.000
<i>chaleira_OFF</i>	.967	.668	.638	1.000	.468	1.000	.684	.855	.746	1.000	1.000	.916	1.000	1.000	1.000
<i>chaleira_ON</i>	.865	.668	.617	.512	.774	.999	.684	.981	1.000	.962	1.000	.916	.981	1.000	.962
<i>refrigerador_OFF</i>	.946	.668	.614	.443	1.000	.971	.684	.403	.253	1.000	.987	.916	.707	.547	1.000
<i>refrigerador_ON</i>	.991	.668	.990	1.000	.981	.989	.684	.990	1.000	.981	.990	.916	.990	1.000	.981
<i>secador_OFF</i>	.999	.668	.800	.667	1.000	1.000	.684	.810	1.000	.680	1.000	.916	1.000	1.000	1.000
<i>secador_ON</i>	.993	.668	.840	.725	1.000	1.000	.684	.980	.962	1.000	1.000	.916	.980	.962	1.000
<i>televisor_OFF</i>	.475	.668	.027	.040	.020	.343	.684	.078	1.000	.041	.958	.916	.339	1.000	.204
<i>televisor_ON</i>	1.000	.668	.891	1.000	.804	.993	.684	.960	.980	.941	.998	.916	.990	.981	1.000
<i>Micro-averaging</i>	.867	.668	.623	.649	.688	.882	.684	.648	.808	.684	.993	.916	.902	.952	.916
<i>Macro-averaging</i>	.8634	.6680	.6131	.6387	.6603	.8785	.6840	.6395	.7941	.6809	.9933	.9160	.8987	.9490	.9147

Predictions #8	KNN					SVM					ANN-MLP				
Eventos SECRBR	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
<i>aspirador_OFF</i>	.898	.630	.714	1.000	.556	.895	.716	.338	1.000	.204	1.000	.814	1.000	1.000	1.000
<i>aspirador_ON</i>	.500	.630	.000	.000	.000	.557	.716	.000	.000	.000	1.000	.814	1.000	1.000	1.000
<i>chaleira_OFF</i>	.787	.630	.638	1.000	.468	1.000	.716	1.000	1.000	1.000	1.000	.814	.157	1.000	.085
<i>chaleira_ON</i>	.882	.630	.377	.333	.434	.999	.716	.981	1.000	.962	1.000	.814	.981	1.000	.962
<i>refrigerador_OFF</i>	.946	.630	.610	.439	1.000	.969	.716	.403	.253	1.000	.987	.814	.603	.431	1.000
<i>refrigerador_ON</i>	.991	.630	.990	1.000	.981	.990	.716	.990	1.000	.981	.988	.814	.990	1.000	.981
<i>secador_OFF</i>	.978	.630	.800	.667	1.000	1.000	.716	1.000	1.000	1.000	1.000	.814	.800	.667	1.000
<i>secador_ON</i>	.998	.630	.769	.625	1.000	1.000	.716	.980	.962	1.000	1.000	.814	.980	.962	1.000
<i>televisor_OFF</i>	.490	.630	.027	.040	.020	.331	.716	.078	1.000	.041	.914	.814	.115	1.000	.061
<i>televisor_ON</i>	.990	.630	.879	1.000	.784	.990	.716	.960	.980	.941	.997	.814	.952	.926	.980
<i>Micro-averaging</i>	.850	.630	.589	.620	.630	.877	.716	.681	.832	.716	.989	.814	.767	.902	.814
<i>Macro-averaging</i>	.8460	.6300	.5804	.6104	.6243	.8731	.7160	.673	.8195	.7129	.9886	.8140	.7578	.8986	.8069

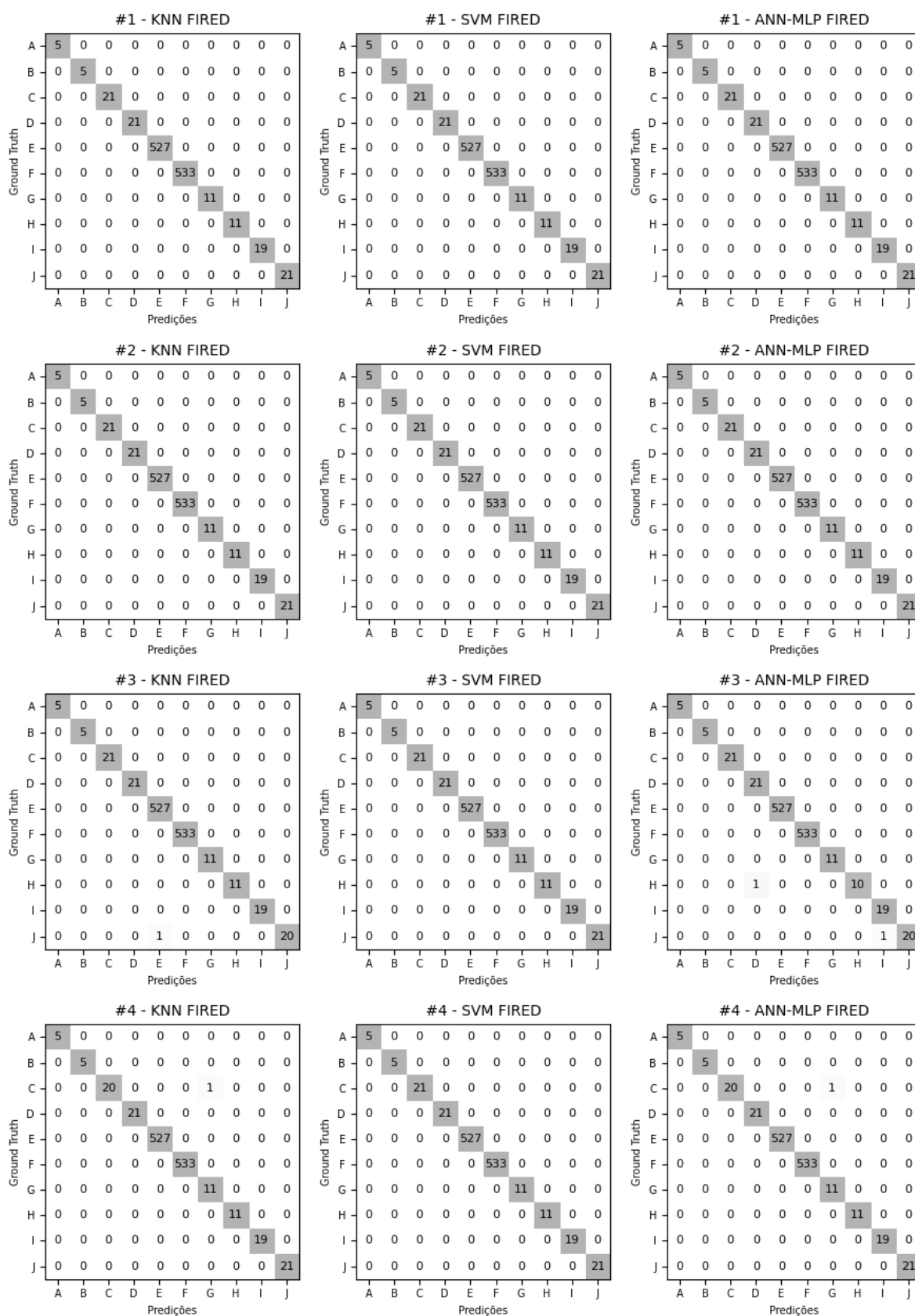
Predictions #9	KNN					SVM					ANN-MLP				
Eventos SECRBR	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
<i>aspirador_OFF</i>	.750	.660	.633	1.000	.463	.754	.698	.000	.000	.000	1.000	.942	1.000	1.000	1.000
<i>aspirador_ON</i>	.500	.660	.000	.000	.000	.557	.698	.000	.000	.000	1.000	.942	1.000	1.000	1.000
<i>chaleira_OFF</i>	.842	.660	.638	1.000	.468	1.000	.698	1.000	1.000	1.000	1.000	.942	1.000	1.000	1.000
<i>chaleira_ON</i>	.865	.660	.617	.512	.774	.999	.698	.981	1.000	.962	.999	.942	.981	1.000	.962
<i>refrigerador_OFF</i>	.946	.660	.618	.448	1.000	.971	.698	.388	.241	1.000	.994	.942	.810	.681	1.000
<i>refrigerador_ON</i>	.991	.660	.990	1.000	.981	.990	.698	.990	1.000	.981	.989	.942	.990	1.000	.981
<i>secador_OFF</i>	.986	.660	.800	.667	1.000	1.000	.698	1.000	1.000	1.000	1.000	.942	1.000	1.000	1.000
<i>secador_ON</i>	.993	.660	.840	.725	1.000	1.000	.698	.980	.962	1.000	1.000	.942	.980	.962	1.000
<i>televisor_OFF</i>	.466	.660	.025	.033	.020	.329	.698	.078	1.000	.041	.987	.942	.639	1.000	.469
<i>televisor_ON</i>	1.000	.660	.903	1.000	.824	.992	.698	.980	.980	.980	.998	.942	.953	.911	1.000
<i>Micro-averaging</i>	.837	.660	.615	.649	.660	.862	.698	.645	.723	.698	.997	.942	.937	.957	.942
<i>Macro-averaging</i>	.8339	.6600	.6064	.6385	.6530	.8592	.6980	.6397	.7183	.6964	.9967	.9420	.9353	.9554	.9412

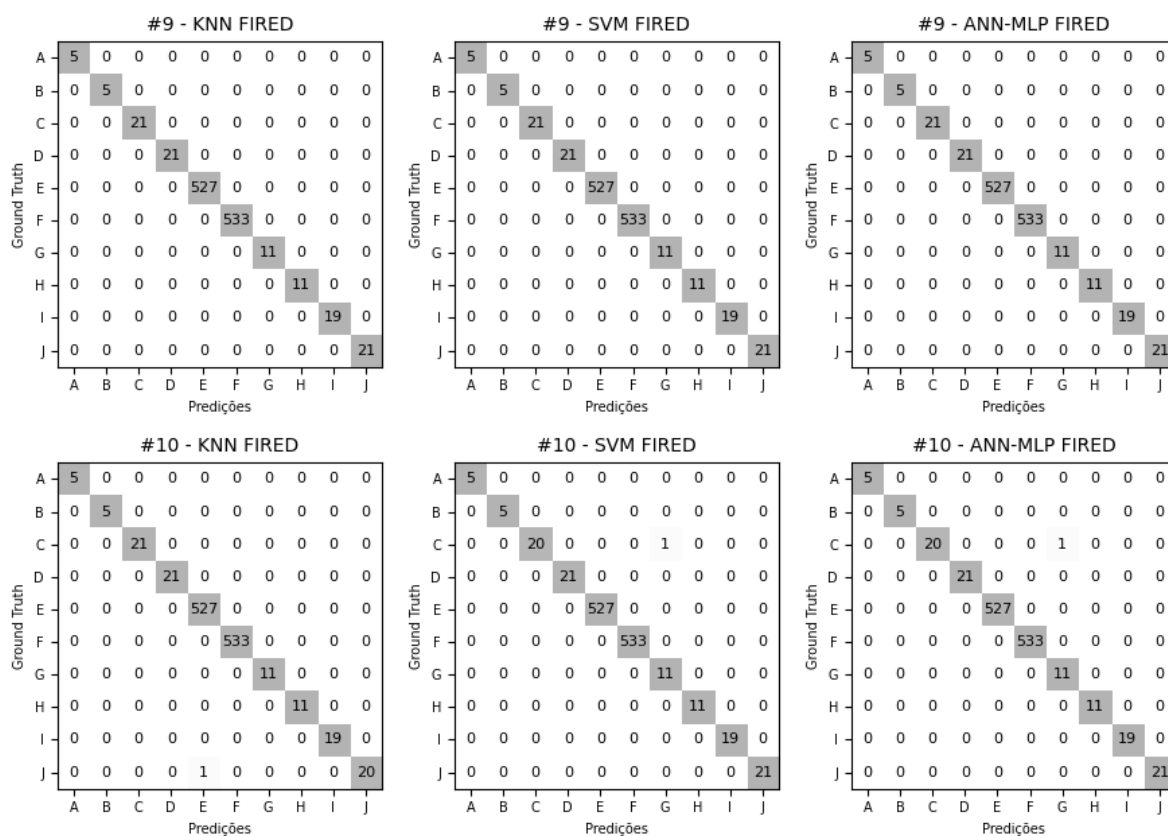
Predictions #10	KNN					SVM					ANN-MLP				
Eventos SECRBR	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre	Rec	Auc	Acc	F1	Pre.	Rec
<i>aspirador_OFF</i>	.750	.634	.667	1.000	.500	.929	.724	.813	1.000	.685	1.000	.734	.590	.419	1.000
<i>aspirador_ON</i>	.500	.634	.000	.000	.000	.857	.724	.000	.000	.000	1.000	.734	.989	1.000	.978
<i>chaleira_OFF</i>	1.000	.634	.638	1.000	.468	1.000	.724	.638	1.000	.468	1.000	.734	.000	.000	.000
<i>chaleira_ON</i>	.902	.634	.596	.477	.792	1.000	.724	.981	1.000	.962	.999	.734	.962	.962	.962
<i>refrigerador_OFF</i>	.928	.634	.560	.388	1.000	.972	.724	.667	.500	1.000	.991	.734	.648	.480	1.000
<i>refrigerador_ON</i>	.991	.634	.990	1.000	.981	.986	.724	.689	.531	.981	.991	.734	.981	.981	.981
<i>secador_OFF</i>	1.000	.634	.800	.667	1.000	1.000	.724	.800	.667	1.000	.944	.734	.000	.000	.000
<i>secador_ON</i>	.998	.634	.901	.820	1.000	1.000	.724	.980	.962	1.000	.999	.734	.960	.960	.960
<i>televisor_OFF</i>	.446	.634	.026	.036	.020	.354	.724	.087	.150	.061	.977	.734	.559	1.000	.388
<i>televisor_ON</i>	.843	.634	.675	1.000	.510	.996	.724	.980	.980	.980	.998	.734	.981	.962	1.000
<i>Micro-averaging</i>	.840	.634	.594	.649	.634	.911	.724	.674	.689	.724	.990	.734	.672	.680	.734
<i>Macro-averaging</i>	.8358	.6340	.5853	.6388	.6271	.9094	.7240	.6635	.6790	.7137	.9899	.7340	.6670	.6764	.7269

Fonte: Autor.

Legenda: **(Auc)** área sob a curva ROC; **(Acc)** Accuracy; **(F1)** F1-score; **(Pre)** Precision; **(Rec)** Recall.

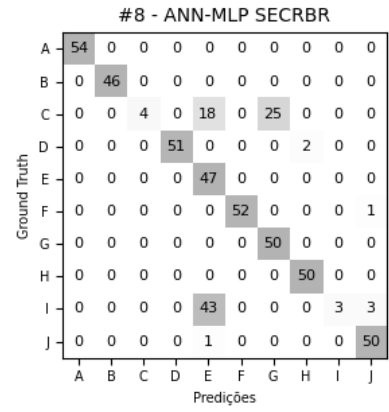
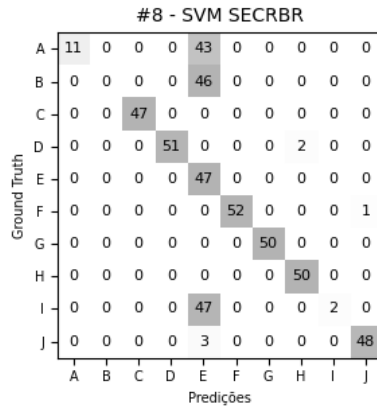
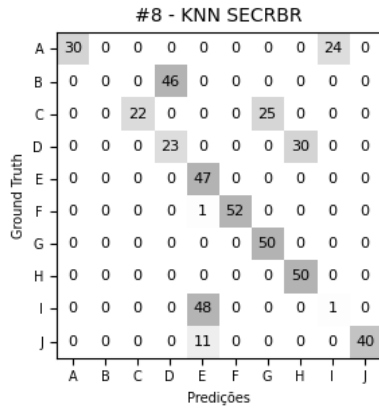
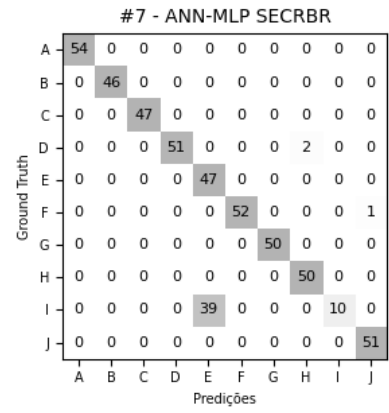
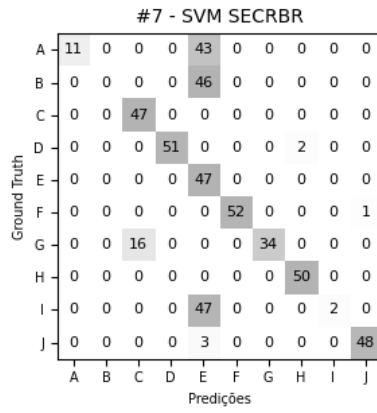
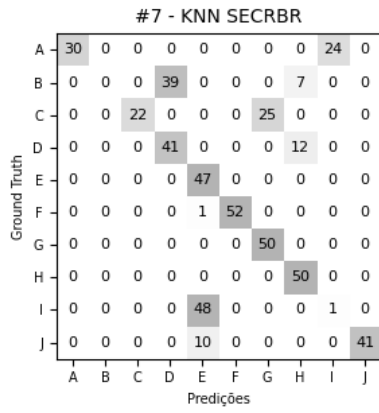
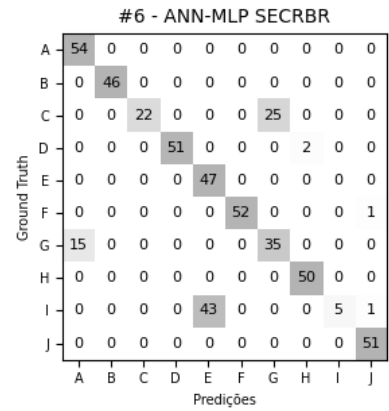
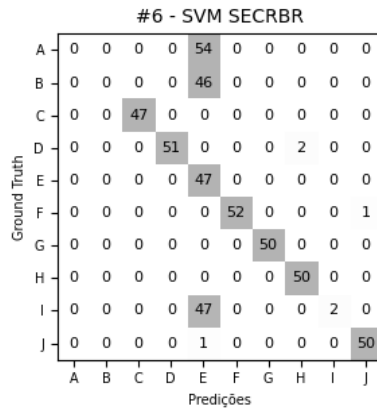
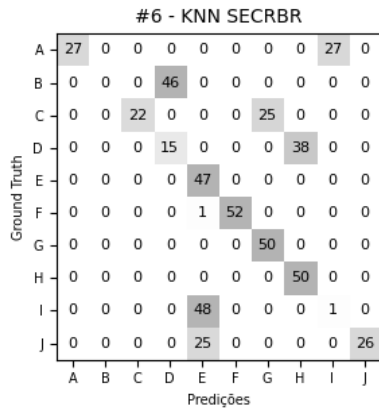
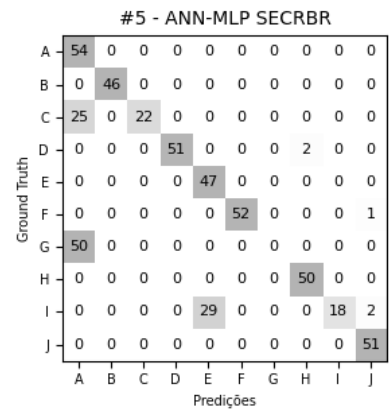
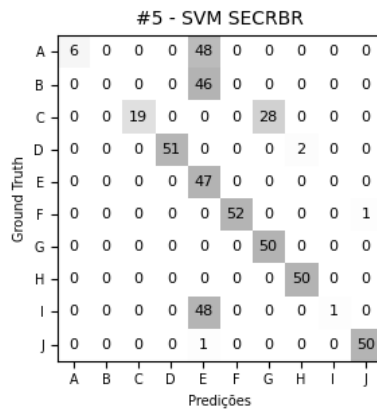
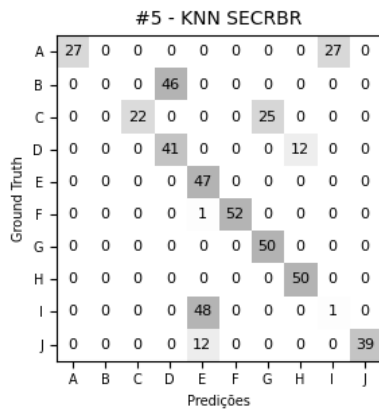
APÊNDICE G – Resultados por matriz de confusão do conjunto público *FIRE*D

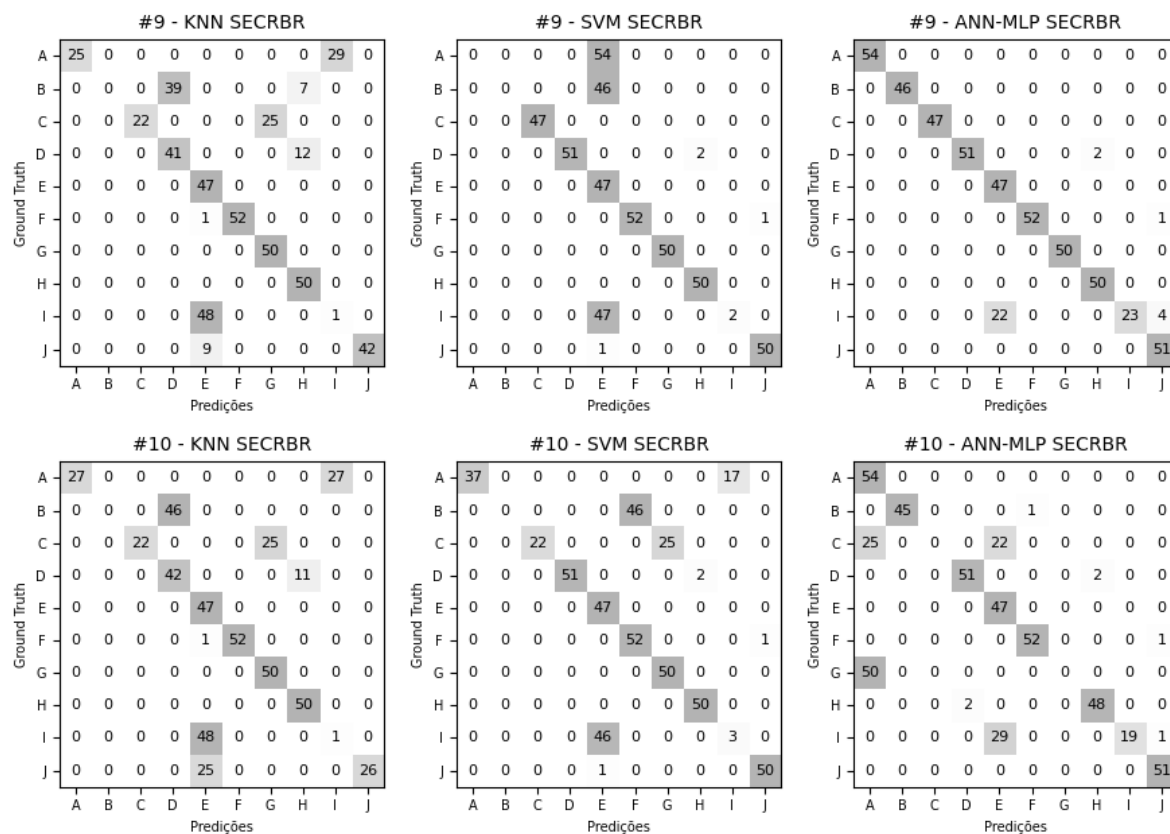




Fonte: Autor.

Legenda: (A) aspirador_OFF; (B) aspirador_ON; (C) chaleira_OFF; (D) chaleira_ON;
 (E) refrigerador_OFF; (F) refrigerador_ON; (G) secador_OFF; (H) secador_ON;
 (I) televisor_OFF; (J) televisor_ON.





Fonte: Autor.

Legenda: (A) aspirador_OFF; (B) aspirador_ON; (C) chaleira_OFF; (D) chaleira_ON;
 (E) refrigerador_OFF; (F) refrigerador_ON; (G) secador_OFF; (H) secador_ON;
 (I) televisor_OFF; (J) televisor_ON.