



RODRIGO BOTELHO DE LIMA

**CLASSIFICADOR NÃO INVASIVO DE CARGAS
ELÉTRICAS RESIDENCIAIS COM ACIONAMENTO
SIMULTÂNEO**

**LAVRAS – MG
2018**

RODRIGO BOTELHO DE LIMA

**CLASSIFICADOR NÃO INVASIVO DE CARGAS ELÉTRICAS
RESIDENCIAIS COM ACIONAMENTO SIMULTÂNEO**

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Automação da Universidade Federal de Lavras como parte dos requisitos necessários a obtenção do título de Mestre em Engenharia de Sistemas e Automação.

Prof. Dr. Wilian Soares Lacerda
Orientador

Prof. Dr. Danton Diego Ferreira
Coorientador

**LAVRAS – MG
2018**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Lima, Rodrigo Botelho de.

Classificador não invasivo de cargas elétricas residenciais com
acionamento simultâneo / Rodrigo Botelho de Lima. - 2018.

74 p. : il.

Orientador(a): Wilian Soares Lacerda.

Coorientador(a): Danton Diego Ferreira.

Dissertação (mestrado acadêmico) - Universidade Federal de
Lavras, 2018.

Bibliografia.

1. Monitoramento não invasivo. 2. Classificador de cargas
elétricas. 3. Estatística de Ordem Superior. I. Lacerda, Wilian
Soares. II. Ferreira, Danton Diego. III. Título.

RODRIGO BOTELHO DE LIMA

**CLASSIFICADOR NÃO INVASIVO DE CARGAS ELÉTRICAS
RESIDENCIAIS COM ACIONAMENTO SIMULTÂNEO**

**NON-INVASIVE CLASSIFIER FOR RESIDENTIAL ELECTRICAL LOADS
SWITCHED AT THE SAME TIME**

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Automação da Universidade Federal de Lavras como parte dos requisitos necessários a obtenção do título de Mestre em Engenharia de Sistemas e Automação.

APROVADA em 20 de setembro de 2018.

Prof. Dr. Danton Diego Ferreira UFLA
Prof.^a Dra. Lane Maria Rabelo Baccarini UFSJ
Prof.^a Dra. Sílvia Costa Ferreira UFLA

Prof. Dr. Wilian Soares Lacerda
Orientador

Prof. Dr. Danton Diego Ferreira
Coorientador

**LAVRAS – MG
2018**

A Deus, por ter feito com que tudo acontecesse e por estar cuidando de cada coisa na minha vida de forma tão especial.

Ao meu pai Álvaro Mendes Lima e minha mãe Dulcineia Botelho Lima, que propiciaram tudo isso, só eu sei o esforço que vocês fizeram, trabalhando até altas horas da madrugada para preparar as coisas e levar para vender na feira, tudo isso para ajudar a me manter na faculdade. Agradeço ainda, pai e mãe, o amor, o carinho e a educação que me deram desde criança e muitas outras coisas que me fizeram o homem que sou.

À minha esposa Krisnanda que me motivou durante o mestrado, em todos os dias foi minha companheira, meu braço direito e me ajudou a fazer muitas coisas durante o percurso.

À minha querida filha Bianca, aos meus irmãos Renato e Graziela, à Nice, ao Sr. Alexandre e às minhas sobrinhas, a estes que fazem parte da minha vida e me motivam a ir adiante.

Dedico

AGRADECIMENTOS

A Deus e à minha família por todo amor e apoio que me dão.

À Universidade Federal de Lavras, ao departamento de engenharia e ao Programa de Pós-graduação em Engenharia de Sistemas e Automação.

Ao meu orientador professor e amigo Wilian Soares Lacerda que me direcionou durante essa trajetória, me orientou e fez importantes contribuições para que desse certo este trabalho.

Ao meu coorientador, professor e amigo Danton Diego Ferreira, que me instruiu a cada etapa, esteve presente e me auxiliou nos momentos de dificuldade com contribuições valiosas.

Aos colegas do laboratório Labsine.

Aos professores e técnicos da UFLA pelo conhecimento que me passaram.

“O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001”.

A todos, minha gratidão!

“Um pouco de ciência nos afasta de Deus. Muito, nos aproxima.”

Louis Pasteur

RESUMO

A preocupação com o futuro dos recursos energéticos mundiais leva os pesquisadores a buscarem alternativas que possam permitir o gerenciamento desses recursos. Essa preocupação se torna maior com o aumento do consumo de energia elétrica pela população. A necessidade de conhecer o perfil dos consumidores é ainda mais proeminente com a entrada da Tarifa Branca, autorizada pela ANEEL, em janeiro de 2018. Essa tarifa se aplica aos consumidores residenciais que, para optar por essa tarifa, precisam conhecer o seu perfil de consumo. Este trabalho de mestrado traz uma proposta para classificar cargas acionadas individualmente ou mesmo quando outra já está em funcionamento. Para isso, é montado um banco de dados com cinco tipos de cargas que compõem cinco classes diferentes. Também são montadas cinco classes que utilizam o acionamento de uma carga enquanto outra está em funcionamento. Como classificador foram implementados algoritmos de aprendizagem de máquina, são eles: Redes Neurais Artificiais, Máquina de Vetor Suporte e Floresta Aleatória. Após o desenvolvimento do modelo, pela análise das métricas utilizadas, pôde ser observado que a acurácia média para o classificador baseado em Máquina de Vetor Suporte foi de 99,8%, a precisão média foi de 99,31%, e a sensibilidade média foi de 99,8%. O classificador baseado em Redes Neurais Artificiais alcançou acurácia média de 98,85%, precisão média de 98,82% e sensibilidade média de 98,5%. Quanto ao classificador baseado em Florestas Aleatórias, foram alcançados os valores médios para acurácia de 98,95%, para precisão de 98,81% e para sensibilidade de 98,8%. Para todos os classificadores, o valor médio para a métrica de *F1 score* foi de 0,99, resultado que mostra um bom desempenho para todos os modelos. Dessa forma, como observado, o SVM mostrou ser melhor em todas as métricas. Mostrou ainda melhor desempenho relativo a tempo de treinamento sugerindo um menor esforço computacional, e conseqüentemente, menor custo para colocar em produção. Outra avaliação a ser realizada é que a Estatística de Ordem Superior mostrou ser uma excelente ferramenta para a extração dos parâmetros de treinamento para fins de classificação de cargas elétricas.

Palavras-chave — Monitoramento não Invasivo, Identificação de Cargas Elétricas, Estatística de Ordem Superior, SVM, Florestas Aleatórias, Redes Neurais Artificiais.

ABSTRACT

Thinking about the future of the world's, especially about energy resources, scientists are seeking alternatives that allow management of these resources. This concern increases with the high consumption of electric energy by the people. Beside of this, in January 2018, electrical energy consumers were able to choose between staying in the normal tariff or switching to a white tariff, but to do this choice the needs to know their profile consume. This study presents a proposal to classify individually driven loads or even when another one is already in operation. For this a dataset was set up with five different types of loads that make up five different classes. Five classes are also assembled which use the triggering of one load while another is in operation. As a classifier, machine learning algorithms were implemented using: Artificial Neural Networks, Vector Support Machine and Random Forests. After developed the model, the results of the collected metrics, showed that the average accuracy for the Classifier based on Vector Support Machine was 99.8%, the average accuracy was 99.31%, and the sensitivity was 99.8%. The Classifier based on Artificial Neural Networks reached the average accuracy of 98.85%, the average precision of 98.82% and the average sensitivity of 98.5%. As for the classifier based on Random Forests, the classifier reached the average values for accuracy of 98.95%, 98.81% for precision and sensitivity of 98.8%. For all classifiers the mean value reached for the *F1 score* metric was 0.99, this result analysis showed that the models performed very well. Thus, as observed, the SVM proved to be better in all metrics. The SVM still presented better performance in training time, suggesting a lower computational effort and hence a lower cost to produce. Another evaluation to be carried out is that the Higher Order Statistics also showed to be efficient in the extraction of parameters in this work.

Keywords: Non-Invasive Monitoring, Electrical Loads Identification, Upper Order Statistics, SVM, Random Forests, Artificial Neural Networks.

LISTA DE FIGURAS

Figura 1: Cumulante de segunda ordem para um sinal senoidal que não possui interferência de harmônicos e/ou ruídos.	23
Figura 2: Cumulante de segunda ordem para um sinal senoidal com ruído gaussiano.	23
Figura 3: Cumulante de segunda ordem para um sinal senoidal com distorções harmônicas.	24
Figura 4: Cumulante de quarta ordem para um sinal senoidal sem presença de ruídos e/ou harmônicos.	25
Figura 5: Cumulante de quarta ordem para um sinal senoidal com ruído gaussiano.	25
Figura 6: Cumulante de quarta ordem para um sinal senoidal com distorções harmônicas.	26
Figura 7: Projeção das amostras em um eixo reduzindo a dimensão dos dados.	27
Figura 8: Representação de um neurônio não linear.	29
Figura 9: Representação de uma rede multicamada, quais sejam uma camada de entrada x, duas camadas ocultas e uma camada de saída y.	30
Figura 10: Representação de uma rede multicamada, quais sejam, uma camada de entrada x, duas camadas ocultas e uma camada de saída y.	30
Figura 11: Representação de uma rede multicamada com recorrências nas camadas ocultas. ...	31
Figura 12: Esquema apresentando o hiperplano ótimo para classificar os dados com as maiores margens.	34
Figura 13: (a) Conjunto de dados linearmente separáveis; (b) Conjunto de dados não separáveis linearmente.	34
Figura 14: Estrutura da árvore de decisão, esta figura representa uma floresta de K árvores contendo o Nó Raiz, Nó ramo e o Nó Folha.	37
Figura 15: Fluxograma indicando as etapas a serem realizadas na fase de desenvolvimento do classificador.	41
Figura 16: Placa de aquisição de NI USB-6251 fabricado por National Instruments.	43
Figura 17: Ponteira de corrente FLUKE i200s.	44
Figura 18: Ponteira para aquisição de dados de tensão.	44
Figura 19: Esquema para coleta de sinais.	45
Figura 20: Multiprocessador de alimentos e Lâmpada de LED conectados no momento da aquisição dos dados.	45
Figura 21: (a) conexão da ponteira de tensão; (b) conexão da ponteira de corrente; (c) Conexão das ponteiras à placa de aquisição de dados NI USB-6251.	46
Figura 22: (a) Aquisição de dados da Lâmpada de LED ligada juntamente com o Multiprocessador de alimentos; (b) aquisição de dados da Lâmpada de LED.	46
Figura 23: (a) Transitório de partida de uma carga acionada isoladamente (TV de Led); (b) transitório de uma TV de Led enquanto um multiprocessador de alimentos está em uso e (c) transitório de uma lâmpada de Led enquanto um chuveiro está em uso.	48
Figura 24: Acionamento da TV de Led após o acionamento do Chuveiro, em azul o sinal filtrado utilizando o filtro Notch.	48
Figura 25: Espectros de corrente elétrica referente ao acionamento de duas cargas (chuveiro e lâmpada de LED) antes e após a utilização do filtro Notch.	49
Figura 26: Processo de identificação dos regimes transitórios referentes ao acionamento das cargas.	50

Figura 27: Exemplo dos transitórios obtidos para as classes C1 – chuveiro, C2 – Televisão de LED, classe C3 – Micro-ondas, C4 - Lâmpada de LED e C5 – Multiprocessador.....	54
Figura 28: Exemplo dos transitórios obtidos para as classes CS1 – Chuveiro + Lâmpada de LED CS2 – Chuveiro + Multiprocessador, CS3 – Lâmpada de LED + Multiprocessador, CS4 – Televisão de LED + Multiprocessador e CS5 - Micro-ondas + Multiprocessador.	55
Figura 29: Gráfico apresentando: (a) cumulantes de segunda ordem para as cargas acionadas individualmente; (b) cumulantes de segunda ordem para as classes com acionamentos simultâneos; (c) cumulantes de quarta ordem para as cargas acionadas individualmente; (d) cumulantes de quarta ordem para as classes com acionamentos simultâneos; O eixo das abcissas são os índices dos cumulantes e o eixo das ordenadas são os cumulantes.....	57
Figura 30: Discriminante de Fisher para todas as classes, do índice 1 ao índice 1024 são os cumulantes de segunda ordem e o restante são os cumulantes de quarta ordem.....	58

SUMÁRIO

1	INTRODUÇÃO.....	13
1.1	Objetivo Geral.....	14
1.2	Objetivo Específico	14
1.3	Motivação	14
1.4	Trabalhos Relacionados	16
1.5	Organização do documento	19
2	REFERENCIAL TEÓRICO	20
2.1	Estatística de Ordem Superior (EOS).....	20
2.2	Relação Discriminante de Fisher	26
2.3	Redes Neurais Artificiais (RNA)	28
2.4	Máquina de Vetor Suporte (SVM)	33
2.5	Florestas Aleatórias	36
2.6	Matriz de Confusão	38
3	MATERIAIS E MÉTODOS	41
3.1	Aquisição dos dados.....	42
3.2	Processamento dos dados	47
3.3	Treinamento dos classificadores propostos	51
4	Resultado e discussão	53
4.1	Processamento dos dados	53
4.2	Extração e Seleção de parâmetros.....	55
4.3	Acurácia.....	59
4.4	Precisão	60
4.5	Sensibilidade.....	60
4.6	F1 score	61
5	CONCLUSÃO.....	67
5.1	Trabalhos futuros	68
	REFERÊNCIAS	69
	ANEXO A - Especificação da placa NI USB-6251.....	73
	ANEXO B – Especificação ponteira para medição de corrente Fluke i200s	74

1 INTRODUÇÃO

O consumo de energia elétrica tem aumentado ao longo dos anos, e com isso a necessidade de se conhecer o perfil dos consumidores. No setor elétrico, o foco está em pesquisas sobre energia alternativa e qualidade de energia, mas atualmente passa-se por um momento em que as concessionárias precisam direcionar mais os investimentos para aplicação efetiva da tecnologia nas redes e na relação com o consumidor.

O consumo de energia elétrica ainda é um pacote sobre o qual se conhece o valor total, mas não é possível um detalhamento dessa fatura assim como é feito nas contas de telefone, por exemplo, que traz as informações de todas as ligações, duração e custo de cada uma. A fatura de energia elétrica não permite identificar o perfil do consumidor sobre como ele gasta essa energia.

Pesquisadores têm demonstrado constante interesse em ambientes e equipamentos eletroeletrônicos inteligentes, é a chamada tecnologia *Smart*. Diversas pesquisas encontradas na literatura, conforme serão apresentadas no item 1.4 desta dissertação, vêm propondo algoritmos e metodologias para que seja possível realizar a discriminação do consumo das cargas e seus comportamentos dentro de uma determinada área física. Essas áreas podem ser industriais, edifícios, casas residenciais, escritórios ou outros aglomerados de cargas nos quais não se consegue determinar a utilização dos equipamentos elétricos sem que se use um método invasivo e pontual para cada carga instalada.

Em Guedes, Ferreira e Barbosa (2016), foi apresentado um classificador baseado em algoritmos genéticos, utilizando parâmetros baseados em Estatística de Ordem Superior para classificar cargas acionadas isoladamente. Dando continuidade a esse trabalho, nesta dissertação propõe-se realizar a classificação de cargas acionadas simultaneamente, ou seja, cargas acionadas enquanto outras cargas já estão ligadas, diferenciando-se quanto ao banco de dados e os classificadores a serem utilizados.

Em uma residência, no dia a dia, cargas são acionadas simultaneamente, mas como o banco de dados montado para o trabalho anterior não contemplava a associação de cargas, faz-se necessária uma nova aquisição de dados. Além disso, também é necessário evoluir os classificadores para que sejam capazes de identificar as associações de cargas de forma correta.

1.1 Objetivo Geral

O objetivo deste trabalho é obter um classificador não invasivo para cargas residenciais acionadas simultaneamente, utilizando Estatística de Ordem Superior para extrair parâmetros das assinaturas dos equipamentos elétricos e comparar técnicas de reconhecimento de padrões aplicadas através de Redes Neurais, Florestas Aleatórias e Máquina de Vetor Suporte.

1.2 Objetivo Específico

- Criar um banco de dados com sinais de cargas isoladas e simultâneas;
- Implementar as técnicas de inteligência computacional (Redes Neurais, Máquina de Vetor Suporte e Florestas Aleatórias) para obter um classificador de cargas ligadas simultaneamente em uma mesma rede residencial;
- Comparar as técnicas de classificação a fim de obter o melhor classificador.

1.3 Motivação

Na última década, os relatórios da Empresa de Pesquisa Energética (EPE) apontaram um aumento no consumo residencial de 54%, no período de 2006 a 2016 devido ao aumento do poder aquisitivo da população e das reduções em IPI (Imposto sobre Produto Industrializado) para eletrodomésticos.

Entretanto, ainda é preciso responder às seguintes questões: Os consumidores conseguem gerenciar o seu consumo? Sabem o que consome mais ou quais os horários de maior consumo em sua residência?

Em janeiro de 2018, entrou em vigor a Tarifa Branca para os consumidores residenciais, que traz uma variação nos valores para o consumo em determinados horários do dia.

O consumidor que optar por essa tarifa pagará um valor menor que a tarifa atual, nos horários fora da faixa de 18 horas às 22 horas, e valores maiores nessa faixa de horário que é chamada de Intermediária (18h e 22h) e Ponta (19h às 21h). Mas a Agência Nacional de Energia Elétrica (ANEEL) adverte que, antes de optar pela Tarifa Branca, o consumidor conheça seu perfil de consumo e a relação entre a Tarifa Branca e a Convencional. Quanto

mais o consumidor deslocar seu consumo para o período fora de ponta e quanto maior for a diferença entre essas duas tarifas, maiores serão os benefícios da Tarifa Branca. Contudo, se ocorrer o consumo maior nos períodos de ponta e intermediário e não houver possibilidade de transferência do uso dessa energia elétrica para o período fora de ponta, a Tarifa Branca não é recomendada. Nesses casos, a Tarifa Branca pode resultar em uma conta maior e torna-se mais vantajoso continuar na Tarifa Convencional (AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA - ANEEL, 2016). A Tarifa Branca já está em vigor para as unidades que consomem mais de 500kWh/mês. Estima-se que mais de 4,5 milhões de unidades se enquadram nesse perfil. Além disso, em 2019, a tarifa branca poderá ser adotada por consumidores que possuem consumo superior a 250kWh/mês e, em 2020, será estendida para todas as unidades.

Atualmente não se tem conhecimento detalhado do consumo nem pelos clientes nem pelas concessionárias. Para as pesquisas de consumo, são necessárias equipes que vão a campo realizar enquetes para o levantamento de dados, embora não seja eficaz devido aos erros do tipo sistemático que tornam os dados não representativos, tais como informações faltantes devido à negativa de resposta por parte do entrevistado ou devido às falhas por parte do entrevistador. Ainda pode-se citar outros erros, como por exemplo, o fator psicológico de cada indivíduo, que quando questionados, podem ceder informações diferentes dos seus hábitos de consumos reais ou, ainda, quando questionado quantas vezes acendem uma lâmpada no dia, não conseguem informar corretamente.

Esse detalhamento do consumo pode ser obtido utilizando técnicas de monitoramento invasivo e não invasivo. No monitoramento invasivo, instalam-se os equipamentos de monitoramento entre a tomada e o eletrodoméstico para que seja monitorada a sua utilização e consumo. Essa abordagem pode se tornar inconveniente dependendo do equipamento utilizado para o monitoramento, pois pode ser necessário que o cliente registre em um diário, de tempo em tempo o consumo do equipamento. Além disso, pode ser dispendioso dependendo da quantidade de equipamentos elétricos a ser monitorada. (XU; MILANOVIC, 2015).

No monitoramento não invasivo (*Non-Intrusive Appliance Load Monitoring - NIALM*), utiliza-se apenas um equipamento na entrada de energia elétrica do consumidor. Essa técnica tem maior privacidade para o consumidor e pode trazer uma redução representativa no custo de hardware, quando comparada com a técnica invasiva para monitoramento de complexos residenciais ou uma casa com grande quantidade de equipamentos elétricos a serem monitorados. Contudo, esse método depende da resolução dos

dados para identificar a assinatura de carga, como por exemplo a forma de onda ou os transitórios (XU; MILANOVIC, 2015). O sistema de monitoramento de cargas não invasivo (NIALM) é uma importante ferramenta quando se fala das casas inteligentes (*Smart Home*) e das redes inteligentes (*Smart Grid*).

Esse sistema funciona como um sensor que permite ao consumidor residencial monitorar a corrente, tensão e potência consumida por seus eletrodomésticos, controlar o gasto mensal com energia elétrica e conhecer o perfil de utilização dos equipamentos eletroeletrônicos a fim de optar por tarifas que possam proporcionar economias nos faturamentos mensais ou configurar uma casa inteligente (*Smart Home*). Para as concessionárias de energia elétrica, esse sistema permite monitorar o perfil dos clientes, discriminar a fatura de consumo, identificar possíveis furtos de energia e ainda contribui significativamente para os programas de eficiência energética. Além do mais, pode auxiliar as empresas concessionárias de energia elétrica a tomar decisões sobre áreas de concentração de cargas (ferramenta para tomada de decisões).

Como principais contribuições deste trabalho podem-se citar o início da montagem de um banco de dados de cargas elétricas residenciais. Dar continuidade ao trabalhado de identificação do acionamento de cargas elétricas visando à utilização da ferramenta no futuro. O comparativo entre as técnicas de classificação: Redes Neurais Artificiais, Florestas Aleatórias e Máquina de Vetor Suporte. Validar a utilização de Estatística de Ordem Superior nesse tipo de trabalho.

1.4 Trabalhos Relacionados

O monitoramento de equipamentos elétricos tem recebido atenção de pesquisadores ao longo dos últimos anos. Um algoritmo para monitoramento não-invasivo de cargas elétricas (*No-Invasive Load Monitoring - NILM*) baseado em *Deep learning* (técnica de aprendizagem de máquinas) foi proposto por Bonfigli et al. (2018). Foram utilizados três datasets, *Almanac of Minutely Power* que contém dados de dois anos referentes a uma residência do Canadá, o *dataset UK-DALE* que é composto por perfis de consumo registrados em cinco casas no Reino Unido ao longo de dois anos e o *dataset REDD* que contém perfis de energia agregados e em nível de circuito de várias residências dos EUA. Os cinco aparelhos-alvo considerados em todos os experimentos são: geladeira, máquina de lavar, lava-louças, chaleira e micro-ondas. Os resultados foram superiores ao algoritmo de *Maximu Posteriori*

Aproximado Factorial (AFAMAP), modelo utilizado como referência. O monitoramento NILM é tratado como um problema de redução de ruído resolvido usando arquitetura de autoencoder de desruidização (dAE).

Em Liu et al. (2018), uma modelagem gráfica probabilística foi proposta. Uma estrutura de rede para padrões espaço-temporal (STPN) foi utilizada com dados multivariados de séries temporais para NILM. A estrutura proposta do STPN é capaz de usar diversos tipos de dados, descobrindo padrões específicos do uso de energia / geração em medições fasoriais e desagregação dos usos das cargas com alta precisão. O *dataset* foi montado com dados de circuitos elétricos colhidos da concessionária de energia da Califórnia. Foram adotados dois períodos de monitoramento diferentes tendo em vista a mudança do consumo de acordo com o período do ano. A técnica supera o estado da arte fatorial escondido Modelos de Markov (FHMM).

No trabalho de Laput, Zhang e Harrison (2017), foi realizada uma abordagem utilizando um sensor multifuncional capaz de detectar o acionamento das cargas elétricas, utilização das torneiras, dentre vários outros tipos de interações em uma residência. Apesar de ser mais invasivo e a proposta ser diferente da que está em proposição por este projeto, os sensores identificam o acionamento das cargas e a duração da sua utilização. Este trabalho demonstra o atual interesse na filosofia de se monitorar a utilização das cargas elétricas e outros consumos em uma residência.

A multiplicidade de métricas, a variedade de conjuntos de dados e a diversidade de metodologias tornam quase impossível uma comparação objetiva, segundo a revisão de Nalmpantis e Vrakas (2018). Um estudo abrangente, conduzido por Hosseini et al. (2017), foi realizado a fim de revelar as significativas inevitabilidades de um NILM, contribuindo para a ideia das casas inteligentes (*Smart Home*). O foco do estudo foi mostrar e discutir questões críticas relacionadas às necessidades e pré-requisitos do NILM, identificando obstáculos e propondo soluções que o torna viável.

Em Dinesh et al. (2017), foi proposto um novo método de Monitoramento Não-Invasivo de Cargas (NILM) para residências com instalações de energia solar instalada. Esse método identifica simultaneamente a quantidade de influxo de energia solar e os dispositivos que estão ligados, seus modos de operação e níveis de consumo de energia. Em Gillis e Morsi (2017), foi realizada uma abordagem baseada em *wavelet* e aprendizado de máquina semi-supervisionado que foram aplicados no monitoramento não-invasivo de cargas elétricas. As cargas utilizadas foram: um carregador de bateria; uma lâmpada fluorescente compacta; um computador pessoal; e uma lâmpada incandescente. Um co-treinamento de dois

classificadores foi usado para automatizar o processo de aprendizagem do padrão das cargas depois de projetar novas *wavelets*. Os resultados mostraram melhoria na precisão média da classificação obtida usando aprendizado semi-supervisionado quando comparado ao aprendizado de máquina supervisionado, atingindo 96,24% e 94,36% para o caso do vizinho mais próximo e da árvore de decisão, respectivamente.

O trabalho de Kim, Le e Kim (2017) apresenta o problema para monitoramento não-invasivo de cargas quando há aparelhos multiestado e consumo de energia similar a esses aparelhos. Para realizar a desagregação das cargas, foi utilizada a metodologia baseada em redes neurais recorrentes de memória de curto prazo (LSTM-RNN) e foi proposta uma nova assinatura para os equipamentos multiestado. A ideia principal da assinatura é separar o sinal original usando uma taxa de reflexão e subtrair um sinal de energia variante de outro sinal de energia variante.

A contribuição de Cominola et al. (2017) foi um algoritmo híbrido, eficiente computacionalmente, denominado Desagregação Iterativa Baseada em Assinatura Híbrida (HSID). Este é baseado nos modelos ocultos de Markov associados ao algoritmo iterativo *Subsequence Dynamic Time Warping*.

Uma abordagem do evento ON-OFF foi realizada por Yang, Soh e Yap (2015), na qual o evento foi monitorado de forma não invasiva com o objetivo de determinar sua duração. A identificação do evento ON-OFF deu-se com os transitórios característicos da natureza de cada equipamento. A partir de um conjunto de dados de um *dataset* foi realizada uma filtragem, utilizando a mediana para redução de ruídos. Para a detecção do evento, foi combinado um método estatístico para verificar a aderência dos dados (*goodness-of-fit* - GOF). Após a aplicação do GOF, o conjunto de dados positivos foi separado do conjunto negativo e calculada a corrente de cada conjunto. Para diferenças menores que 10%, o evento positivo identifica o evento ON e o conjunto negativo o evento OFF. Para finalizar, foi aplicada a metodologia do K-means (método de aprendizado de máquina não supervisionado) para agrupar os eventos semelhantes, obtendo a informação de quanto tempo o equipamento ficou ligado.

Outras abordagens sobre o assunto foram realizadas por Wang e Srinivasan (2015) que utilizaram o K-means para classificar diferentes ciclos de operação de equipamentos eletrodomésticos. Agyeman, Han e Han (2015) realizou uma abordagem diferente, utilizando o modelo escondido de Markov (HMM) para detectar um aparelho e seu estado de operação.

Um estudo sobre a assinatura do campo eletromagnético dos equipamentos foi realizado por Kulkarni, Harnett e Welch (2015). Uma antena a 30 cm do equipamento em uso capta o campo eletromagnético do equipamento e o identifica, por exemplo, um secador de cabelo tem um campo eletromagnético gerado pela corrente elétrica que o percorre, já um ventilador tem um campo eletromagnético distinto devido à corrente elétrica possuir componentes harmônicas. Após a coleta de dados referentes ao campo eletromagnético foi implementado um algoritmo baseado em árvore de decisão para classificar uma lâmpada incandescente, uma geladeira, um secador de cabelo, uma lâmpada fluorescente e uma lâmpada fluorescente compacta.

Outros autores como Lai et al. (2012), Nascimento (2012) e Xu e Milanovic (2015) apresentaram trabalhos relacionados a desmembramento de agrupamento de cargas, utilizando técnicas não invasivas associadas a um algoritmo de aprendizagem de máquina para identificar o consumo dos equipamentos.

1.5 Organização do documento

Este trabalho está dividido em partes independentes que se complementam. No capítulo 2, são abordados os conceitos teóricos relevantes ao desenvolvimento dos classificadores de cargas elétrica. No capítulo 3, são abordados os materiais necessários e a forma como ocorreu a implementação de todas as etapas para a realização deste trabalho. Adiante, com o capítulo 4, são expostos os resultados alcançados e suas análises. Finalizando, tem-se a conclusão e as referências bibliográficas utilizadas para embasamento desta dissertação.

2 REFERENCIAL TEÓRICO

Alguns conceitos da literatura serão apresentados neste capítulo a fim de proporcionar um melhor entendimento do trabalho realizado.

2.1 Estatística de Ordem Superior (EOS)

Um sinal gaussiano é caracterizado por sua variância e por sua média, seu gráfico possui característica simétrica e é representado pelo formato da “curva do sino”. A Estatística de Ordem Superior - EOS é menos afetada por ruídos Gaussianos que pela estatística de segunda ordem, pois as EOS de um sinal Gaussiano são nulas ou redundantes. Nas últimas décadas, a Estatística de Ordem Superior tem sido utilizada com frequência na literatura, uma vez que as aplicações práticas possuem ruído aditivo gaussiano (MENDEL, 1991).

A Estatística de Ordem Superior pode levar a resultados mais representativos quando utilizada em sistemas não lineares. Além disso em termos de Momentos que são utilizados para sinais determinísticos e cumulantes para sinais aleatórios (MENDEL, 1991).

O uso de EOS tem sido reportado na literatura por diversos autores. Com foco em energia elétrica podem ser citados os trabalhos realizados por Alameda-Hernández et al. (2016) que utilizaram cumulantes para criar um novo método que pode facilitar novas abordagens para análise da ergodicidade em sistemas de energia. Em Quirós-Olozábal et al. (2016), foi utilizada curtose para detecção de harmônicos. A curtose foi utilizada por Zhang et al. (2014) para discriminação da corrente inrush de magnetização e falha interna em transformadores de potência. Em seu trabalho, Zhang et al. (2014) utilizou cumulantes para propor um esquema de proteção de linhas de transmissão. No trabalho de Guedes, Ferreira e Barbosa (2016), eles utilizaram cumulante e método não invasivo para um classificador de carga não simultânea.

Como o sistema elétrico alimenta cargas não lineares, o uso de métodos baseados em cumulantes é o que será adotado neste trabalho. As expressões dos cumulantes de segunda, terceira e quarta ordem de um sinal aleatório $E\{x[n]\}=0$, estão representados nas equações de 1 a 3 respectivamente (MENDEL, 1991):

$$C_{2,x}[i] = E\{x[n] x[n+i]\}, \quad (1)$$

$$C_{3,x}[i] = E\{x[n] x^2[n+i]\}, \quad (2)$$

$$C_{4,x}[i] = E\{x[n] x^3[n+i]\} - 3C_{2,x}[i] C_{2,x}[0], \quad (3)$$

em que i é o i -ésimo atraso e $x[n]$ é o N -ésimo elemento do vetor \mathbf{x} . Para um vetor com comprimento finito (N), as aproximações estocásticas oferecem as seguintes expressões:

$$\hat{C}_{2,x}[i] := \frac{2}{N} \sum_{n=0}^{N-1} x[n] x[n+i], \quad (4)$$

$$\hat{C}_{3,x}[i] := \frac{2}{N} \sum_{n=0}^{N-1} x[n] x^2[n+i], \quad (5)$$

$$\hat{C}_{4,x}[i] := \frac{2}{N} \sum_{n=0}^{N-1} x[n] x^3[n+i] - \frac{2}{N^2} \sum_{n=0}^{N-1} x[n] x[n+i] \sum_{n=0}^{N-1} x^2[n], \quad (6)$$

em que $i = 0, 1, \dots, N/2 - 1$.

No entanto, para $i > N/2 - 1$ as equações (4) à (6) não podem ser usadas devido a $n+i$ ser maior que N . Desse modo, as informações no cálculo dos cumulantes serão perdidos. A proposta de Ribeiro et al. (2007) é uma abordagem alternativa, em que cada cumulante é calculado utilizando todas as N amostras de sinal. Basicamente forma-se um *buffer* circular nos sinais de tal maneira que, o valor de $n+i$ seja maior que N em k unidades. Esse valor é substituído por k , tomando com isso as amostras antecessoras ao valor i , que anteriormente não eram utilizadas. Na verdade, é como se houvesse uma continuidade da última amostra para a primeira, fazendo como que o sinal seja virtualmente circular. Matematicamente, a substituição de $n+i$, quando $n+i > N$, pode ser dada por:

$$\text{mod}(n+i, N) = [n+i] - bN, \quad (7)$$

em que b é o número inteiro obtido quando se desconsideram as casas decimais da divisão de $n+i$ por N , o que nada mais é que, o resto da divisão $n+i$ por N . Dessa forma, as expressões (4) a (6) podem ser estimadas, para N finito, por:

$$\hat{C}_{2,x}[i] := \frac{1}{N} \sum_{n=0}^{N-1} x[n] x[\text{mod}[n+i, N]] \quad (8)$$

$$\hat{C}_{3,x}[i] := \frac{1}{N} \sum_{n=0}^{N-1} x[n]x^2[\text{mod}[n+i, N]] \quad (9)$$

$$\hat{C}_{4,x}[i] := \frac{1}{N} \sum_{n=0}^{N-1} x[n]x^3[\text{mod}[n+i, N]] - \frac{3}{N^2} \sum_{n=0}^{N-1} x[n]x[\text{mod}[n+i, N]] \sum_{n=0}^{N-1} x^2[n], \quad (10)$$

em que $\text{mod}[n+i, N]$ é o resto inteiro da divisão de $n+i$ por N .

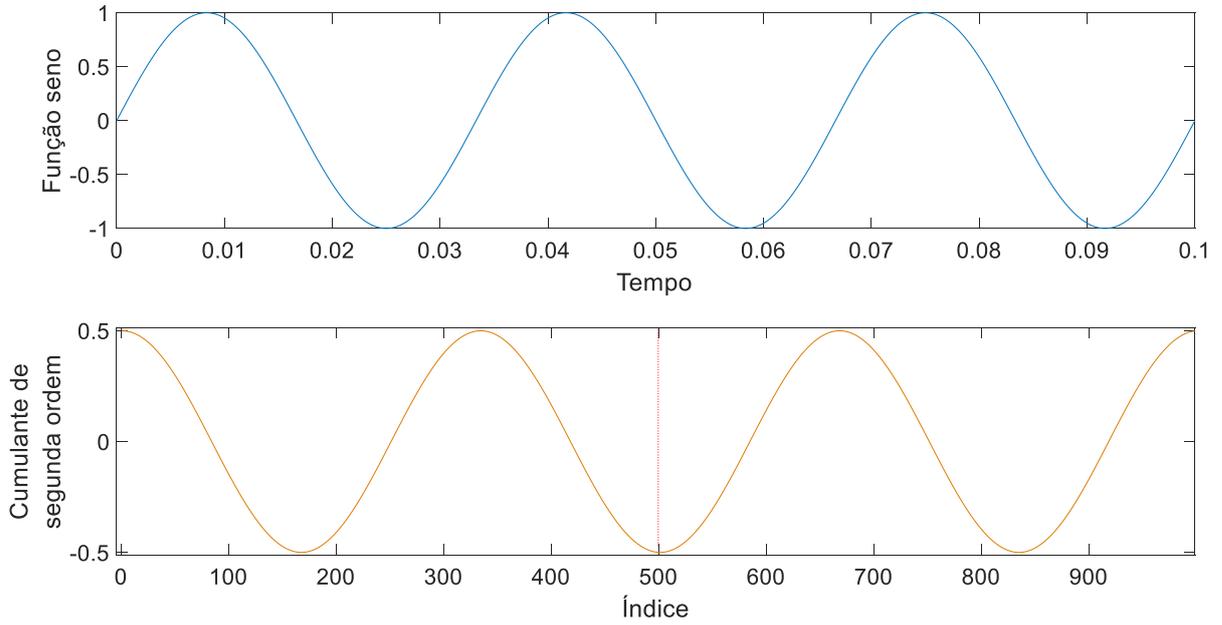
Para um vetor de comprimento finito, as equações (8) a (10) apresentam uma boa simplificação para o problema. Essa aproximação é indicada para sinais periódicos, nos quais pode-se incluir os sinais de tensão e corrente.

Em um sinal em que o número de amostras é igual a N , tem-se N cumulantes para cada ordem da EOS. Conforme mostrado por Ferreira et al. (2009) e Ribeiro et al. (2007), a combinação de poucos desses cumulantes são suficientes para obtenção de um bom desempenho quando se trata de detecção e classificação. Neste trabalho será utilizado apenas o cumulante de segunda e quarta ordem, haja vista que o cumulante de terceira ordem pode não apresentar boa representatividade para alguns sinais eletrônicos (FERREIRA et al., 2009), isso ocorre por ser esse cumulante equivalente à assimetia e os sinais serem simétricos.

O cumulante de segunda ordem de um sinal é igual a sua variância, ou seja, $C_2 = \sigma^2$. A Figura 1 mostra o cumulante de segunda ordem para um sinal senoidal sem interferências de harmônicos e ruído. Na Figura 2, é acrescentado apenas ruído gaussiano com taxa Sinal-Ruído de 20,09dB, um vez que nas aplicações práticas é comum a presença de ruídos. Na Figura 3, são inseridas distorções harmônicas no sinal senoidal, os harmônicos/amplitude são: $3^\circ/2$, $5^\circ/4$, $7^\circ/1$, $9^\circ/1$ e $11^\circ/3$.

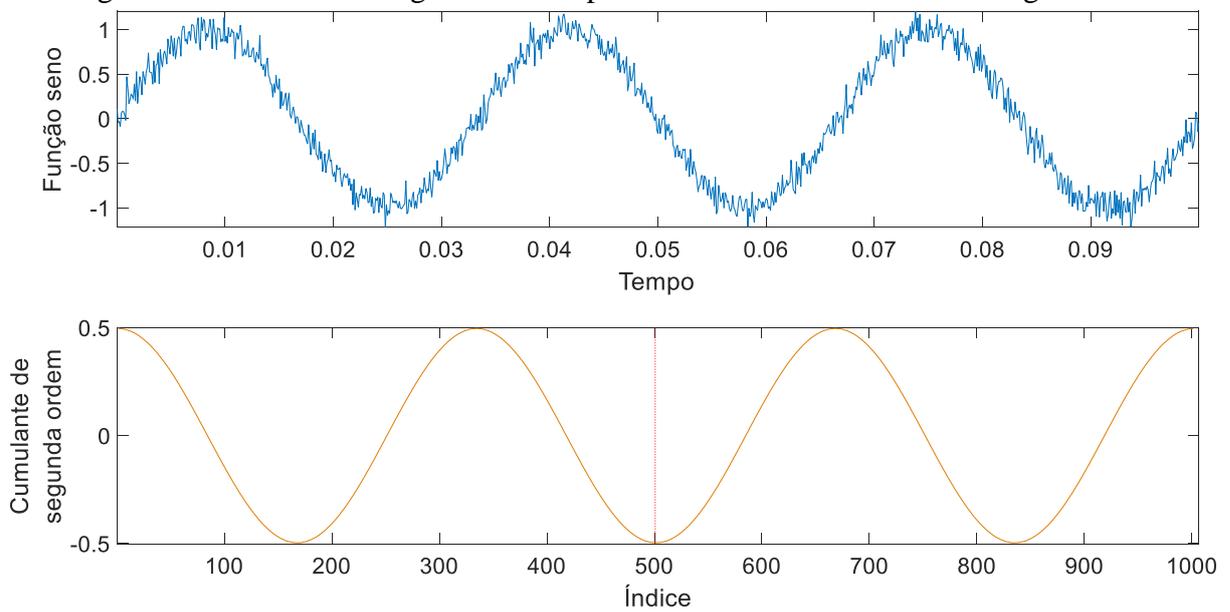
Observando as Figuras de 1 a 3, é possível visualizar o comportamento do cumulante de segunda ordem na extração de parâmetros. A primeira observação que pode ser feita é que o cumulante de segunda ordem para sinais que possuem distribuição normal, ou é nulo ou é redundante, nesse caso apresentou redundância. Isto é, a primeira metade do gráfico é simétrica à segunda metade para todos os casos apresentados. Outra observação a ser feita é que o cumulante de segunda ordem apresentou uma certa “imunidade” ao ruído gaussiano que foi inserido no sinal senoidal, conforme Figura 1 e Figura 2.

Figura 1: Cumulante de segunda ordem para um sinal senoidal que não possui interferência de harmônicos e/ou ruídos.



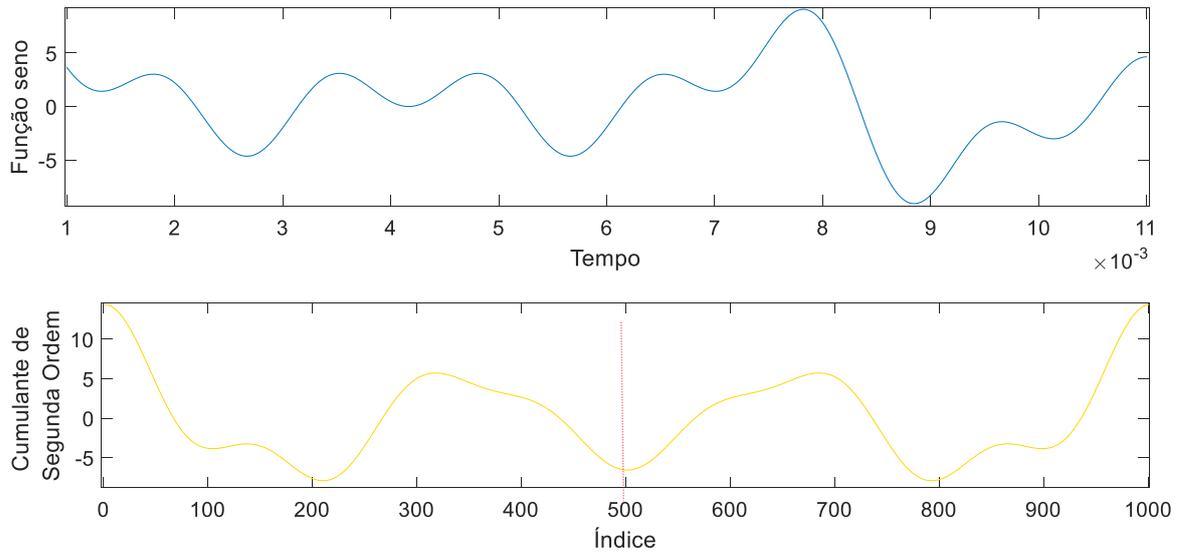
Fonte: Do autor (2018).

Figura 2: Cumulante de segunda ordem para um sinal senoidal com ruído gaussiano.



Fonte: Do autor (2018).

Figura 3: Cumulante de segunda ordem para um sinal senoidal com distorções harmônicas.

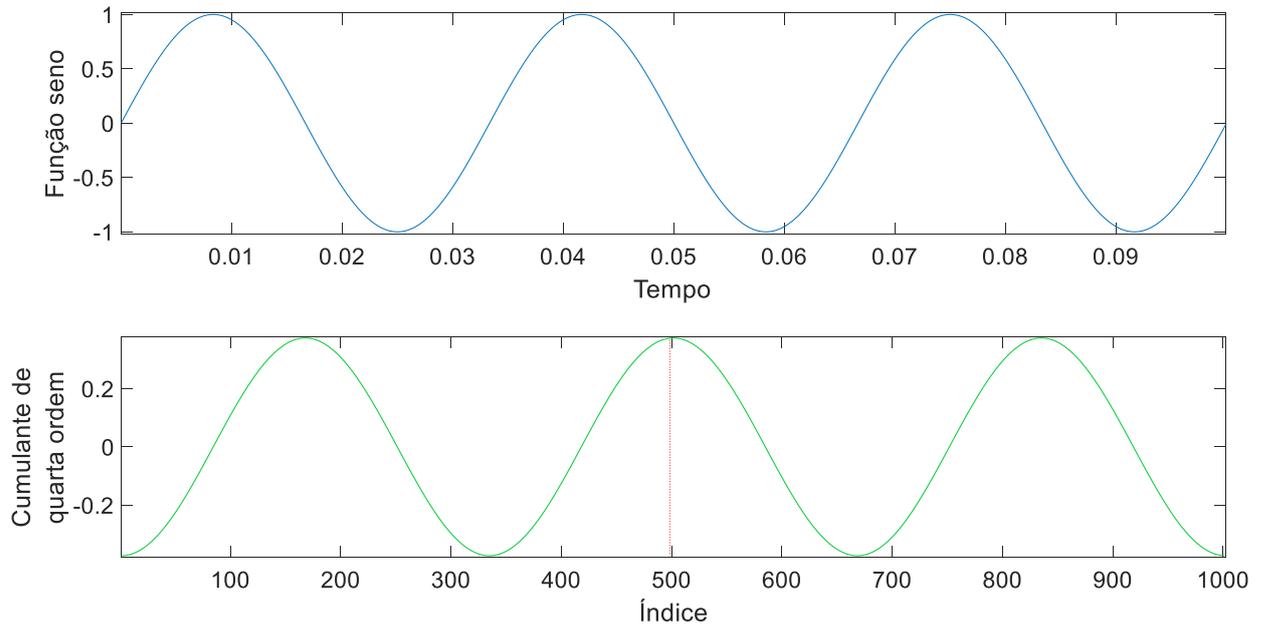


Fonte: Do autor (2018).

O cumulante de primeira ordem é equivalente a média do sinal, como já mencionado, o cumulante de segunda ordem é equivalente a variância, o cumulante de terceira ordem é equivalente à assimetria do sinal. A partir do cumulante de quarta ordem as equivalências não são diretas, o mesmo equivale à curtose do sinal.

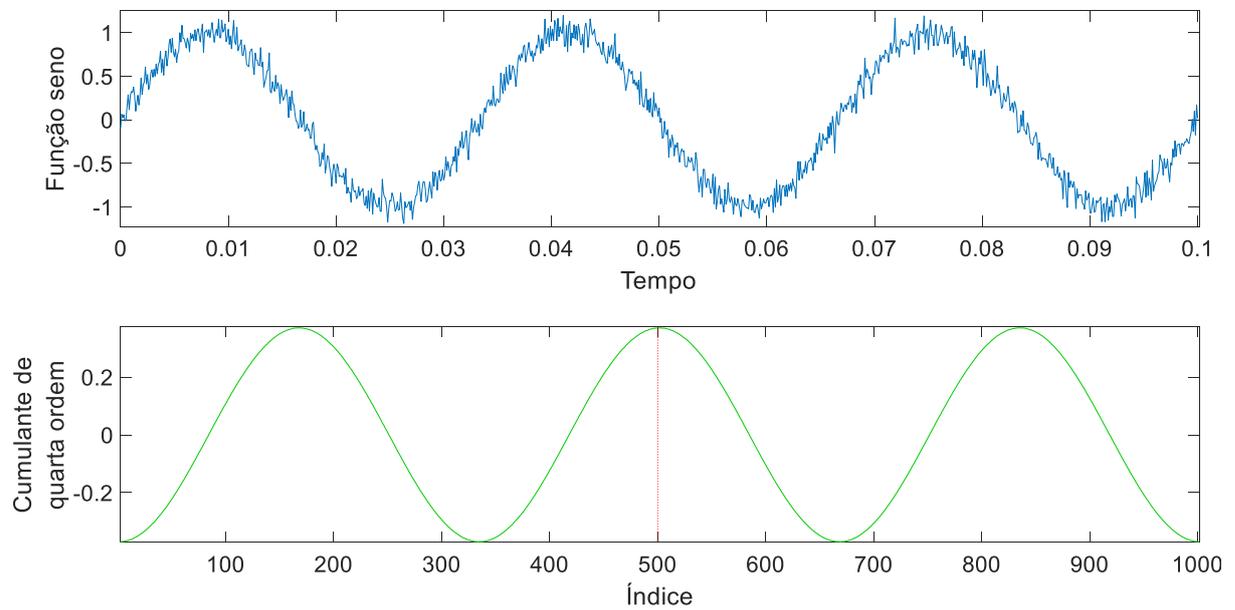
Nas Figuras de 4 a 6, estão representados os cumulantes de quarta ordem. É possível visualizar o comportamento desse cumulante na extração de parâmetros. A primeira observação que pode ser feita é que o cumulante de quarta ordem para sinais que possuem distribuição normal não é necessariamente redundante, ou seja, não é constante a simetria conforme pode ser visto na Figura 6, na qual a segunda metade do gráfico difere da primeira. Outra observação a ser feita é que o cumulante de quarta ordem manteve a “imunidade” ao ruído gaussiano que foi inserido no sinal senoidal, conforme Figura 5, o que permite dizer que a Estatística de Ordem Superior é menos afetada por ruídos gaussianos presentes no sistema de amostragem do sinal.

Figura 4: Cumulante de quarta ordem para um sinal senoidal sem presença de ruídos e/ou harmônicos.



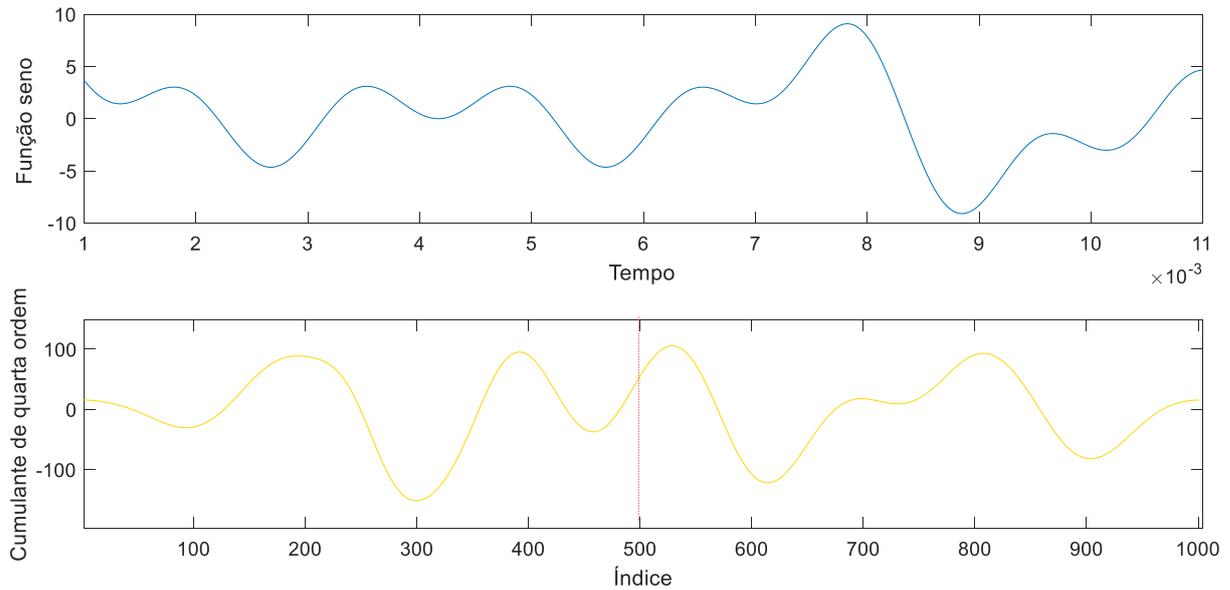
Fonte: Do autor (2018).

Figura 5: Cumulante de quarta ordem para um sinal senoidal com ruído gaussiano.



Fonte: Do autor (2018).

Figura 6: Cumulante de quarta ordem para um sinal senoidal com distorções harmônicas.



Fonte: Do autor (2018).

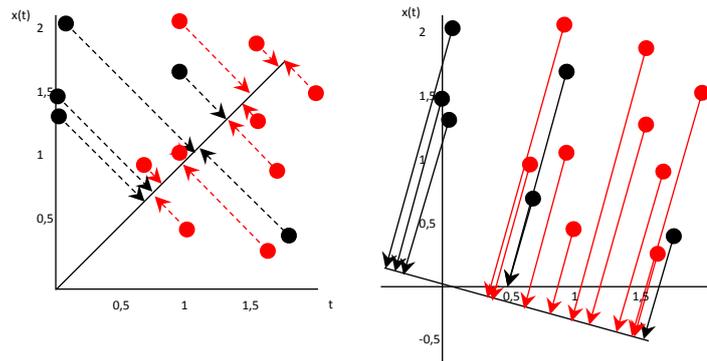
2.2 Relação Discriminante de Fisher

Um dos problemas encontrados ao aplicar técnicas de estatística em trabalhos de reconhecimento de padrões é conhecido como “maldição da dimensionalidade”, trata-se do fenômeno em que o número de elementos requeridos para que um determinado classificador tenha um bom desempenho é uma função crescente, em alguns casos cresce exponencialmente, esse problema de aumento da dimensão dos dados torna impraticável os trabalhos para dimensões de 50 ou 100 (DUDA; HART; STORK, 2000).

É possível reduzir a dimensionalidade de d dimensões para uma dimensão projetando os dados d -dimensionais em uma linha. Mesmo que as amostras formem clusters compactos bem separados no espaço d e a projeção em uma linha produza uma mistura confusa de amostras de todas as classes, se rotacionar a linha conforme Figura 7, pode-se encontrar uma orientação para a qual as amostras projetadas estejam bem separadas e isso é exatamente o objetivo da análise discriminante clássica.

Duda, Hart e Stork (2000) e Theodoridis e Koutroumbas (2009) demonstram o discriminante de Fisher (*Fisher Discriminant Ratio – FDR*) tanto para classificação quanto para seleção de parâmetros, por ser uma técnica simples para discriminação de dados multidimensionais. Porém, este trabalho adotará apenas a aplicação do discriminante como seleção de parâmetros.

Figura 7: Projeção das amostras em um eixo reduzindo a dimensão dos dados.



Fonte: Duda, Hart e Stork (2000).

A média das amostras d -dimensionais da classe i é:

$$\boldsymbol{\mu}_i = \frac{1}{n_i} \sum_{x \in C_i} x \quad (11)$$

Então a função custo do critério do FDR como ferramenta de seleção de parâmetro é:

$$J_c = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^2 \circ \frac{1}{\boldsymbol{\sigma}_1^2 - \boldsymbol{\sigma}_2^2} \quad (12)$$

$J_c = [J_1 \dots J_{L_l}]^T$, L_l é o número total de parâmetros, $\boldsymbol{\mu}_1$ e $\boldsymbol{\mu}_2$ são os vetores médias e $\boldsymbol{\sigma}_1$ e $\boldsymbol{\sigma}_2$ são os vetores variância dos dados das classes 1 e 2. O símbolo \circ é referente ao produto de Hadamard, ou seja, $\mathbf{r} \circ \mathbf{s} = [r_1 s_1 \dots r_{L_r-1} s_{L_r-1}]^T$ (RIBEIRO et al., 2007).

Para os maiores valores de J_c , os i -ésimos elementos do vetor de parâmetros representam os parâmetros com melhor separabilidade entre classes. Essa técnica é utilizada em classificações mais complexas por permitir reduzir a quantidade de parâmetros e verificar os melhores a serem utilizados na classificação. Essa redução é importante para se ter uma menor complexidade computacional em sistemas que operam em tempo real.

Uma outra vantagem em reduzir a quantidade de parâmetros é o desenvolvimento de classificadores mais simples e com melhor desempenho, uma vez que os parâmetros referentes à fronteira de decisão entre as classes foram retirados.

2.3 Redes Neurais Artificiais (RNA)

Entre as década de 80 e 90, as Redes Neurais Artificiais (RNA) começaram a ganhar destaque na comunidade acadêmica e atualmente são amplamente utilizadas em diversas aplicações, tais como tarefas de classificação e regressão ou até em soluções de problemas relacionados a otimização.

As RNAs são modelos matemáticos aplicados a técnicas computacionais capazes de generalizar informações após os treinamentos, que podem ser supervisionados ou não. Atribui-se o sucesso das redes neurais a sua flexibilidade de síntese de mapeamentos multidimensionais não-lineares de entrada e saída, refletida em sua capacidade de aproximação universal (HAYKIN, 2008).

2.3.1 Neurônio Artificial

A RNA é baseada nos neurônios humanos, o primeiro modelo de neurônio artificial foi proposto por Mcculloch e Pitts (1943). Mas apenas em Rosenblatt (1957) foi desenvolvido o modelo de neurônio artificial que é utilizado até hoje, este é chamado de *Perceptron*. O neurônio artificial está baseado em receber diferentes sinais de entrada x_n (dendritos) que são multiplicados por um peso, todos estes sinais então são somados e, utilizando uma função de ativação, o neurônio devolve um único sinal de saída y (axônio), conforme Figura 8.

A função que representa esse neurônio é dada pela Equação (13).

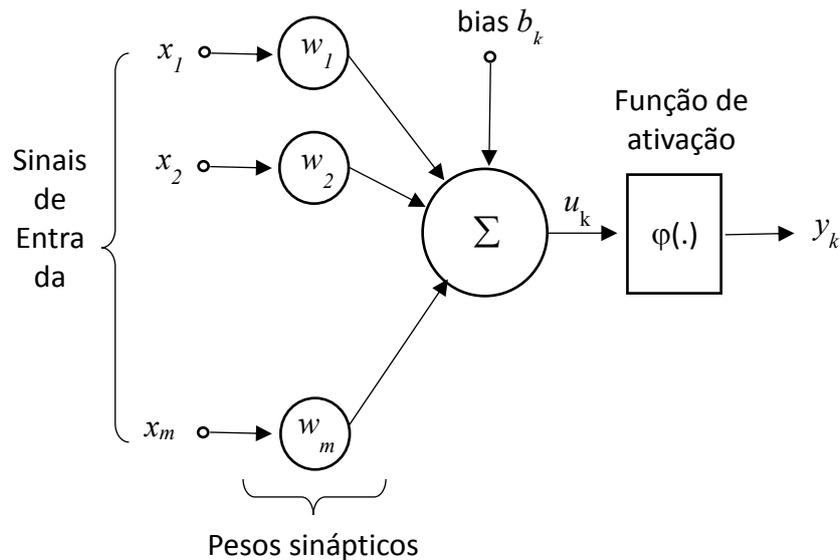
$$y = \sum_{i=1}^m w_i * x_i + b \quad (13)$$

A normalização dos dados é uma etapa de pré-processamento fundamental para ter um bom resultado com as RNAs. A normalização linear considera os valores máximos e mínimos de cada variável durante a transformação em um intervalo fechado, que pode ser entre 0 e 1 ou entre -1 e 1, mantendo a distância entre os valores normalizados proporcionais às distâncias entre os dados originais, conforme Equação (14) (VALENÇA, 2007).

$$x' = \frac{(x - x_{min})(b - a)}{(x_{max} - x_{min})} + a \quad (14)$$

em que x' é o valor normalizado, x é o valor original, x_{min} é o valor mínimo da variável, x_{max} é o valor máximo da variável e a é o limite inferior do intervalo de normalização 0 ou -1; e b é o limite superior do intervalo de normalização 1.

Figura 8: Representação de um neurônio não linear.



Fonte: Haykin (2008)

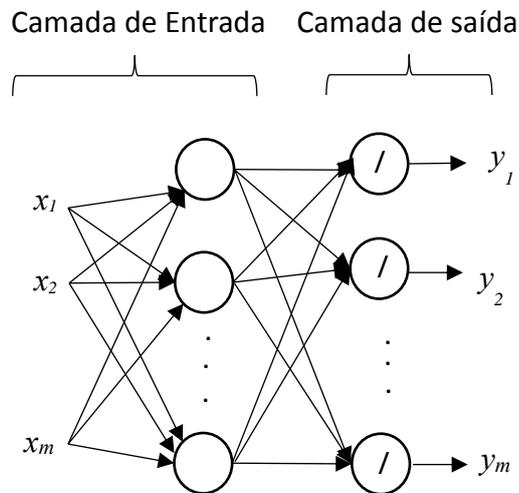
A função de ativação define a saída do neurônio em termos do potencial de ativação, seu objetivo é limitar a saída entre 0 e 1 ou entre -1 e 1. Diversas funções de ativação poderão ser utilizadas, no entanto as mais comuns são as funções sigmoidal, tangente hiperbólica e identidade.

2.3.2 Arquitetura das RNA's

Existem diversas arquiteturas de RNA's, a forma com que os neurônios são conectados uns aos outros define como deverão ser implementados os algoritmos de aprendizado de máquina para cada rede.

Nas redes *feedforward* de camada única, os neurônios da camada de entrada são conectados diretamente aos neurônios da camada de saída. O fluxo de informação é somente para frente, por isso o nome *feedforward*. A Figura 9 mostra um exemplo desse tipo de rede.

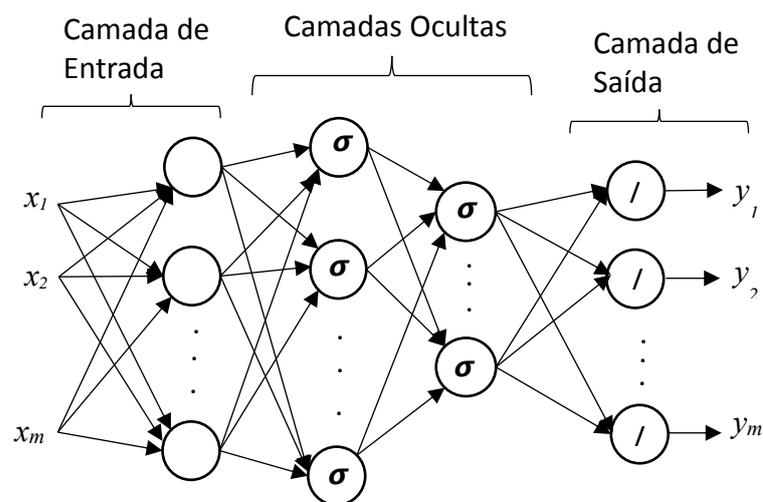
Figura 9: Representação de uma rede multicamada, quais sejam uma camada de entrada x , duas camadas ocultas e uma camada de saída y .



Fonte: Do autor (2018).

A topologia de RNA mais utilizada em aplicações de reconhecimento de padrões é a rede com múltiplas camadas. Haykin (2008) definiu essa rede como um conjunto de unidades sensoriais que constituem a camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Essas redes são chamadas de perceptrons de múltiplas camadas (*Multilayer Perceptron - MLP*) (ROSENBLATT, 1962). A Figura 10 ilustra um perceptron com quatro camadas, duas camadas são ocultas.

Figura 10: Representação de uma rede multicamada, quais sejam, uma camada de entrada x , duas camadas ocultas e uma camada de saída y .

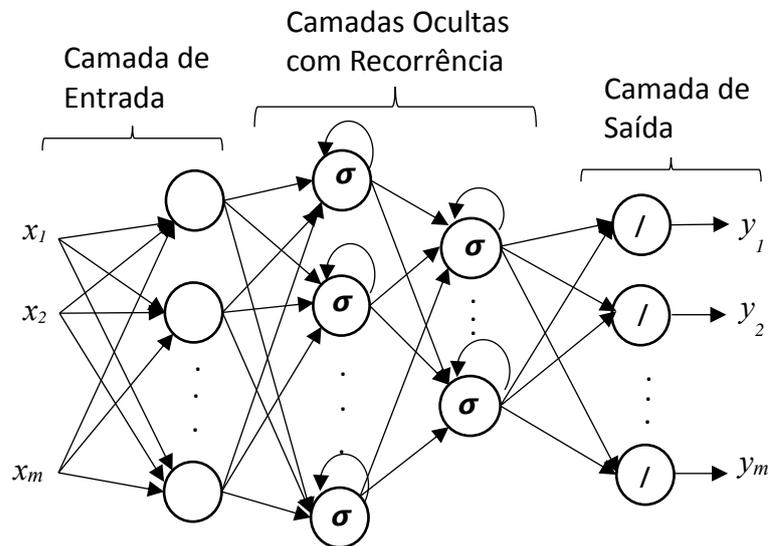


Fonte: Do autor (2018).

Existem ainda as redes recorrentes, estas podem ser de camada única ou multicamadas, no entanto existe uma retroalimentação nos neurônios. Esse novo fluxo de informação no neurônio é chamado de *feedback loop* e pode ocorrer a conexão de um neurônio com neurônios das camadas anteriores ou com ele mesmo. A Figura 11 representa uma rede multicamada com recorrências nas camadas ocultas.

As redes multicamadas têm sido aplicadas com sucesso para a resolução de diversos problemas complexos e não-lineares, por meio de treinamentos supervisionados com um algoritmo de retropropagação de erro (*error back-propagation*) (HAYKIN, 2008).

Figura 11: Representação de uma rede multicamada com recorrências nas camadas ocultas.



Fonte: Do autor (2018).

2.3.3 Formas de Treinamento

Quanto às formas de treinamento da RNA, estas podem ser classificadas conforme o algoritmo de aprendizado utilizado. Existem três formas de treinamento que são mais utilizadas: aprendizado supervisionado, aprendizado não-supervisionado e aprendizado semi-supervisionado.

No aprendizado supervisionado, existe um “professor” que apresenta as respostas corretas ao algoritmo, e este por sua vez, as compara com as respostas obtidas pela rede neural artificial. Então, o erro entre o valor obtido e o valor desejado é calculado e o algoritmo realiza os ajustes dos pesos de cada sinapse da rede. Isso ocorre repetidas vezes até conseguir o resultado satisfatório.

O aprendizado não-supervisionado não tem o “professor” e o treinamento ocorre por meio do comportamento estatístico dos dados. Já o aprendizado semi-supervisionado ocorre quando, para uma parte dos dados, se conhece a resposta desejada.

2.3.4 *Back-Propagation* (Aprendizado por retropropagação)

As RNAs possuem uma etapa de aprendizagem ou treinamento que consiste em um processo iterativo de ajustes dos pesos de cada conexão entre as unidades de processamento que são armazenadas ao final do processo. Além disso, extraem informações relevantes dos padrões de informação e criam assim uma representação própria.

As redes multicamadas têm sido aplicadas com sucesso para a resolução de diversos problemas complexos e não-lineares, por meio de treinamentos supervisionados com um algoritmo de retropropagação de erro (*error back-propagation*) (HAYKIN, 2008).

Esse algoritmo é baseado na regra de aprendizagem por correção do erro, ou seja, são fornecidas informações na entrada da rede, essas informações são processadas e passam para a camada da frente até alcançar a saída da rede. O algoritmo compara a saída obtida com a desejada e, então, por meio de uma regra, calcula-se o erro. Com base nesse erro obtido, o processo de retropropagação consiste em ajustar os pesos sinápticos da RNA começando na saída da rede passando para camada anterior até alcançar a camada de entrada da rede.

Em Haykin (2008) tem-se que a aproximação é obtida por meio da redução do índice de desempenho conforme Equação (17).

$$\xi(n) = \frac{1}{2} e_k^2(n). \quad (17)$$

A redução do índice de desempenho continua ocorrendo até que o critério de parada seja alcançado, podendo ser o máximo de épocas ou iterações, ou que seja obtida a resposta desejada, ou ainda que um limite pré-definido para o índice de desempenho seja alcançado. Para os ajustes de redução do índice $\xi(n)$, utiliza-se a regra de aprendizado conhecida como regra delta (WIDROW; HOFF, 1960). Na regra delta, o peso sináptico w_{kj} no tempo n recebe o ajuste Δw_{kj} calculado conforme Equação (18).

$$\Delta w_{kj}(n) = \eta e_k(n) x_j(n), \quad (18)$$

sendo η a taxa de aprendizado e w_{kj} o peso sináptico da sinapsis j pertencente ao neurônio k . Como a regra delta presume que o sinal de erro seja mensurável, é necessário que a saída desejada para a rede seja informada previamente por uma fonte externa acessível ao neurônio k (HAYKIN, 2008).

Assim, conforme Haykin (2008), os ajustes dos pesos a partir da regra delta é denominado Regra Delta Generalizada, e pode ser obtido por meio da Equação (19).

$$w_{ji}^{(l)}(n+1) = w_{ji}^{(l)}(n) + \alpha[w_{ji}^{(l)}(n-1) + \eta\delta_j^{(l)}(n)y_i^{(l-1)}(n)]. \quad (19)$$

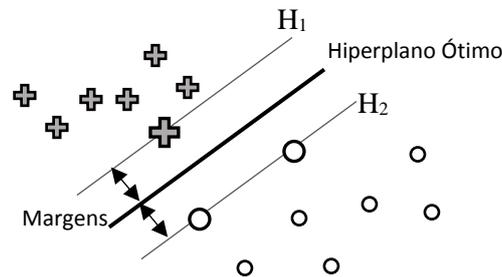
2.4 Máquina de Vetor Suporte (SVM)

A Máquina de Vetor Suporte (*Support Vector Machines - SVM*) é um algoritmo que tem sido aplicado a diversas áreas das ciências. Em Hassan, Javed e Arshad (2014), o algoritmo foi utilizado para avaliar um novo tipo de perfil de assinatura bidimensional para cargas elétricas residenciais, visando ao monitoramento não invasivo. No trabalho de Eseye, Zhang e Zheng (2018), eles propuseram um modelo híbrido combinando transformada *wavelet*, otimização de enxame de partículas e máquina de vetores de suporte para previsão de curto prazo (um dia a frente) da geração de um sistema fotovoltaico de uma microrrede real.

A máquina de vetor suporte é uma técnica de aprendizagem de máquina que usa conceitos de lógica binária, ou seja, que usa separação automática entre classes diferentes, reconhece os padrões utilizando técnicas de estatística de regressão (VAPNIK, 2013).

Em sua concepção inicial, o SVM foi desenvolvido para a classificação binária e para encontrar um hiperplano capaz de separar os dados em dois grupos. Contudo, existem diversos classificadores que fazem isso, então o SVM tenta encontrar o melhor hiperplano de modo que cada classe ou grupo de dados possa ter a maior área que os represente. Essas áreas são chamadas de hiperplanos de margem máxima ou hiperplanos ótimos, que são traçados a partir dos hiperplanos auxiliares H1 e H2 conforme Figura 12 (VAPNIK, 2013).

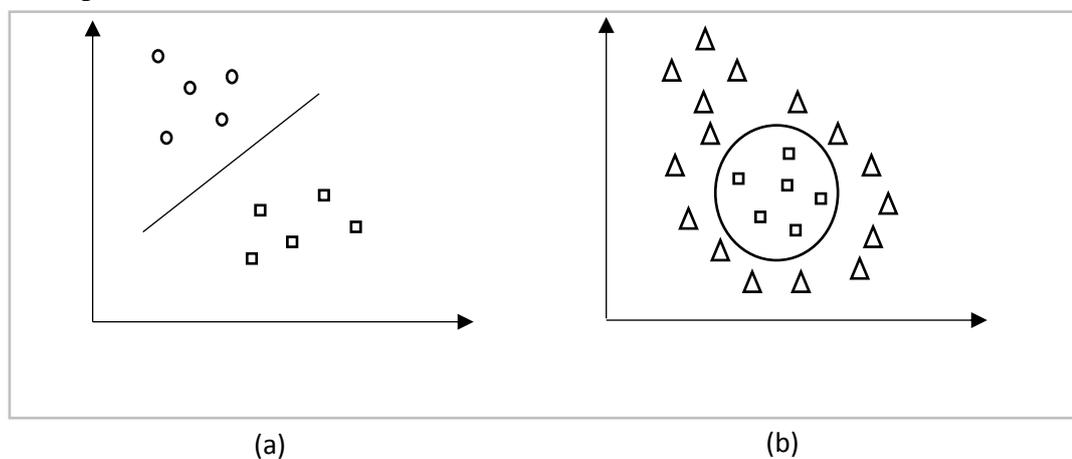
Figura 12: Esquema apresentando o hiperplano ótimo para classificar os dados com as maiores margens.



Fonte: Do autor (2018).

Porém há muitos casos em que não se podem dividir as classes de dados por um hiperplano, um exemplo poderá ser visto na Figura 13. Em problemas não lineares, as máquinas de vetor suporte mapeiam o conjunto de treinamento de seu espaço original, para um novo espaço de maior dimensão, denominado espaço de características (HEASRT et al., 1998).

Figura 13: (a) Conjunto de dados linearmente separáveis; (b) Conjunto de dados não separáveis linearmente.



Fonte: Do autor (2018).

A alta capacidade de generalização da SVM e a capacidade de modelar problemas não lineares, faz com que seja utilizado para problemas de regressão e séries temporais.

Para um conjunto de N amostras de treinamento no espaço R^n , por exemplo, (x_1, y_1) , (x_2, y_2) , ..., (x_n, y_n) , sendo $x_i \in R^n$ e $y_i \in [-1, +1]$, e ainda x conjunto de entrada e y é usado para representar a que classe os dados pertencem. A otimização resolvida pelo algoritmo pode ser dada pela Equação (20).

$$\min\left(\frac{1}{2w^T}w + C \sum_{i=1}^l \xi_i\right), \quad (20)$$

tendo a restrição demonstrada na equação (21):

$$\begin{cases} y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i, \\ \xi_i \geq 0, \end{cases} \quad (21)$$

onde ϕx_i mapeia x_i em um hiperplano. A constante C é positiva e responsável pelo balanço entre a simplicidade do modelo e a capacidade de generalizar (CORTES; VAPNIK, 1995), quanto maior o valor, maior a quantidade de vetores suportes serão utilizados e mais complexo é para determinar as fronteiras de decisão do hiperplano.

Uma vez que a variável w pode apresentar mais de uma dimensão, então a otimização pode ser dada pela Equação (22).

$$\min\left(\frac{1}{2\alpha^T}Q\alpha - e^T\alpha\right) \quad (22)$$

tendo a restrição demonstrada na equação (23):

$$\begin{cases} y^T\alpha = 0 \\ 0 \leq \alpha_i \leq C, \end{cases} \quad (23)$$

sendo $e = [1,1,1,1,\dots,1]^T$, e Q é uma matriz positiva de tamanho x , calculada por meio da Equação (24):

$$Q_{ij} \equiv y_i y_j K(x_i, x_j) \quad (24)$$

onde K é uma função conhecida como *kernel*. Escolher o melhor *kernel* é um dos parâmetros a serem determinados pelo projetista. Uma das funções utilizadas como *kernel* é a Função de Base Radial (RBF), essa função consegue mapear os dados que não são linearmente separáveis, por isso recomenda-se iniciar com essa função (HSU; CHANG; LIN, 2003), e pode ser definida conforme Equação (25).

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (25)$$

sendo γ correspondente ao grau de influência que uma única amostra do treino pode ter. Logo, quanto maior for o valor de γ menor será sua influência. Outra função utilizada é a polinomial definida conforme Equação (26).

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0 \quad (26)$$

No entanto, quando a dimensão dos dados é muito grande, a função mais indicada é a *Kernel Linear*, que está definida na Equação (27) (HSU; CHANG; LIN, 2003).

$$K(x_i, x_j) = x_i^T x_j \quad (27)$$

Definindo a função *Kernel* e otimizando o problema de minimização apresentado na Equação (22), obtém-se a equação geral de classificação apresentada (28), sendo *sgn* a função sinal.

$$y_i(w^T \phi(x_i) + b) = \text{sgn} \left(\sum_{i=1}^l \alpha_i y_i K(x_i, x) + b \right) \quad (28)$$

2.5 Florestas Aleatórias

Floresta Aleatória (do inglês *Random Florest*) é um método de classificação proposto por Breiman (2001). Esse método é baseado na técnica de agregação de classificadores do tipo árvore de decisão, construído para que sua estrutura seja composta de forma aleatória. Para determinar as classes, o método utiliza a combinação de várias árvores de decisão por meio do mecanismo de votação, no qual cada árvore dá a sua classificação para a classe. A classificação final se dá pela classe que mais foi votada entre as demais árvores da floresta (BREIMAN, 2001).

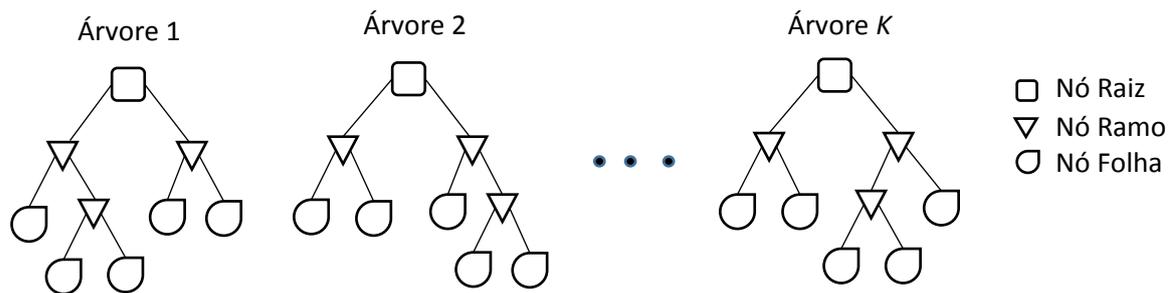
Esse método é um grupo de árvores de decisão $\{h(X, \theta_k)$, em que $k = 1, 2, \dots, K\}$ onde $\{\theta_k\}$ são vetores randômicos distribuídos identicamente e independentemente, e K é o número de árvores de decisão. Em uma variável de um sinal elétrico X , cada classificador das árvores votam para determinar qual é o melhor resultado de classificação (BREIMAN, 2001).

A partir de um conjunto de dados de treinamento, e utilizando o método estatístico de *bootstrap*, z novas amostras independentes e definidas aleatoriamente são selecionadas. Então são construídas K árvores de decisão.

Em um conjunto de N características, seleciona-se aleatoriamente m características ($m \leq N$). Ao calcular a quantidade de informação contida em cada característica, seleciona-se a que possui melhor capacidade de classificação para executar a divisão de nó.

A Figura 14 representa uma floresta de K árvores de decisão compostas pelo Nó Raiz, Nó Ramo e Nó Folha. O Nó Raiz é o ponto de início da classificação, o qual representa as características mais apropriadas para um determinado sinal elétrico. Os Nós Ramos dividem esses dados em dois conjuntos que possuem regras diferentes e os Nós Folha obtêm os resultados da classificação desses dados.

Figura 14: Estrutura da árvore de decisão, esta figura representa uma floresta de K árvores contendo o Nó Raiz, Nó ramo e o Nó Folha.



Fonte: Do autor (2018).

A função margem dada na Equação (29) é usada para mensurar o grau da média de classificação correta e o erro médio de classificação. Dado um conjunto de classificadores $h_1(x), h_2(x), \dots, h_k(x)$, e o conjunto de dados de treino selecionado aleatoriamente da distribuição de vetores aleatórios Y e X , define-se a margem (BREIMAN, 2001):

$$mg(X, Y) = av_k I(h_k(X) = Y) - \max_{j \neq Y} av_k I(h_k(X) = j) \quad (29)$$

na qual $I(\cdot)$ é a função de indicação. A margem mensura a medida em que a média dos números de votos em X e Y para a classe certa excede a média de votos em qualquer outra classe. O erro de generalização pode ser dado pela Equação (30) (BREIMAN, 2001):

$$PE^* = P_{X,Y}(mg(X,Y) < 0) \quad (30)$$

Na Equação (30), os subscritos X e Y indicam que a probabilidade P abrange o espaço X e Y. À medida que o número de árvores de decisão na floresta aleatória aumenta, todas as sequências $\theta_1, \theta_2, \dots, PE^*$ (uma vez que θ_k é uma variável aleatória independente e distribuída de forma igual), convergem para a Equação (31) (BREIMAN, 2001):

$$P_{X,Y}(P_{\theta}(h(X, \theta) = Y) - \max_{j \neq Y} P_{\theta}(h(X, \theta) = j) < 0) \quad (31)$$

Essas equações demonstram que as Florestas Aleatórias não sofrem o excesso de treinamento (*overfit*) quando mais árvores são adicionadas, mais produzem um valor limite do erro de generalização.

2.6 Matriz de Confusão

Em aprendizado de máquina, a matriz de confusão é uma das métricas mais utilizadas para avaliar os resultados obtidos por um classificador. Cada coluna da matriz de confusão representa a instância real dos dados e cada linha representa as instâncias que foram classificadas pelo algoritmo de aprendizado de máquina (MATRIZ..., 2017).

Para melhor visualização, a Tabela 1 apresenta um exemplo de matriz de confusão, com duas instâncias, uma positiva e outra negativa. Os Verdadeiros Positivos (VP) são os dados positivos que foram classificados como positivo, os Falsos Positivos (FP) são os dados negativos que foram classificados como positivos, os Falsos Negativos (FN) são os dados positivos que foram classificados como negativos e os Verdadeiros Negativos são os dados negativos classificados como negativos.

Com base na Tabela 1, para verificar a qualidade dos resultados obtidos poderão ser calculadas algumas métricas tais como: Acurácia, sensibilidade, precisão, F1 *score*, dentre outras.

Tabela 1: Matriz de Confusão para Positivos e Negativos.

	Positivo	Negativo
Positivo	VP	FP
Negativo	FN	VN

Fonte: Do autor (2018).

2.6.1 Acurácia

A acurácia nada mais é que a proporção de predições corretas e não leva em consideração a qual instância pertence o dado. A acurácia é calculada pela Equação (32).

$$Acurácia = \frac{\text{Total de acertos}}{\text{Total de dados do conjunto}} = \frac{VP + VN}{VP + FP + VN + FN} \quad (32)$$

2.6.2 Sensibilidade

A sensibilidade, também conhecida como revocação, é a capacidade do sistema classificar corretamente os dados que realmente pertencem à instância. É calculada pela Equação (33).

$$Sensibilidade = \frac{\text{Total de positivos corretamente classificados}}{\text{Total de dados positivos reais}} = \frac{VP}{VP + FN} \quad (33)$$

2.6.3 Precisão

Em algumas vezes confundida com a acurácia, a precisão é a taxa de acertos na classificação de dados como positivo em relação ao total de dados que foram classificados como positivo, pode ser calculada por meio da Equação (34).

$$Precisão = \frac{\text{Total de positivos corretamente classificados}}{\text{Total de dados classificados como positivo}} = \frac{VP}{VP + FP} \quad (34)$$

2.6.4 F1 - Score

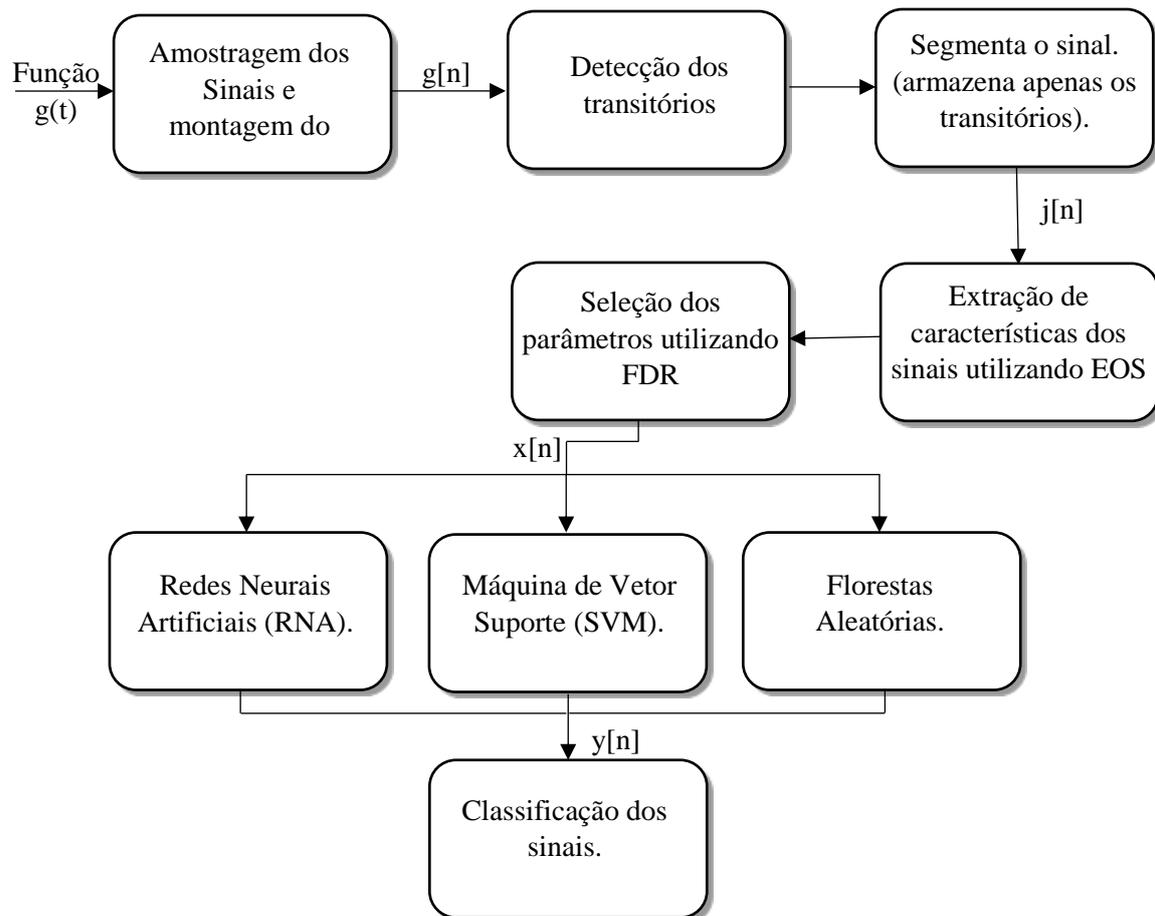
A pontuação F1 pode ser interpretada como uma média ponderada da precisão e sensibilidade, em que uma pontuação F1 atinge seu melhor valor em 1 e pior em 0. A taxa F1-Score balanceada pode ser dada pela Equação (35):

$$F_1 = 2 \cdot \frac{\textit{precisão} \cdot \textit{sensibilidade}}{\textit{precisão} + \textit{sensibilidade}} \quad (35)$$

3 MATERIAIS E MÉTODOS

Neste capítulo são apresentadas as etapas de implantação do projeto, desde a aquisição do sinal à sua classificação quanto à classe que pertence. Na Figura 15, verificam-se as principais fases para o desenvolvimento dos classificadores.

Figura 15: Fluxograma indicando as etapas a serem realizadas na fase de desenvolvimento do classificador.



Fonte: Do autor (2018).

Pode-se dividir este trabalho em duas partes, a primeira delas é a aquisição dos dados, momento em que é montado um banco de dados e a segunda parte o processamento desses sinais, que têm como produto final o classificador de acionamento de cargas.

3.1 Aquisição dos dados

A aquisição dos dados foi realizada utilizando cargas residenciais para obtenção do sinal de corrente elétrica que caracteriza melhor o tipo de carga acionada.

A coleta de dados utilizou cinco classes distintas que são combinadas em outras cinco classes. Foram coletados os dados com 50 repetições para cada classe, estas estão listadas na Tabela 2 e para cada conjunto de cargas ligadas simultaneamente conforme Tabela 3, dos 50 eventos registrados para cada classe de carga acionada junto a outra carga em operação, foram realizados 25 eventos com uma das cargas acionadas primeiro (Ex: Chuveiro acionado após a TV de LED estar em operação) e 25 com a segunda carga acionada antes (Ex: TV de LED acionada após o Chuveiro estar em operação). Os eletrodomésticos estão especificados na Tabela 4.

Tabela 2: Classes (C) para os equipamentos residenciais isolados a serem monitorados.

Ordem	Equipamento	Evento
C1	Chuveiro	50
C2	Televisão de LED 40 polegadas	50
C3	Micro-ondas	50
C4	Lampada de Led	50
C5	Multiprocessador	50

Fonte: Do autor (2018).

Tabela 3: Classes Simultâneas (CS) de equipamentos residenciais a serem monitorados.

Ordem	Equipamento	Evento
CS1	Chuveiro + Lâmpada de Led	50
CS2	Chuveiro + Multiprocessador	50
CS3	Lâmpada de Led + Multiprocessador	50
CS4	Televisão de LED 40 polegadas + Multiprocessador	50
CS5	Micro-ondas + Multiprocessador	50

Fonte: Do autor (2018).

A coleta dos dados foi realizada experimentalmente em uma residência, com frequência de amostragem de 15.360 Hz, utilizando a placa de aquisição de dados NI USB-6251 mostrada na Figura 16 e especificada no Anexo A. Em conjunto foi utilizada uma ponteira de corrente Fluke i200s, conforme Figura 17, e especificação no Anexo B.

Tabela 4: Especificação dos eletrodomésticos

Equipamento	Modelo	Fabricante	Tensão	Potência
Chuveiro	Bela Ducha 4T	Lorenzeti	127V	5500W
Televisão de LED	UN40C7000	SAMSUNG	120/240V	150W
Micro-ondas	MEF33	Eletrolux	127V	1300W
Lâmpada de Led	Pera	Avant	127/220V	7W
Multiprocessador	Mega Master	Walita	127V	500W

Fonte: Do autor (2018).

Figura 16: Placa de aquisição de NI USB-6251 fabricado por National Instruments



Fonte: http://www.ni.com/cms/images/devzone/tut/USB-6251_BNC.jpg

Figura 17: Ponteira de corrente FLUKE i200s



Fonte: Do autor (2018).

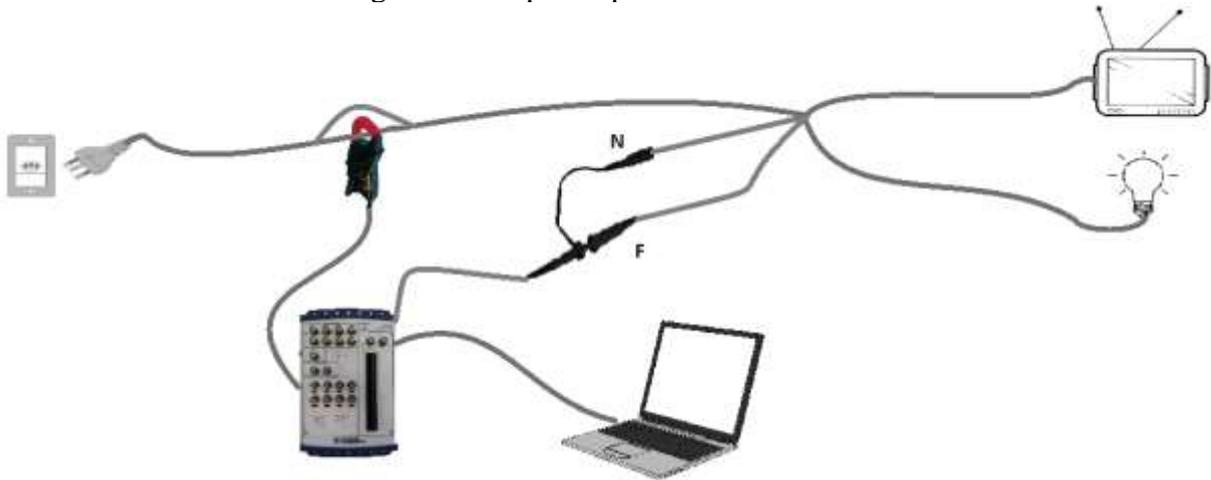
Também foram utilizados uma ponteira Tektronix P5100A 100X 500 MHz para aquisição dos dados relativos a tensão (Figura 18) e um cabo paralelo 2x4mm² para auxiliar as conexões dos equipamentos e aquisição dos sinais conforme esquema apresentado na Figura 19.

Figura 18: Ponteira para aquisição de dados de tensão.



Fonte: Do autor (2018).

Figura 19: Esquema para coleta de sinais



Fonte: Do autor (2018).

A Figura 20 apresenta a disposição dos equipamentos (multiprocessador de alimentos + lâmpada de LED) ligados. Para a lâmpada de LED, foi instalado um interruptor visando facilitar o evento liga/desliga. Adotou-se um multiplicador de tomada (conhecido como T) para fazer a conexão da lâmpada junto do multiprocessador. Uma das saídas do multiplicador de tomadas foi utilizada para fazer o monitoramento da tensão. Então o cabo de extensão foi ligado a uma tomada 127V, próximo ao padrão de entrada.

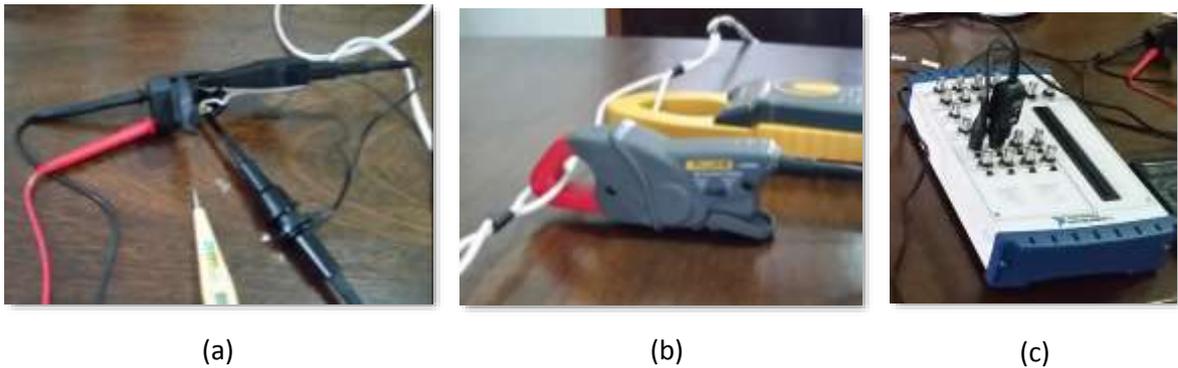
Figura 20: Multiprocessador de alimentos e Lâmpada de LED conectados no momento da aquisição dos dados.



Fonte: Do autor (2018).

A ponteira de corrente foi colocada em uma das fases do cabo de extensão, conforme mostra a Figura 21 (b). A ponteira de tensão foi conectada conforme apresentado na Figura 21 (a) e ambos foram conectados à Placa de Aquisição de dados NI USB-6251 conforme Figura 21 (c). Foram utilizados 2 multímetros para aferição dos dados coletados, conforme Figura 21.

Figura 21: (a) conexão da ponteira de tensão; (b) conexão da ponteira de corrente; (c) Conexão das ponteiras à placa de aquisição de dados NI USB-6251.



Fonte: Do autor (2018).

Utilizando um notebook Dell Vostro com 4GB de memória RAM, processador Intel i5 e com sistema operacional Windows 8.1, foi montada uma interface no software Labview, e foram então realizadas as coletas de dados conforme Figura 22. Para cada equipamento foram realizados 50 eventos de liga e desliga, assim como para cada associação de cargas e os dados foram armazenados no *Hard Disk* (HD) do notebook.

Figura 22: (a) Aquisição de dados da Lâmpada de LED ligada juntamente com o Multiprocessador de alimentos; (b) aquisição de dados da Lâmpada de LED.



Fonte: Do autor (2018).

3.2 Processamento dos dados

Após a etapa de coleta de dados, é necessário tratá-los para serem utilizados na classificação das cargas. Os dados foram importados para o software Matlab (deste ponto em diante todas as etapas foram implementadas utilizando o Matlab) e foram agrupados em um único arquivo para cada classe. Então foram realizadas as detecções das assinaturas das cargas, neste caso a energia dos transitórios. Quando não há carga nenhuma ligada, facilmente é identificado o transitório de entrada de uma carga em operação, todavia quando já existe uma carga em operação e uma segunda carga entra em operação, o regime transitório de entrada da segunda carga pode não ser facilmente identificado, conforme ilustrado na Figura 23.

Na Figura 23(a), tem-se a entrada de uma carga em operação sem nenhuma outra ligada, nesse caso, pode-se perceber que a energia do transitório é superior a do sinal antes e depois de sua ocorrência, facilitando sua identificação. A Figura 23(b) apresenta um transitório de partida de uma carga enquanto outra já está ligada, porém a energia desse transitório também é superior a do sinal antes e depois de sua ocorrência. O mesmo não ocorre na Figura 23(c), na qual o transitório de partida da segunda carga é muito inferior ao sinal da carga já ligada, o que torna complexa a identificação desse transitório.

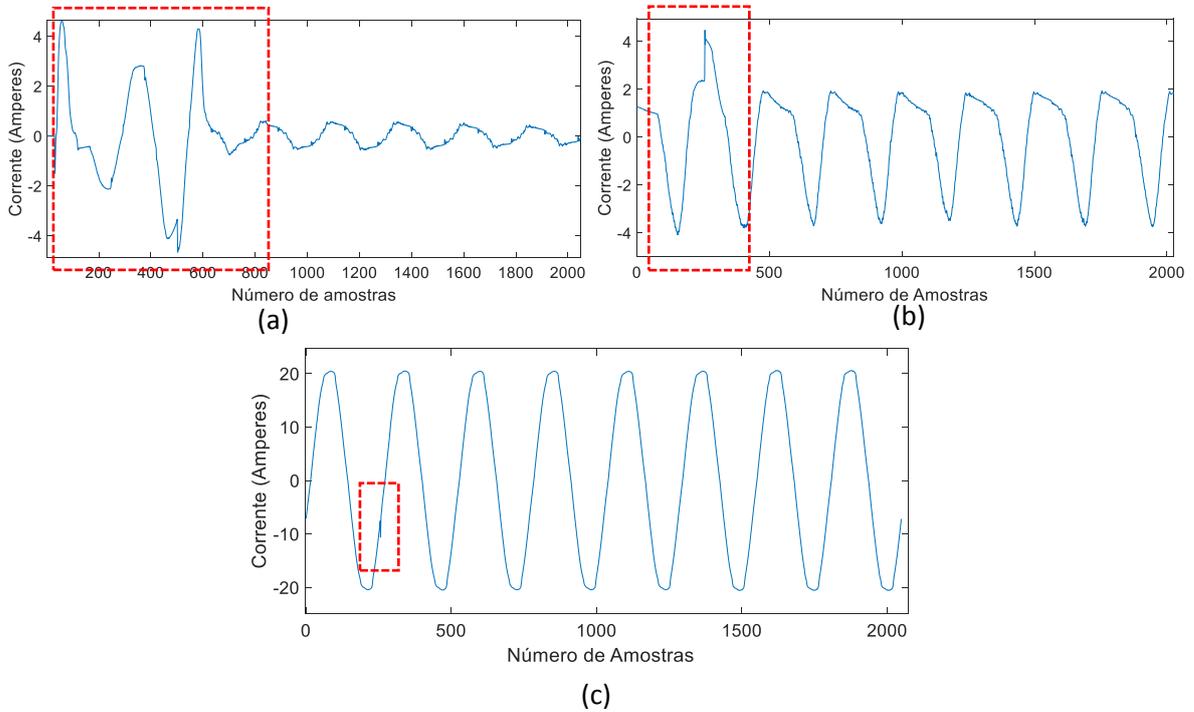
Para resolver o problema apresentado na Figura 23(c), foi implementado um filtro *Notch* e aplicado aos dados. Como resultado, a frequência de 60Hz foi atenuada, conforme ilustrado nas Figuras 24 e 25. Ressalta-se que o uso do filtro *Notch* para remover o componente em 60Hz é bastante favorável ao processo de classificação das cargas, visto que este componente é informação redundante, pois está presente em todos os sinais de acionamento das cargas.

O filtro *Notch* é um filtro rejeita banda, ou seja, permite passar todas as frequências exceto as que estão dentro da estreita banda de corte, nesse caso, 60Hz. O filtro *Notch* é um filtro IIR de 2ª ordem e a sua transformada Z está representada na Equação 36 (MITRA, 2005).

$$H_0(z) = \frac{1 + (-2\cos\omega_0)z^{-1} + z^{-2}}{1 + \rho(-2\cos\omega_0)z^{-1} + \rho_0^2 z^{-2}}, \quad (36)$$

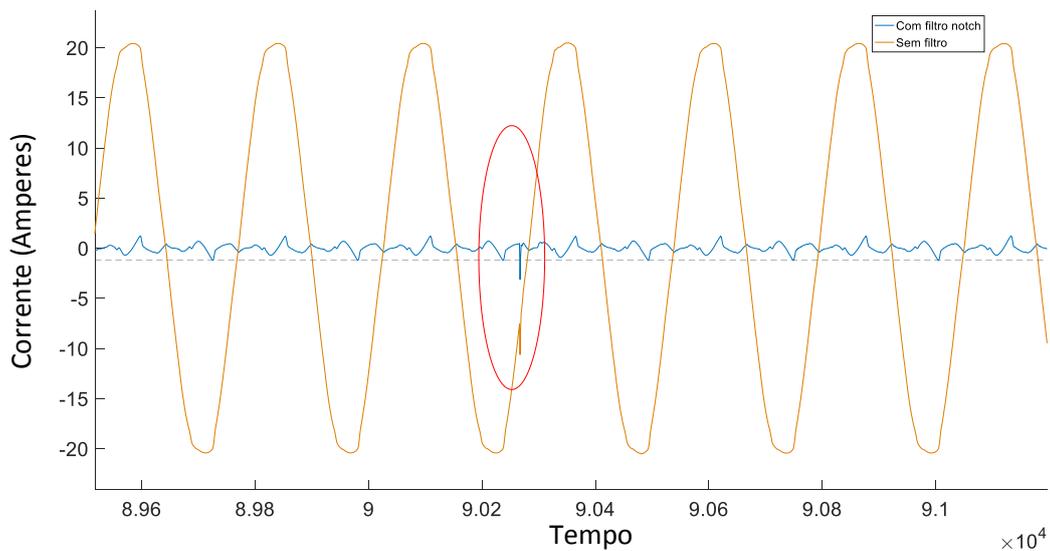
na qual ρ_0 é o fator *Notch* com $0 \leq \rho_0 \leq 1$. Por meio das Figuras 24 e 25, nota-se que os transitórios sobressaíram após a aplicação do filtro, o que facilitou a sua identificação.

Figura 23: (a) Transitório de partida de uma carga acionada isoladamente (TV de Led); (b) transitório de uma TV de Led enquanto um multiprocessador de alimentos está em uso e (c) transitório de uma lâmpada de Led enquanto um chuveiro está em uso.



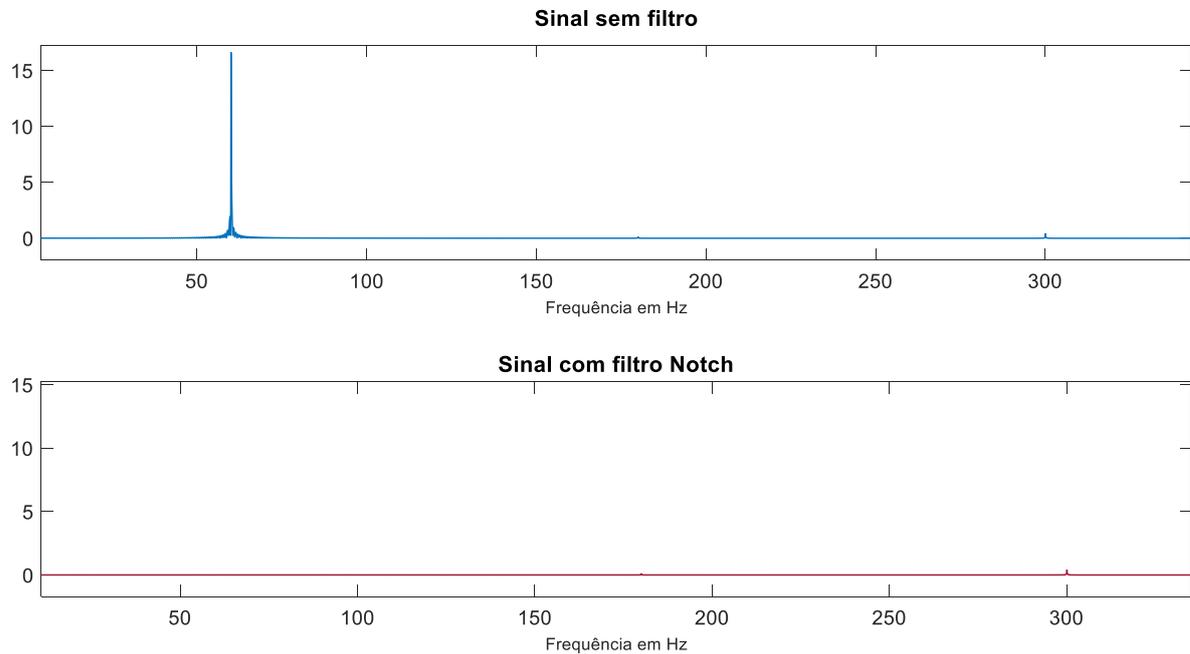
Fonte: Do autor (2018).

Figura 24: Acionamento da TV de Led após o acionamento do Chuveiro, em azul o sinal filtrado utilizando o filtro *Notch*.



Fonte: Do autor (2018).

Figura 25: Espectros de corrente elétrica referente ao acionamento de duas cargas (chuveiro e lâmpada de LED) antes e após a utilização do filtro *Notch*.



Fonte: Do autor (2018).

Após aplicar o filtro *Notch*, para se identificar o início do regime transitório de cada equipamento foi utilizado um detector de acionamento baseado na inovação da energia do sinal, que armazena um ciclo anterior ao instante de detecção do acionamento e três ciclos após o instante do acionamento detectado. Para a detecção de um evento discreto no tempo $x[n]$, o detector de inovação realiza a operação descrita na Equação (37) (GUEDES; FERREIRA; BARBOSA, 2016).

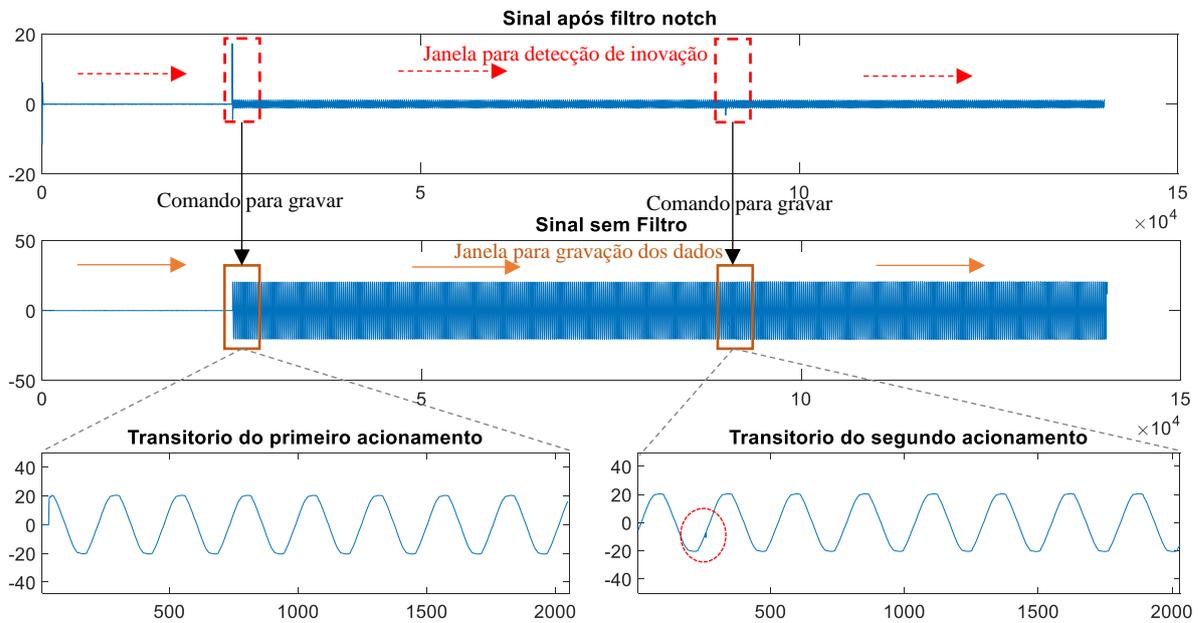
$$d = E_x[n + 5] - E_x[n - 5], \quad (37)$$

Sabendo que E_x é a energia do evento x em uma dada amostra n . Então para d maior que um limite previamente estabelecido para o conjunto de treinamento, o acionamento é detectado. Assim, a uma frequência de amostragem de 15.360Hz, tem-se 256 amostras por ciclo, como foi utilizada uma janela deslizante de 8 ciclos, foram armazenadas 2048 amostras por acionamento.

O detector de inovação foi aplicado ao sinal filtrado e a janela deslizante para a recorte dos transitórios foi aplicada ao sinal sem estar filtrado. Ao identificar o transitório, os dados

contidos na janela deslizante foram recortados e armazenados, conforme ilustrado na Figura 26.

Figura 26: Processo de identificação dos regimes transitórios referentes ao acionamento das cargas.



Fonte: Do autor (2018).

Após a etapa de identificação do transitório, foi realizada a etapa de extração de parâmetros. Para isso, foi aplicada a Estatística de Ordem Superior para obtenção dos cumulantes de segunda e quarta ordens. Em seguida, em cada classe, os dados foram embaralhados e o Discriminante de Fisher foi aplicado para selecionar os dois parâmetros que melhor representam os sinais de cada classe.

Uma forma diferente de extração de parâmetro foi utilizada, foi aplicada a Transformada de Fourier para a extração dos espectros de frequência de cada classe. Para fins de comparação, os dados foram divididos em conjunto de treino e conjunto de teste e classificados pela Rede Neural Artificial descrita em 3.3.1.

3.3 Treinamento dos classificadores propostos

Na etapa de classificação foram utilizados três tipos de classificadores: Redes Neurais Artificiais, Florestas Aleatórias e Máquina de Vetor Suporte. O projeto de cada um deles é apresentado a seguir.

3.3.1 Redes Neurais Artificiais (RNA)

Um algoritmo foi elaborado para realizar as classificações para cargas acionadas isoladamente, e a partir desse protótipo foram realizados diversos testes até se alcançar uma configuração que conseguisse obter um resultado satisfatório na classificação tanto de cargas individuais quanto para cargas simultâneas.

Foram montadas redes sem camada escondida, com 1 camada escondida, com 2 camadas escondidas e 3 camadas escondidas. A quantidade de neurônios em cada camada foi variada e a RNA escolhida foi montada com 20 entradas (Dimensão do vetor de parâmetros), 200 neurônios na primeira camada escondida, 150 neurônios na segunda camada escondida, 50 neurônios na terceira camada escondida e 10 saídas, ou seja, uma para cada classe.

Os dados foram divididos em 60% para treino e 40% para validação. Os dados de treinamento da rede foram subdivididos em Treino e Teste utilizando a função *net.divideParam.trainRatio* do MatLab, seguindo a mesma proporção de 60% e 40%, respectivamente, e como critério de parada foi adotado o erro médio quadrático de 4×10^{-3} ou 20.000 épocas.

As redes foram treinadas utilizando o algoritmo de treinamento de retropropagação com ajustes dos pesos pelo método do Gradiente Conjugado em Escala (*Scaled Conjugate Gradient - Backpropagation*) e como função de ativação a tangente hiperbólica. As redes foram validadas utilizando o conjunto de dados para validação e a acurácia foi calculada utilizando a média dos resultados obtidos com o treinamento e validação de 20 redes diferentes. Para os demais parâmetros foi mantido o *default* da biblioteca de redes neurais do Matlab.

3.3.2 Florestas aleatórias

O método de classificação utilizando Florestas Aleatórias foi implementado utilizando a função *TreeBagger* no MatLab. Foram realizados alguns testes com configurações

diferentes e o que foi adotado para o resultado final foi o ajuste dos parâmetros para 5 árvores de decisão. A função *OOBPrediction* foi acionada e foi selecionado o método de classificação no parâmetro *Method*.

3.3.3 Máquina de Vetor Suporte -SVM

Utilizando Máquina de Vetor Suporte, foi implementado um classificador utilizando a *Função de Base Radial-RBF* como *Kernel*. A função *fitcecoc* foi utilizada para parametrizar o *coding* em *allpairs*, uma vez que os parâmetros foram passados para o SVM em colunas então ele fará a comparação em linha. Após os ajustes de parâmetros, todos os classificadores foram treinados 20 vezes com dados escolhidos aleatoriamente para o conjunto de treino e de validação.

3.4 Métricas comparativas.

Após a implementação e treinamento dos classificadores, foi realizada uma análise de eficiência dos modelos propostos. Os classificadores foram treinados 20 vezes, com dados sorteados aleatoriamente e os resultados foram armazenados. As métricas utilizadas foram a Acurácia, Sensibilidade, Precisão e *F1 score*.

4 RESULTADO E DISCUSSÃO

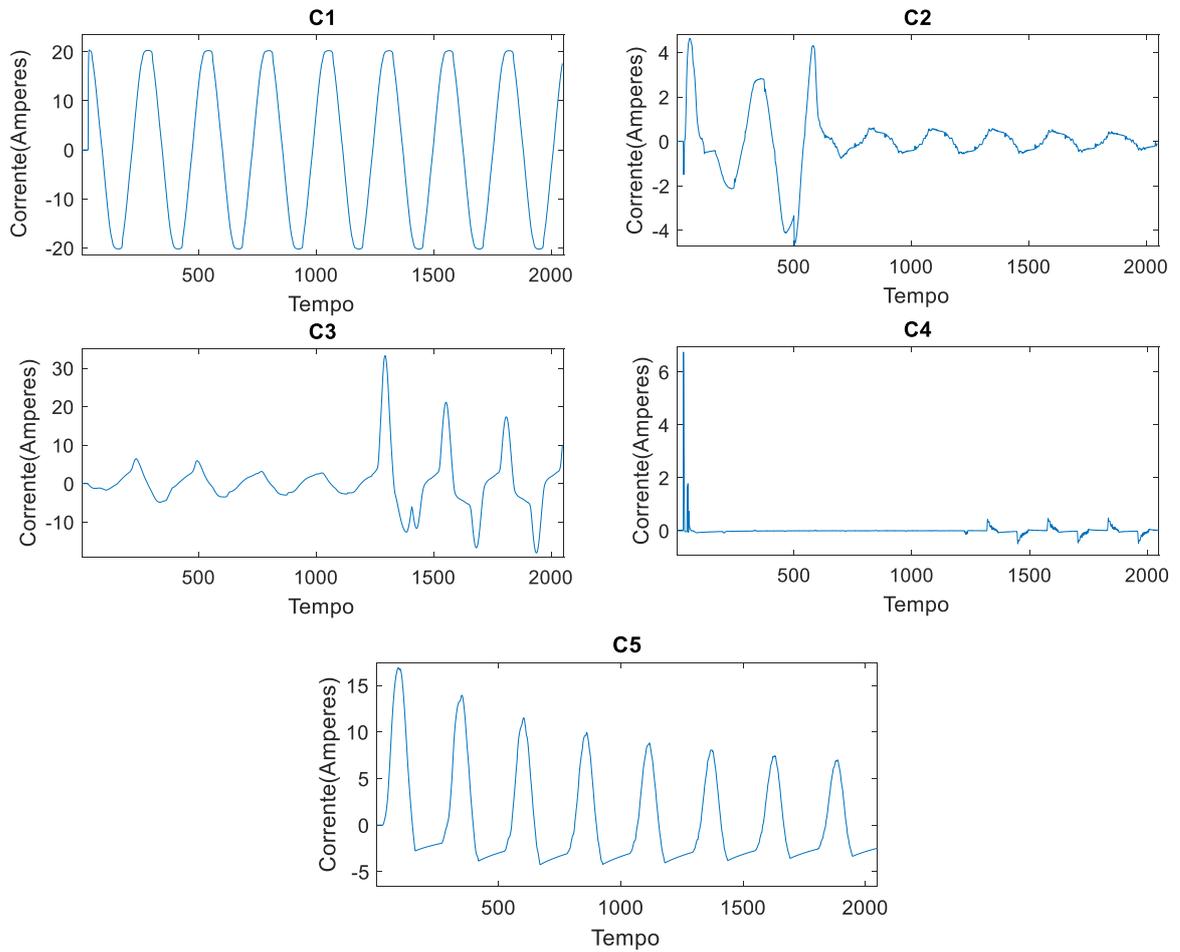
Após realizar a etapa inicial de aquisição dos dados, processamento dos dados, extração de características com Estatística de Ordem Superior e proceder com a seleção de parâmetros, os classificadores foram treinados e os resultados são apresentados neste capítulo.

4.1 Processamento dos dados

Os dados foram filtrados utilizando um filtro do tipo *Notch* e recortados conforme descrito na Seção 3.2. Nas Figuras de 27 e 28 podem ser observados exemplos dos transitórios recortados para cada classe. Analisando as figuras, as diferenças entre os formatos de ondas para quase todas as classes são perceptíveis, como por exemplo as duas classes que mais se destacam são a classe C2 – Tv de LED e a C4 – Lâmpada de LED, nas quais os formatos de ondas são consideravelmente distorcidos quando comparados a senoide fundamental. Outra característica dessas classes é a amplitude do sinal muito baixa, em ambas as classes a amplitude do sinal fica em torno de 0,5 Amperes, enquanto nas demais classes fica em torno de 15 a 20 Amperes.

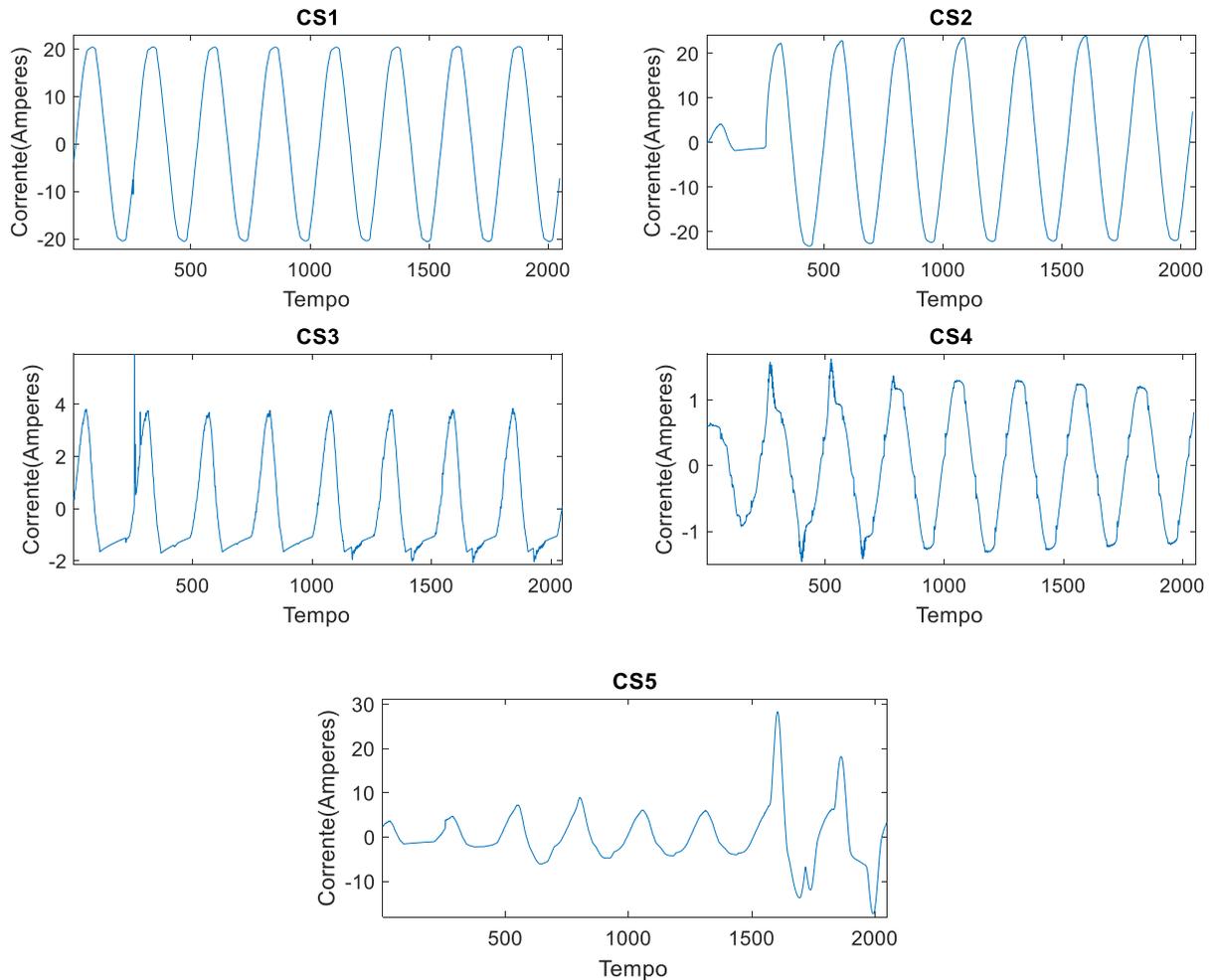
A classe C1-chuveiro, a CS1 chuveiro + Lâmpada de LED e a classe CS2 – chuveiro + multiprocessador possuem as formas de ondas muito parecidas, isso ocorre devido a diferença de amplitude entre as ondas das classes C1 e as das classes C4 e C5 que compõem a classe CS1 e CS2 respectivamente. De uma forma geral, as classes com acionamento simultâneos em que a Lâmpada de LED ou a Televisão de LED foi associada a uma outra carga, verificou-se que tanto a lâmpada de LED quanto a Televisão pouco influenciaram ou não alteraram o formato de onda da carga da outra carga.

Figura 27: Exemplo dos transitórios obtidos para as classes C1 – chuveiro, C2 – Televisão de LED, classe C3 – Micro-ondas, C4 - Lâmpada de LED e C5 – Multiprocessador.



Fonte: Do autor (2018).

Figura 28: Exemplo dos transitórios obtidos para as classes CS1 – Chuveiro + Lâmpada de LED CS2 – Chuveiro + Multiprocessador, CS3 – Lâmpada de LED + Multiprocessador, CS4 – Televisão de LED + Multiprocessador e CS5 - Microondas + Multiprocessador.



Fonte: Do autor (2018).

4.2 Extração e Seleção de parâmetros

A extração de parâmetros teve início ao aplicar Estatística de Ordem Superior para obter o segundo e o quarto cumulantes para cada um dos sinais. Na Figura 29 estão representados os cumulantes de segunda e quarta ordem para cada uma das classes. Observa-se que o cumulante de segunda ordem é simétrico, ou seja, será necessário avaliar apenas a metade dos parâmetros referentes a esse cumulante. Assim, o vetor de parâmetros será formado com os índices de 1 a 1024 com parâmetros do cumulante de segunda ordem e de 1025 a 3072 com os cumulantes de quarta ordem.

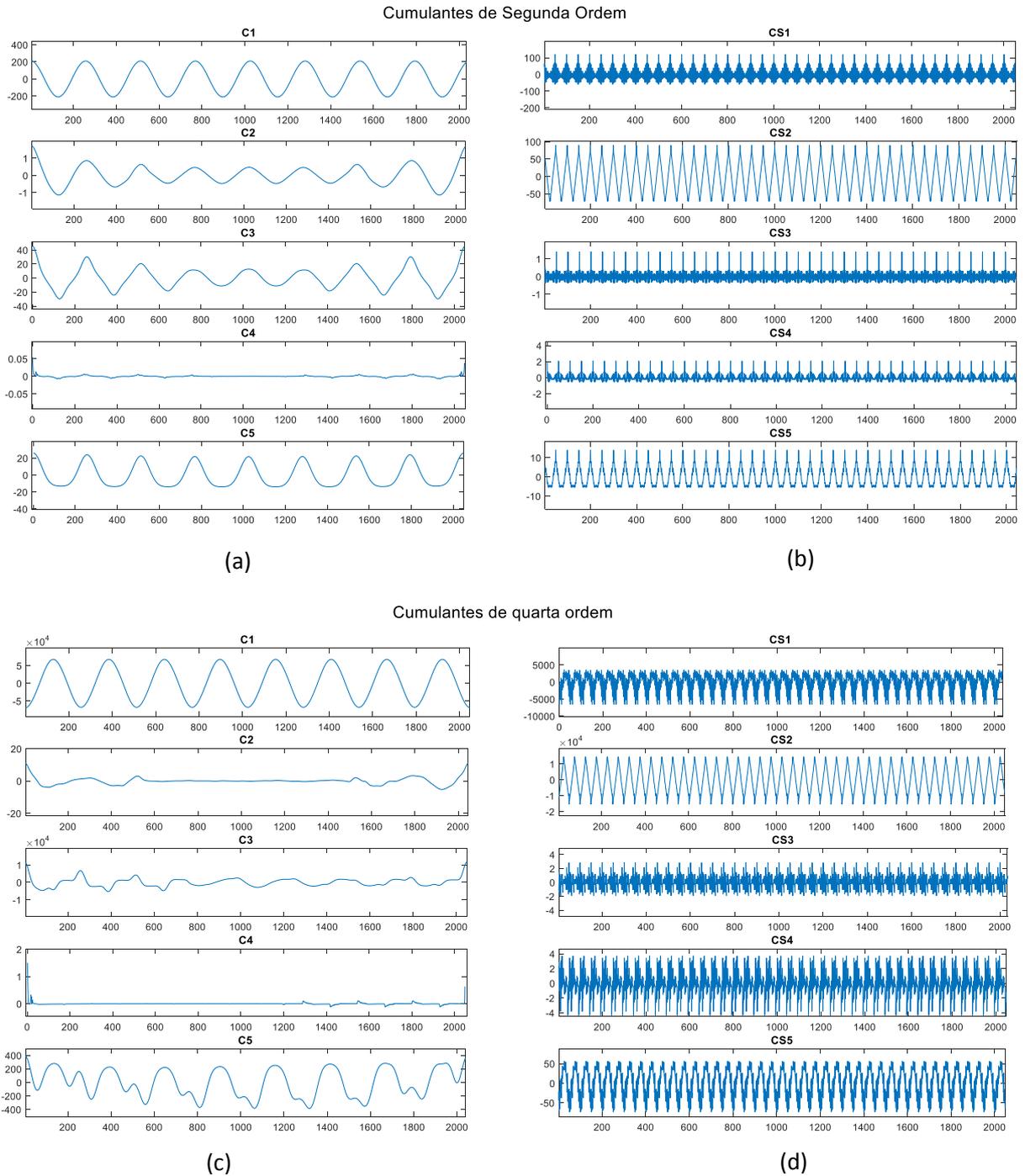
A Figura 30 mostra o Discriminante de Fisher (FDR) para cada uma das classes. É possível verificar que apenas a classe C1 obteve seus parâmetros mais representativos com o

cumulante de quarta ordem, a classe C3 obteve um de seus parâmetros mais representativos um no cumulante de segunda ordem e um no cumulante de quarta ordem. As classes que tiveram ou Lâmpada de LED ou Televisão de LED, mostraram ter os resultados do discriminante de Fisher muito semelhantes, caracterizado pelo cumulante de segunda ordem. A classe CS1, ao contrário da classe C1, foi caracterizada predominantemente pelo cumulante de segunda ordem. Esse resultado demonstra que ao associar a lâmpada de LED ao chuveiro, mesmo que o formato de onda não apresente alterações visuais, estatisticamente o sinal passou a ter outra característica. As classes C5 e CS2 tiveram seus cumulantes mais representativos distribuídos entre os de segunda e quarta ordens.

Na Tabela 5, estão os dez índices dos parâmetros de cada classe que melhor as representam, de acordo com o critério de Fisher. O discriminante de Fisher faz uma avaliação de capacidade de discriminação dos parâmetros a parâmetros, ou seja, o cumulante de índice 100 é comparado com os cumulantes das demais classes cujo índice é 100. Para a classe C2, o índice do parâmetro que a melhor representa é o índice 1, conforme Tabela 5, e este é comparado ao índice 1 das demais classes. As Classes C4 e CS3 também possuem o índice 1 como parâmetro que melhor as discrimina, no entanto, usar o mesmo parâmetro significa usar informação redundante. Portanto, foram selecionados outros índices para as classes C4 (o índice 501) e CS3 (o índice 251), no lugar do índice 1. Os valores representados nessa Tabela são referentes ao índice da posição no vetor onde encontra-se o valor do discriminante de maior representatividade para uma determinada classe, este valor do discriminante é referente aos atrasos aplicados na equação dos cumulantes conforme item 2.1 deste trabalho.

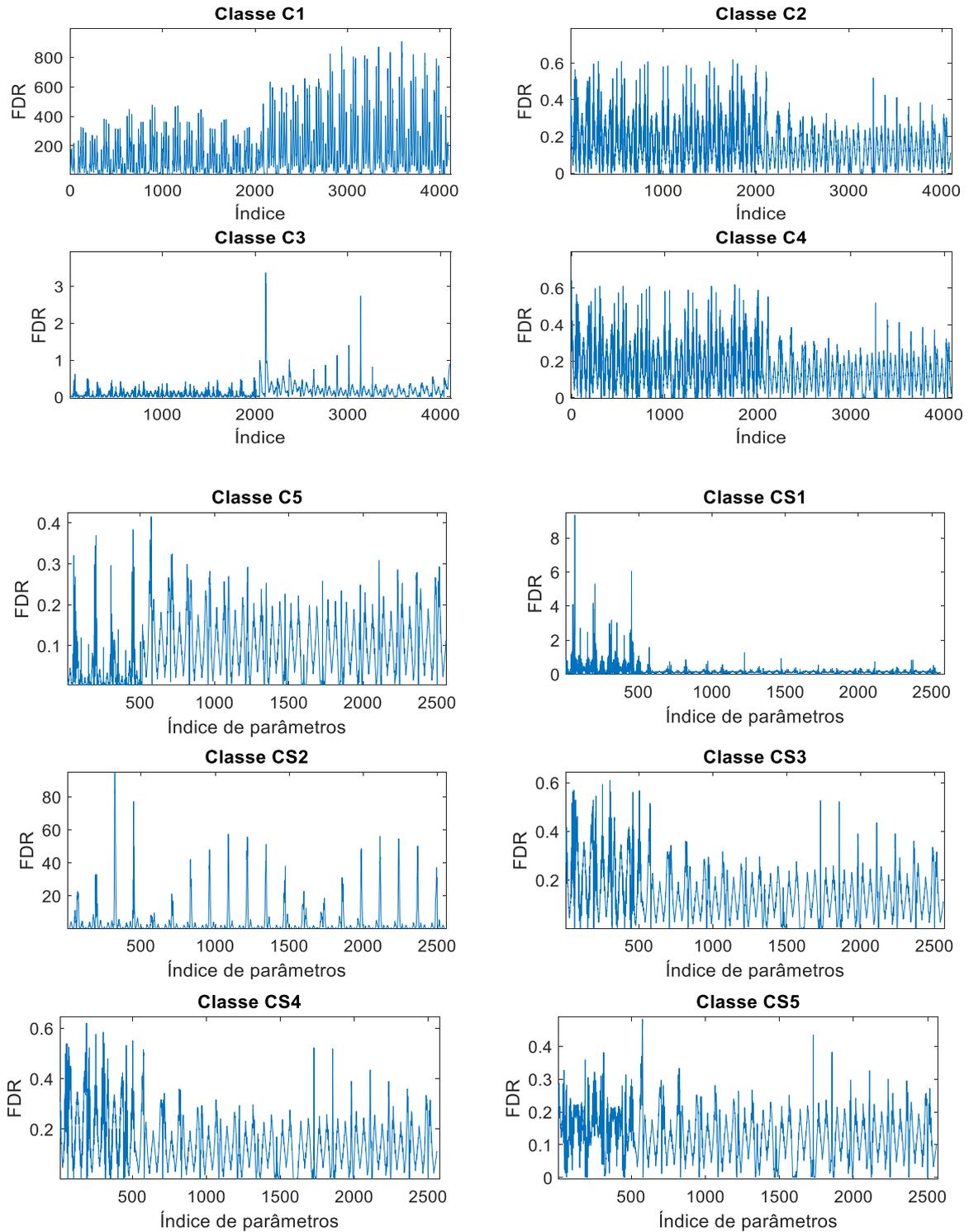
Mesmo com os valores dos cumulantes diferentes conforme mostrado na Figura 29, quando é realizado o Discriminante de Fisher, conforme demonstrado na Figura 30 e na Tabela 5, é possível observar que existe uma semelhança muito grande entre as cargas da classe C2 e C4. Isso pode ocorrer, devido às cargas utilizarem fonte chaveada e serem formadas por LED's. Assim, quando realizado FDR, os índices dos cumulantes que as melhor caracterizam, quando comparadas com as demais cargas, são basicamente os mesmos.

Figura 29: Gráfico apresentando: (a) cumulantes de segunda ordem para as cargas acionadas individualmente; (b) cumulantes de segunda ordem para as classes com acionamentos simultâneos; (c) cumulantes de quarta ordem para as cargas acionadas individualmente; (d) cumulantes de quarta ordem para as classes com acionamentos simultâneos; O eixo das abcissas são os índices dos cumulantes e o eixo das ordenadas são os cumulantes.



Fonte: Do autor (2018).

Figura 30: Discriminante de Fisher para todas as classes, do índice 1 ao índice 1024 são os cumulantes de segunda ordem e o restante são os cumulantes de quarta ordem.



Fonte: Arquivo do autor (2018).

Tabela 5: Índices dos dez parâmetros mais significativos de cada classe ordenado do mais significativo para o menos significativo.

	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°
C1	2052	1802	1402	1927	2054	1404	2177	1804	2302	1677
C2	1	301	251	501	51	573	68	307	184	57
C3	578	580	1602	577	579	581	1601	576	582	1603
C4	1	301	251	501	51	573	68	307	184	207
C5	573	451	575	195	61	571	201	457	574	455
CS1	64	189	63	451	191	201	62	51	190	61
CS2	322	324	323	321	449	325	320	1088	1089	1218
CS3	1	251	301	573	501	307	68	57	188	207
CS4	188	1	573	251	186	66	301	307	68	501
CS5	573	575	1728	574	571	1	1854	188	45	307

Fonte: Do autor (2018).

Uma das características importantes da etapa de seleção de parâmetros é a redução da dimensão dos dados, pois antes da seleção, a dimensão era 2.048 para cada uma das dez classes, isso implicaria em uma rede com 20.480 entradas. Com a seleção de parâmetros, essa dimensão passou a ser de 2 por classe, implicando em uma rede com 20 entradas e em um menor custo computacional. Concluindo a fase de pré-processamento dos dados, iniciou-se a etapa de classificação.

4.3 Acurácia

Avaliando a acurácia dos classificadores pela Tabela 6, verifica-se que os três modelos tiveram um desempenho global muito bom. No entanto ao fazer uma análise por classes, percebe-se que o modelo RNA foi o que alcançou o menor índice, na classe C2 – TV de Led, com acurácia média de $89 \pm 27,65\%$. Quando comparados os dados apenas entre classes dentro do mesmo modelo, o SVM também obteve o índice mais baixo para a classe C2, alcançando $98,75 \pm 2,22\%$ de acurácia para essa classe. O motivo da classe C2 apresentar resultados inferiores aos das outras classes pode estar ligado aos parâmetros utilizados para classificá-la ou devido à semelhança com outras classes, conforme visto na Figura 30, embora o classificador baseado em Florestas Aleatórias tenha obtido 100% das classificações em C2.

Tanto para a classe C1 -Chuveiro quanto para a CS1- Chuveiro junto à Lâmpada de Led, o classificador baseado em Florestas Aleatórias obteve índices inferiores aos obtidos para outras classes. Nesse caso, aparentemente, o classificador teve dificuldade em classificar a carga resistiva (Chuveiro). Quando comparados os três classificadores, o SVM apresentou maior acurácia para a classificação das classes utilizadas neste trabalho.

4.4 Precisão

Na Tabela 7, apresenta-se a taxa de acertos na classificação de dados como sendo de uma determinada classe em relação ao total de dados que foram classificados como sendo daquela classe (positivos + falso positivo), ou seja, a precisão dos classificadores. O classificador RNA obteve o índice mais baixo entre todos os classificadores, para a classe C4- Lâmpada de Led, com índice de $94 \pm 15,17\%$ e para a classe C2 $95 \pm 22,36\%$ de precisão. Quando comparadas as classes dentro de um único modelo, o modelo de classificação baseado em Florestas Aleatórias foi o que obteve o maior número de índices abaixo de 100%, o menor deles foi obtido para a classe CS1, na qual alcançou $95 \pm 22,33\%$. Esse modelo de classificador apresentou a menor precisão, quando comparado com os demais classificadores. Já o SVM apresentou maior precisão na classificação, quando comparado com os demais.

4.5 Sensibilidade

Na Tabela 8, estão representados os dados referentes à sensibilidade dos classificadores. Essa métrica nada mais é que a capacidade do sistema classificar corretamente os dados que realmente pertencem a uma determinada classe. A Rede Neural Artificial obteve o índice de $86\% \pm 30,47$ em sensibilidade para a classe C2 e de 100% para as demais classes. O SVM também mostrou ter uma sensibilidade um pouco menor para a classe C2, alcançando $99 \pm 30,47\%$ contra 100% para as demais classes. Ao contrário destes, o classificador baseado em Florestas Aleatórias alcançou 100% de sensibilidade para a classe C2, no entanto para a classe CS1 o classificador alcançou $89\% \pm 30,76$.

Dessa forma, os 3 classificadores apresentaram uma boa sensibilidade para a classificação dessas cargas, uma vez que o SVM foi o que demonstrou maior sensibilidade na classificação.

4.6 F1 score

A última métrica avaliada foi o F1- *score*, apresentada na Tabela 9, que pode ser interpretada como uma média ponderada da precisão e sensibilidade, em que uma pontuação F1 atinge seu melhor valor em 1 e pior em 0. Para essa métrica, o modelo RNA obteve o índice mais baixo, alcançando $0,91 \pm 0,25$ para a classe C2 e o modelo SVM obteve o índice $0,94 \pm 0,22$ para essa mesma classe. O classificador baseado em Florestas Aleatórias obteve $0,94 \pm 0,22$ para a classe CS1. Mais uma vez o SVM foi o classificador que obteve os melhores índices.

De forma geral, o classificador SVM demonstrou o melhor resultado em todas as métricas quando comparado com os demais. O classificador baseado em florestas aleatórias demonstrou ser menos preciso e também o que possui menor sensibilidade para cargas predominantemente resistivas, neste trabalho, no entanto são necessários testes com outras cargas dessa característica. O SVM junto à Floresta Aleatória foram mais eficientes quando considerado o tempo de treinamento, contudo durante o processo de validação o tempo de execução para os 3 modelos foi praticamente o mesmo.

O método de extração de dados utilizando a Estatística de Ordem Superior obteve bom desempenho para este trabalho. Para verificação, foi utilizada a Transformada Rápida de Fourier como parâmetro. Na Tabela 10, estão os resultados obtidos utilizando Estatística de Ordem Superior junto do classificador RNA, que alcançou a média geral na métrica acurácia de 98,85% e a Transformada Rápida de Fourier (FFT) com a mesma RNA foi de 86,26%, o melhor resultado alcançado por uma RNA com FFT foi de 94,6% (treinamento de número 6).

Tabela 6: Acurácia dos classificadores baseado em Máquina de Vetor Suporte, Redes Neurais Artificiais e Florestas aleatórias.

	Treino	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Média	Devio \pm		
SVM	C1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	C2	100	100	100	100	100	100	95	95	100	100	100	95	100	100	95	100	95	100	100	100	100	99	2,22	
	C3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	C4	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	C5	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	95	100	100	100	100	100	1,12
	CS3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS4	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS5	100	100	100	90	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	2,24
RNA	C1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	C2	100	100	95	100	100	100	100	100	0	100	100	100	30	100	100	100	100	55	100	100	100	89	27,65	
	C3	100	95	100	100	100	100	100	100	100	100	100	100	100	100	100	95	100	100	100	100	100	100	1,54	
	C4	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	C5	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS4	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS5	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
Florestas Aleatórias	C1	95	100	95	95	60	100	100	100	95	100	100	100	100	100	100	100	90	100	100	100	97	9,05		
	C2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	C3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	95	100	100	100	100	1,12	
	C4	100	100	100	100	100	95	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	1,12	
	C5	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	CS1	100	100	100	100	100	100	100	100	100	100	100	100	0	100	80	100	100	100	100	100	100	94	22,57	
	CS2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	CS3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS4	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS5	100	100	100	100	90	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	2,24

Fonte: Do autor (2018).

Tabela 7: Precisão dos classificadores baseado em Máquina de Vetor Suporte, Redes Neurais Artificiais e Florestas aleatórias.

	Treino	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Média	Desvio \pm		
SVM	C1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	C2	100	100	100	100	100	100	100	100	100	0	100	100	100	100	100	100	100	100	100	100	100	100	95	22,36
	C3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	C4	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	C5	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS4	100	100	100	91	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	2,03
	CS5	100	100	100	100	100	100	100	95	95	100	100	100	95	100	100	95	100	91	100	100	100	99	2,65	
RNA	C1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	C2	100	100	100	100	100	100	100	100	100	0	100	100	100	100	100	100	100	100	100	100	100	100	95	22,36
	C3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	C4	100	100	95	100	100	100	100	100	100	50	100	100	100	59	100	100	100	100	69	100	100	94	15,17	
	C5	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS3	100	95	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	1,06
	CS4	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS5	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	95	100	100	100	100	100	100	1,06
Florestas Aleatórias	C1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	C2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	C3	100	100	100	100	67	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	98	7,45	
	C4	100	100	100	100	100	100	100	100	100	100	100	100	50	100	83	100	100	100	100	100	100	97	11,60	
	C5	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	95	100	100	100	100	1,06	
	CS1	100	100	100	100	100	95	100	100	100	100	100	100	100	0	100	100	100	100	100	100	100	95	22,33	
	CS2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS4	100	100	95	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	1,06
	CS5	95	100	100	95	100	100	100	100	100	95	100	100	100	100	100	100	100	91	100	100	100	99	2,55	

Fonte: Do autor (2018).

Tabela 8: Sensibilidade dos classificadores baseado em Máquina de Vetor Suporte, Redes Neurais Artificiais e Florestas aleatórias.

	Treino	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Média	Devio \pm		
SVM	C1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	C2	100	100	100	100	100	100	95	95	100	100	100	95	100	100	95	100	95	100	100	100	100	99	2,22	
	C3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	C4	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	C5	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	95	100	100	100	100	100	1,12
	CS3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS4	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS5	100	100	100	90	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	2,24
RNA	C1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	C2	100	100	95	100	100	100	100	100	0	100	100	100	30	30	100	100	100	55	100	100	100	86	30,47	
	C3	100	95	100	100	100	100	100	100	100	100	100	100	100	100	100	95	100	100	100	100	100	100	1,54	
	C4	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	C5	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	CS1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS4	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS5	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
Florestas Aleatórias	C1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	C2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00	
	C3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	95	100	100	100	100	1,12	
	C4	100	100	100	100	100	95	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	1,12
	C5	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS1	100	100	100	100	100	100	100	100	100	100	100	100	0	0	80	100	100	100	100	100	100	89	30,76	
	CS2	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS3	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS4	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS5	100	100	100	100	90	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	2,24

Fonte: Do autor (2018).

Tabela 9: F1 *score* dos classificadores baseado em Máquina de Vetor Suporte, Redes Neurais Artificiais e Florestas aleatórias.

		Treino	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Média	Devio ±			
SVM	C1	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00		
	C2	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,97	0,97	0,00	1,00	1,00	0,97	1,00	1,00	0,97	1,00	0,97	1,00	1,00	1,00	1,00	0,94	0,22		
	C3	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	
	C4	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	
	C5	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	
	CS1	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00
	CS2	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,97	1,00	1,00	1,00	1,00	1,00	0,01	
	CS3	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	
	CS4	1,00	1,00	1,00	0,95	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,01	
	CS5	1,00	1,00	1,00	0,95	1,00	1,00	0,98	0,98	1,00	1,00	1,00	0,98	1,00	1,00	0,98	1,00	0,95	1,00	1,00	1,00	1,00	1,00	0,99	0,02		
RNA	C1	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00		
	C2	1,00	1,00	0,97	1,00	1,00	1,00	1,00	1,00	0,00	1,00	1,00	1,00	0,46	1,00	1,00	1,00	1,00	0,71	1,00	1,00	1,00	0,91	0,25			
	C3	1,00	0,97	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,97	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,01		
	C4	1,00	1,00	0,98	1,00	1,00	1,00	1,00	1,00	0,67	1,00	1,00	1,00	0,74	1,00	1,00	1,00	1,00	0,82	1,00	1,00	1,00	0,96	0,10			
	C5	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00		
	CS1	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	
	CS2	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	
	CS3	1,00	0,98	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,01	
	CS4	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	
	CS5	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,98	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,01	
Florestas Aleatórias	C1	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00		
	C2	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00		
	C3	1,00	1,00	1,00	1,00	0,80	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,97	1,00	1,00	1,00	0,99	0,04			
	C4	1,00	1,00	1,00	1,00	1,00	0,97	1,00	1,00	1,00	1,00	1,00	1,00	0,67	1,00	0,91	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,98	0,08		
	C5	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,98	1,00	1,00	1,00	1,00	1,00	1,00	0,01		
	CS1	1,00	1,00	1,00	1,00	1,00	0,98	1,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00	0,89	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,94	0,22		
	CS2	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00		
	CS3	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	
	CS4	1,00	1,00	0,98	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,01	
	CS5	0,98	1,00	1,00	0,98	0,95	1,00	1,00	1,00	0,98	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,95	1,00	1,00	1,00	1,00	0,99	0,02			

Fonte: Do autor (2018).

Tabela 10 Valores obtidos com o treinamento de uma RNA utilizando a extração de parâmetro por FFT e por EOS.

		1	2	3	4	5	6	7	8	9	10	Média	Devio ±
FFT	C1	100	100	100	100	100	100	100	100	100	100	100	51,30
	C2	100	87	87	93	100	93	93	100	87	73	91	47,20
	C3	93	100	93	93	93	100	100	100	87	93	95	49,00
	C4	87	73	93	93	73	100	67	87	67	80	82	42,84
	C5	93	100	93	100	100	100	100	80	100	100	97	49,79
	CS1	93	80	100	100	93	80	93	93	87	93	91	47,10
	CS2	93	60	73	80	93	100	67	87	53	87	79	42,08
	CS3	60	67	47	53	73	87	73	67	60	80	67	35,21
	CS4	80	73	87	93	87	87	67	87	60	87	81	42,02
	CS5	100	73	100	87	87	100	100	87	93	93	92	47,58
EOS	C1	100	100	100	100	100	100	100	100	100	100	100	0,00
	C2	100	100	95	100	100	100	100	100	0	100	89	27,65
	C3	100	95	100	100	100	100	100	100	100	100	100	1,54
	C4	100	100	100	100	100	100	100	100	100	100	100	0,00
	C5	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS1	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS2	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS3	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS4	100	100	100	100	100	100	100	100	100	100	100	0,00
	CS5	100	100	100	100	100	100	100	100	100	100	100	0,00

Fonte: Do autor (2018).

5 CONCLUSÃO

Este trabalho apresentou o desenvolvimento de um classificador para cargas elétricas residenciais, acionadas isoladamente e para cargas acionadas enquanto outra estava em funcionamento (acionamentos simultâneos). Para o desenvolvimento do método, os dados amostrados foram pré-processados utilizando Estatística de Ordem Superior para realizar a extração das características de cada sinal. A extração se deu utilizando os cumulantes de segunda e quarta ordens. Utilizando o Discriminante de Fisher, foram selecionados os parâmetros mais significativos para cada uma das classes, então foram avaliados utilizando uma tabela. À vista disto, os dois parâmetros mais significativos de cada classe, foram selecionados e o vetor de características foi montado com dimensão 20, ou seja, uma redução de 99% da dimensão dos dados, que antes era de 2048, impactando na redução do esforço computacional.

Após a etapa de pré-processamento foram implementados os classificadores utilizando as técnicas de aprendizagem de máquina Redes Neurais Artificiais, Máquina de Vetor Suporte e Florestas Aleatórias. Para todos os classificadores, foram realizadas 20 execuções, nas quais o conjunto de dados eram misturados e redivididos. Os resultados foram expressos em forma de matriz de confusão e foram utilizadas as métricas acurácia, sensibilidade, precisão e *F1 score* para comparar os classificadores.

Para todos os classificadores, os resultados alcançados foram bons, atingindo 100% de acerto para a maioria dos dados. No entanto, para as classes C2 – TV de Led, C4-Lâmpada de Led e a Casse CS1- Chuveiro + Lâmpada de Led, os classificados alcançaram índices inferiores, chegando a 86% para a classe C2 quando avaliada a sensibilidade via RNA.

Como considerações finais, pode-se dizer que o classificador baseado em Máquina de Vetor de Suporte foi o que obteve melhor resultado em todas as métricas utilizadas, embora a variação entre os resultados obtidos pelos três classificadores serem pequenas. A acurácia média geral para o SVM foi de 99,8%; para a Rede Neural Artificial foi de 98,85% e para o classificador baseado em Florestas Aleatórias foi de 98,95%. O *F1 score* médio geral foi de 0,99 para todos os classificadores. No fator tempo gasto para treinamento, a Rede Neural Artificial gastou em média 3 minutos para cada treinamento, já os classificadores SVM e Florestas Aleatórias gastaram 0,09 segundos em média (medidos usando a função tic toc do Matlab). Essa diferença de tempo mostra que o esforço computacional para a RNA é muito superior ao esforço computacional dos outros dois classificadores. No entanto, na fase de

validação, o tempo gasto foi praticamente o mesmo para todos os classificadores, em torno de 1 segundo.

5.1 Trabalhos futuros

A proposta final é ter um equipamento que seja capaz de discriminar todo o consumo de uma residência, permitindo que uma conta de luz possa vir detalhada conforme é realizado nas contas do serviço de telefonia atualmente.

Mas para isso é necessário que o software possa identificar as cargas, calcular o consumo, e então, embarcar o sistema e realizar testes em campo. Deverá ser realizado um levantamento do custo de implementar o projeto. Um desafio é conseguir montar um banco de dados que contemple todas as cargas existentes em uma residência, ou ao menos a maior parte e treinar o classificador para um número maior de cargas.

REFERÊNCIAS

- AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA - ANEEL. **ANEEL aprova Tarifa Branca, nova opção para os consumidores a partir de 2018**. Brasília: ANEEL 2016. Disponível em: <http://www.aneel.gov.br/sala-de-imprensa-exibicao/-/asset_publisher/XGPXSqdMFHrE/content/aneel-aprova-tarifa-branca-nova-opcao-para-os-consumidores-a-partir-de-2018/656877?inheritRedirect=false>. Acesso em: 26 set. 2017.
- AGYEMAN, K.; HAN, S.; HAN, S. Real-time recognition non-intrusive electrical appliance monitoring algorithm for a residential building energy management system. **Energies**, Basel, v. 8, n. 9, p. 9029-9048, Aug. 2015.
- ALAMEDA-HERNÁNDEZ, E. et al. Higher-order statistics for power systems: effects of the sampling frequency on ergodicity. **Applied Mathematical Modelling**, Guildford, v. 40, n. 15/16, p. 6924-6933, Aug. 2016.
- BONFIGLI, R. et al. Denoising autoencoders for non-intrusive load monitoring: improvements and comparative evaluation. **Energy and Buildings**, Lausanne, v. 158, p. 1461-1474, Jan. 2018.
- BREIMAN, L. Random forests. **Machine Learning**, Boston, v. 45, n. 1, p. 5-32, Apr. 2001.
- COMINOLA, A. et al. A hybrid signature-based iterative disaggregation algorithm for non-intrusive load monitoring. **Applied Energy**, London, v. 185, p. 331-344, Jan. 2017.
- CORTES, C.; VAPNIK, V. Support-vector networks. **Machine Learning**, Boston, v. 20, n. 3, p. 273-297, Sept. 1995.
- DINESH, C. et al. Non-intrusive load monitoring under residential solar power influx. **Applied Energy**, London, v. 205, p. 1068-1080, Nov. 2017.
- DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern classification**. 2. ed. New York: Wiley-Interscience, 2000. 654 p.
- EMPRESA DE PESQUISA ENERGÉTICA - EPE. **Consumo mensal de energia elétrica por classe (regiões e subsistemas) – 2004-2017**. Brasília: EPE, 2012. Disponível em: <[http://www.epe.gov.br/mercado/Paginas/Consumomensaldeenergiael%C3%A9tricaporclasse\(regi%C3%B5esesubistemas\)%E2%80%932011-2012.aspx](http://www.epe.gov.br/mercado/Paginas/Consumomensaldeenergiael%C3%A9tricaporclasse(regi%C3%B5esesubistemas)%E2%80%932011-2012.aspx)>. Acesso em: 16 maio 2017.
- ESEYE, A. T.; ZHANG, J.; ZHENG, D. Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and Meteorological information. **Renewable Energy**, Oxford, v. 118, p. 357-367, Apr. 2018.
- FERREIRA, D. D. et al. HOS-based method for classification of power quality disturbances. **Electronics Letters**, London, v. 45, n. 3, p. 183-185, Jan. 2009.

GILLIS, J. M.; MORSI, W. G. Non-intrusive load monitoring using semi-supervised machine learning and wavelet design. **IEEE Transactions on Smart Grid**, Piscataway, v. 8, n. 6, p. 2648-2655, 2017.

GUEDES, J. D. S.; FERREIRA, D. D.; BARBOSA, B. H. G. A non-intrusive approach to classify electrical appliances based on higher-order statistics and genetic algorithm: a smart grid perspective. **Electric Power Systems Research**, Lausanne, v. 140, p. 65-69, July 2016.

HASSAN, T.; JAVED, F.; ARSHAD, N. An empirical investigation of VI trajectory based load signatures for non-intrusive load monitoring. **IEEE Transactions on Smart Grid**, Piscataway, v. 5, n. 2, p. 870-878, 2014.

HAYKIN, S. **Neural networks and learning machines**. 3. ed. New Jersey: Pearson, 2008. 906 p.

HEASRT, M. et al. Trends and controversies-support vector machines. **IEEE Intelligent Systems**, Los Alamitos, v. 13, n. 4, p. 18-28, 1998.

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

HSU, C.-W.; CHANG, C.-C.; LIN, C.-J. **A practical guide to support vector classification**. Taiwan: National Taiwan University, 2016. 16 p.

KIM, J.; LE, T.-T.-H.; KIM, H. Nonintrusive load monitoring based on advanced deep learning and novel signature. **Computational Intelligence and Neuroscience**, Essex, v. 2017, p. 1-22, 2017.

KULKARNI, A. S.; HARNETT, C. K.; WELCH, K. C. EMF signature for appliance classification. **IEEE Sensors Journal**, Manchester, v. 15, n. 6, p. 3573-3581, June 2015.

LAI, Y.-X. et al. Multi-appliance recognition system with hybrid SVM/GMM classifier in ubiquitous smart home. **Information Sciences**, New York, v. 230, p. 39-55, May 2012.

LAPUT, G.; ZHANG, Y.; HARRISON, C. Synthetic Sensors: towards general-purpose sensing. CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2017, Denver. **Proceedings...** New York: ACM, 2017. p. 3986-3999.

LIU, C. et al. Multivariate exploration of non-intrusive load monitoring via spatiotemporal pattern network. **Applied Energy**, London, v. 211, p. 1106-1122, Feb. 2018.

MATRIZ de classificação (Analysis Services - Mineração de dados). **Microsoft Corporation**, [S.l.], jun. 2017. Disponível em: <<https://docs.microsoft.com/pt-br/sql/analysis-services/data-mining/classification-matrix-analysis-services-data-mining?view=sql-server-2014>>. Acesso em: 30 ago. 2018.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The Bulletin of Mathematical Biophysics**, Chicago, v. 5, n. 4, p. 115-133, Dec. 1943.

MENDEL, J. M. Tutorial on higher-order statistics (Spectra) in signal processing and system theory: theoretical results and some applications. **Proceedings of the IEEE**, New York, v. 79, n. 3, p. 278-305, Mar. 1991.

MITRA, S. K. **Digital signal processing: a computer-based approach**. 3. ed. Nova York: McGraw-Hill, 2005. 972 p.

NALMPANTIS, C.; VRAKAS, D. Machine learning approaches for non-intrusive load monitoring: from qualitative to quantitative comparison. **Artificial Intelligence Review**, Dordrecht, p. 1-27, Jan. 2018.

NASCIMENTO, D. M. D. **Identificação não invasiva de cargas elétricas residenciais**. 2012. 94 p. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012.

NATIONAL INSTRUMENTS - NI. **Data sheet**. Austin: NI, 2014. Disponível em: <<http://www.ni.com/datasheet/pdf/en/ds-21>>. Acesso em: 26 set. 2017.

QUIRÓS-OLOZÁBAL, Á. et al. A novel FPGA-based system for real-time calculation of the Spectral Kurtosis: a prospective application to harmonic detection. **Measurement**, Grattan, v. 86, p. 101-113, Feb. 2016.

RIBEIRO, M. V. et al. Detection of disturbances in voltage signals for power quality analysis using HOS. **EURASIP Journal on Advances in Signal Processing**, New York, v. 2007, p. 13, Dec. 2007.

ROSENBLATT, F. The perceptron, a perceiving and recognizing automaton. Report: cornell aeronautical laboratory, v. 85, n. 460/461, New York, p. 85-460, 1957.

_____. **Principles of neurodynamics: perceptrons and the theory of brain mechanisms**. Virginia: Spartan Books, 1962. 621 p.

THEODORIDIS, S.; KOUTROUMBAS, K. **Pattern Recognitions**. 4. ed. Boston: Academic Press, 2009. 984 p.

VALENÇA, M. J. **Fundamentos das redes neurais: exemplos em Java**. Olinda: Livro Rápido, 2007. 382 p.

VAPNIK, V. **The nature of statistical learning theory**. Berkeley: Springer Science & Business Media, 2013. 314 p.

WANG, Z.; SRINIVASAN, R. Classification of household appliance operation cycles: a case-study approach. **Energies**, Basel, v. 8, n. 9, p. 10522-10536, Sept. 2015.

WIDROW, B.; HOFF, M. E. Adaptive switching circuits. In: IRE WESCON CONVENTION RECORD, 1960, New York. **Proceedings...** New York: Institute of Radio Engineers, 1960. p. 96-104.

XU, Y.; MILANOVIC, J. V. Artificial-intelligence-based methodology for load disaggregation at bulk supply point. **IEEE Transactions on Power Systems**, New York, v. 30, n. 2, p. 795-803, Mar. 2015.

YANG, C.; SOH, C.; YAP, V. A systematic approach to ON-OFF event detection and clustering analysis of non-intrusive appliance load monitoring. **Frontiers in Energy**, Beijing, v. 9, n. 2, p. 231-237, June 2015.

ZHANG, L. L. et al. Skewness-based differential protection scheme for EHV/UHV transmission lines. **IEEE Transactions on Power Delivery**, New York, v. 29, n. 3, p. 1518-1520, June 2014.

ANEXO A - ESPECIFICAÇÃO DA PLACA NI USB-6251

Especificação	
ENTRADA	
Número de canais	8 diferencial ou 16 de extremidade única
Resolução ADC	16 bits
DNL	Não há códigos faltantes garantidos
Taxa de amostragem	1,25M amostras/s canal simples 1M amostras/s Multi-canal.
Precisão de Tempo	50 ppm da taxa de amostragem
Resolução de Tempo	50ns
Acoplamento de entrada	DC
Faixa de entrada	$\pm 10\text{ V}$, $\pm 5\text{ V}$, $\pm 2\text{ V}$, $\pm 1\text{ V}$, $\pm 0.5\text{ V}$, $\pm 0.2\text{ V}$, $\pm 0.1\text{ V}$
Tensão Máxima	$\pm 11\text{ V}$ de AI GND
Impedância de entrada equipamento ligado	
AI+ a AI GND	$>10\text{ G}\Omega$ em paralelo com 100 pF
AI- a AI GND	$>10\text{ G}\Omega$ em paralelo com 100 pF
Impedância de entrada equipamento desligado	
AI+ a AI GND	820 Ω
AI- a AI GND	820 Ω
Saída	
Número de canais	2
Resolução do DAC	16bits
DNL	$\pm 1\text{ LSB}$
Monotonicidade	16 bit
Precisão	2,080 μV na escala de 10V 1,045 μV na escala de 5V
Taxa máxima de aquisição	
1 canal	2,86M amostras/s
2 canais	2,00M amostras/s
3 canais	1,54M amostras/s
4 canais	1,25M amostras/s
Impedância de saída	0.2 Ω
Corrente de saída	$\pm 5\text{ mA}$
Faixa de saída	$\pm 10\text{ V}$, $\pm 5\text{ V}$

Fonte: National Instruments (2017)

**ANEXO B – ESPECIFICAÇÃO PONTEIRA PARA MEDIÇÃO DE CORRENTE
FLUKE I200S**

Especificação	
Faixa nominal de corrente	20 A 200 A
Faixa de corrente contínua	0.1 A - 24 A 0.5 A - 200 A
Corrente máxima não destrutiva	240 A
Corrente mensurável mais baixa	0.1 A
Precisão básica	1.5% + 0.5 A (48-65 Hz) (% reading + floorspec)
Frequência de utilização	40 Hz - 10 kHz
Níveis de saída	100 mV/A 10mV/A
Segurança	CAT III, 600 V
Tensão máxima	600 V AC
Garantia	1 year
Diâmetro máximo do condutor	20 mm
Tamanho máximo do condutor (barramento)	20 x 5 mm
Comprimento do cabo de saída	2 m
Adaptador BNC	Yes
Adaptador BNC para banana	Yes

Fonte: Manual do equipamento Fluke i200s (2017)