



UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Projeto de Classificadores para Sistema de Reconhecimento Automático de Modulações

Myke Douglas de Medeiros Valadão

Manaus – Amazonas

Setembro de 2019

Myke Douglas de Medeiros Valadão

Projeto de Classificadores para Sistema de Reconhecimento Automático de Modulações

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração de Processamento Digital de Imagens.

Orientador: Prof. D.Sc. Waldir Sabino da Silva Júnior

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

V136p Valadão, Myke Douglas de Medeiros
Projeto de Classificadores para Sistema de Reconhecimento Automático de Modulações / Myke Douglas de Medeiros Valadão. 2019
89 f.: il. color; 31 cm.

Orientador: Waldir Sabino da Silva Júnior
Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Amazonas.

1. Rádio Cognitivo . 2. Classificação Automático de Modulações.
3. Redes Neurais Artificiais . 4. Random Forest. I. Silva Júnior, Waldir Sabino da II. Universidade Federal do Amazonas III. Título

MYKE DOUGLAS DE MEDEIROS VALADÃO

PROJETO DE CLASSIFICADORES PARA SISTEMA DE RECONHECIMENTO
AUTOMÁTICO DE MODULAÇÕES

Dissertação apresentada ao Programa de Pós-
Graduação em Engenharia Elétrica da
Universidade Federal do Amazonas, como
requisito parcial para obtenção do título de
Mestre em Engenharia Elétrica na área de
concentração Controle e Automação de Sistemas.

Aprovado em 11 de setembro de 2019.

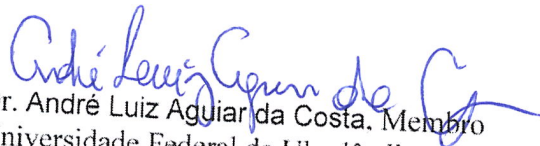
BANCA EXAMINADORA



Prof. Dr. Waldir Sabino da Silva Júnior, Presidente
Universidade Federal do Amazonas



Prof. Dr. Celso Barbosa Carvalho, Membro
Universidade Federal do Amazonas



Prof. Dr. André Luiz Aguiar da Costa, Membro
Universidade Federal de Uberlândia

Agradecimentos

Meu muitíssimo obrigado a Rose (mãe), Thiago e Matheus (irmãos) pelo inquestionável suporte.

Agradeço também ao meu orientador Professor Dr. Waldir Sabino pela paciência e preocupação com minha formação, muito obrigado.

Agradeço imensamente os amigos que estavam comigo durante essa jornada, Leonardo Neto (Leona), Filipe Mory (Lilica), Eduardo, Gabi, Yamille, Luiz, Adriana, Osvaldo, Ronaldo (Bate-Bate), Davi (Magra), Tayná, Lennon, Alay, Breno, Cesário, Ingrid, Bruna, Lucas, Caio, Arlen, Ruben, Maria (Peão), Áurea, Thiago (Assassina), Anselmo, Andreza, Elton (Sykora), é, acho que não vai dar pra colocar todos, mas agradeço a todos que passaram e ainda estão em contato, obrigado gente.

Agradeço a UFAM, onde se tornou minha segunda casa, obrigado.

Agradeço a CAPES pelo apoio financeiro nesses pouco mais de dois anos de mestrado, obrigado.

E por último, agradeço Miquéias Ferrão, obrigado por pelo apoio e incentivo.

Resumo da Dissertação apresentada à UFAM como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia Elétrica

PROJETO DE CLASSIFICADORES PARA SISTEMA DE
RECONHECIMENTO AUTOMÁTICO DE MODULAÇÕES

Myke Douglas de Medeiros Valadão

Orientador: Waldir Sabino da Silva Júnior

Programa: Pós-Graduação em Engenharia Elétrica

A subutilização do espectro de frequência é uma problema recorrente atualmente e, com o aumento da demanda de usuários que utilizam sistemas de comunicação remota, foi necessário buscar uma maneira mais eficiente de alocar usuários no espectro, surgindo assim, as técnicas que aplicam o rádio cognitivo. O rádio cognitivo, detecta *buracos espectrais* e aloca de maneira dinâmica usuários nestes espaços inutilizados. Com isso em vista, as técnicas de classificação automática de modulação vieram para fornecer informações a priori que auxiliam no sensoriamento do espectro. Nesta dissertação propõe-se, para classificar sinais modulados, utilizar uma gama de classificadores multiclasse supervisionados baseados em *aprendizado de máquina* e *aprendizado profundo*, com seus parâmetros pré-estabelecidos. Dentre os classificadores englobados em *aprendizado de máquina*, abordamos algoritmos baseados em árvore de decisão e algoritmo de classificação probabilística, *Naive Bayes*. Dentro do *aprendizado profundo*, aplicou-se redes neurais artificiais através de uma rede perceptron multicamada totalmente conectada com retropropagação utilizando algoritmo de *Levenberg-Marquardt* para atualização dos pesos da rede. Foram obtidos taxas de acurácia de 95,2866% e 93,1253% nos classificadores baseados em árvore de decisão, 87,4% na rede neural e 74,7845% no *Naive Bayes*. Na literatura

foi encontrado um trabalho com base de dados semelhante qualitativamente a utilizada nesta dissertação e sua acurácia foi de 89,72%, enquanto a melhor acurácia apresentada nesta dissertação foi de 95,2866%.

Palavras-chave: Rádio Cognitivo, Classificação Automática de Modulações, Redes Neurais Artificiais, *J48*, *Random Forest*, *Naive Bayes*.

Abstract of Dissertation presented to UFAM as a partial fulfillment of the requirements for the degree of Master in Electrical Engineering

PROJETO DE CLASSIFICADORES PARA SISTEMA DE
RECONHECIMENTO AUTOMÁTICO DE MODULAÇÕES

Myke Douglas de Medeiros Valadão

Advisor: Waldir Sabino da Silva Júnior

Department: Postgraduate in Electrical Engineering

Underutilization of the frequency spectrum is a recurring problem today, and with the increasing demand of users using remote communication systems, it was necessary to search for a more efficient way to allocate users in the spectrum, appearing thus, the techniques that apply cognitive radio. Cognitive radio detects spectral holes and dynamically allocates users in these unused spaces. With this in view, automatic modulation classification techniques have come to provide a priori information that aid in spectrum sensing. In this dissertation it is proposed to classify modulated signals using a range of supervised multiclass classifiers based on machine learning and deep learning, with its pre-established parameters. Among the classifiers encompassed in machine learning, we approach algorithms based on decision tree and probabilistic classification algorithm, Naive Bayes. Within deep learning, artificial neural networks were applied through a multilayer perceptron network fully connected with backpropagation using Levenberg-Marquardt algorithm to update the network weights. Accuracy rates of 95.2866% and 93.1253% were obtained in the decision tree-based classifiers, 87.4% in the neural network, and 74.7845% in the *Naive Bayes*. In the literature, we found a study with a similar database qualitatively to that used in this dissertation and its accuracy was 89.72%, while the best accuracy presented in this dissertation was 95.2866%.

Keywords: Cognitive Radio, Automatic Modulation Classification, Artificial Neural Network, *J48*, *Random Forest*, *Naive Bayes*.

Sumário

1	Introdução	1
1.1	Objetivos da Dissertação	4
1.2	Estrutura do Trabalho	5
2	Fundamentos Teóricos	6
2.1	Conceitos de Telecomunicações	6
2.2	Conceitos de Reconhecimento de Padrões	7
2.3	Classificação Automática de Modulação	8
2.4	Tipos de Modulações	10
2.4.1	Modulações Analógicas	10
2.4.2	Modulações Digitais	11
2.4.3	Técnicas de Múltiplo Acesso	12
2.5	Extração de características	13
2.5.1	Características Espectrais	14
2.5.2	Características por Transformada	16
2.5.3	Características Estatísticas	17
2.5.4	Característica de Teste de Distribuição	17
2.5.5	Características Cicloestacionárias	18
2.6	Classificadores	19
3	Trabalhos Relacionados	22
3.1	Técnicas para Classificação Automática de Modulação	22
3.2	Síntese dos Trabalhos Relacionados	23
3.3	Comparação entre Literatura e Trabalho Proposto	25

4	Metodologia	27
4.1	Geração de Sinais	28
4.2	Transmissão de Sinais	30
4.3	Amostragem e Determinação da Envoltória Complexa	32
4.4	Extração de Características	32
4.5	Projeto dos Classificadores	33
4.5.1	Redes Neurais Artificiais	33
4.5.2	J48	38
4.5.3	Random Forest	40
4.5.4	Naive Bayes	43
5	Experimentos e Resultados	45
5.1	<i>Setup</i> dos Experimentos	45
5.2	Base de Dados	45
5.3	Métricas de Desempenho	46
5.4	Experimentos	48
5.4.1	Rede Neural Artificial	48
5.4.2	J48	53
5.4.3	<i>Random Forest</i>	57
5.4.4	<i>Naive Bayes</i>	61
5.4.5	Análise Comparativa	64
6	Conclusões	67
6.1	Trabalhos Futuros	69
	Referências Bibliográficas	70

Lista de Figuras

2.1	Diagrama em blocos de um sistema de comunicação típico.	7
2.2	Projeto de um sistema de reconhecimento de padrões.	8
2.3	Taxa de ocupação do espectro de frequência na cidade do Rio de Janeiro na faixa de 144 a 2690 MHz em 2013.	9
2.4	Modelo de neurônio idealizado por McCulloch.	20
4.1	Visão geral da metodologia de classificação constituída pelas seguintes etapas: geração de sinais em banda base, acréscimo de ruído, deslocamento na frequência, amostragem, envoltória complexa, extração de características e classificação.	27
4.2	Diagrama da rede <i>perceptron</i> multicamadas totalmente conectada com <i>backpropagation</i>	34
4.3	Estrutura de votação ponderada da <i>Random Forest</i>	41
4.4	Divisão dos dados de treinamento por <i>bootstrap</i>	41
5.1	Gráfico do desempenho MSE do processo de treinamento e teste da RNA dado a época.	51
5.2	Gráfico do comportamento dos parâmetros da RNA no processo de treinamento dada a época.	51
5.3	Gráfico da acurácia em (%) pela variação da SNR (<i>dB</i>) para a RNA.	52
5.4	Gráfico da acurácia em (%) pela variação da SNR (<i>dB</i>) para o classificador <i>J48</i>	56
5.5	Gráfico da acurácia em (%) pela variação da SNR (<i>dB</i>) para o classificador <i>Random Forest</i>	60
5.6	Gráfico da acurácia em (%) pela variação da SNR (<i>dB</i>) para o classificador <i>Naive Bayes</i>	63

5.7	Gráfico da acurácia em (%) geral da rede por cada classificador proposto.	64
5.8	Gráfico da acurácia em (%) pela variação da SNR (dB) para todos os métodos propostos.	65

Lista de Tabelas

3.1	Apresenta a diversidade de técnicas de classificação, tamanhos de base de dados, classificadores, modelagem de canal e grupos de modulação.	24
4.1	Descrição e simbologia matemática das características extraídas para os sistemas de classificação automática de modulação propostos.	33
4.2	Descrição de todos os parâmetros aplicáveis ao algoritmo <i>J48</i> .	40
4.3	Descrição de todos os parâmetros aplicáveis ao algoritmo <i>Random Forest</i> .	43
5.1	Quantidade de vetores de características por classe de modulação e nível de SNR.	47
5.2	Parâmetros da RNA proposta.	49
5.3	Matriz de confusão de treinamento da classificação realizada pela RNA proposta.	49
5.4	Matriz de confusão de teste da classificação realizada pela RNA proposta.	50
5.5	Precisão por SNR das modulações classificadas pela RNA proposta.	50
5.6	Parâmetros experimentais do algoritmo <i>J48</i> .	54
5.7	Matriz de confusão de treinamento da classificação realizada pelo algoritmo <i>J48</i> .	55
5.8	Matriz de confusão de teste da classificação realizada pelo algoritmo <i>J48</i> .	55
5.9	Precisão por SNR das modulações classificadas pelo algoritmo <i>J48</i> .	55
5.10	Parâmetros experimentais do algoritmo <i>Random Forest</i> .	58
5.11	Matriz de confusão de treinamento da classificação realizada pelo algoritmo <i>Random Forest</i> .	58

5.12	Matriz de confusão de teste da classificação realizada pelo algoritmo <i>Random Forest</i>	58
5.13	Precisão por SNR das modulações classificadas pelo algoritmo <i>Random Forest</i>	59
5.14	Matriz de confusão de treinamento da classificação realizada pelo classificador <i>Naive Bayes</i>	61
5.15	Matriz de confusão de teste da classificação realizada pelo classificador <i>Naive Bayes</i>	62
5.16	Precisão por SNR das modulações classificadas pelo classificador <i>Naive Bayes</i>	62
5.17	Tabela comparativa entre métodos propostos e literaturas.	66

Capítulo 1

Introdução

Reconhecimento de padrões é uma área de pesquisa que tem como objetivo discriminar objetos em uma ou mais categorias com base em características únicas do objeto em questão [1]. Dentro da literatura existem algumas definições, como em [2], que define reconhecimento de padrões como parte integral da maioria dos sistemas de máquina inteligente construídos para tomada de decisão. Em [3], o propósito final do reconhecimento de padrões é projetar máquinas para resolver a discrepância entre teoria e aplicação. Em 1996, [4], definiu reconhecimento de padrões como uma disciplina em que se aprende algumas teorias e métodos para projetar máquinas que podem reconhecer padrões em dados ruidosos ou em ambientes complexos. Já em [5], o objetivo do reconhecimento de padrões é compreender complicados mecanismos de tomada de decisão e automatizar funções utilizando o computador.

A classificação automática de modulações (do inglês *automatic modulation classification*, AMC) é uma subárea do reconhecimento de padrões, que utiliza dos seus princípios e métodos para, de maneira automática, indicar que modulação o sinal capturado está empregando. As técnicas de AMC tiveram início nas décadas de 80 e 90 [6], inicialmente voltadas para propósitos militares e, com a popularização do uso dos dispositivos de comunicação remota, houve ampliação em aplicações de uso civil. Para fins militares, as informações de modulação são úteis para formação de estratégias de ataque, suporte e defesa eletrônica [7]. Em aplicações para uso civil, AMC é utilizada para aumentar a eficiência do espectro de frequência. Um dos objetivos da pesquisa e desenvolvimento de técnicas de AMC é utilizar as informações prévias das modulações para, através do rádio cognitivo, alocar de maneira

dinâmica usuários secundários nos espaços inutilizados do espectro de frequência, denominados *buracos espectrais* [8].

Existe uma subutilização do espectro de frequência devido a política de alocação fixa de espectro a usuários, esta subutilização é sanada com o uso de rádios cognitivos. O rádio cognitivo aplica técnicas de sensoriamento para mapear o espectro e identificar *buracos espectrais* [9], uma subclasse da técnica de sensoriamento por *feature detection* (em português, detecção por características), o detector por forma de onda, depende de informações a priori do usuário primário para determinar se aquele espaço de espectro está sendo utilizado ou não por este usuário. Uma das informações a priori requeridas para aplicação deste método de detecção é a modulação, assim, a AMC é necessária para auxiliar no sensoriamento de espectro realizada pelo rádio cognitivo. Dentre os trabalhos em AMC, temos o proposto em [10], que utilizando o classificador SVM (do inglês *support vector machine*), foi capaz de classificar sete modulações monoportadora, dentre elas, modulações analógicas e digitais. Ainda, utilizando SVM, temos o trabalho de [11], que abrangeu mais modulações adicionando técnicas de múltiplo acesso, com diferenças também na metodologia de aquisição dos sinais. Há, também, trabalhos utilizando redes neurais artificiais propostos por [12] e [8], que obtiveram um bom desempenho na classificação de modulações.

A maior dificuldade nos processo de classificação de modulações está na etapa de geração de sinais, boa parte das literaturas trabalham com sinais em banda base [12] [13] [14] [15], o que não retrata os sistemas de comunicações reais. Outros fatores como formatação de pulso, modelagem do canal de comunicação, processo de amostragem e obtenção da envoltória complexa, que se ignorados, podem acabar comprometendo a validação e comparação dos métodos. Outra dificuldade na área de AMC está na classificação de modulações com forte nível de correlação entre elas. Alguns trabalhos como [11] [10] [16] [12] obtiveram boas taxas de acertos em determinadas modulações, apesar de métodos de obtenção de características distintos. Entretanto, notamos que, nestes trabalhos, as modulações em quadratura, como 16-QAM e 64-QAM, têm uma menor taxa de verdadeiros positivos justamente devido a semelhança entre suas propriedades, dificultando o processo de classificação.

Dada as dificuldades, procurou-se, então, simular sinais modulados condizen-

tes com encontrados em sistemas de comunicação reais, e com essa base de dados realizar uma investigação do comportamento dos classificadores para a problemática da classificação de modulações. Além da questão qualitativa da base de dados é proposto também uma investigação quantitativa, avaliando o impacto de uma grande base de dados com diversas modulações no comportamento destes classificadores. Investiga-se também, neste trabalho, quais modulações sofrem mais os impactos do canal ruidoso, e quais os classificadores tem melhor resposta na classificação destas modulações. E por último, é feita uma comparação entre os classificadores, apontando qual deles apresenta melhor nível de generalização e quais apresentam melhores resultados na classificação de modulações dada as variações da relação sinal-ruído (do inglês *signal-noise rate*, SNR). Esta proposta tem relevância devido a não presença na literatura de uma investigação do comportamento de diversos classificadores dada uma base de dados com estes parâmetros quantitativos e qualitativos.

Para o trabalho proposto, foram extraídas 29 características de 376800 sinais, totalizando 47100 sinais para cada modulação. Os sinais gerados são em banda base acrescidos de ruído modelado, também em banda base, pelos canais Rayleigh e AWGN (do inglês *additive white gaussian noise*) em 5 níveis de SNR: 0, 5, 10, 20 e 30 dB. É realizado a etapa de formatação de pulso para então serem elevados, pelo *upconverter*, para uma frequência mais elevada. É realizado o processo de amostragem e a obtenção da envoltória complexa, que facilita a extração de informações de amplitude, frequência e fase. As modulações abordadas são as AM, FM, BPSK, QPSK, 16-QAM, 64-QAM, GMSK e WCDMA. O processo de classificação é dado por quatro classificadores multiclasse: uma rede neural artificial (RNA), duas árvores de decisão e um classificador probabilístico. A rede neural é uma rede *perceptron* multicamada totalmente conectada com algoritmo de retropropagação (do inglês *backpropagation*) de Levenberg-Marquardt para atualização dos pesos. Dois classificadores por árvore de decisão, um deles utilizando o algoritmo *J48*, que utiliza a entropia da informação para decidir quais características de entrada tem mais relevância na segregação das modulações, e outro utilizando o algoritmo *Random Forest*, que cria uma árvore de decisão para cada subconjunto de treinamento criado pela técnica *bootstrap*. E por último, o classificador *Naive Bayes* que utiliza o

teorema de Bayes, probabilidade condicional, para classificação.

Na literatura, o trabalho de [17] tem uma base de dados semelhante qualitativamente a apresentada neste trabalho e, utilizando um classificador SVM hierárquico, obteve uma taxa de 89,72 % de correta classificação de modulações. Em [18] é utilizado duas bases de dados diferentes e uma rede neural com otimização Bayesiana e obteve taxas de 80,81 % e 90,64 % com uma SNR de 18 dB. Já em [12] a taxa de acurácia obtida é de mais de 90 %. Dentre os quatro classificadores abordados neste trabalho, a árvore de decisão com algoritmo *J48* obteve 93,1253 % de acurácia, a árvore de decisão com algoritmo *Random Forest* obteve taxa de 95,2866 %, já o *Naive Bayes* obteve 74,7845 % de acurácia e, por último, a RNA obteve taxa de 87,4 %. Entretanto, como dito anteriormente, não existe uma padronização na etapa de geração de sinais o que resulta numa discrepância dos vetores de características utilizados no treinamento e teste dos classificadores, comprometendo, assim, a validação e comparação dos resultados com diversas literaturas.

1.1 Objetivos da Dissertação

O principal objetivo desta dissertação é a investigação do comportamento de diversos classificadores para a problemática da classificação automática de modulações. Investigar a resposta de quais classificadores apresentam melhor desempenho conforme a variação da SNR. Investigar quais das modulações abordadas (AM, FM, BPSK, QPSK, 16-QAM, 64-QAM, GMSK e WCDMA) são mais susceptíveis ao canal ruidoso e quais classificadores os discriminam melhor. Investigar o comportamento dos classificadores na discriminação de modulações mais correlacionadas.

Objetivos Específicos

- Estender a base de dados, os classificadores tender a ter um nível de generalização maior quando treinados com vários exemplos de cada classe.
- Definir e parametrizar quais classificadores abordar para a investigação da AMC.
- Investigar o comportamento destes classificadores para as modulações abor-

dadas e analisar individualmente o desempenho de cada classificador.

- Análise comparativa entre os classificadores abordados conforme a variação da SNR, tempo de treinamento e nível de generalização.

1.2 Estrutura do Trabalho

Este trabalho está organizado da seguinte maneira:

- No Capítulo 2, é exposto os fundamentos teóricos utilizados neste trabalho. É apresentado conceitos de telecomunicações, conceitos de reconhecimento de padrões e uma fundamentação da classificação automática de modulações. É informado sobre os tipos de modulações e descrito de forma matemática as características que foram extraídas do sinais modulados. E por fim, é dado uma visão sobre o funcionamento dos classificadores.
- No Capítulo 3 temos a revisão bibliográfica apresentando as técnicas atualmente mais abordadas nas literaturas para AMC, elencando os tipos de modulações, canais de comunicação, classificadores e base de dados utilizados. Por fim, a comparação entre as literaturas e o trabalho proposto nesta dissertação.
- No Capítulo 4 é apresentado a metodologia proposta, desde a geração dos sinais, transmissão, extração de características e finalmente e mais importante os projetos de classificadores, onde é demonstrado como funciona o processo de treinamento de cada classificador proposto.
- É Capítulo 5 é mostrado os detalhes do procedimento experimental como as especificações da máquina utilizada, da base de dados e as métricas utilizadas. Por fim, os resultados dos classificadores são apresentados em forma de matrizes de confusão, gráficos e análises de seus comportamentos.
- No Capítulo 6 é apresentado as considerações finais e as propostas de trabalhos futuros.

Capítulo 2

Fundamentos Teóricos

2.1 Conceitos de Telecomunicações

A telecomunicação é um ramo da Engenharia que engloba o projeto, implantação, manutenção e controle de sistemas de comunicação. O objetivo mais importante das telecomunicações é suprir a necessidade da humanidade de se comunicar a distância [19], [20]. Os sistemas de comunicação são projetados para transmitir sinais de comunicação de um emissor a um receptor que seja capaz de converter esses sinais em informações inteligíveis aos seres humanos [21]. Em um sistema de comunicação, os sinais transmitidos são representados pelo espectro de frequência, sendo assim, cada sinal ocupa o que é designado como faixa do espectro de frequência ou *banda* [21], [7]. Em uma transmissão, os sinais podem ser transmitidos em banda base ou banda passante [22]. Os sinais em banda base possuem uma gama de frequência muito estreita (sinais em baixa frequência) e no processo de comunicação, para serem transmitidos sem haja um deslocamento de frequência, utiliza-se meios confinados na transmissão. Os sinais que passam pelo processo de modulação, sinais em banda passante, são transmitidos, normalmente, em meios não confinados (espaço livre), apesar de também serem transmitidos em meios confinados [22].

Temos ainda, em um sistema de comunicação, o meio de transmissão, conforme Figura 2.1, onde o sinal transmitido é contaminado por ruído e distorção [23]. O ruído é uma interferência estatística e imprevisível e, normalmente, não pode ser removida devido a sua aleatoriedade [23]. A distorção, a princípio, pode ser eliminada ou reduzida, pois trata-se de uma modificação no sinal através de uma

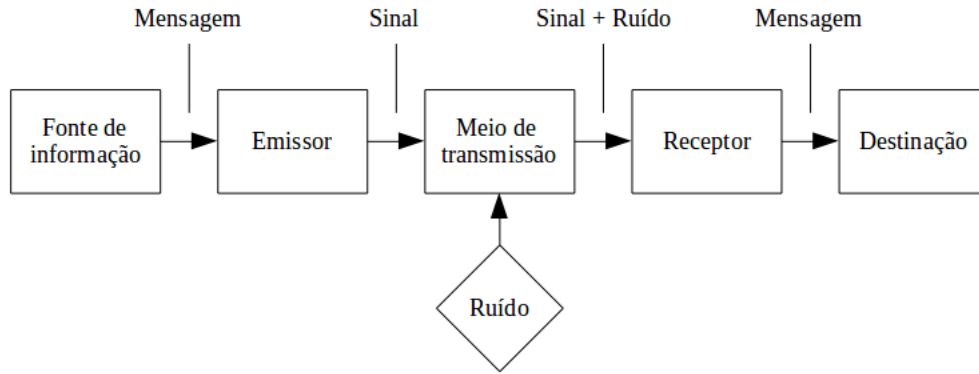


Figura 2.1: Diagrama em blocos de um sistema de comunicação típico.

operação fixa [23]. A relação sinal-ruído (*signal-to-noise ratio*, SNR) é uma métrica utilizada para medir o quanto o ruído interfere no sinal transmitido, sendo a razão entre a potência do sinal e a potência do ruído expressos em decibéis.

Desde o início do uso das comunicações remotas alguns órgãos ficaram responsáveis pela política de alocação de espectro de frequência. A união internacional de telecomunicações (UIT), a nível internacional, é responsável por ditar diretrizes de utilização do espectro [6] que são seguidas pelas nações, a agência nacional de telecomunicações (Anatel) é o órgão responsável no Brasil. A alocação de espectro é, desde então, feita de maneira fixa, cada usuário recebe uma determinada e exclusiva fatia do espectro.

2.2 Conceitos de Reconhecimento de Padrões

Reconhecimento de padrões é uma área da pesquisa com propósito de classificar objetos em diferentes classes ou categorias. Assim, reconhecimento de padrões pode ser definido como parte integrante na maioria dos sistemas de inteligência de máquina construídos para tomada de decisão [24]. Um sistema de reconhecimento de padrões envolve, conforme a Figura 2.2, uma coleta de dados do objeto a ser classificado, a extração de características, a definição de um modelo de classificador, o treinamento do classificador e, por fim, uma avaliação do desempenho do classificador.

Na primeira etapa de um sistema de reconhecimento de padrões, conforme mostrado na Figura 2.2, temos a coleta de dados, nesta etapa é adquirido as infor-

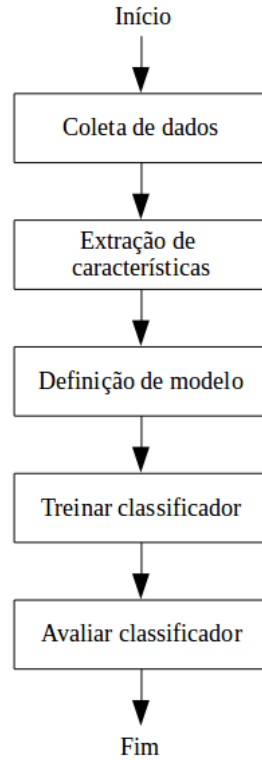


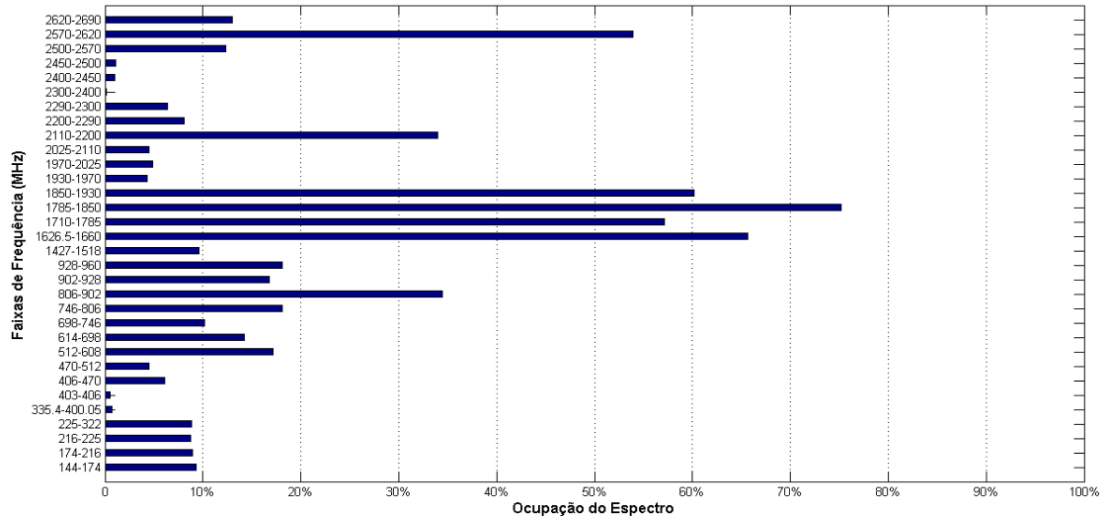
Figura 2.2: Projeto de um sistema de reconhecimento de padrões.

mações referente ao objeto a ser classificado. A segunda fase temos a extração de características, nesta parte é retirado das informações coletadas as características mais singulares do objeto, deixando-o, com isso, o mais distinguível possível. Após essas etapas é preciso definir um modelo de classificador que seja mais adequado para a tarefa de classificação proposta. Em seguida, treina-se o classificador utilizando as características extraídas do objeto e sequencialmente avalia-se se o classificador cumpriu com o objetivo de discriminar o objeto em questão.

2.3 Classificação Automática de Modulação

O espectro de radiofrequência é um recurso natural finito e engloba uma faixa de frequências que vai de 3 kHz a 300 GHz. Este espectro é utilizado em sistemas de comunicação remota, via rádio, e com chegada de novas tecnologias a demanda por serviços de comunicação vem crescendo drasticamente nos últimos anos. E devido a política de alocação fixa de espectro sugerida pela UIT, através do setor de radiocomunicação (UIT-R), a eficiência na alocação de usuários no espectro

Figura 2.3: Taxa de ocupação do espectro de frequência na cidade do Rio de Janeiro na faixa de 144 a 2690 MHz em 2013.



Fonte: DE LIMA, 2013 [25].

de frequência fica comprometida, apesar desta política garantir a não interferência entre serviços. A subutilização deste espectro fica evidente na pesquisa realizada por [25] no Rio de Janeiro, onde a média de ocupação do espectro de frequência foi de 19,6 %, o gráfico mostrado na Figura 2.3 mostra claramente as faixas de frequência mais subutilizadas. Por exemplo, neste estudo de 2013, na faixa de 335 a 400 MHz, que é utilizada para radionavegação aeronáutica (conforme o plano de atribuição de faixa de frequência do Brasil em 2018), menos de 3% do espectro é utilizado, resultando numa subutilização de aproximadamente 97%. Nesta problemática surgiram os rádios cognitivos que realizam o mapeamento do espectro de radiofrequência e alocam de maneira dinâmica usuários nos espaços inutilizados do espectro, denominados buracos espectrais. Porém, é necessário informações a priori destes usuários para aloca-los nestes buracos espectrais sem causar interferências entre os usuários. A classificação automática de modulação nasceu neste contexto, onde, para fazer o sensoriamento do espectro, o rádio cognitivo precisa de informações a priori do usuário primário [9].

Os sistemas de classificação automática de modulação tiveram início nas décadas de 80 e 90 [6], com propósito de atender as demandas do setor militar e posteriormente ampliado para o uso civil. Dentro da esfera militar a AMC era aplicada para formação de estratégias de ataque, suporte e defesa eletrônica [11]. No uso civil a AMC é aplicada para aumentar a eficiência do uso do espectro de frequência.

Existe uma variedade de métodos de AMC: por verossimilhança (*likelihood-based, LB*), características (*feature-based, FB*), baseados teste de distribuição e método de classificação cega de modulações [7] [16]. A abordagem por *feature-based* consiste na extração de um vetor de características dos sinais modulados (banda passante). Assim, com uma grande base de sinais é possível extrair os vetores de características e treinar um classificador para identificar modulações automaticamente. O classificador tem o papel de aprender da forma mais eficiente possível como distinguir essas modulações com base nas características extraídas. Essa informação é utilizada por rádios cognitivos que dinamizam o uso do espectro de frequência para torná-lo mais eficiente.

2.4 Tipos de Modulações

A modulação é uma técnica onde as propriedades de um sinal portador (sinal modulado) são alterados com o propósito de transmitir informação, as modificações ocorrem de maneira combinada em frequência, amplitude ou fase [21]. Um sinal portador (portadora) é o sinal que irá portar o sinal que contém a informação a ser transmitida, que normalmente se encontra em banda base, e possui uma frequência bem mais elevada que o sinal de informação. Existem modulações analógicas, digitais e as técnicas de múltiplo acesso.

2.4.1 Modulações Analógicas

As modulações analógicas surgiram da necessidade de se transmitir sinais de informação de natureza analógica a longas distâncias por espaço livre [26]. Devido os sinais desta natureza possuírem baixas frequências, como a voz, os deixariam quase impossíveis de serem transmitidos a longas distâncias por espaço livre, pois, o tamanho da antena que recebe sinais de rádio é proporcional ao comprimento de onda transmitida, conseqüentemente, seria necessário antenas de tamanhos absurdos para captar sinais em banda base [22]. As modulações analógicas mais utilizadas são as modulações por amplitude, frequência e fase.

Na modulação por amplitude (do inglês *amplitude modulation, AM*) a amplitude da portadora se modifica proporcionalmente à amplitude do sinal em banda

base [22]. No domínio da frequência é possível notar a presença de uma componente de frequência próxima da origem, referente ao sinal modulante, e outras duas componentes ao redor da frequência da portadora ω_c , denominadas banda lateral superior e banda lateral inferior. Na modulação por frequência (do inglês *frequency modulation*, FM) a frequência da portadora se modifica proporcionalmente à amplitude do sinal em banda base [22]. A representação no espectro de frequência do sinal FM é composto por intermináveis frequências $n f_m$ localizadas simetricamente à componente ω_c , sendo f_m a frequência do sinal em banda base. Assim, a modulação FM mantém a amplitude da portadora constante e altera sua frequência, enquanto a modulação AM não altera a frequência da portadora somente sua amplitude.

2.4.2 Modulações Digitais

As modulações digitais são métodos onde a informação a ser transmitida faz parte de um agrupamento de valores discretos que representam um símbolo [27]. Estes símbolos são representados por componentes complexas obtidas após o processo de quantização do sinal em banda base. São criados, então, pulsos retangulares com as mesmas amplitudes das componentes complexas, tanto parte real quanto parte imaginária. O sinal em banda passante é gerado somando-se a multiplicação destes pulsos retangulares por uma portadora em fase (*in fase*, I) e outra deslocada de 90° (*quadrature*, Q), as duas com a mesma frequência. Não existe interferência entre as componentes I-Q devido a ortogonalidade entre as portadoras.

O processo de quantização indica a ordem M da modulação que constitui-se no número de pontos na constelação do sinal. A constelação é forma de representação dos pontos que os símbolos da modulação podem assumir no plano complexo formado pelas componentes I-Q. O formato da constelação dita o comportamento em amplitude, fase e frequência dos sinais digitais, sendo assim, fundamental na discriminação das classes de modulação [22]. Foi utilizado, neste trabalho, as modulações em ordem 2, 4, 16 e 64, respectivamente BPSK (do inglês *binary phase shift keying*), QPSK (do inglês *quadrature phase shift keying*), 16-QAM e 64-QAM (do inglês *quadrature amplitude modulation*). Além destas modulações, utilizou-se a modulação GMSK (*gaussian minimum shift keying*) que é semelhante a modulação QPSK com o diferencial da forma de onda modulante, que é um pulso gaussiano em

vez de um pulso retangular.

2.4.3 Técnicas de Múltiplo Acesso

Entre as técnicas de múltiplo acesso há aquelas que utilizam múltiplas portadoras para transmitir informação e consistem em dividir os dados a serem transmitidos em diversos e intercalados fluxos de *bits*, e utilizando esses fluxos para modular várias portadoras [28]. Esta técnica surgiu da necessidade de se reduzir a interferência simbólica (do inglês *inter-symbol interference*, ISI) que as modulações monoportadora geram com o aumento da taxa de símbolos. A utilização de técnicas multiportadora, então, busca aumentar o tempo do símbolo e tornar o processo de transmissão mais resistente a atrasos de multipercurso e, assim, reduzir a interferência simbólica. A transmissão multiportadora tem como ideia central um processo denominado *conversão serial-paralelo*, onde uma sequência de símbolos de alta taxa é dividido em várias subsequências paralelas com taxas menores para então serem moduladas por modulações digitais monoportadora [29]. Assim, no momento da transmissão é alocado para cada monoportadora um subcanal onde o tempo de símbolo maior.

Entre as técnicas multiportadora mais comuns são as de multiplexação por divisão em frequência chamada OFDM (do inglês *orthogonal frequency division multiplexing*) [30]. A principal diferença entre a tradicional FDM (do inglês *frequency division multiplexing*) e a OFDM, é que na OFDM o espectro das individuais portadoras se sobrepõem mutuamente, resultando numa eficiência de espectro melhor [31], deixando assim o tempo de símbolo praticamente igual a taxa de *Nyquist*¹ [32]. É possível eliminar, na OFDM, a ISI e a interferência entre canais (do inglês *inter-channel interference*, ICI) inserindo entre os símbolos um pequeno tempo no intervalo de guarda. Com um tamanho do intervalo de guarda igual ou maior que o tempo de propagação do canal, a ICI é eliminada, dentro do intervalo de guarda, devido a manutenção da ortogonalidade das portadoras durante o período do símbolo mesmo se a forma de onda do sinal do símbolo se estender periodicamente [31]. Também, já que símbolos sucessivos não se sobrepõem por causa do intervalo de guarda, a

¹Teorema da amostragem de *Nyquist* diz que a frequência de amostragem (f_s) deve ser maior que duas vezes a maior frequência do sinal em questão para que não haja problemas em sua reconstrução.

ISI é eliminada. Assim, uma das principais vantagens do OFDM é seu desempenho quase ideal em canais não equalizados e em canais de múltiplos caminhos.

Outra técnica de múltiplo acesso é que a realizam espalhamento espectral, e consistem em aumentar propositalmente a largura da banda de transmissão de um sinal em banda base de maneira que seja maior que a mínima banda necessária para transmitir a informação [32]. O espalhamento de bandas é realizado por meio de um código que é independente dos dados transmitidos, e uma recepção sincronizada com esse mesmo código no receptor é utilizada para o processo inverso ao espalhamento e sequencialmente a recuperação dos dados [33]. Dentre os métodos de espalhamento espectral nos temos a DS-CDMA (do inglês *direct sequence - code division multiple access*), FH-CDMA (do inglês *frequency hopping - code division multiple access*) e TH-CDMA (do inglês *time hopping - code division multiple access*). No DS-CDMA, uma sequência gerada pseudo-aleatoriamente, denominada *pseudo-noise*, provoca transições de fase na portadora que carrega a informação. Na recepção, a recuperação do sinal de informação se dá pela multiplicação do sinal recebido pelo *pseudo-noise* aplicado na transmissão. Na técnica FH-CDMA a portadora a mudar de frequência de maneira pseudo-aleatória, e na recepção o sinal de informação é recuperado acessando as frequências pseudo-aleatoriamente escolhidas na transmissão. E, por fim, na técnica TH-CDMA as rajadas de sinais são iniciadas em tempos pseudo-aleatórios [33]. Combinações híbridas destas técnicas são frequentemente usadas, permitindo que grande quantidade de usuários de sistemas de comunicação tenham acesso simultâneo ao meio.

2.5 Extração de características

O processo de extração de características tem como objetivo gerar o vetor de características que evidencia propriedades importantes para discriminação dos sinais. Neste trabalho, propõe-se classificar sinais modulados por modulações analógicas e digitais, como também, sinais transmitidos utilizando-se técnicas de multiplexação. Para isso, extraiu-se, então, dos sinais modulados, cinco categorias de características, sendo elas: características espectrais, de transformada, estatísticas, de teste de distribuição e cicloestacionárias [7].

2.5.1 Características Espectrais

Dentre as características espectrais existe sete que são largamente utilizadas na classificação automática de modulações [8]. Estas características avaliam propriedades espectrais únicas dos sinais modulados que apresentam diferenciações em amplitude, fase e frequência, e são:

- Valor máximo da densidade espectral de potência (do inglês *power spectrum density*, PSD) da amplitude instantânea normalizada e centralizada (γ_{max}) é dada pela PDS da amplitude instantânea normalizada e centralizada:

$$\gamma_{max} = \frac{\max |DFT(a_{nc}(n))|^2}{N_s} \quad (2.1)$$

sendo N_s o número de amostras por segmento e $a_{nc}(n)$ é a amplitude instantânea normalizada e centralizada, onde $a_{nc}(n) = \frac{|\mathcal{H}\{x(n)\}e^{i2\pi f_c n}|}{m_a} - 1$. Sendo \mathcal{H} a transformada de Hilbert, $x(n)$ é o sinal de entrada amostrado em instantes de tempo $t = \frac{n}{f_s}$ e m_a é a média do valor da amplitude instantânea sobre determinado seguimento do sinal, dado por $\frac{1}{N_s} \sum_{n=1}^{N_s} |\mathcal{H}\{x(n)\}e^{i2\pi f_c n}|$.

- Desvio padrão da amplitude instantânea normalizada e centralizada (σ_{aa}):

$$\sigma_{aa} = \sqrt{\frac{1}{n} \left(\sum_{n=1}^n (a_{nc}(n) - \overline{a_{nc}(n)})^2 \right)} \quad (2.2)$$

sendo $a_{nc}(n)$ a amplitude instantânea normalizada e centralizada mostrado no item anterior e $\overline{a_{nc}(n)}$ a média da amplitude instantânea normalizada e centralizada.

- Desvio padrão da fase instantânea absoluta não linear centralizada (σ_{ap}) é avaliado sobre intervalos não fracos do segmento do sinal, definiu-se, então, a região onde $a_n(n) \geq 0,1$ como segmentos fracos. O σ_{ap} é expresso a seguir:

$$\sigma_{ap} = \sqrt{\frac{1}{C} \left(\sum_{a_n(n) > 0,1}^N \phi_{NL}^2(n) \right) - \left(\sum_{a_n(n) > 0,1} |\phi_{NL}(i)| \right)^2} \quad (2.3)$$

sendo $a_n(n) = \frac{|\mathcal{H}\{x(n)\}e^{i2\pi f_c n}|}{m_a}$ e C é o total de amostras no segmento não fraco do sinal. ϕ_{NL} é a fase não linear expressa pela angulação entre as componentes real e imaginária da transformada de Hilbert do sinal de entrada, $\mathcal{H}\{x(n)\}$. Já $\phi_{NL}(n)$ é o valor da componente não linear da fase instantânea em instantes de tempo $t = \frac{n}{f_s}$.

- Desvio padrão da fase não linear direta centralizada (σ_{dp}):

$$\sigma_{dp} = \sqrt{\frac{1}{C} \left(\sum_{a_n(n)>0,1}^N \phi_{NL}^2(i) \right) - \left(\sum_{a_n(n)>0,1} \phi_{NL}(n) \right)^2} \quad (2.4)$$

- Desvio padrão da frequência instantânea normalizada e centralizada (σ_{af}) é avaliado sobre intervalos não fracos de um segmento do sinal, σ_{af} é obtido segundo a seguinte expressão:

$$\sigma_{af} = \sqrt{\frac{1}{C} \left(\sum_{a_n(n)>0,1}^N f_N^2(n) \right) - \left(\sum_{a_n(n)>0,1} |f_N(n)| \right)^2} \quad (2.5)$$

sendo $f_N(n) = \frac{f(n)-m_f}{r_s}$, onde r_s é a taxa de símbolos da sequência digital, $m_f = \frac{1}{N_s} \sum_{n=1}^{N_s} f(n)$ e $f(n)$ é a frequência instantânea dada por pela derivada em relação ao tempo de $\phi_{NL}(n)$ dividida por 2π , $\frac{1}{2\pi} \frac{d\phi}{dt}$.

- Desvio padrão do valor absoluto da frequência instantânea normalizada e centralizada (σ_f):

$$\sigma_f = \sqrt{\frac{1}{C} \left(\sum_{a_n(n)>0,1}^N f_N^2(n) \right) - \left(\sum_{a_n(n)>0,1} |f_N(i)| \right)^2} \quad (2.6)$$

- Valor máximo da PSD da frequência instantânea normalizada e centralizada (γ_{maxf}) é dado pela equação:

$$\gamma_{maxf} = \frac{\max |DFT(f_N(n))|^2}{N_s} \quad (2.7)$$

sendo f_N a frequência normalizada e centralizada mostrada nos itens anteriores.

2.5.2 Características por Transformada

A representação dos sinais modulados em diversos domínios também fornece propriedades peculiares que auxiliam na classificação automática de modulações. A representação dos sinais modulados nos domínios do tempo, frequência e espaço é realizado pelas transformadas, assim, temos:

- Valor máximo da transformada discreta de cosseno (max_{dct}), que pode ser descrita pela equação:

$$C_x(k) = \begin{cases} \sum_{n=0}^{N-1} 2x(n) \cos\left(\frac{\pi}{2N}k(2n+1)\right), & \text{para } 0 \leq k \leq N \\ 0, & \text{caso contrário} \end{cases} \quad (2.8)$$

O valor máximo resultante da utilização da transformada discreta de cosseno sobre a envoltória complexa do sinal, dada pela $\mathcal{H}\{x(n)\}$, representa a característica.

- Desvio padrão da transformada de Walsh-Hadamard do sinal (σ_{wht}) é dada pela seguinte expressão:

$$\mathbf{WTH}_N = \overbrace{DFT_2 \otimes \dots \otimes DFT_2}^{n \text{ vezes}} \quad (2.9)$$

sendo $DFT_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ que é a matriz DFT de 2 pontos e \otimes é o produto de Kronecker. A característica é obtida calculando-se o desvio padrão dos coeficientes da transformada de Walsh-Hadamard sobre o sinal modulado.

- Desvio padrão da transformada de Wavelet discreta (σ_{dwt}) é dado pela expressão a seguir:

$$\sigma_{dwt} = \max |DWT(n)| \quad (2.10)$$

sendo DWT a transformada discreta de Wavelet.

- O valor máximo dos coeficientes ceptrais na frequência mel:

$$M(f) = 1125 \ln \left(\frac{f}{700} \right) \quad (2.11)$$

2.5.3 Características Estatísticas

As características estatísticas são representadas pelo cálculo das cumulantes das partes reais e imaginárias da envoltória complexa. As cumulantes são amplamente utilizadas como características de discriminação de modulações por apresentarem robustez a influência de ruído [10].

- Cálculo de cumulantes de segunda ordem é dado pela seguinte equação:

$$C_{z_1, z_2} = \frac{1}{N} \sum_{n=1}^N z_1^n z_2^n \quad (2.12)$$

- Cálculo de cumulantes de terceira ordem:

$$C_{z_1, z_2, z_3} = \frac{1}{N} \sum_{n=1}^N z_1^n z_2^n z_3^n \quad (2.13)$$

- Cálculo de cumulantes de quarta ordem:

$$C_{z_1, z_2, z_3, z_4} = \frac{1}{N} \sum_{n=1}^N z_1^n z_2^n z_3^n z_4^n - C_{z_1, z_2} C_{z_3, z_4} - C_{z_1, z_3} C_{z_2, z_4} - C_{z_1, z_4} C_{z_2, z_3} \quad (2.14)$$

sendo z_i um vetor que descreve a envoltória complexa do sinal recebido.

2.5.4 Característica de Teste de Distribuição

Cada sinal modulado demonstra uma diferenciação a cada tipo de teste de distribuição devido ao formato da constelação. Os testes de distribuição mais comuns são: Hinich, Kolmogorov-Smirnov, Jarque-Bera, Giannakis-Tsatsanis, Anderson-Darling, D'agostino-Pearson, Shapiro-Wilk, Cramer-von-Mises e Lilliefors. Apesar de haver vários testes, eles apresentam deficiências na presença de ruído e por esse motivo não são adequados para classificação de sinais digitais [34].

- Função de Teste de Gaussianidade ($X_{m,n}$) é dado pela expressão a seguir:

$$X_{m,n} = (N^{1-4c} Q_{m,n}^{-1/2}) \left[S_x \left(2\pi(2m-1) \frac{M}{2N} \right) S_x \left(2\pi(2n-1) \frac{M}{2N} \right) \right] S_x \left(2\pi(m+n-1) \frac{M}{N} \right) \hat{B}_x(m,n) \quad (2.15)$$

sendo $Q(m,n)$ o número de (j,k) de quadrados que estão em D , não incluindo as fronteiras, mais duas vezes a quantidade de quadrados da fronteira, onde $D = \{0 < j \leq N/2, 0 < k \leq j, 2j+k \leq N\}$. S_x é o espectro de $x(n)$ e c é uma variável de regularização entre *bias* e variância. E $\hat{B}_x(m,n)$ é um estimador do biespectro.

2.5.5 Características Cicloestacionárias

Por fim, as características cicloestacionárias, que computam a curtose (k_{SCD}), desvio padrão (σ_{SCD}) e a variância (σ_{SCD}^2) da densidade de correlação espectral (do inglês *spectral correlation density*, SCD). A SCD calcula a similitude entre versões deslocadas da PSD sobre a frequência cíclica α que é adquirida com base na teoria da cicloestacionariedade. A cicloestacionariedade é usada para discriminar processos estocásticos não estacionários e achar periodicidades ocultas. Assim, devido a deficiência da PSD quando se trata da procura de informações sobre a distribuição na faixa de frequência, utiliza-se a cicloestacionariedade na procura destes ciclos de operação.

- A SCD é definida como a transformada de Fourier da função de autocorrelação cíclica (CAF) (do inglês *Cyclic Autocorrelation Function*):

$$S_x^\alpha(f) = \int_{-\infty}^{\infty} R_x^\alpha(\tau) e^{-j2\pi f\tau} = \left\langle X \left(t, f + \frac{\alpha}{2} \right) X \left(t, f - \frac{\alpha}{2} \right) \right\rangle \quad (2.16)$$

sendo $X \left(n, f + \frac{\alpha}{2} \right)$ e $X \left(n, f - \frac{\alpha}{2} \right)$ são envelopes complexos de banda estreita, componentes de em banda passante de $x(n)$.

- Cumulantes cíclicos de terceira e quarta ordem:

$$m_{3x}(t, \tau_1, \tau_2) = \frac{1}{T} \sum_{t=0}^{T-1} [x(t)e^{-j2\pi\alpha t}] [x(t + \tau_1)e^{-j2\pi\alpha(t+\tau_1)}] [x(t + \tau_2)e^{-j2\pi\alpha(t-\tau_2)}] e^{j2\pi\alpha t} \quad (2.17)$$

sendo $m_{3x}(t, \tau_1, \tau_2)$ é a cumulante cíclico de terceira ordem. A cumulante de quarta ordem é obtida de maneira semelhante considerando-se um momento de quarta ordem τ_3 .

2.6 Classificadores

A tarefa do classificador de um sistema completo é utilizar o vetor de características, adquirido na extração de características, para determinar a qual classe ou categoria pertence o objeto [1]. No processo de classificação é normalmente improvável se obter uma classificação perfeita, o que se propõe usualmente é a probabilidade de ocorrência de uma classe ou outra. A grande dificuldade do problema de classificação é a variabilidade nos valores de características nos objetos de mesma categoria em relação a diferença entre valores de características de objetos em categorias diferentes. Dentre os métodos de classificação mais populares estão os classificadores probabilísticos, como *Naive Bayes*, os baseados em árvore de decisão, como *random forest* e as RNAs.

O classificador *Naive Bayes* é um classificador probabilístico baseado no Teorema de *Bayes* e atualmente se tornou muito popular na área de aprendizado de máquina (*machine learning*). Esse algoritmo recebe o nome de *naive* pelo fato de considerar uma total independência entre as características, o que é improvável nas aplicações reais, por isso o nome *naive* (ingênuo) [35]. Por causa do pressuposto de independência assumido pelo algoritmo, os parâmetros para cada característica podem ser aprendidos separadamente, e isso simplifica o aprendizado, especialmente se a quantidade de características for grande [36]. O *Naive Bayes* classifica objetos calculando a probabilidade a condicional de que o objeto pertence a determinada classe, sabendo-se da ocorrência de uma determinada característica, $P(C = c_k | X = x)$, sendo c_k a classe e x a característica. Assim, dentre todo o conjunto de classes, a

classe que obteve a maior probabilidade é selecionada como a classe do objeto.

Árvore de decisão são técnicas de aprendizado de máquina supervisionado vastamente utilizadas em trabalhos de classificação e regressão. Os algoritmos baseados em árvore de decisão utilizam a estratégia de dividir para conquistar, onde um problema complexo é decomposto em um conjunto de subproblemas mais simples de se solucionar [37]. Uma árvore de decisão estruturalmente é constituída por nós, ramos e folhas. Cada nó contém um teste para algum atributo, obedecendo a regra de *se-então*, gerando ramos que levam para novos nós ou folhas, o nó raiz é a origem da árvore e as folhas estão associadas a uma classe. Assim, cada percurso da árvore, do nó raiz à folha corresponde a uma regra de classificação. Na maioria dos algoritmos baseados em árvore é utilizado o ganho de informação ou a razão do ganho de informação, através do cálculo de entropia da informação, como método de determinação de quais características tem mais relevância na etapa de classificação.

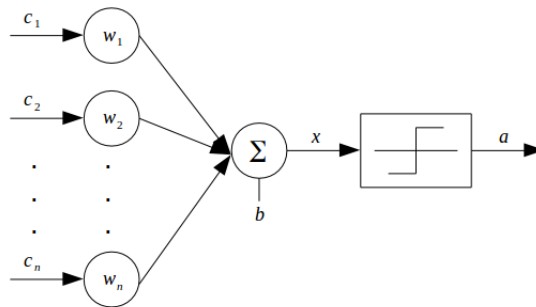


Figura 2.4: Modelo de neurônio idealizado por McCulloch.

A primeira concepção de neurônio, realizada na área de biologia, ocorreu no século 19, quando o neurologista (espanhol) Ramón Cajal identificou e descreveu anatomicamente um neurônio biológico. Neste século, DuBois Reymond, observou pela primeira vez manifestações elétricas de neurônios biológicos com o auxílio de galvanômetros. Na década de 1920, Erlanger e Gasse, observaram a atividade elétrica nervosa. Essas principais descobertas e os estudos, em 1943, de Turing e Neumann sobre a natureza booleana da inteligência permitiram Warren McCulloch e Walter Pitts conceber o primeiro trabalho sobre RNA [38]. O dispositivo idealizado por McCulloch é um classificador linear com entradas binárias. De maneira genérica um classificador linear de várias entradas e uma saída pode ser descrito pela seguinte expressão:

$$a = \operatorname{sgn} \left(\sum_{i=1}^N w_i c_i + b \right) \rightarrow a \in [-1; 1] \quad (2.18)$$

sendo os pesos w_i são associados a cada característica de entrada c_i , b é o valor de limiar, e sgn^2 é a função de ativação do neurônio que retorna como saída um valor a contínuo no intervalo $[-1; 1]$. Na Figura 2.4 temos a representação gráfica do neurônio idealizado por McCulloch e foi a partir dessa ideia que surgiram outras configurações de RNAs mais robustas como: adaline e madaline, redes perceptron multicamadas, redes neurais convolutivas [38] [16].

²Função sinal, que retorna valor 1 se a entrada for maior que zero e -1 se a entrada for menor que zero.

Capítulo 3

Trabalhos Relacionados

3.1 Técnicas para Classificação Automática de Modulação

Os sistemas de AMC tiveram início nas décadas de 80 e 90 [6], com propósito de atender as demandas do setor militar e posteriormente ampliado para o uso civil. Dentro da esfera militar a AMC era aplicada para formação de estratégias de ataque, suporte e defesa eletrônica [11]. No uso civil a AMC é aplicada para aumentar a eficiência do uso do espectro de frequência. Existe uma variedade de métodos de AMC: por verossimilhança (*likelihood-based*, *LB*), características (*feature-based*, *FB*), baseados teste de distribuição e método de classificação cega de modulações [7] [16]. Os classificadores mais populares baseados em *feature-based*, são: RNA, redes neurais convolucionais (do inglês *convolutional neural network*, *CNN*), classificador probabilístico (*Naive Bayes*), as árvores de decisão e a SVM.

A abordagem por *feature-based* consiste na extração de um vetor de características dos sinais modulados. O intuito é que haja uma certo nível de independência entre as características extraídas de cada modulação para que o classificador possa discriminar as modulações. Para *feature-based*, quanto às características, existe uma grande variedade aplicável aos sinais modulados, como: amplitude, fase e frequência instantâneas; cumulantes; transformadas de *Fourier* e *Wavelet*; cicloestacionárias e outras. Assim, com uma grande base de sinais é possível extrair os vetores de características e treinar um classificador para identificar modulações automaticamente.

O classificador tem o papel de aprender da forma mais eficiente possível como distinguir essas modulações com base nas características extraídas. Essa informação é utilizada por rádios cognitivos que dinamizam o uso do espectro de frequência para torná-lo mais eficiente.

Mais recentemente, conforme os trabalhos de [13] e [14], outro método de classificação de modulações vem sendo implementado. Para segregar modulações digitais, que possuem mapa de constelação¹, foram utilizadas imagens destes mapas como base de dado de entrada do classificador. Estas imagens são denominadas, nas literaturas, como matriz de constelações de grade (do inglês *grid constellation matrix*, CGM) e, em outros casos, constelação de grade aprimorada com contrastes (do inglês *contrast enhanced grid constellation*, CGC). Em [13] foi utilizado, para classificar estas imagens, uma rede convolutiva totalmente contrastante (do inglês *contrastive fully convolutional network*, CFCN), já em [14] foi utilizada uma rede neural convolucional compressiva (do inglês *compressive convolutional neural network*, CCNN).

3.2 Síntese dos Trabalhos Relacionados

Atualmente, as técnicas mais abordadas para classificar modulações são as baseadas em FB e as baseadas em imagens, para as modulações digitais, do mapa de constelação. Além disso, existem outros aspectos que devem ser levados em consideração e que possuem uma gama de possibilidades de aplicações, como: o tamanho da base de dados, o tipo de classificador, o modelo de canal de transmissão e as modulações abordadas. Diante disso, apresentamos uma síntese dos trabalhos de maior relevância na literatura atualmente. Em relação à base de dados, temos em [11] e [41] uma base de 400 amostras, na casa dos milhares temos [12], [39], [40], [13], [14], [42], [18] e [43], e por fim, [15] que utilizou uma base de 1200000 amostras. Quanto aos classificadores, utilizaram o SVM [11] e [41], redes neurais [12], [39] e [40], e as redes convolucionais [13], [14], [42], [15], [18] e [43]. Pode-se notar que maioria dos trabalhos utilizaram somente um modelo de canal de comunicação, o AWGN, só em [11], [12], [13] e [14] há a utilização de outro modelo

¹Mostra o sinal como um diagrama de dispersão bidimensional no plano complexo em instantes de amostragem de símbolo.

Trabalhos	Tipo	Base de dados	Classificador	Canal	Modulações
Amoedo <i>el al</i> [11]	FB	400	SVM	AWGN e Rayleigh	AM, FM, BPSK, QPSK, 16-QAM, 64-QAM, GMSK, WCDMA e OFDM
Kim <i>el al</i> [12]	FB	20000	RNA	AWGN e Rician	BPSK, QPSK, 8-PSK, 16-QAM e 64-QAM
Han <i>el al</i> [39]	FB	10000	RNA	AWGN	BPSK, 8-PSK, 4-QAM, 16-QAM e 64-QAM
Ali <i>el al</i> [40]	FB	4096	RNA	AWGN	BPSK, 4-QAM, 16-QAM e 64-QAM
Wu <i>el al</i> [41]	FB	400	SVM	AWGN	2-ASK, 4-ASK, 8-ASK, 2-FSK, 4-FSK, 8-FSK, 2-PSK, 4-PSK, 8-PSK e 16-QAM
Huang <i>el al</i> [13]	GCM	35000	CFCN	AWGN e Rayleigh	BPSK, QPSK, 8PSK, 16-QAM e 64-QAM
Triantaris <i>el al</i> [15]	FB	1200000	CNN	AWGN	8-PSK, AM-DSB, BPSK, CPFSK, GFSK, 4-PAM, 16-QAM, 64-QAM, QPSK e WBFM
Ramjee <i>el al</i> [42]	FB	160000	CNN	AWNG	BPSK, QPSK, 8-PSK, 16-QAM, 64-QAM, BFSK, CPFSK, 4-PAM, WB-FM e AM-DSB
Huang <i>el al</i> [14]	CGC	150000	CCNN	AWGN e Rayleigh	QPSK, 16-QAM e 64-QAM
Wei <i>el al</i> [18]	FB	220000	CNN	AWGN	8-PSK, AM-DSB, AM-SSB, BPSK, CPFSK, GFSK, 4-PAM, 16-QAM, 64-QAM, QPSK e WBFM
Wang <i>el al</i> [43]	FB	6000	CNN	AWGN	BPSK, QPSK, 8-PSK, GFSK, CPFSK, 4-PAM, 16-QAM e 64-QAM

Tabela 3.1: Apresenta a diversidade de técnicas de classificação, tamanhos de base de dados, classificadores, modelagem de canal e grupos de modulação.

de canal de comunicação além do AWGN. Na Tabela 3.1 temos uma perspectiva da diversidade de aspectos aplicados em diversos trabalhos na literatura, como também, a gama de modulações abordadas em cada literatura.

3.3 Comparação entre Literatura e Trabalho Proposto

Devido a falta de padronização entre os métodos de geração de sinal e modelagem do canal de comunicação a comparação entre as metodologias abordadas fica comprometida. Na geração de sinal, por exemplo, a formatação de pulso (*pulse shaping*) e o deslocamento de frequência (*upconverter*) não são usualmente implementados. Em [13], [14] e [12] as informações das modulações são extraídas dos sinais em banda base, o que não retrata a realidade dos processos de comunicação reais. Em aplicações reais os sinais são amostrados diretamente em banda passante, com isso, o formato da envoltória complexa é determinado por estas amostras. Assim, nestes trabalhos onde a classificação é realizada com sinais recebidos em banda base, a avaliação do desempenho é afetada. A etapa de formatação de pulso também afeta a avaliação dos trabalhos, a envoltória complexa também é influenciada pela formatação de pulso, já que, a formatação é aplicada para restringir a necessidade de banda para transmissão dos sinais e reduz a interferência entre símbolos.

Outro aspecto importante no desempenho dos classificadores é a base de dados. Os classificadores, no processo de treinamento, aprendem com base na quantidade de exemplos de cada classe que lhes foram fornecidos na entrada. Assim, quanto mais informações temos na entrada dos classificadores mais genéricos e eficientes na classificação de novas instâncias eles serão. Assim, notamos que os trabalhos [11] e [41] têm uma pequena base de dados, o que reflete numa falta de generalização do classificador, e em sistemas reais, onde uma grande quantidade de sinais é recebida, a classificação pode ficar comprometida. O modelo de canal de comunicação também é um aspecto bastante relevante. Tendo em vista isso, a modelagem inadequada do comportamento do ruído pode prejudicar a validação do método. Na literatura a maioria dos trabalhos modelam seu canal utilizando somente o AWGN [39], [40], [41], [15], [42], [18] e [43], o que não condiz com toda a

complexidade do processo de transmissão, tendo em vista degradação do sinal por efeitos naturais ou artificiais e múltiplos percursos que podem danificar a envoltória complexa. Então, uma modelagem mais elaborada do canal de comunicação aumenta a probabilidade de criar classificadores com maior coesão com os sistemas de comunicação reais.

Para este trabalho, com o intuito de criar um ambiente mais próximo possível dos sistemas de comunicações praticáveis, busca-se utilizar sinais em banda passante e com formatação de pulso, como também, uma base de dados grande o suficiente para que o classificador atinja altos níveis de generalização. A modelagem do canal pelo AWGN e pelo Rayleigh, que tende a dar mais similaridade com os sistemas de comunicação reais. Serão abordados, nesta dissertação, as modulações AM, FM, BPSK, QPSK, 16-QAM, 64-QAM GMSK, além da técnica de múltiplo acesso WCDMA.

Capítulo 4

Metodologia

Na literatura, é possível encontrar diversos métodos de classificação automática de modulações. Na grande maioria destas pesquisas a metodologia é dividida em três etapas: simulação de sinais modulados, extração de características e a classificação. Conforme descrito no Capítulo 3, temos uma diversidade de métodos que podem ser abordados. Como exemplo, temos literaturas que extraem características de sinais em banda base [13] e outros de sinais em banda passante [11]. O modelo adotado para o canal de comunicação também varia nas literaturas, alguns trabalhos utilizam somente o canal AWGN [40], [15] e [43], outras literaturas já procuram outros modelos de canais além do AWGN, como em [11], [12] e [13] que utilizam os modelos Rayleigh e Rician. As características utilizadas para classificação deixam, também, os métodos mais singulares, podemos ver até trabalhos com classificação por imagem [13] e [14]. As modulações que cada trabalho aborda na classificação também é um diferencial entre as metodologias, o que explicita ainda mais a diversidade de métodos praticáveis para classificar automaticamente modulações.

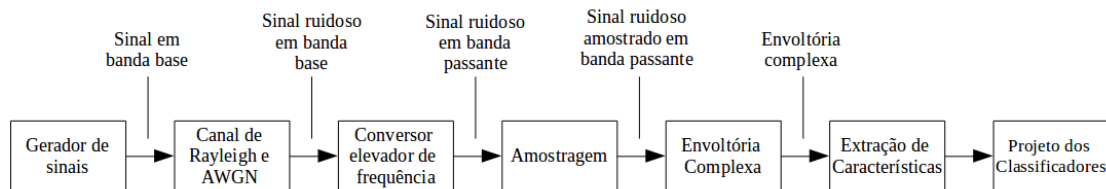


Figura 4.1: Visão geral da metodologia de classificação constituída pelas seguintes etapas: geração de sinais em banda base, acréscimo de ruído, deslocamento na frequência, amostragem, envoltória complexa, extração de características e classificação.

Neste trabalho, propõe-se uma investigação dos métodos de classificação automática de uma gama de modulações que são bastante utilizadas em sistemas de comunicação reais. Os sinais, para este trabalho, são gerados em banda base e acrescidos de ruído modelados pelos canais AWGN e Rayleigh também em banda base. O *upconverter* realiza a elevação de frequência dos sinais ruidosos em banda base, para então, serem amostrados e suas envoltórias complexas determinadas. As características destes sinais são extraídas e serão as entradas dos classificadores, conforme Figura 4.1. Então, propõe-se, quatro classificadores multiclasse, uma rede *perceptron* multicamadas com utilização do algoritmo de Levenberg-Marquardt para atualização dos pesos, os algoritmos *J48* e *Random Forest* de construção de árvores de decisão [44], [45] [37], [46] e um classificador probabilístico, *Naive Bayes*.

4.1 Geração de Sinais

Nesta sessão serão demonstrados os métodos utilizados para geração, em banda base, dos sinais modulados por modulações analógicas, digitais e por técnica de múltiplo acesso.

Modulações Analógicas

As modulações analógicas abordadas neste trabalho são as AM (do inglês *amplitude modulation*) e FM (do inglês *frequency modulation*) e seus sinais em banda base são representados por [17]:

$$x_{AM}(t) = A + k_a m(t) \quad (4.1)$$

$$x_{FM}(t) = \int_0^t 2\pi f_c + k_f m(t) dt \quad (4.2)$$

sendo $m(t) = \cos(2\pi f_m t)$, com f_m sendo a frequência modulante fixa; f_c é um valor constante igual a frequência da portadora; k_a e k_f são, respectivamente, a sensibilidade do receptor e a sensibilidade de frequência e A é a constante onde o sinal $m(t)$ ficará em torno.

Modulações Digitais

As modulações digitais abordadas são a BPSK (do inglês *binary phase shift keying*), QPSK (do inglês *quadrature phase shift keying*), 16-QAM (do inglês *16 quadrature amplitude modulation*), 64-QAM (do inglês *64 quadrature amplitude modulation*) e a GMSK (do inglês *gaussian minimum shift keying*), e são formados por símbolos de modulação. Os símbolos de modulação podem ser formados por diferentes quantidades de *bits* e representam os componentes em quadratura (Q) e fase (I) de um envelope complexo, $I + jQ$. As modulações BPSK, QPSK, 16-QAM e 64-QAM possuem respectivamente 2, 4, 16 e 64 pontos no mapa de constelações, já que, seus símbolos possuem 1, 2, 4 e 6 *bits* respectivamente. As componentes $I - Q$, por serem representadas de maneira discreta, não podem ser multiplicadas pelo elevador de frequência, já que é um sinal contínuo. É preciso, realizar a codificação de linha para transformar as componentes I e Q em trens de pulsos contínuos, conforme equação a seguir [17]:

$$x_{BPSK,QPSK,16QAM,64QAM}(t) = \begin{cases} I(t) = \sum_{n=1}^N I_n r(t - nT_s) \\ Q(t) = \sum_{n=1}^N Q_n r(t - nT_s) \end{cases} \quad (4.3)$$

sendo I_n e Q_n as sequências discretas de símbolos; $r(t - nT_s)$ é um pulso retangular com duração de tempo T .

A modulação GMSK é uma variante da modulação MSK (do inglês *minimum shift keying*) que codifica dois *bits* por símbolo. A modulação GMSK difere-se da modulação MSK devido a aplicação de um filtro gaussiano aos símbolos I_n e Q_n e a aplicação de um modulador FM em banda base [17], a expressão a seguir demonstra como é obtido o sinal GMSK em banda base:

$$x_{GMSK}(t) = \begin{cases} I(t) = 2\pi \left(\frac{k_m}{T_s} \right) \int_0^t \sum_{n=0}^N I_n p_{gmsk}(\tau - nT_s) d\tau \\ Q(t) = 2\pi \left(\frac{k_m}{T_s} \right) \int_0^t \sum_{n=0}^N Q_n p_{gmsk}(\tau - nT_s) d\tau \end{cases} \quad (4.4)$$

sendo p_{gmsk} a convolução entre o pulso de meia onda $\cos(\frac{\pi t}{2T})$ utilizado na modulação MSK e a resposta ao impulso do filtro gaussiano $K \sqrt{\frac{2\pi}{\ln(2)}} B e^{-2\frac{(B\pi)^2}{\ln(2)} t^2}$.

Técnica de Múltiplo Acesso

Para este trabalho foi explorado a técnica de multiplexação WCDMA (do inglês *wide-band code division multiple access*), onde as componentes I e Q são espalhadas na frequência pela multiplicação por um código denominado pseudo-ruído [47]. Nesta dissertação foi utilizado o modelo da telefonia de terceira geração *downlink*, onde vários usuários são considerados. O pseudo-ruído para este caso é definido conforme a equação abaixo [17]:

$$c^{(k)}(t) = \sum_{l=1}^L c_l^{(k)} r(t - lT_c) \quad (4.5)$$

sendo $\{k = 0, \dots, K - 1\}$, com K sendo o número de clientes ativos; $r(t - lT_c)$ é 1 para a região de existência $0 \leq t \leq T_c$ e 0 caso contrário; L é o comprimento do pseudo-ruído; T_c é a duração do *chip* e $c_l^{(k)}$ são os *chips* do pseudo-ruído específico do cliente k . Foi utilizado a modulação QPSK para mapear os símbolos de cada cliente, assim, o espalhamento espectral é dado segundo a seguinte equação:

$$x_{WCDMA}(t) = \begin{cases} I^{(k)}(t) = \sum_{n=1}^N I_n^{(k)} \sum_{l=1}^L c_l^{(k)} r(t - lT_c) \\ Q^{(k)}(t) = \sum_{n=1}^N Q_n^{(k)} \sum_{l=1}^L c_l^{(k)} r(t - lT_c) \end{cases} \quad (4.6)$$

sendo $I_n^{(k)}$ e $Q_n^{(k)}$ as seqüências de símbolos de cada cliente k .

4.2 Transmissão de Sinais

Como os sinais AM e FM são sinais contínuos em banda base a aplicação do *upconverter* é realizado de maneira direta. Entretanto, os sinais BPSK, QPSK, 16-QAM, 64-QAM, WCDMA e GMSK, utilizam pulsos retangulares e gaussiano (no caso do GMSK) para codificar os símbolos das componentes I - Q [17]. E para reduzir interferência e dispersão entre símbolos, aplica-se o filtro de formatação de pulso *raised* cosseno sobre estas componentes em banda base. O pulso *raised* cosseno é expresso conforme a seguinte equação [48]:

$$FP_{rc}(t) = \text{sinc} \left(\frac{t}{T} \right) \frac{\cos \left(\frac{\pi \alpha t}{T} \right)}{1 - \left(\frac{4\alpha^2 t^2}{T^2} \right)} \quad (4.7)$$

sendo $\text{sinc}(x) = \frac{\sin(x)}{x}$; T é a duração do pulso e α é fator de *roll-off*¹.

Adição de ruídos

Neste trabalho, como já foi mencionado, os canais de comunicação utilizados são os canais AWGN e Rayleigh. O canal AWGN modela efeitos aleatórios com igual intensidade em diferentes frequências, o que lhe dá uma densidade espectral de potência constante [49]. Já o canal Rayleigh modela o ruído que causa atenuações devido aos efeitos do desvanecimento multipercurso [50]. Estes ruídos em ambientes práticos são aplicados em sinais banda passante, porém, por questões práticas, neste trabalho, foram acrescentados, aos sinais modulados em banda base, ruídos também em banda base. Então, como dois modelos de canais são aleatórios e independentes entre si, foi adicionado primeiro o ruído AWGN e em seguida o desvanecimento causado pelo canal Rayleigh [17].

Upconverter

Na etapa de elevação de frequência, o sinal em banda base (sinal de baixa frequência) é deslocado para uma frequência mais elevada, tornando-se um sinal em banda passante (sinal em alta frequência). Para os sinais modulados em AM e FM em banda base o *upconverter* realiza a multiplicação destes sinais por um sinal portadora (sinal senoidal) de maior frequência. No caso do sinal AM em banda passante a senoide portadora, $\cos(2\pi f_c t)$ sendo f_c a frequência da portadora, é multiplicado pelo sinal $x_{AM}(t)$, equação (4.1). Já na modulação FM em banda passante, a frequência da portadora se altera conforme o sinal $x_{FM}(t)$, equação (4.2) [17].

Para as modulações digitais BPSK, QPSK, 16-QAM, 64-QAM, GMSK e WCDMA onde existe a presença das componentes *I-Q* o processo de elevação de frequência é realizado de maneira diferente. É utilizado, para estas modulações, um modulador em quadratura, composto por dois sinais portadoras de mesma frequência f_c defasados entre si de 90° , sendo cada sinal portadora multiplicado a uma das componentes *I-Q* dos sinais modulados em banda base, equação (4.3).

¹Fator de decaimento, que mostra a largura de banda que excede a banda do canal ideal

4.3 Amostragem e Determinação da Envoltória Complexa

Para o processo de extração de características é necessário que os sinais recebidos em banda passante sejam amostrados, já que os sinais na recepção são analógicos e contínuos no tempo [17]. O processo de amostragem é realizado segundo a seguinte equação:

$$x_a = \sum_{n=-\infty}^{\infty} x_{bp}(nT_s)\delta(t - nT_s) \quad (4.8)$$

sendo x_{bp} o sinal em banda passante na recepção; T_s o tempo de amostragem e $\delta(t)$ a função delta de Dirac².

Para facilitar a extrair, dos sinais, informações importantes de amplitude, frequência e fase instantâneas é obtido a envoltória complexa dos sinais amostrados em banda passante. A envoltória complexa é obtida aplicando a seguinte expressão:

$$\alpha(n) = z(n)e^{j2\pi f_c n} \quad (4.9)$$

sendo $z(n) = x(n) + j\mathcal{H}\{x(n)\}$, com $x(n)$ sendo o sinal amostrado em banda passante e $\mathcal{H}\{\cdot\}$ a transformada de Hilbert.

4.4 Extração de Características

O processo de extração de características é fundamental para a correta classificação das modulações por fornecer informações de amplitude, frequência e fase dos sinais modulados. Para este trabalho foram extraídas características espectrais, estatísticas, baseadas em transformada, baseadas em teste de distribuição e ciclo-estacionárias, totalizando 29 características. Na Tabela 4.1 temos a descrição das características extraídas e no Capítulo 2 temos as expressões matemáticas de cada uma das características.

²Função onde no ponto 0 tende ao infinito e vale 0 em todos os outros pontos

Características	
Sigla	Descrição
γ_{max}	Valor máximo da densidade espectral de potência da amplitude instantânea normalizada e centralizada
σ_{aa}	Desvio padrão da amplitude instantânea normalizada e centralizada
σ_{ap}	Desvio padrão da fase instantânea absoluta não linear centralizada
σ_{dp}	Desvio padrão da fase não linear direta centralizada
σ_{af}	Desvio padrão da frequência instantânea normalizada e centralizada
σ_f	Desvio padrão do valor absoluto da frequência instantânea normalizada e centralizada
γ_{maxf}	Valor máximo da PDS da frequência instantânea normalizada e centralizada
max_{dct}	Valor máximo da transformada discreta de cosseno
σ_{wht}	Desvio padrão da transformada de Walsh-Hadamard
σ_{dwt}	Desvio padrão da transformada de Wavelet discreta
Mel	Valor máximo dos coeficientes ceprais na frequência mel
C_{z_1, z_2}	Cumulantes de segunda ordem
C_{z_1, z_2, z_3}	Cumulantes de terceira ordem
C_{z_1, z_2, z_3, z_4}	Cumulantes de quarta ordem
$X_{m, n}$	Função de teste Gaussianidade
k_{SCD}	Curtose da densidade de correlação espectral
σ_{SCD}	Desvio padrão da densidade de correlação espectral
σ_{SCD}^2	Variância da densidade de correlação espectral
$CC - 3$	Cumulantes cíclicas de terceira ordem
$CC - 4$	Cumulantes cíclicas de quarta ordem

Tabela 4.1: Descrição e simbologia matemática das características extraídas para os sistemas de classificação automática de modulação propostos.

4.5 Projeto dos Classificadores

A proposta deste trabalho é uma investigação de uma gama de classificadores para o processo de classificação automática de modulações. Com isso em mente, esta etapa, projeto dos classificadores, é a mais importante dentro do processo metodológico. Nesta etapa vamos descrever o funcionamento dos classificadores abordados apresentando diagramas, equações e os parâmetros que podem ser configurados para melhorar o desempenho do classificador. Os classificadores escolhidos estão entre os mais populares dentro da área de pesquisa do *machine learning*, e são: rede neural artificial [8], árvores de decisão [51] com implementação de algoritmo *J48* e *Random Forest* e um classificador *Naive Bayes* [52].

4.5.1 Redes Neurais Artificiais

A RNA proposta é uma *perceptron* multicamadas supervisionada com aplicação de retropropagação (*backpropagation*) como algoritmo de aprendizado da rede, conforme Figura 4.2. As camadas da *perceptron* idealizada são totalmente conectadas, todos os nós da camada atual estão conectados com cada um dos nós da imediata camada seguinte [53]. A retropropagação é responsável por determinar as atualizações dos pesos da rede de maneira interativa, recursiva e eficiente até que

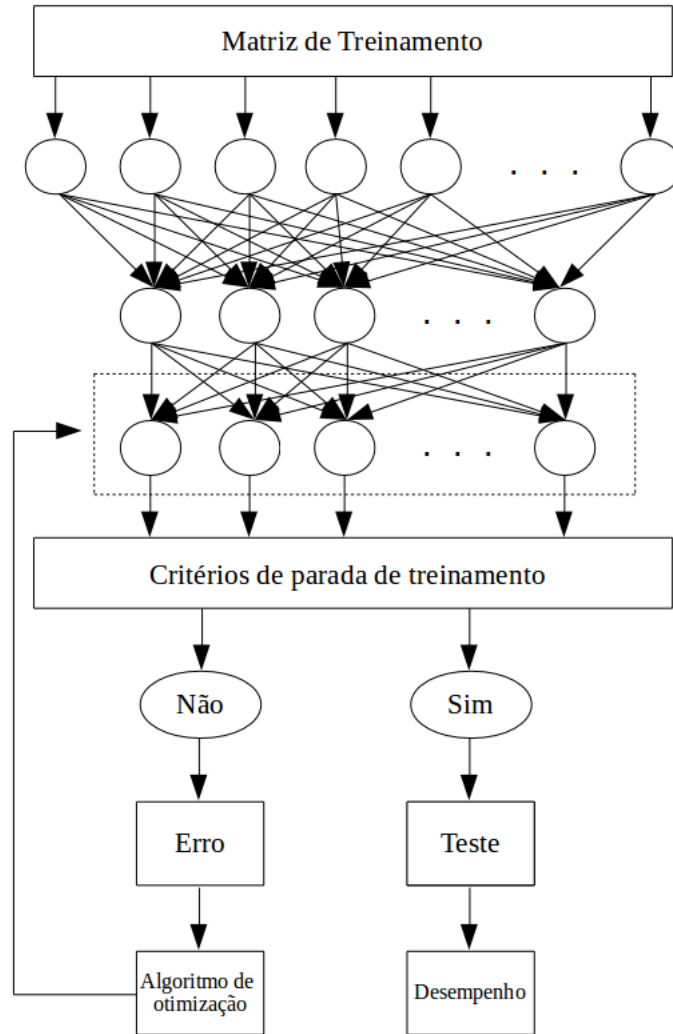


Figura 4.2: Diagrama da rede *perceptron* multicamadas totalmente conectada com *backpropagation*.

a rede seja capaz de realizar a tarefa pela qual foi treinada [38]. O algoritmo de atualização dos pesos utilizado na retropropagação, para metodologia proposta, é o Levenberg-Marquardt.

Na entrada da rede temos a matriz de características e a matriz de rótulos (*target matrix*), na saída temos a matriz de previsões. A matriz de características é composta por vetores, que representam cada classe, o qual a RNA pretende discriminar. A matriz de características é dividida em matrizes de treinamento, teste e validação [12]. A matriz de rótulos atribui valor aos vetores de cada classe e a matriz de previsões é a matriz resultante do teste da rede e quando comparada à matriz de rótulos é possível aferir a sua acurácia e precisão da RNA.

O processo de treinamento é a etapa mais crucial no projeto de classificação

por RNA, já que os pesos dos neurônios são definidos nesse estágio. A operação realizada em cada neurônio das camadas é descrita conforme a seguir:

$$x_j = \sum_{i=1}^N w_i c_i + b \quad (4.10)$$

sendo, na equação (4.10), N o número de características de entrada, w_i são os pesos atribuídos a cada característica de entrada, c_i são as características e b é a constante do neurônio. A função de ativação é um dos componentes do neurônio artificial e é a operação realizada ao final do somatório de cada neurônio.

Para o processo de treinamento foi projetada uma rede totalmente conectada de duas camadas, uma intermediária e outra de saída. A primeira camada (*camada intermediária*) da rede possui j neurônios e a operação realizada por esta camada é descrita na equação a seguir:

$$\mathbf{X} = \left(\left(\begin{pmatrix} w_1^1 & w_2^1 & \dots & w_N^1 \\ \vdots & \vdots & \ddots & \vdots \\ w_1^j & w_2^j & \dots & w_N^j \end{pmatrix} \cdot \begin{pmatrix} c_1^1 & c_1^2 & \dots & c_1^M \\ \vdots & \vdots & \ddots & \vdots \\ c_N^1 & c_N^2 & \dots & c_N^M \end{pmatrix} \right) + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_j \end{pmatrix} \right) \quad (4.11)$$

sendo, na equação (4.11), N o número de características extraídas e M é a quantidade de vetores de características. Os elementos do vetor coluna na equação (4.11) são as constantes de cada neurônio da rede. A função de ativação desta camada é a *tangente hiperbólica* descrita conforme a expressão a seguir:

$$a_j = \frac{2}{1 + e^{-2x_j}} - 1 \quad (4.12)$$

Na entrada da segunda camada (*camada de saída*) temos as saídas a_j da camada anterior. A equação (4.10) é aplicada novamente pelos j neurônios desta segunda camada. A representação matricial do cálculo executado nos neurônios da camada de saída da RNA proposta é demonstrado na equação a seguir:

$$\mathbf{X} = \left(\left(\begin{pmatrix} w_1^1 & w_2^1 & \dots & w_N^1 \\ \vdots & \vdots & \ddots & \vdots \\ w_1^j & w_2^j & \dots & w_N^j \end{pmatrix} \cdot \begin{pmatrix} a_1^1 & a_1^2 & \dots & a_1^M \\ \vdots & \vdots & \ddots & \vdots \\ a_N^1 & a_N^2 & \dots & a_N^M \end{pmatrix} \right) + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_j \end{pmatrix} \right) \quad (4.13)$$

$$y_j = \frac{e^{x_j}}{\sum e^{x_j}} \quad (4.14)$$

sendo, na equação (4.14), a função de ativação (função *softmax*) aplicada a cada um dos neurônios da camada de saída da RNA proposta. Onde temos em y_j as saídas da rede neural, que dizem a que classe pertence as características de entrada.

Ao final do processo de treinamento é avaliado se os critérios de parada de treinamento foram alcançados. Para que o treinamento encerre um dos seguintes critérios deve ser atendido: número de máximo de épocas³ é atingido; o limite de tempo de treinamento é alcançado; máximo valor do parâmetro de Marquardt (μ) [54] é atingido; o desempenho seja alcançado ou o gradiente atinge seu valor mínimo. Caso um dos critérios seja alcançado a rede é testada e a performance é calculada conforme a seguir:

$$D = \gamma \left(\frac{1}{W} \sum_{i=1}^W w_i^2 \right) + (1 - \gamma) \left(\frac{1}{K} \sum_{i=1}^K (t_i - y_i)^2 \right) \quad (4.15)$$

sendo, na equação (4.15), $\frac{1}{W} \sum_{i=1}^W w_i^2$ a média dos quadrados dos pesos, com W sendo o número de elementos da matriz de pesos. E $\frac{1}{K} \sum_{i=1}^K (t_i - y_i)^2$ é o erro quadrático médio, sendo K o número de elementos das matrizes de rótulos (*target matrix*) t_i , y_i as previsões e γ a taxa de regularização.

Caso nenhum dos critérios seja atingido calcula-se o erro e com isso aplica-se o algoritmo de otimização de Levenberg-Marquardt no processo de retropropagação para atualização dos pesos da rede. O erro é calculado da seguinte maneira:

$$E = \frac{1}{K} \sum_{i=1}^K w_i^e (t_i - y_i)^2 \quad (4.16)$$

³Uma época é definida como um ciclo onde todos os exemplos de treinamento passam, individualmente, através da rede ajustando os pesos depois de cada exemplo.

sendo, na equação (4.16), w_i^e a matriz de pesos individuais de cada erro quadrático calculada pela distância euclidiana entre os valores esperados (matriz de rótulos) e os preditos (matriz de predições) pela rede, $\sum_{i=1}^K \sqrt{(t_i - y_i)^2}$.

$$\Delta x = \frac{-J^T E}{J^T J + \mu I} \quad (4.17)$$

$$J = \begin{pmatrix} \frac{\partial E_1}{\partial w_1^e} & \frac{\partial E_1}{\partial w_2^e} & \cdots & \frac{\partial E_1}{\partial w_K^e} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial E_K}{\partial w_1^e} & \frac{\partial E_K}{\partial w_2^e} & \cdots & \frac{\partial E_K}{\partial w_K^e} \end{pmatrix} \quad (4.18)$$

A rede é atualizada conforme a equação (4.17), sendo J a matriz jacobiana que contém as primeiras derivadas dos erros da rede em relação aos pesos, conforme equação (4.18), μ é o parâmetro de Marquardt e I é a matriz identidade. Todo o processo de treinamento descrito, devido a retropropagação, acontece de maneira recursiva até que um dos critérios de parada seja atingido. Além dos critérios de parada, existem outros parâmetros que influenciam no desempenho da rede e são configurados de acordo com os vetores de entrada para que a RNA classifique-os com mais eficiência. Dentre eles estão: parâmetro de regularização γ ; parâmetro de Marquardt μ inicial e fator de decrescimento e crescimento de μ .

A regularização Bayesiana é uma técnica de treinamento que acrescenta um elemento de regularização à função objetivo, convergindo parâmetros pouco relevantes para zero e com isso reduzindo a quantidade de parâmetros efetivos utilizados no processo [55]. O algoritmo de Levenberg-Marquardt foi escolhido como método de otimização por resolver o problema de instabilidade do método de Gauss-Newton [56], retornando o mínimo da função erro independente do estado inicial. Definido os métodos de treinamento e otimização é preciso configurar os parâmetros de regularização e de Marquardt. O parâmetro de regularização γ foi escolhido de maneira a deixar a RNA mais eficiente e utilizar somente a quantidade necessária de elementos da matriz de pesos. É configurado um valor inicial para o parâmetro de Marquardt μ e um valor limite, assim como, também, fatores de decrescimento e crescimento para o valor de μ se adequar para retornar melhor resultado possível.

4.5.2 J48

O *J48* é um algoritmo de classificação supervisionado baseado em árvore de decisão e surgiu da necessidade de recodificar o algoritmo *C4.5*, que foi desenvolvido inicialmente em C, para a linguagem JAVA [37]. A árvore de decisão é constituída por nós que representam os atributos, por ligações (*ramos*) provenientes dos nós e os nós folha, que representam as classes. O modelo *J48* adota o método de "dividir para conquistar", com base no conceito de razão de ganho de informação. A ideia de razão de ganho de informação baseia-se no quanto uma característica é informativa por meio da redução da entropia, com isso, espera-se uma redução na entropia caso um determinado nó seja selecionado para realizar a divisão dos dados [57].

Algorithm 1 *J48(D)*

```
1: procedure J48(D) ▷ Input: um conjunto de dados D
2:   Árvore = {} ;
3:   if D é puro OU existe outro critério de parada then
4:     encerrar;
5:   for each atributo a ∈ D do
6:     calcular a razão do ganho de informação;
7:   end
8:   amelhor = Melhor atributo de acordo com o cálculo da razão do ganho de
   informação;
9:   Árvore = Cria um nó baseado no amelhor;
10:  Dv = Divide o subconjunto de D baseado no amelhor;
11:  for each Dv do
12:    Árvore = J48(Dv);
13:  Fixe a Árvorev no galho obtido no passo anterior;
14:  end
15:  return [Árvore]
```

O algoritmo *J48*, no tratamento de características, ordena todos os valores do conjunto de treinamento em ordem crescente com proposito de escolher o atributo que favorecerá o valor da entropia [58]. As seguintes etapas, descritas em [57] e [59], são realizadas para determinar a característica atribuída ao nó raiz:

- Cálculo da entropia:

$$E(D) = - \sum_{i=1}^k p_i \log_2(p_i) \quad (4.19)$$

sendo p_i a probabilidade relativa de cada classe i pertencente ao conjunto de características D e k é a quantidade de classes.

- Cálculo da entropia da informação esperada, $E_x(D)$, para uma característica x pertencente ao conjunto D :

$$E_x(D) = - \sum_{i=1}^n \left(\frac{|D_i|}{|D|} E(D_i) \right) \quad (4.20)$$

sendo n o número de nós-filhos, ou seja, os valores possíveis que a característica x pode assumir.

- O ganho de informação é dado por:

$$G(x) = E(D) - E_x(D) \quad (4.21)$$

Devido a possíveis grandes variações do critério de ganho de informação, o que pode resultar em erros, é recomendado o uso do critério de razão de ganho de informação para escolha do nó raiz [57]. No critério de razão de ganho de informação o denominador realiza uma normalização do conjunto de amostras de características que apresentam grandes variações, contornando o problema da utilização do ganho de informação.

- Razão do ganho de informação:

$$RG = \frac{G(x)}{\sum_{i=1}^n \left(\frac{|D_i|}{|D|} \log_2 \left(\frac{|D_i|}{|D|} \right) \right)} \quad (4.22)$$

Tendo em perspectiva o processo acima descrito podemos identificar as próximas etapas do processo de construção da árvore: o crescimento, constituído pelo treinamento e teste; e a poda. O crescimento da árvore é ditado pelos dados de entrada divididos em treinamento, aprendizado da estrutura, e teste, que avalia o quão precisa é a predição da árvore. O pseudocódigo, descrito no Algoritmo 1, descreve o processo de construção e poda da árvore. Para o processo de poda é preciso definir dois parâmetros, o fator de confiança e número de *folds* para reduzir erro. Além destes dois parâmetros é preciso indicar o número mínimo de instâncias por folha e a quantidade de sementes para embaralhamento aleatório de dados. A Tabela 4.2 elenca e descreve todos os possíveis e aplicáveis parâmetros do algoritmo *J48*.

Parâmetros	
Sigla	Descrição
-U	Usar árvore não podada
-O	Não colapse a árvore
-C	Limite de confiança para poda
-M	Mínimo número de instâncias por folha
-R	Usar erro de poda reduzido
-N	Número de <i>folds</i> para reduzir erro de poda
-B	Usar somente divisão binária
-S	Não realizar crescimento de subárvores
-L	Não limpar depois que árvore estiver construída
-A	Suavização por Laplace para probabilidades previstas
-J	Não usar correção MDL para ganho de informação em atributos numéricos
-Q	Número de sementes para embaralhamento aleatório de dados

Tabela 4.2: Descrição de todos os parâmetros aplicáveis ao algoritmo *J48*.

4.5.3 Random Forest

A *Random Forest* é um algoritmo de classificação constituído por uma quantidade de árvores de decisão, sendo que cada árvore cresce utilizando alguma forma de aleatorização [60]. Foi proposto inicialmente em 2001 por Breiman [61], baseando-se no algoritmo *CART* (sem poda), sendo categorizado como classificador se a resposta for discreta e trabalha como algoritmo de regressão se sua resposta for contínua. Este algoritmo foi escolhido por apresentar diversas vantagens descritas em [44], como: alto desempenho com grande quantidade de base de dados, manipula altíssimo número de variáveis de entrada sem que sejam excluídas, fornece uma estimativa de quais variáveis são relevantes na classificação, gera uma estimativa interna imparcial do erro de generalização, é relativamente robusto a presença de ruído e é computacionalmente mais leve que outros algoritmos de árvore de decisão.

O algoritmo de classificação *Random Forest* é aplicado em duas etapas. Na primeira fase o algoritmo extrai um subconjunto de dados do conjunto de características original utilizando *bootstrap* como método de redistribuição de características e cria uma árvore de decisão para cada subconjunto [45], conforme Figura 4.3. Depois, o algoritmo discrimina as árvores de decisão com base em uma votação simples, com o maior voto da classificação como resposta final da predição [45], conforme o Algoritmo 2 apresentado em [62]. O *bootstrap* é um método de reamostragem do conjunto original dos dados de treinamento. Ele gera um novo conjunto de dados de treinamento com mesma dimensão do conjunto original, a partir do sorteio aleatório

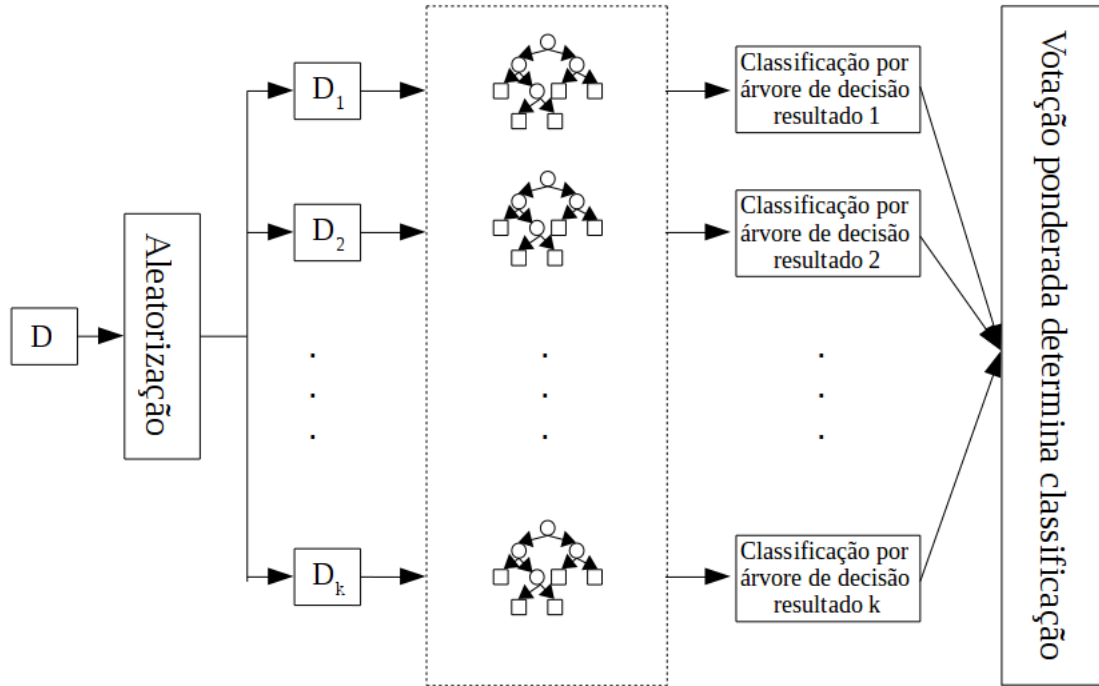


Figura 4.3: Estrutura de votação ponderada da *Random Forest*.

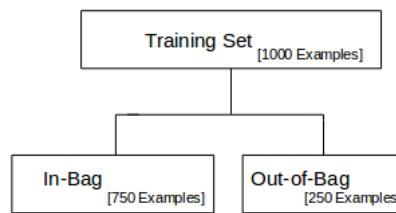


Figura 4.4: Divisão dos dados de treinamento por *bootstrap*.

com reposição de seus elementos, conforme mostra a Figura 4.4.

A votação majoritária demonstrada no Algoritmo 2 pode ser realizada por três critérios com o intuito de minimizar o nível de impureza do nó [63]. A métrica Gdi descrita na equação (4.23) é o *default* do algoritmo *CART* e retorna zero se o nó for puro (nó com uma classe somente), senão, retorna um valor positivo.

$$Gdi = 1 - \sum_{i=1}^k P^2(i) \quad (4.23)$$

sendo P a proporção de observação que a classe i chega no nó que discrimina a classe i .

A métrica Gdi tem deficiência quando se trata de problemas multiclasse, para contornar isso é aplicado as métricas *Toving* e *Entropia*. A regra de *Toving*,

Algorithm 2 *RandomForest*(D)

1: **procedure** *RandomForest*(D) ▷ **Input:** um conjunto de dados D
2: **for** $D_i, i = 1, \dots, D$
3: Extrair um *bootstrap* S de dimensão N do conjunto de treinamento
4: Crescer uma árvore *RandomForest* não podada, T_d , usando
5: *bootstrapped data*, até o tamanho mínimo do nó terminal,
6: n_{min} , ser obtido recursivamente pelo sub-algoritmo a seguir:

- Selecionar aleatoriamente p variáveis do conjunto total de P variáveis.
- Selecionar a melhor variável/ponto-de-divisão entre as p variáveis
- Dividir nó em dois nós filhos.

7: **end for**
8: Conjunto de saída de árvores $\{T_d\}_1^D$
9: Prever novas observações, ou observações *out-of-bag*:

- Para regressão: $f_{rf}^D(x) = \frac{1}{D} \sum_{d=1}^D T_D(x)$
- Para classificação: Sendo $C_d(x)$ a previsão de classe d^{th} árvore *Random Forest* $\rightarrow C_{rf}^D(x) = \text{majority vote } \{C_d(x)\}_1^D$

equação (4.24), funciona para árvores binárias onde a quantidade de nós é dois e a partição T é dividida em dois nós, t_G e t_D .

$$Twoing = p(t_G)p(t_D) \left(\sum_{i=1}^k |p_i(t_G) - p_i(t_D)| \right) \quad (4.24)$$

sendo $p(t_G)$, depois da divisão das partição, o número de elementos da classe i no nó filho esquerdo e $p(t_D)$, também depois da divisão das partição, o número de elementos da classe i no nó filho direito. Entretanto, a métrica *Twoing* não calcula o nível de pureza do nó, a *Entropia*, equação (4.19), é utilizada para determinar o nível de pureza do nó.

Existem ainda alguns parâmetros que devem ser configurados para a execução do algoritmo *Random Forest*, como: o número de atributos para investigação aleatória; número mínimo de instâncias por folha; definição da proporção de variância da classe numérica; número de sementes para o gerador aleatório e a profundidade máxima da árvore, além de outros possíveis parâmetros ajustáveis conforme mostrado na Tabela 4.3.

Parâmetros	
Sigla	Descrição
-P	Tamanho de cada bagagem, como % do tamanho do conjunto de treinamento
-O	Calcular o erro fora da bagagem
-stop-out-of-bag-predictions	Armazenamento de previsões fora da bagagem no objeto de avaliação interna
-output-out-of-bag-complexity-statistics	Gerar estatísticas baseadas na complexidade quando a avaliação do <i>out-of-bag</i> é realizada
-Print	Mostrar os classificadores individuais na saída
-attribute-importance	Computar e mostre importância de atributos
-I	Número de interações
-num-slots	Número de <i>slots</i> de execução
-K	Número de atributos para investigação aleatória
-M	Número mínimo de instâncias por folha
-V	Valor mínimo de variância proporcional de classes
-S	Sementes para gerador de números aleatórios
-depth	Máxima profundidade da árvore
-N	Número de <i> folds</i> para <i>backfitting</i>
-U	Permitir instâncias não classificadas
-B	Quebrar ramos aleatoriamente quando vários atributos são semelhantes
-output-debug-info	Adicionar informação em caso de debug para o <i>console</i>
-do-not-check-capabilities	Não checar classificador antes de estar construído
-num-decimal-places	Número de espaços decimais para a saída de números no modelo
-batch-size	Número desejado de tamanho de <i>batch</i> para previsão de <i>batch</i>

Tabela 4.3: Descrição de todos os parâmetros aplicáveis ao algoritmo *Random Forest*.

4.5.4 Naive Bayes

O classificador *Naive Bayes* é algoritmo de classificação por probabilística, baseado no teorema de *Bayes*, chamado de *naive* por pressupor uma grande independência entre as características [35]. O teorema de *Bayes* é dado pela equação a seguir:

$$P(C = c_k | X = x) = \frac{P(X = x | C = c_k)P(C = c_k)}{P(x)} \quad (4.25)$$

sendo, na equação (4.25), C uma variável aleatória com valor de uma das classes c_k e X é um vetor de variáveis aleatórias cujos valores são vetores de características $x = (x_1, x_2, \dots, x_d)$.

Assim, $P(C = c_k | X = x)$ é a probabilidade a condicional de que C pertence a classe c_k sabendo da ocorrência do vetor de características x . Sabendo que os valores c_k e x são valores das variáveis aleatórias C e X é possível simplificar a equação (4.25) [52], assim, temos:

$$P(c_k | x) = \frac{P(c_k)P(x|c_k)}{P(x)} \quad (4.26)$$

Quando se sabe o valor da probabilidade $P(c_k|x)$, podemos assumir que dado um vetor de características x pertence a classe c_k quando $P(c_k|x)$ é alta. Porém não se sabe o valor de $P(c_k|x)$, então a regra de *Bayes* propõe a estimativa de $P(x|c_k)$, $P(c_k)$ e $P(x)$ individualmente, para então realizar a uma combinação para

obter $P(c_k|x)$ [52]. Assim, a distribuição x condicional a c_k é decomposta e escrita conforme a seguinte expressão:

$$P(x|c_k) = \prod_{j=1}^d P(x_j|c_k) \quad (4.27)$$

Com essa suposição, equação (4.27), temos, que dada uma característica c_k , existe uma independência estatística entre a ocorrência de um valor particular x_j e um outro valor $x_{j'}$. Substituindo a equação (4.27) na equação (4.26), temos:

$$P(c_k|x) = \frac{P(c_k) \prod_{j=1}^d P(x_j|c_k)}{P(x)} \quad (4.28)$$

Então, se o objetivo da classificação é de minimizar o número de erros, um vetor de características x é atribuído a classe c_k quando a probabilidade condicional $P(c_k|x)$ é a maior. Normalmente, o denominador da equação (4.28) não é computado para erro mínimo de classificação, pois é igual para todas as classes c_k . Assim, a classificação correta é dada pelo mais alto valor da operação $P(c_k) \prod_{j=1}^d P(x_j|c_k)$.

Capítulo 5

Experimentos e Resultados

Neste capítulo demonstramos as etapas experimentais e as diferentes abordagens propostas no Capítulo 4 de modo a verificar a metodologia proposta. O capítulo está organizado da seguinte maneira: é realizado, na primeira sessão, o *setup* dos experimentos; na segunda sessão é informado de maneira detalhada como a base de dados é confeccionada; em seguida é realizado uma descrição das métricas utilizadas para avaliar cada método proposto na metodologia e por último uma sessão mostrando resultados e análises de cada método proposto como também uma análise comparativa dos mesmos.

5.1 *Setup* dos Experimentos

Os experimentos foram implementados na linguagem de programação MATLAB, nas versões R2017b e R2014b, e no *software* de *machine learning* *Weka*, em ambiente Linux, Ubuntu versão 16.04.4 LTS, processador Intel(R)i7-3770 CPU @ 3,4 GHz e memória RAM 16 GB. Para geração da base de dados foi utilizado o MATLAB/Simulink versão R2014b e os métodos de classificação foram implementados na versão R2017b do MATLAB, utilizando a *toolbox neural network*, e no *Weka*.

5.2 Base de Dados

A base de dados aplicada a um classificador é de fundamental importância para a discriminação correta dos sinais modulados. Existe uma carência na litera-

tura de bases de dados que retratam, ou se aproximam, sistemas de comunicação reais. Com intuito de simular situações reais de comunicação, a base de dados, abrange as etapas de formação de pulso por *raised* cosseno, transmissão em banda passante com *upconverter* e adição de ruídos pelos canais Rayleigh e AWGN. Foi desenvolvido um algoritmo de geração da base de dados podendo-se alterar a quantidade de sinais gerais e seus parâmetros, como: número de amostras do sinal em banda passante, frequência da portadora, caminhos de propagação Rayleigh, níveis de ruído do canal AWGN, número de sinais gerados para cada tipo de modulação e número de características nos vetores de características. A base utilizada contém 29 características para os sinais modulados.

As 29 características são extraídas a partir da amostragem dos sinais modulados em banda passante com 1000 amostras por sinal modulado. A frequência da portadora no *upconverter* é de 2,412 GHz. Para todos os tipos de modulação, há adição de ruído pelo canal de Rayleigh, considerando, além do caminho de propagação principal, quatro caminhos com atraso fixo: $\tau = 4 \times 10^{-9}$, $\tau = 12,5 \times 10^{-9}$, $\tau = 25 \times 10^{-9}$ e $\tau = 50 \times 10^{-9}$. No canal AWGN, considera cinco níveis de SNR: 0, 5, 10, 20 e 30dB. Para cada relação sinal-ruído foram gerados 9420 sinais modulados, totalizando 47100 vetores de características para cada uma das classes de modulação, conforme a Tabela 5.1. Aleatórios 75% dos 47100 vetores de características foram destinados para o processo de treinamento dos classificadores propostos, restando, 25% para os testes dos classificadores.

5.3 Métricas de Desempenho

Considerou-se para a avaliação do desempenho dos classificadores propostos as seguintes métricas: taxa de verdadeiros positivos (*TP*)(também conhecido como *recall*), taxa de falsos positivos (*FP*), precisão (*precision*) e acurácia (*accuracy*). Essas métricas são expostas nas matrizes de confusão de treinamento e teste. Essas métricas são definidas conforme as seguintes equações:

Classes	SNR (AWGN)	Quantidade de amostras
AM	0	9420
	5	9420
	10	9420
	20	9420
	30	9420
FM	0	9420
	5	9420
	10	9420
	20	9420
	30	9420
BPSK	0	9420
	5	9420
	10	9420
	20	9420
	30	9420
QPSK	0	9420
	5	9420
	10	9420
	20	9420
	30	9420
16-QAM	0	9420
	5	9420
	10	9420
	20	9420
	30	9420
64-QAM	0	9420
	5	9420
	10	9420
	20	9420
	30	9420
GMSK	0	9420
	5	9420
	10	9420
	20	9420
	30	9420
WCDMA	0	9420
	5	9420
	10	9420
	20	9420
	30	9420

Tabela 5.1: Quantidade de vetores de características por classe de modulação e nível de SNR.

- Taxa de verdadeiros positivos (*recall*):

$$TP(\%) = 100 \frac{\text{Total de exemplos de teste preditos como pertencentes à classe } k}{\text{Total de exemplos de teste pertencentes à classe } k} \quad (5.1)$$

- Taxa de falsos positivos:

$$FP(\%) = 100 \frac{\text{Total de exemplos de teste preditos como pertencentes à classe } k}{\text{Total de exemplos de teste que não pertencem à classe } k} \quad (5.2)$$

- Precisão:

$$precision(\%) = \frac{TP}{TP + FP} \quad (5.3)$$

- Acurácia:

$$accuracy(\%) = \frac{TP + TN}{TP + TN + FP + FN} \quad (5.4)$$

sendo TN a taxa de verdadeiros negativos e FN a taxa de falsos negativos.

5.4 Experimentos

Nesta seção, serão apresentados os parâmetros apontados nas metodologias, como também, os resultados e as análises para cada método, finalizando com uma análise comparativa entre os classificadores abordados.

5.4.1 Rede Neural Artificial

Para o método de treinamento da RNA proposta alguns parâmetros foram configurados conforme o *default* do MatLab. O número de camadas e a quantidade de neurônios por camada foram definidos de maneira a deixar a rede o mais eficiente possível. O tempo de treinamento da rede é proporcional a quantidade de interações, assim, quanto mais neurônios mais tempo a rede iria requerer para realizar o treinamento. As funções de ativação das camadas são a *tansig* (tangente hiperbólica), para a camada intermediária, e a função *softmax*, para a camada de saída. O algoritmo de otimização, utilizado na retropropagação, que atualiza os pesos da rede é o de Levenberg-Marquardt e seus parâmetros são mostrados, também, na Tabela 5.2 conforme os *default* do MatLab.

Parâmetros	Valores
Número de características	29
Número de camadas internas	1
Número de neurônios da camada interna	29
Número de neurônios da camada de saída	7
Função de ativação da camada interna	<i>Tansig</i>
Função de ativação da camada de saída	<i>Softmax</i>
Número máximo de épocas	1000
Desempenho ideal	0
Método de treinamento	Retropropagação por Regularização Bayesiana
Parâmetro de Regularização (γ)	0
Função de desempenho	MSE
Parâmetro de Marquardt inicial (μ)	$5e - 3$
Fator de decrescimento do parâmetro de Marquardt (μ)	$1e - 1$
Fator de crescimento do parâmetro de Marquardt (μ)	10
Máximo valor do parâmetro de Marquardt (μ)	$1e + 10$
Máximo valor de falhas de validação	∞
Valor mínimo do gradiente de desempenho	$1e - 7$
Limite de tempo de treinamento	∞

Tabela 5.2: Parâmetros da RNA proposta.

Resultados

Os resultados apresentados estão em forma de matrizes de confusão de treinamento e teste, Tabela 5.3 e Tabela 5.4 respectivamente, e a Tabela 5.5 que demonstra a variação da precisão de cada modulação conforme a alteração da SNR. As matrizes de confusão mostram o número de sinais treinados e classificados corretamente e incorretamente, como também, as porcentagens de verdadeiros positivos, precisão e acurácia. A tabela apresenta a influência do ruído na precisão dos sinais de cada modulação. O gráfico, Figura 5.3, destaca a influência do ruído na acurácia total da RNA para cada SNR.

	AM	FM	BPSK	QPSK	16-QAM	64-QAM	GMSK	WCDMA	
AM	35411	0	0	0	0	0	0	0	100(%)
FM	0	353570	2	0	0	0	0	0	100(%)
BPSK	0	0	32888	367	284	209	509	0	96(%)
QPSK	0	0	817	30214	3416	3149	96	0	80,2(%)
16-QAM	0	0	293	1735	20431	7882	84	0	67,2(%)
64-QAM	0	0	371	2942	11110	23995	149	0	62,2(%)
GMSK	0	0	848	111	96	105	34429	0	96,7(%)
WCDMA	0	0	0	1	0	0	0	35295	100(%)
	100(%)	100(%)	93,4(%)	85,4(%)	67,8(%)	67,9(%)	97,6(%)	100(%)	87,8(%)

Tabela 5.3: Matriz de confusão de treinamento da classificação realizada pela RNA proposta.

Na Tabela 5.3 temos a matriz de confusão do processo de treinamento da rede. Notamos que na diagonal principal temos o número de modulações treinadas corretamente e nas redondezas da diagonal principal temos o número de sinais trei-

nados de maneira incorreta. A última linha representa os valores em porcentagens dos verdadeiros positivos, a última coluna temos os valores em porcentagem da precisão e a acurácia total de treinamento da rede é demonstrada no último elemento inferior direito da matriz de confusão. A Tabela 5.4 é a matriz de confusão do processo de teste da rede, e a distribuição dos seus elementos da matriz é semelhante a da Tabela 5.3, mostrando verdadeiros positivos, precisão e acurácia para o teste da RNA.

	AM	FM	BPSK	QPSK	16-QAM	64-QAM	GMSK	WCDMA	
AM	11689	0	1	0	0	0	0	0	100(%)
FM	0	11743	0	0	0	0	0	0	100(%)
BPSK	0	0	10928	140	93	92	186	0	95,5(%)
QPSK	0	0	291	9965	1125	1053	32	0	79,9(%)
16-QAM	0	0	127	650	6780	2666	31	0	66,1(%)
64-QAM	0	0	163	939	3730	7910	43	1	61,9(%)
GMSK	0	0	371	34	35	39	11541	0	96(%)
WCDMA	0	0	0	2	0	0	0	11806	100(%)
	100(%)	100(%)	92(%)	85(%)	67,6(%)	67,3(%)	97,5(%)	100(%)	87,4(%)

Tabela 5.4: Matriz de confusão de teste da classificação realizada pela RNA proposta.

	0(dB)	5(dB)	10(dB)	20(dB)	30(dB)
AM	100(%)	100(%)	100(%)	100(%)	100(%)
FM	100(%)	100(%)	100(%)	100(%)	100(%)
BPSK	79,7(%)	96,6(%)	99,7(%)	99,5(%)	98,9(%)
QPSK	50,4(%)	71,7(%)	81,9(%)	97,3(%)	98,2(%)
16-QAM	36,9(%)	52,7(%)	60,3(%)	78,3(%)	80,1(%)
64-QAM	40,1(%)	49(%)	58,4(%)	75,1(%)	87,3(%)
GMSK	82,4(%)	96,5(%)	99,7(%)	100(%)	100(%)
WCDMA	100(%)	100(%)	100(%)	99,8(%)	100(%)

Tabela 5.5: Precisão por SNR das modulações classificadas pela RNA proposta.

Na Tabela 5.5 temos a taxa em porcentagem da precisão de cada modulação por SNR. Cada linha da tabela mostra a variação da precisão de cada classe de modulação conforme a alteração da SNR, assim, pode-se avaliar a influência do ruído na classificação correta das modulações. Também, nesta tabela, fica mais evidente em quais modulações o ruído tem mais impacto. Para adquirir estas informações foi necessário testar a RNA, já anteriormente treinada, com uma nova base de dados de teste contendo 2355 sinais de cada SNR para cada modulação, equivalente a 25% da quantidade de sinais por SNR da base de dados.

Na Figura 5.3, o gráfico mostra a acurácia total da RNA testada com sinais em todas as relações sinal-ruído. No eixo das abscissas (eixo horizontal do plano

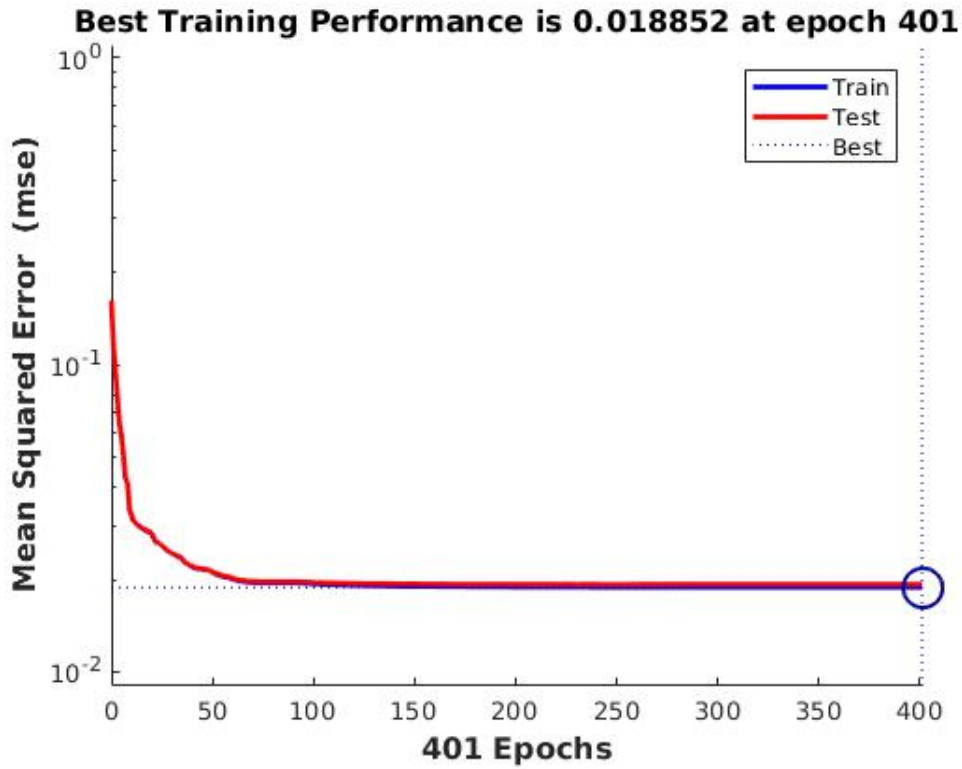


Figura 5.1: Gráfico do desempenho MSE do processo de treinamento e teste da RNA dado a época.

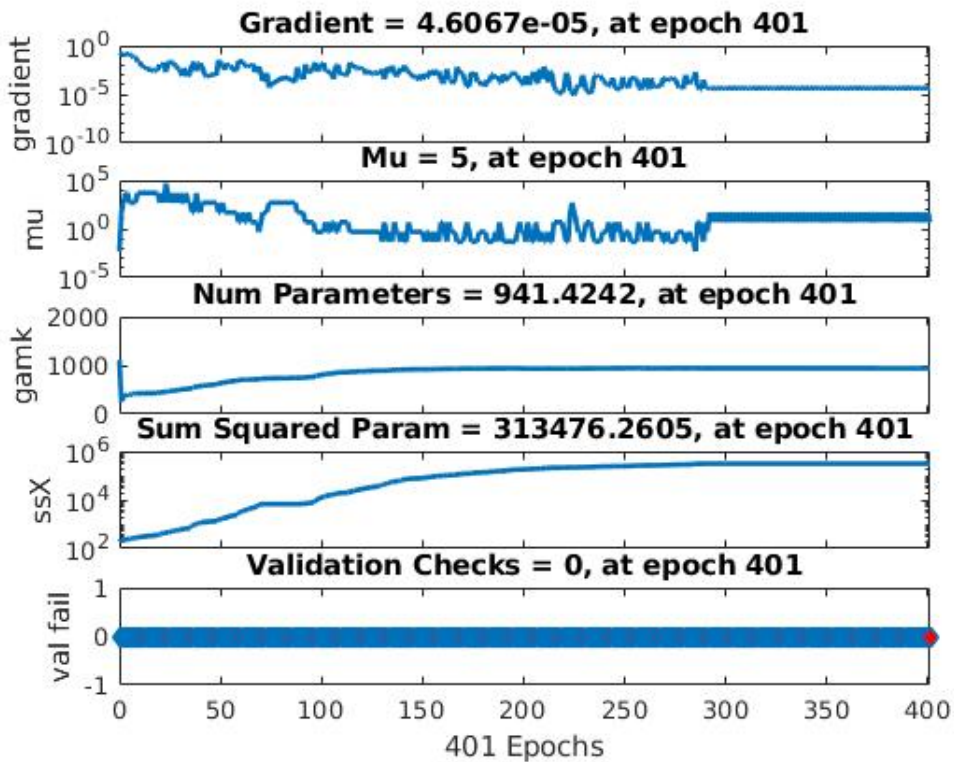


Figura 5.2: Gráfico do comportamento dos parâmetros da RNA no processo de treinamento dada a época.

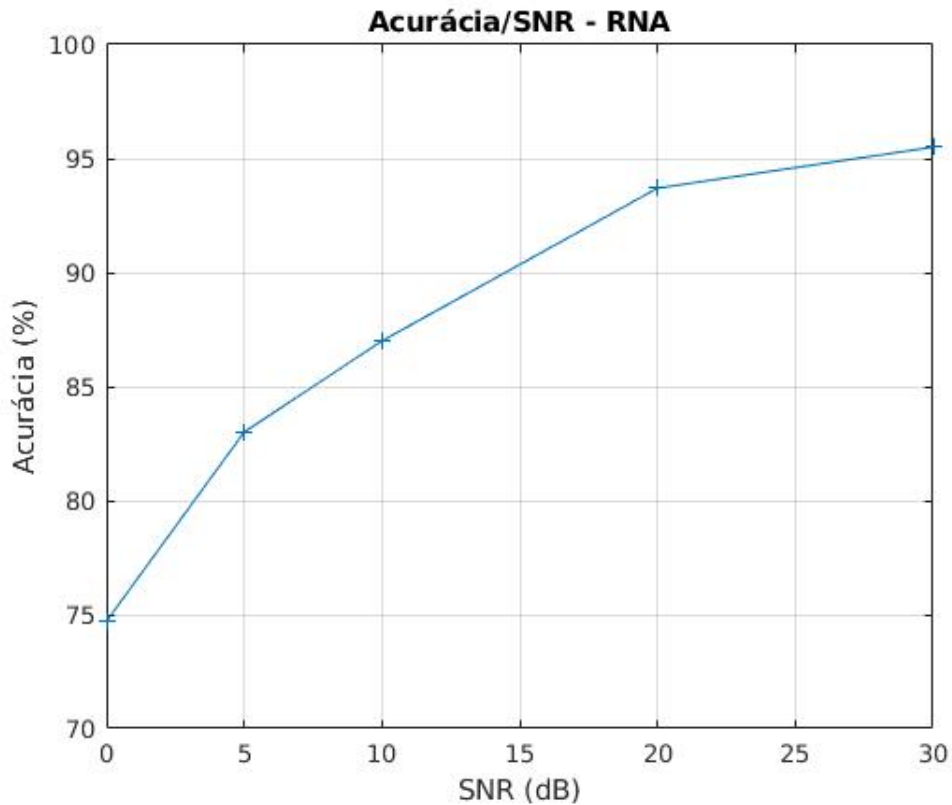


Figura 5.3: Gráfico da acurácia em (%) pela variação da SNR (dB) para a RNA.

cartesiano) temos os valores de SNR e no eixo das coordenadas (eixo vertical do plano cartesiano) temos os valores em porcentagem da acurácia total correspondentes a cada SNR. Este gráfico dá uma visão mais genérica de como a RNA reage à influência do ruído em seus diferentes níveis.

Análise

A RNA proposta levou 27h06min43seg para realizar o processo de treinamento. Isso devido ao processo de treinamento ser interrompido na 401ª época. Porém, essa interrupção não compromete o nível de aprendizado da rede. Notamos na Figura 5.1 que aproximadamente a partir da época 120 o desempenho, MSE, da rede se mantém constante. Outro indício, fica evidente, na Figura 5.2 onde aproximadamente após a 290ª época os parâmetros se comportam de maneira constante. Por exemplo, o parâmetro de Marquardt (μ), tem um comportamento semelhante a uma reta em determinado momento. Tendo isso em vista, podemos reparar que a continuidade do treinamento da rede era desnecessário. Outra informação relevante

sobre a rede é que o número de parâmetros efetivamente utilizados pela rede diminuiu dos iniciais 1110 para aproximadamente 941, conforme gráfico da Figura 5.2, isso prova a efetividade do método de regularização Bayesiana.

Notamos, conforme a Tabela 5.3, que no processo de treinamento a RNA teve acurácia de 87,8 %. No processo de teste temos, conforme Tabela 5.4, uma acurácia de 87,4 % na correta identificação das modulações, o que expressa a boa generalização do classificador. Repara-se que em todas as modulações temos taxas de verdadeiros positivos acima de 90 %, com exceção das modulações digitais QPSK, 16-QAM e 64-QAM que tiveram taxas de verdadeiros positivos de 79,9 %, 66,1 % e 61,9 % respectivamente. Na Tabela 5.5 temos a taxa de precisão das modulações conforme a variação da SNR, notamos então, que as modulações AM, FM e WCDMA obtiveram taxas de precisão superiores a 99 % para todas as SNRs abordadas, o que retrata, para o processo de classificação, que o nível de ruído não é uma influência para a correta identificação destas modulações. Ainda na Tabela 5.5, notamos que as modulações BPSK e GMSK obtiveram respostas relativamente inferiores para a SNR igual a 0, porém, obtiveram taxas de precisão acima de 96 % para SNRs acima de 0. Em contra partida, notamos que as modulações QPSK, 16-QAM e 64-QAM são muito susceptíveis a presença de ruído, conforme os dados mostrados na Tabela 5.5. O gráfico da Figura 5.3 mostra o nível de acurácia do classificador para cada SNR, notamos então, que a RNA tende a ter melhor acurácia conforme o aumento da SNR.

5.4.2 J48

Utilizou-se o *software Weka* no experimento envolvendo o algoritmo baseado em árvore de decisão *J48*. Todos os possíveis parâmetros estão descritos na Tabela 4.2 do Capítulo 4. Nesta etapa configuramos alguns daqueles parâmetros para utilização do classificador. A árvore de decisão baseada no algoritmo *J48* é treinada com poda, conforme a metodologia proposta, então, os parâmetros aplicados, para este experimento, são: limiar de poda, com valor de 0,25; o número mínimo de instâncias por nó folha é de 2; o número de *folds* é 3 e a quantidade de sementes é 1. Estes parâmetros são o *default* do *Weka* e apresentados na Tabela 5.6.

Resultados

Os resultados apresentados estão em forma de matrizes de confusão de treinamento e teste, Tabela 5.7 e Tabela 5.8 respectivamente, e a Tabela 5.9 que demonstra a variação da precisão de cada modulação conforme a alteração da SNR. As matrizes de confusão mostram o número de sinais treinados e classificados corretamente e incorretamente, como também, as porcentagens de verdadeiros positivos, precisão e acurácia. A tabela apresenta a influência do ruído na precisão dos sinais de cada modulação. O gráfico, Figura 5.4, destaca a influência do ruído na acurácia total da árvore de decisão baseada no algoritmo $J48$ para cada SNR.

Na Tabela 5.7 temos a matriz de confusão do processo de treinamento da árvore de decisão. Notamos que na diagonal principal temos o número de modulações treinadas corretamente e nas redondezas da diagonal principal temos o número de sinais treinados de maneira incorreta. A última linha representa os valores em porcentagens dos verdadeiros positivos, a última coluna temos os valores em porcentagem da precisão e a acurácia total de treinamento da árvore de decisão é demonstrada no último elemento inferior direito da matriz de confusão. A Tabela 5.8 é a matriz de confusão do processo de teste da árvore, e a distribuição dos seus elementos da matriz é semelhante a da Tabela 5.7, mostrando verdadeiros positivos, precisão e acurácia para o teste da árvore.

Na Tabela 5.9 temos a taxa em porcentagem da precisão de cada modulação por SNR. Cada linha da tabela mostra a variação da precisão de cada classe de modulação conforme a alteração da SNR, assim, pode-se avaliar a influência do ruído na classificação correta das modulações. Também, nesta tabela, fica mais evidente em quais modulações o ruído tem mais impacto. Para adquirir estas informações foi necessário testar a árvore de decisão $J48$, já anteriormente treinada, com uma nova base de dados de teste contendo 2355 sinais de cada SNR para cada modulação,

Parâmetros	Valores
-C	0,25
-M	2
-N	3
-Q	1

Tabela 5.6: Parâmetros experimentais do algoritmo $J48$.

	AM	FM	BPSK	QPSK	16-QAM	64-QAM	GMSK	WCDMA	
AM	35272	0	0	0	0	0	0	0	100(%)
FM	0	35169	0	0	0	0	0	0	100(%)
BPSK	0	0	35139	109	56	58	105	1	99,1(%)
QPSK	0	0	84	34676	251	219	13	2	98,4(%)
16-QAM	0	0	72	265	33581	1297	15	3	95,3(%)
64-QAM	0	0	56	263	875	34280	12	3	96,6(%)
GMSK	0	0	101	17	27	24	35221	0	99,6(%)
WCDMA	0	10	0	2	2	0	35329	0	100(%)
	100(%)	100(%)	99,1(%)	98,4(%)	95,3(%)	96,6(%)	99,5(%)	100(%)	98,6083(%)

Tabela 5.7: Matriz de confusão de treinamento da classificação realizada pelo algoritmo *J48*.

	AM	FM	BPSK	QPSK	16-QAM	64-QAM	GMSK	WCDMA	
AM	11825	0	2	0	0	0	1	0	100(%)
FM	0	11931	0	0	0	0	0	0	100(%)
BPSK	0	1	11170	149	90	63	158	1	95,6(%)
QPSK	0	0	137	10738	535	420	25	0	90,1(%)
16-QAM	0	0	82	549	9363	1850	21	2	81,1(%)
64-QAM	0	0	87	456	1527	9521	18	2	80,1(%)
GMSK	0	3	207	23	29	31	11417	0	98,1(%)
WCDMA	0	1	1	3	2	0	0	11759	100(%)
	100(%)	100(%)	96(%)	90,6(%)	78,9(%)	82(%)	97,5(%)	99,9(%)	93,1253(%)

Tabela 5.8: Matriz de confusão de teste da classificação realizada pelo algoritmo *J48*.

equivalente a 25% da quantidade de sinais por SNR da base de dados.

Na Figura 5.4, o gráfico mostra a acurácia total da árvore de decisão testada com sinais em todas as relações sinal-ruído. No eixo das abscissas (eixo horizontal do plano cartesiano) temos os valores de SNR e no eixo das coordenadas (eixo vertical do plano cartesiano) temos os valores em porcentagem da acurácia total correspondentes a cada SNR. Este gráfico dá uma visão mais genérica de como a RNA reage a influência do ruído em seus diferentes níveis.

	0(dB)	5(dB)	10(dB)	20(dB)	30(dB)
AM	100(%)	100(%)	100(%)	100(%)	100(%)
FM	99,9(%)	99,9(%)	100(%)	100(%)	99,9(%)
BPSK	70,4(%)	94,4(%)	97(%)	98,5(%)	97,4(%)
QPSK	43,4(%)	77,5(%)	79,5(%)	95,4(%)	97,2(%)
16-QAM	35,8(%)	61,6(%)	48,3(%)	65,2(%)	67(%)
64-QAM	38,3(%)	62(%)	50,5(%)	64,8(%)	68,5(%)
GMSK	79,5(%)	96,8(%)	99,2(%)	99,9(%)	100(%)
WCDMA	99,8(%)	100(%)	99,8(%)	99,7(%)	100(%)

Tabela 5.9: Precisão por SNR das modulações classificadas pelo algoritmo *J48*.

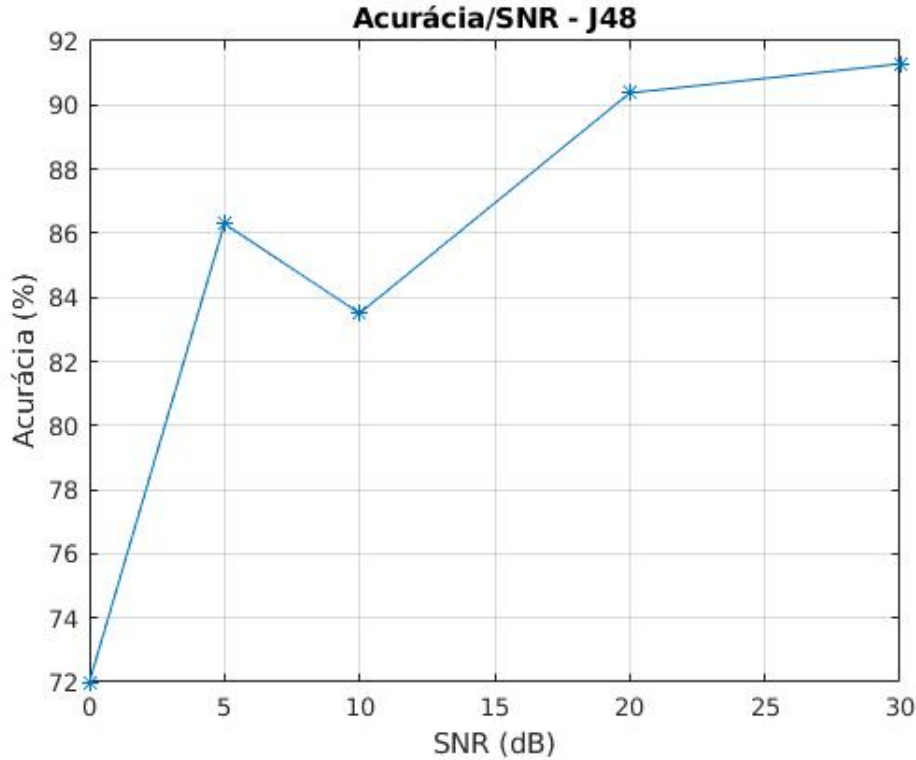


Figura 5.4: Gráfico da acurácia em (%) pela variação da SNR (dB) para o classificador *J48*.

Análise

A árvore de decisão construída pelo algoritmo *J48* teve, para este experimento, um número de 9511 nós-folha e o tamanho da árvore foi de 19021. O tempo necessário para construir o modelo da árvore de decisão foi de 147,87 segundos, o tempo gasto para testar o modelo com os dados de treinamento foi de 1,67 segundos, levando em consideração que a base de dados de treinamento possui 282600 amostras. Para o processo de validação do método, o tempo de teste do modelo com a base de dados de teste foi de 0,55 segundos, levando em consideração que a base de dados de teste possui 94200 amostras. Notamos, conforme a Tabela 5.7, que no processo de treinamento o algoritmo *J48* teve acurácia de 98,6083 %. No processo de teste temos, conforme Tabela 5.8, uma acurácia de 93,1253 % na correta identificação das modulações, o que expressa a boa generalização do classificador. Repara-se que em todas as modulações temos taxas de verdadeiros positivos acima de 90 %, com exceção das modulações digitais 16-QAM e 64-QAM que tiveram taxas de verdadeiros positivos de 78,9 % e 82 % respectivamente. Na Tabela 5.9 temos a taxa de precisão das modulações conforme a variação da SNR, notamos então, que

as modulações AM, FM e WCDMA obtiveram taxas de precisão superiores a 99 % para todas as SNRs abordadas, o que retrata, para o processo de classificação, que o nível de ruído não é uma influência para a correta identificação destas modulações. Ainda na Tabela 5.9, notamos que as modulações BPSK e GMSK obtiveram respostas relativamente inferiores para a SNR igual a 0, porém, obtiveram taxas de precisão acima de 94 % para SNRs acima de 0. Em contra partida, notamos que as modulações QPSK, 16-QAM e 64-QAM são muito susceptíveis a presença de ruído, conforme os dados mostrados na Tabela 5.9. O gráfico da Figura 5.4 mostra o nível de acurácia do classificador para cada SNR, notamos então, que o classificador $J48$ tende a ter melhor acurácia conforme o aumento da SNR, com exceção de um ponto, onde a acurácia foi inferior em comparação ao ponto anterior. Neste caso, notamos que para a SNR igual a 10, o classificador teve dificuldades de diferenciar modulações, principalmente as 16-QAM e 64-QAM, aumentando assim o número de falsos positivos e negativos, impactando no cálculo da acurácia dada pela equação (5.4).

5.4.3 *Random Forest*

Utilizou-se o *software Weka* no experimento envolvendo o algoritmo de árvore de decisão baseado no *Random Forest*. Todos os possíveis parâmetros estão descritos na Tabela 4.3 do Capítulo 4. Nesta etapa configuramos alguns daqueles parâmetros para utilização do classificador. A árvore de decisão baseada no algoritmo *Random Forest* é treinada conforme o *default* do *Weka*, e os valores dos parâmetros estão expostos na Tabela 5.10.

Resultados

Os resultados apresentados estão em forma de matrizes de confusão de treinamento e teste, Tabela 5.11 e Tabela 5.12 respectivamente, e a Tabela 5.13 que demonstra a variação da precisão de cada modulação conforme a alteração da SNR. As matrizes de confusão mostram o número de sinais treinados e classificados corretamente e incorretamente, como também, as porcentagens de verdadeiros positivos, precisão e acurácia. A tabela apresenta a influência do ruído na precisão dos sinais de cada modulação. O gráfico, Figura 5.5, destaca a influência do ruído na acurácia total da árvore de decisão baseada no algoritmo *Random Forest* para cada SNR.

Parâmetros	Valores
-P	100
-I	100
-num-slots	1
-K	0
-M	1
-V	0,001
-S	1
-depth	0 (ilimitado)
-N	0
-num-decimal-places	2
-do-not-check-capabilities	usado
-batch-size	100

Tabela 5.10: Parâmetros experimentais do algoritmo *Random Forest*.

	AM	FM	BPSK	QPSK	16-QAM	64-QAM	GMSK	WCDMA	
AM	35272	0	0	0	0	0	0	0	100(%)
FM	0	35169	0	0	0	0	0	0	100(%)
BPSK	0	0	35468	0	0	0	0	0	100(%)
QPSK	0	0	0	35245	0	0	0	0	100(%)
16-QAM	0	0	0	0	35233	0	0	0	100(%)
64-QAM	0	0	0	0	1	35488	0	0	100(%)
GMSK	0	0	0	0	0	0	35390	0	100(%)
WCDMA	0	0	0	0	0	0	0	35334	100(%)
	100(%)	100(%)	100(%)	100(%)	100(%)	100(%)	100(%)	100(%)	99,9996(%)

Tabela 5.11: Matriz de confusão de treinamento da classificação realizada pelo algoritmo *Random Forest*.

Na Tabela 5.13 temos a taxa em porcentagem da precisão de cada modulação por SNR. Cada linha da tabela mostra a variação da precisão de cada classe de modulação conforme a alteração da SNR, assim, pode-se avaliar a influência do ruído na classificação correta das modulações. Também, nesta tabela, fica mais evidente em quais modulações o ruído tem mais impacto. Para adquirir estas informações foi necessário testar a árvore de decisão *Random Forest*, já anteriormente treinada, com uma nova base de dados de teste contendo 2355 sinais de cada SNR para cada

	AM	FM	BPSK	QPSK	16-QAM	64-QAM	GMSK	WCDMA	
AM	11828	0	0	0	0	0	0	0	100(%)
FM	0	11931	0	0	0	0	0	0	100(%)
BPSK	0	0	11334	96	32	38	132	0	98,6(%)
QPSK	0	0	38	11112	343	350	12	0	93,1(%)
16-QAM	0	0	44	374	10108	1331	10	0	86,7(%)
64-QAM	0	0	33	352	1163	10050	10	3	85,3(%)
GMSK	0	0	50	5	7	17	11631	0	98,6(%)
WCDMA	0	0	0	0	0	0	0	11766	100(%)
	100(%)	100(%)	97,4(%)	93,7(%)	85,2(%)	86,6(%)	99,3(%)	100(%)	95,2866(%)

Tabela 5.12: Matriz de confusão de teste da classificação realizada pelo algoritmo *Random Forest*.

	0(dB)	5(dB)	10(dB)	20(dB)	30(dB)
AM	100(%)	100(%)	100(%)	100(%)	100(%)
FM	100(%)	100(%)	100(%)	100(%)	100(%)
BPSK	83(%)	98,6(%)	99,8(%)	99,8(%)	99,5(%)
QPSK	48,7(%)	81,9(%)	82,6(%)	96,4(%)	98,6(%)
16-QAM	37,2(%)	68,9(%)	55,8(%)	73,1(%)	74(%)
64-QAM	40(%)	65,9(%)	55,4(%)	71,3(%)	77,5(%)
GMSK	80,2(%)	97,1(%)	99,5(%)	100(%)	100(%)
WCDMA	100(%)	100(%)	99,8(%)	99,7(%)	100(%)

Tabela 5.13: Precisão por SNR das modulações classificadas pelo algoritmo *Random Forest*.

modulação, equivalente a 25% da quantidade de sinais por SNR da base de dados.

Na Figura 5.5, o gráfico mostra a acurácia total da *Random Forest* testada com sinais em todas as relações sinal-ruído. No eixo das abscissas (eixo horizontal do plano cartesiano) temos os valores de SNR e no eixo das coordenadas (eixo vertical do plano cartesiano) temos os valores em porcentagem da acurácia total correspondentes a cada SNR. Este gráfico dá uma visão mais genérica de como a RNA reage a influência do ruído em seus diferentes níveis.

Análise

A árvore de decisão baseada no algoritmo *Random Forest* necessitou, para construir o modelo da árvore de decisão, um tempo de 407,09 segundos, o tempo gasto para testar o modelo com os dados de treinamento foi de 15,32 segundos, levando em consideração que a base de dados de treinamento possui 282600 amostras. Para o processo de validação do método, o tempo de teste do modelo com a base de dados de teste foi de 5.11 segundos, levando em consideração que a base de dados de teste possui 94200 amostras. Notamos, conforme a Tabela 5.11, que no processo de treinamento o algoritmo *Random Forest* teve acurácia de 99,9996 %. No processo de teste temos, conforme Tabela 5.12, uma acurácia de 95,2866 % na correta identificação das modulações, o que expressa a boa generalização do classificador. Repara-se que em todas as modulações temos taxas de verdadeiros positivos acima de 93 %, com exceção das modulações digitais 16-QAM e 64-QAM que tiveram taxas de verdadeiros positivos de 85,2 % e 86,6 % respectivamente. Na Tabela 5.13 temos a taxa de precisão das modulações conforme a variação da SNR, notamos então, que as modulações AM, FM e WCDMA obtiveram taxas de precisão

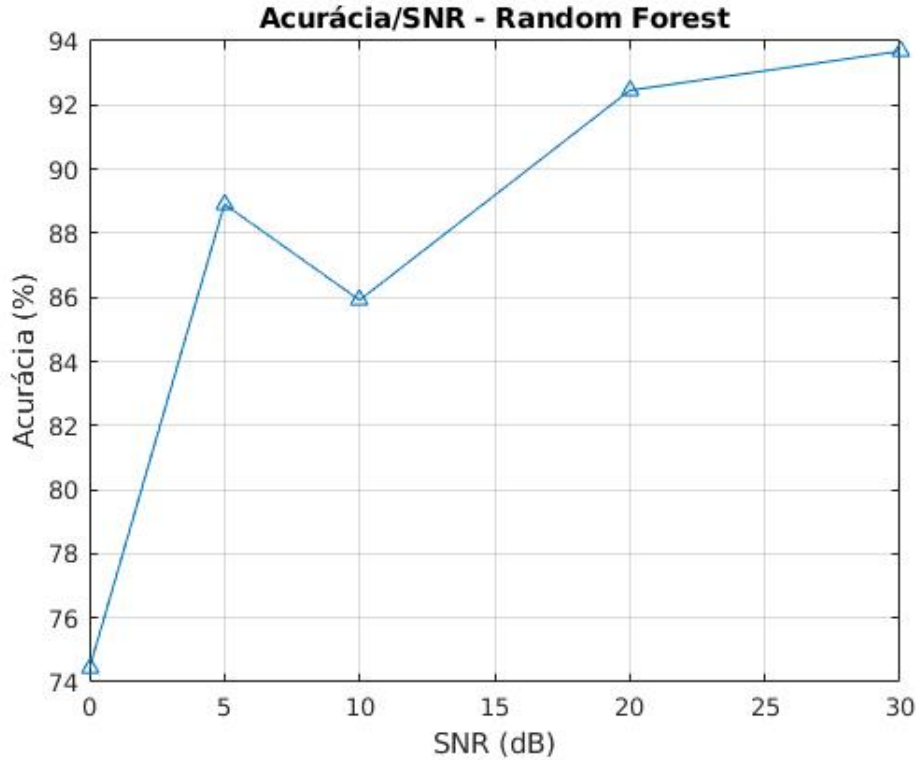


Figura 5.5: Gráfico da acurácia em (%) pela variação da SNR (dB) para o classificador *Random Forest*.

superiores a 99 % para todas as SNRs abordadas, o que retrata, para o processo de classificação, que o nível de ruído não é uma influência para a correta identificação destas modulações. Ainda na Tabela 5.13, notamos que as modulações BPSK e GMSK obtiveram respostas relativamente inferiores para a SNR igual a 0 (83 % e 80,2 % respectivamente), porém, obtiveram taxas de precisão acima de 97 % para SNRs acima de 0. Em contra partida, notamos que as modulações QPSK, 16-QAM e 64-QAM são muito susceptíveis a presença de ruído, conforme os dados mostrados na Tabela 5.13. O gráfico da Figura 5.5 mostra o nível de acurácia do classificador para cada SNR, notamos então, que o classificador *Random Forest* tende a ter melhor acurácia conforme o aumento da SNR, com exceção de um ponto, onde a acurácia foi inferior em comparação ao ponto anterior. Neste caso, notamos que para a SNR igual a 10, o classificador teve dificuldades de diferenciar modulações, principalmente as 16-QAM e 64-QAM, aumentando assim o número de falsos positivos e negativos, impactando no cálculo da acurácia dada pela equação (5.4).

5.4.4 *Naive Bayes*

O classificador probabilístico *Naive Bayes* é, em tese, o menos robusto em comparação aos já apresentados anteriormente, e por esse motivo não existe, pelo menos no *software Weka*, muitos atributos que podem ser alterados. Dentre as opções parâmetros deste classificador temos a possibilidade de aplicar: o uso do estimador de densidade de *kernel* em vez da distribuição normal para atributos numéricos; o uso da discretização supervisionada para processar atributos numéricos e a opção de mostrar o modelo no formato antigo (bom quando tem muitas classes). Para o treinamento do classificador *Naive Bayes* não foi implementado nenhuma destes parâmetros, utilizando então o formato *default* do *Weka*.

Resultados

Os resultados apresentados estão em forma de matrizes de confusão de treinamento e teste, Tabela 5.14 e Tabela 5.15 respectivamente, e a Tabela 5.16 que demonstra a variação da precisão de cada modulação conforme a alteração da SNR. As matrizes de confusão mostram o número de sinais treinados e classificados corretamente e incorretamente, como também, as porcentagens de verdadeiros positivos, precisão e acurácia. A tabela apresenta a influência do ruído na precisão dos sinais de cada modulação. O gráfico, Figura 5.6, destaca a influência do ruído na acurácia total do classificador *Naive Bayes* para cada SNR.

	AM	FM	BPSK	QPSK	16-QAM	64-QAM	GMSK	WCDMA	
AM	34558	0	313	0	0	0	401	0	100(%)
FM	0	35149	10	0	0	0	10	0	99,7(%)
BPSK	0	32	20673	4566	1273	1384	7471	69	88,9(%)
QPSK	0	0	638	187950	3662	10501	285	1364	68,3(%)
16-QAM	0	0	489	2430	9325	21180	340	1469	44,3(%)
64-QAM	0	0	490	942	6111	25945	289	1712	42,6(%)
GMSK	0	82	625	546	698	1637	31760	42	78,3(%)
WCDMA	0	0	6	250	0	214	1	34863	88,2(%)
	98(%)	99,9(%)	58,3(%)	53,3(%)	26,5(%)	73,1(%)	89,7(%)	98,7(%)	74,6879(%)

Tabela 5.14: Matriz de confusão de treinamento da classificação realizada pelo classificador *Naive Bayes*.

Na Tabela 5.16 temos a taxa em porcentagem da precisão de cada modulação por SNR. Cada linha da tabela mostra a variação da precisão de cada classe de modulação conforme a alteração da SNR, assim, pode-se avaliar a influência do ruído na classificação correta das modulações. Também, nesta tabela, fica mais evidente

	AM	FM	BPSK	QPSK	16-QAM	64-QAM	GMSK	WCDMA	
AM	11574	0	109	0	0	0	145	0	100(%)
FM	0	11927	3	0	0	0	1	0	99,7(%)
BPSK	0	11	6814	1404	399	475	2507	22	88,5(%)
QPSK	0	0	205	6356	1285	3473	91	455	69,6(%)
16-QAM	0	0	169	810	3149	7124	109	506	44,5(%)
64-QAM	0	0	180	308	2023	8478	94	528	42,1(%)
GMSK	0	25	215	176	222	528	10538	6	78,1(%)
WCDMA	0	0	3	82	0	70	0	11611	88,5(%)
	97,9(%)	100(%)	58,6(%)	53,6(%)	26,5(%)	73(%)	90(%)	98,7(%)	74,7845(%)

Tabela 5.15: Matriz de confusão de teste da classificação realizada pelo classificador *Naive Bayes*.

	0(dB)	5(dB)	10(dB)	20(dB)	30(dB)
AM	100(%)	100(%)	100(%)	100(%)	100(%)
FM	100(%)	99,6(%)	99,7(%)	99,8(%)	99,5(%)
BPSK	53(%)	95,9(%)	96,9(%)	92(%)	92,9(%)
QPSK	13,5(%)	88,5(%)	87,4(%)	65,3(%)	63,5(%)
16-QAM	30,4(%)	22,9(%)	33,6(%)	59,8(%)	60,2(%)
64-QAM	33,2(%)	36,3(%)	39,1(%)	54,3(%)	60,6(%)
GMSK	50,6(%)	78,7(%)	98,9(%)	99,8(%)	99,5(%)
WCDMA	98(%)	79,3(%)	75(%)	94(%)	98,6(%)

Tabela 5.16: Precisão por SNR das modulações classificadas pelo classificador *Naive Bayes*.

em quais modulações o ruído tem mais impacto. Para adquirir estas informações foi necessário testar o classificador *Naive Bayes*, já anteriormente treinada, com uma nova base de dados de teste contendo 2355 sinais de cada SNR para cada modulação, equivalente a 25% da quantidade de sinais por SNR da base de dados.

Na Figura 5.6, o gráfico mostra a acurácia total do classificador *Naive Bayes* testada com sinais em todas as relações sinal-ruído. No eixo das abscissas (eixo horizontal do plano cartesiano) temos os valores de SNR e no eixo das coordenadas (eixo vertical do plano cartesiano) temos os valores em porcentagem da acurácia total correspondentes a cada SNR. Este gráfico dá uma visão mais genérica de como a RNA reage a influência do ruído em seus diferentes níveis.

Análise

O classificador probabilístico *Naive Bayes* necessitou, para construir o modelo de classificação, um tempo de 3,15 segundos, o tempo gasto para testar o modelo com os dados de treinamento foi de 16,82 segundos, levando em consideração que a base de dados de treinamento possui 282600 amostras. Para o processo de validação do método, o tempo de teste do modelo com a base de dados de teste foi de 5.51

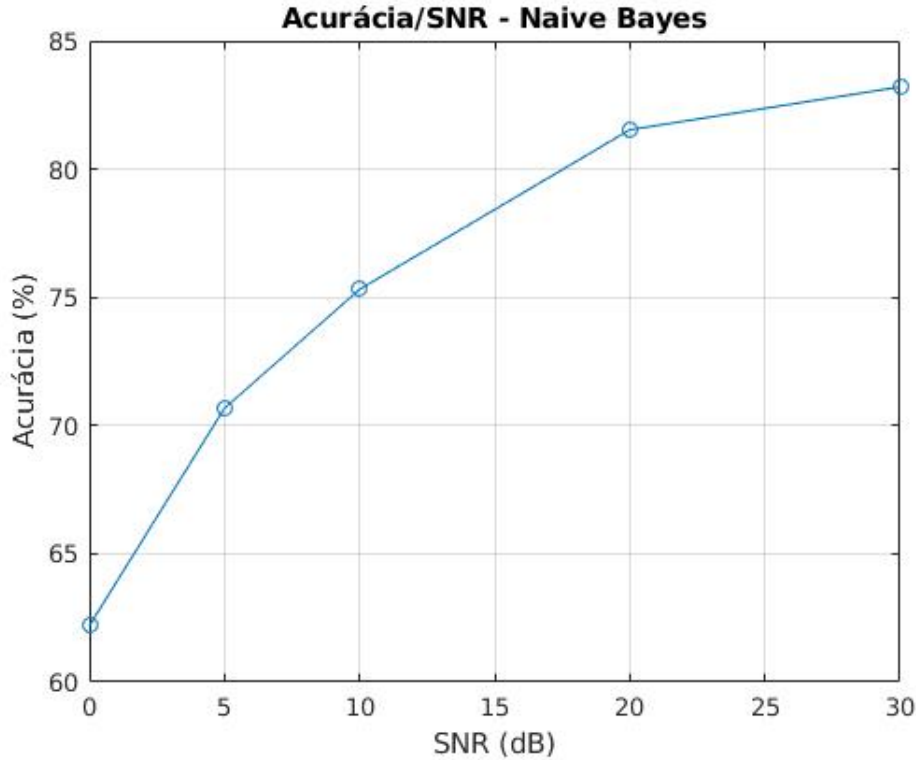


Figura 5.6: Gráfico da acurácia em (%) pela variação da SNR (dB) para o classificador *Naive Bayes*.

segundos, levando em consideração que a base de dados de teste possui 94200 amostras. Notamos, conforme Tabela 5.14, que no processo de treinamento o classificador *Naive Bayes* teve uma acurácia de 74,6879 %. No processo de teste temos, conforme Tabela 5.15, uma acurácia de 74,7848 % na correta identificação das modulações, o que expressa a boa generalização do classificador. Repara-se que as modulações AM, FM, GMSK e WCDMA obtiveram taxas de verdadeiros positivos igual ou superior a 90 %, já as modulações digitais BPSK, QPSK e 16-QAM obtiveram taxas de verdadeiros positivos inferior a 60 %, a modulação 64-QAM obteve 73 % de verdadeiros positivos. Na Tabela 5.16, notamos que as únicas modulações que mantiveram uma constância na correta classificação independente da SNR foram as modulações AM e FM. Ainda na Tabela 5.16, nota-se, que o *Naive Bayes*, é muito susceptível a variação de ruído nas modulações BPSK, QPKS, 16-QAM e 64-QAM, as modulações GMSK e WCDMA sofrem a ação do ruído de maneira mais suave. No gráfico mostrado na Figura 5.6 temos a acurácia do classificador conforme a variação da SNR, notamos então, que a acurácia do *Naive Bayes* responde de forma crescente ao aumento da SNR.

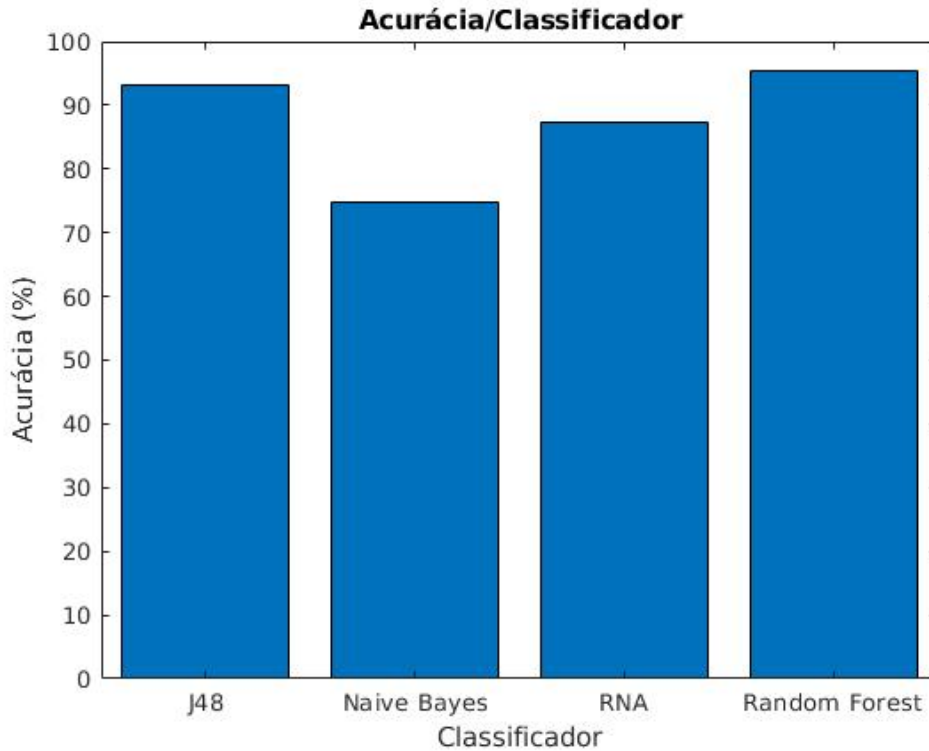


Figura 5.7: Gráfico da acurácia em (%) geral da rede por cada classificador proposto.

5.4.5 Análise Comparativa

Com a base de dados inicial para treinamento e teste, divididas aleatoriamente 75 % para treino e 25 % para teste, obtivemos acurácia geral de 95,2866 % para a árvore de decisão com algoritmo *Random Forest*, 93,1253 % para a árvore de decisão com algoritmo *J48*, 87,4 % para a RNA e 74,7845 % para o classificador *Naive Bayes*, conforme Figura 5.7. Notamos que as árvores de decisão, com essa base de dados, obtiveram um desempenho melhor em comparação aos outros dois métodos. Além da acurácia geral, temos a análise da precisão de cada modulação para cada classificador. Na RNA as modulações AM, FM, BPSK, GMSK e WCDMA obtiveram taxas de precisão superiores a 90 %, entretanto, as modulações QPSK, 16-QAM e 64-QAM obtiveram taxas inferiores a 80 %. Nas classificações utilizando algoritmos *J48* e *Random Forest*, conforme Tabela 5.8 e Tabela 5.12, todas as modulações obtiveram taxas de precisão acima de 80 %, as modulações AM, FM, BPSK, QPSK, GMSK e WCDMA obtiveram ainda melhores precisões, ultrapassando os 90 %. O classificador *Naive Bayes* obteve taxas superiores a 80 % para as modulações AM, FM, BPSK e WCDMA, e inferiores a 70 % nas outras modulações abordadas,

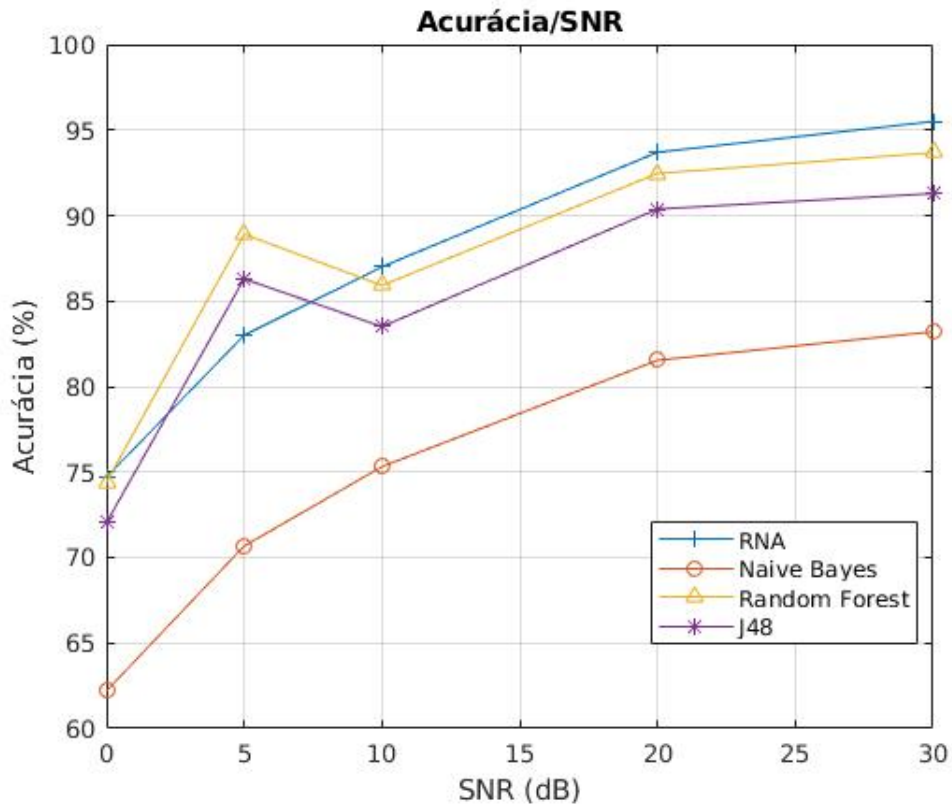


Figura 5.8: Gráfico da acurácia em (%) pela variação da SNR (dB) para todos os métodos propostos.

conforme Tabela 5.15.

Na segunda etapa dos experimentos, as redes já treinadas com a base de dados de treinamento inicial, foram testadas com novas bases de dados separadas por SNR. Na Figura 5.8 temos o comportamento da acurácia dos classificadores conforme a alteração da SNR. Notamos que, para essa nova base de dados de teste, os classificadores baseados em árvore de decisão obtiveram uma acurácia similar, inclusive na transição de 5 para 10 dB onde houve uma queda na acurácia, o que não era esperado. A RNA e o *Naive Bayes* obtiveram comportamento semelhante, operando de maneira crescente conforme o aumento da SNR. Podemos notar, também, que a RNA obteve um melhor desempenho nas mais altas SNRs em comparação aos outros métodos, porém, na média das acurácias, o *Random Forest* tem melhor desempenho geral. Podemos concluir então, conforme Figura 5.8, que nas SNRs mais elevadas a RNA possui melhor acurácia em relação aos outros classificadores, em contrapartida, a *Random Forest* tem desempenho geral superior aos demais. A árvore de decisão utilizando algoritmo *J48* obteve, nesta análise, comportamento se-

melhante ao *Random Forest* porém com desempenho inferior. O classificador *Naive Bayes*, por ser menos robusto que os demais, obteve taxas bem inferiores aos demais classificadores propostos.

Trabalhos	Classificador	Acurácia (%)
[17]	SVM	89,72
[18]	CNN	90,64
Propostos	RNA	87,4
	<i>Naive Bayes</i>	74,7845
	<i>J48</i>	93,1253
	<i>Random Forest</i>	95,2866

Tabela 5.17: Tabela comparativa entre métodos propostos e literaturas.

Na Tabela 5.17 temos uma comparação entre as acurácias dos métodos propostos neste trabalho e alguns presentes na literatura. O trabalho de [17] é o que apresenta a base de dados mais semelhante a proposta, assim, é o mais compatível para a comparação de desempenho. Notamos que [17] obteve 89,72% de acurácia, e que na proposta deste trabalho obtivemos dois métodos com acurácia significativamente superiores, *J48* e *Random Forest*, conforme apresentado na Tabela 5.17. O método por RNA obteve acurácia levemente inferior a apresentada [17]. Notamos também, em comparação a [18], que os métodos por árvore de decisão propostos obtiveram melhor desempenho.

Capítulo 6

Conclusões

Nesta dissertação foi investigado o comportamento de diversos classificadores e alguns de seus parâmetros para a o processo de classificação automática de modulações. A metodologia empregada, na geração dos sinais modulados, se buscou uma similaridade com os sistemas de comunicação reais. Então, os sinais modulados foram gerados em banda base e acrescidos de ruído também em banda base, passaram por formatação de pulso e em seguida foram elevados na frequência. Foi realizado então a amostragem e a obtenção da envoltória complexa dos sinais em banda passante para, então, serem extraídas as características. A base de dados criada contém 376800 vetores de características e foram divididos aleatoriamente para treinamento (75 %) e teste (25 %) dos classificadores propostos. Foram definidos quatro classificadores multiclasse para o processo de classificação: uma RNA, duas árvores de decisão e um classificador probabilístico. A RNA definida foi uma *perceptron* com duas camadas totalmente conectadas, uma primeira com 29 neurônios e uma de saída com 8, e o algoritmo de otimização empregado no *backpropagation* foi o de Levenberg-Marquardt. Os algoritmos empregados nas árvores de decisão foram os *J48*, que utiliza razão do ganho de informação pra determinar quais características tem mais relevância na segregação das modulações, e o *Random Forest*, que através da aplicação da técnica *bootstrap* cria diversas árvores de decisão para então utilizar uma votação majoritária para definir quais árvores distinguem determinadas modulações. O classificador probabilístico é o *Naive Bayes* e aplica o teorema de Bayes, probabilidade condicional, para classificar objetos.

O foco principal do trabalho foram os experimentos com os quatro classifica-

dores propostos. O primeiro experimento foi com a RNA que obteve uma acurácia geral 87,4 %, onde as modulações AM, FM, BPSK, GMSK e WCDMA alcançaram precisões superiores a 95 %, as modulações QPSK, 16-QAM e 64-QAM obtiveram precisões inferiores a 80 %. No segundo experimento foi avaliado o desempenho da árvore de decisão com algoritmo *J48*, onde obtivemos uma acurácia geral de 93,1253 %, sendo as modulações AM, FM, BPSK, GMSK e WCDMA obtiveram precisões superiores a 95 %, enquanto as outros modulações obtiveram precisões acima dos 80 %. O terceiro classificador abordado foi a *Random Forest* com acurácia geral de 95,2866 %, sendo as modulações AM, FM, BPSK, GMSK e WCDMA obtiveram precisões superiores a 98 %, enquanto as outros modulações obtiveram precisões acima dos 85 %. E o quarto experimento, o classificador *Naive Bayes* que obteve acurácia geral de 74,7845 %, as modulações AM e FM alcançaram precisões superiores a 99 %, as modulações BPSK, GMSK e WCDMA obtiveram taxas acima de 78 % e as demais modulações precisões inferiores a 70 %. Numa segunda etapa, uma nova base de dados de teste dividida por SNR foi aplicada aos classificadores já treinados. Nestes experimentos foram avaliados os desempenhos dos classificadores para cada SNR, e notamos que a RNA obteve melhor resultado para as melhores SNRs, em contrapartida, de maneira geral, o *Random Forest* obteve um desempenho médio superior o da RNA.

Em vista da aplicabilidade destas técnicas, a RNA requer um custo computacional alto, conforme mostrado no Capítulo 5, onde o tempo de treinamento para a base de dados proposta foi de 27h06min43seg considerando também o interrompimento do treinamento na 401^a época, assim, o custo computacional seria mais elevado com o aumento da base de dados. Com isso em vista, seria aplicável em sistemas reais somente redes treinadas previamente, e em caso de modificação do sistema de comunicação a RNA deverá ser treinada novamente com a nova base de dados. Nos métodos por árvore de decisão o custo computacional é bem menor, os modelos foram gerados em questão de minutos e os testes em questão de segundos ou até menos no caso da *J48*, semelhante a RNA as árvores seriam aplicadas em sistemas reais empregando seus modelos previamente treinados, porém, em caso de atualização dos sistemas os modelos por árvore poderiam ser retreinados muito mais rápido que a RNA resultando num custo computacional muito menor. O *Naive*

Bayes no quesito custo computacional é semelhante aos métodos por árvore e consequentemente melhor que a RNA, a maior questão sobre a aplicabilidade do *Naive Bayes* é devido ao seu baixo desempenho na classificação de modulações digitais.

6.1 Trabalhos Futuros

- Este trabalho os classificadores foram investigados classificando várias modulações de maneira simultânea. Então como proposta de trabalho futuro, seria a realização de um estudo similar ao proposto nesta dissertação para classificação binária, com uma estrutura hierárquica, e assim investigar quais classificadores são mais eficientes para segregar duas modulações específicas.
- Não foi realizado neste trabalho um estudo mais detalhado da parametrização dos classificadores abordados. Então, como trabalho futuro, propõe-se uma investigação do comportamento dos classificadores com a exploração dos parâmetros dos classificadores abordados.
- Foram abordados neste trabalho os classificadores mais populares dentro do *machine learning*. Propõe-se como trabalho futuro a aplicação de classificadores mais robustos, como as redes neurais convolucionais, com outras técnicas de classificação, com intuito de deixar a proposta de investigação mais atual e abrangente.

Referências Bibliográficas

- [1] DUDA, R. O., HART, P. E., STORK, D. G., *Pattern classification*. John Wiley & Sons, 2012.
- [2] THEODORIDIS, S., PIKRAKIS, A., KOUTROUMBAS, K., et al., *Introduction to pattern recognition: a matlab approach*. Academic Press, 2010.
- [3] CAELLI, T., *Structural, Syntactic, and Statistical Pattern Recognition: Joint IAPR International Workshops SSPR 2002 and SPR 2002, Windsor, Ontario, Canada, August 6-9, 2002. Proceedings*. v. 2396. Springer Science & Business Media, 2002.
- [4] SRIHARI, S., “Covindaraju”, *Pattern recognition, Chapman & Hall, London*, pp. 1034–1041, 1993.
- [5] FUKUNAGA, K., *Introduction to statistical pattern recognition*. Elsevier, 2013.
- [6] POPOOLA, J. J., *Sensing and detection of a primary radio signal in a cognitive radio environment using modulation identification technique*, Ph.D. Thesis, 2012.
- [7] ZHU, Z., NANDI, A. K., *Automatic modulation classification: principles, algorithms and applications*. John Wiley & Sons, 2015.
- [8] AZZOUZ, E. E., NANDI, A. K., “Modulation recognition using artificial neural networks”, In: *Automatic Modulation Recognition of Communication Signals*, pp. 132–176, Springer, 1996.
- [9] ALI, A., HAMOUDA, W., “Advances on spectrum sensing for cognitive radio networks: Theory and applications”, *IEEE communications surveys & tutorials*, v. 19, n. 2, pp. 1277–1304, 2016.

- [10] PETROVA, M., MÄHÖNEN, P., OSUNA, A., “Multi-class classification of analog and digital signals in cognitive radios using Support Vector Machines”. In: *2010 7th International Symposium on Wireless Communication Systems*, pp. 986–990, 2010.
- [11] D. A., A., W. C., M., L. C., C., et al., “Classificação Automática de Modulações Mono e Multiportadoras utilizando Método de Extração de Características e Classificadores SVM”. In: *Proc. Simp. Brasileiro de Telecom. e Processamento de Sinais SBrT*, pp. 393–397, São Pedro, Brazil, 2017.
- [12] KIM, B., KIM, J., CHAE, H., et al., “Deep neural network-based automatic modulation classification technique”. In: *2016 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 579–582, 2016.
- [13] HUANG, S., JIANG, Y., GAO, Y., et al., “Automatic modulation classification using contrastive fully convolutional network”, *IEEE Wireless Communications Letters*, 2019.
- [14] HUANG, S., CHAI, L., LI, Z., et al., “Automatic Modulation Classification Using Compressive Convolutional Neural Network”, *IEEE Access*, v. 7, pp. 79636–79643, 2019.
- [15] TRIANTARIS, P., TSIMBALO, E., CHIN, W. H., et al., “Automatic Modulation Classification in the Presence of Interference”. In: *2019 European Conference on Networks and Communications (EuCNC)*, pp. 549–553, 2019.
- [16] HONG, D., ZHANG, Z., XU, X., “Automatic modulation classification using recurrent neural networks”. In: *Proc. Int. Conf. on Computer and Communications (ICCC)*, pp. 695–700, Dec 2017.
- [17] AMOEDO, D. A., OTHERS, “Classificação automática de modulações mono e multiportadoras utilizando método de extração de características e classificadores SVM”, 2017.

- [18] WEI, S., ZOU, S., LIAO, F., et al., “Automatic Modulation Recognition Using Neural Architecture Search”. In: *2019 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)*, pp. 151–156, 2019.
- [19] LINDSEY, W. C., SIMON, M. K., *Telecommunication systems engineering*. Courier Corporation, 1991.
- [20] MENDIS, G. J., WEI, J., MADANAYAKE, A., “Deep learning-based automated modulation classification for cognitive radio”. In: *2016 IEEE International Conference on Communication Systems (ICCS)*, pp. 1–6, 2016.
- [21] HAYKIN, S., MOHER, M., *Sistemas de Comunicação-5*. Bookman Editora, 2011.
- [22] LATHI, B. P., *Modern digital and analog communication systems*. Oxford University Press, Inc., 1995.
- [23] SHANNON, C. E., “Communication in the presence of noise”, *Proceedings of the IEEE*, v. 86, n. 2, pp. 447–457, 1998.
- [24] DUTT, V., CHAUDHRY, V., KHAN, I., “Pattern recognition: an overview”, *American Journal of Intelligent Systems*, v. 2, n. 1, pp. 23–27, 2012.
- [25] DE LIMA, M. V., *Medição de ocupação espectral e estudo das oportunidades de espectro para uso de sistemas rádio cognitivo*, Ph.D. Thesis, PUC-Rio, 2013.
- [26] LINHARES, A., SANTOYO, R. F., BIZERRA, R., et al., “Aspectos técnicos e regulatórios para implementação da tecnologia de rádios cognitivos no Brasil”, *Revista de Estudios para el Desarrollo Social de la Comunicación*, n. 9, pp. 325, 2014.
- [27] SIMON, M. K., HINEDI, S. M., LINDSEY, W. C., *Digital communication techniques: signal design and detection*. PTR Prentice Hall Englewood Cliffs, 1995.

- [28] BINGHAM, J. A. C., “Multicarrier modulation for data transmission: an idea whose time has come”, *IEEE Communications Magazine*, v. 28, n. 5, pp. 5–14, May 1990.
- [29] TELLADO, J., *Multicarrier modulation with low PAR: applications to DSL and wireless*. v. 587. Springer Science & Business Media, 2006.
- [30] CHAKRAVARTHY, V., NUNEZ, A. S., STEPHENS, J. P., et al., “TDCS, OFDM, and MC-CDMA: a brief tutorial”, *IEEE Communications Magazine*, v. 43, n. 9, pp. S11–S16, 2005.
- [31] MOOSE, P. H., “A technique for orthogonal frequency division multiplexing frequency offset correction”, *IEEE Transactions on communications*, v. 42, n. 10, pp. 2908–2914, 1994.
- [32] FAZEL, K., KAISER, S., *Multi-carrier and spread spectrum systems: from OFDM and MC-CDMA to LTE and WiMAX*. John Wiley & Sons, 2008.
- [33] PICKHOLTZ, R., SCHILLING, D., MILSTEIN, L., “Theory of spread-spectrum communications-a tutorial”, *IEEE transactions on Communications*, v. 30, n. 5, pp. 855–884, 1982.
- [34] AZARMANESH, O., BILÉN, S. G., “IQ diagram utilization in a novel modulation classification technique for cognitive radio applications”, *EURASIP Journal on Wireless Communications and Networking*, v. 2013, n. 1, pp. 289, 2013.
- [35] RISH, I., OTHERS, “An empirical study of the naive Bayes classifier”. In: *IJCAI 2001 workshop on empirical methods in artificial intelligence*, v. 3, n. 22, pp. 41–46, 2001.
- [36] MCCALLUM, A., NIGAM, K., OTHERS, “A comparison of event models for naive bayes text classification”. In: *AAAI-98 workshop on learning for text categorization*, v. 752, n. 1, pp. 41–48, 1998.
- [37] LIBRELOTTO, S. R., MOZZAQUATRO, P. M., “Análise dos algoritmos de mineração J48 e Apriori aplicados na detecção de indicadores da qualidade

de vida e saúde”, *Revista Interdisciplinar de Ensino, Pesquisa e Extensão*, v. 1, n. 1, 2014.

- [38] KOVÁCS, Z. L., *Redes neurais artificiais*. Editora Livraria da Física, 2002.
- [39] HAN, L., GAO, F., LI, Z., et al., “Low complexity automatic modulation classification based on order-statistics”, *IEEE Transactions on Wireless Communications*, v. 16, n. 1, pp. 400–411, 2016.
- [40] ALI, A., YANGYU, F., LIU, S., “Automatic modulation classification of digital modulation signals with stacked autoencoders”, *Digital Signal Processing*, v. 71, pp. 108–116, 2017.
- [41] WU, Z., ZHOU, S., YIN, Z., et al., “Robust automatic modulation classification under varying noise conditions”, *IEEE Access*, v. 5, pp. 19733–19741, 2017.
- [42] RAMJEE, S., JU, S., YANG, D., et al., “Fast deep learning for automatic modulation classification”, *arXiv preprint arXiv:1901.05850*, 2019.
- [43] WANG, Y., LIU, M., YANG, J., et al., “Data-driven deep learning for automatic modulation recognition in cognitive radios”, *IEEE Transactions on Vehicular Technology*, v. 68, n. 4, pp. 4074–4077, 2019.
- [44] RODRIGUEZ-GALIANO, V. F., GHIMIRE, B., ROGAN, J., et al., “An assessment of the effectiveness of a random forest classifier for land-cover classification”, *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 67, pp. 93–104, 2012.
- [45] GAO, X., WEN, J., ZHANG, C., “An Improved Random Forest Algorithm for Predicting Employee Turnover”, *Mathematical Problems in Engineering*, v. 2019, 2019.
- [46] GOGUL, I., KUMAR, V. S., “Flower species recognition system using convolution neural networks and transfer learning”. In: *2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)*, pp. 1–6, 2017.

- [47] TSAMPASIS, E., GKONIS, P., TRAKADAS, P., et al., “On the Performance Evaluation of a MIMO–WCDMA Transmission Architecture for Building Management Systems”, *Sensors*, v. 18, n. 1, pp. 155, 2018.
- [48] BARACAT, G. H., JUNIOR, A. C. S., DE SOUZA, R. A. A., “Estudo sobre os Pulsos de Nyquist”, .
- [49] OLIVEIRA, M. Q. D. A., OTHERS, “Teste para verificação da hipótese de ruído branco utilizando teoria da informação”, 2017.
- [50] FERRARI, A. J., DE ANDRADE, A. A., “Uma nota sobre reticulados rotacionados para o canal com desvanecimento do tipo Rayleigh”, *Proceeding Series of the Brazilian Society of Computational and Applied Mathematics*, v. 3, n. 1, 2015.
- [51] SAFAVIAN, S. R., LANDGREBE, D., “A survey of decision tree classifier methodology”, *IEEE transactions on systems, man, and cybernetics*, v. 21, n. 3, pp. 660–674, 1991.
- [52] LEWIS, D. D., “Naive (Bayes) at forty: The independence assumption in information retrieval”. In: *European conference on machine learning*, pp. 4–15, 1998.
- [53] HAYKIN, S., *Redes neurais: princípios e prática*. Bookman Editora, 2007.
- [54] BENATTI, K. A., “O Método de Levenberg-Marquardt para o Problema de Quadrados Mínimos não Linear”, 2017.
- [55] ARAÚJO, E., GAMA, C., OTHERS, *Replicando características de ciclos econômicos: um estudo comparativo entre Redes Neurais Artificiais e modelos ARIMA*, Tech. rep., Insper Working Paper, Insper Instituto de Ensino e Pesquisa, 2004.
- [56] YU, H., WILAMOWSKI, B. M., “Levenberg-marquardt training”, *Industrial electronics handbook*, v. 5, n. 12, pp. 1, 2011.
- [57] DE ARAUJO VIEIRA, E. M., DE AT NEVES, N. T., DE OLIVEIRA, A. C. C., et al., “Avaliação da performance do algoritmo J48 para constru-

ção de modelos baseados em árvores de decisão”, *Revista Brasileira de Computação Aplicada*, v. 10, n. 2, pp. 80–90, 2018.

- [58] RAMYA, M., LOKESH, V., MANJUNATH, T., et al., “A predictive model construction for mulberry crop productivity”, *Procedia Computer Science*, v. 45, pp. 156–165, 2015.
- [59] CARVALHO, D. R., “Árvore de decisão/ algoritmo genético para tratar o problema de pequenos disjuntos em classificação de dados”, *Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil. Doctor Thesis. 162pp*, 2005.
- [60] BOSCH, A., ZISSERMAN, A., MUNOZ, X., “Image classification using random forests and ferns”. In: *2007 IEEE 11th international conference on computer vision*, pp. 1–8, 2007.
- [61] BREIMAN, L., “Random forests”, *Machine learning*, v. 45, n. 1, pp. 5–32, 2001.
- [62] SIROKY, D. S., OTHERS, “Navigating random forests and related advances in algorithmic modeling”, *Statistics Surveys*, v. 3, pp. 147–163, 2009.
- [63] DAHO, M. E. H., SETTOUTI, N., LAZOUNI, M. E. A., et al., “Weighted vote for trees aggregation in random forest”. In: *2014 International Conference on Multimedia Computing and Systems (ICMCS)*, pp. 438–443, 2014.