

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA ELÉTRICA
ENGENHARIA ELETRÔNICA E DE TELECOMUNICAÇÕES

SAMUEL ALVES TAVARES

OTIMIZAÇÃO DE CARACTERÍSTICAS PARA CLASSIFICAÇÃO
AUTOMÁTICA DE MODULAÇÃO DE SINAIS AFETADOS PELOS CANAIS
AWGN E RAYLEIGH.

UBERLÂNDIA
2022

SAMUEL ALVES TAVARES

OTIMIZAÇÃO DE CARACTERÍSTICAS PARA CLASSIFICAÇÃO AUTOMÁTICA
DE MODULAÇÃO DE SINAIS AFETADOS PELOS CANAIS AWGN E
RAYLEIGH.

Trabalho apresentado como um dos requisitos para aprovação na disciplina de Trabalho de Conclusão de Curso do curso de Engenharia Eletrônica e de Telecomunicações da Universidade Federal de Uberlândia – Campus Santa Mônica.

Orientador: Prof. Dr. André Luiz Aguiar da Costa

UBERLÂNDIA
2022

SAMUEL ALVES TAVARES

OTIMIZAÇÃO DE CARACTERÍSTICAS PARA CLASSIFICAÇÃO AUTOMÁTICA
DE MODULAÇÃO DE SINAIS AFETADOS PELOS CANAIS AWGN E
RAYLEIGH.

Trabalho apresentado como um dos
requisitos para aprovação na disciplina de
Trabalho de Conclusão de Curso do curso de
Engenharia Eletrônica e de Telecomunicações
da Universidade Federal de Uberlândia –
Campus Santa Mônica.

Uberlândia, 29 de janeiro de 2022.

COMISSÃO EXAMINADORA:

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

Prof. Dr. Éderson Rosa da Silva
Universidade Federal de Uberlândia
Examinador

Prof. Dr. Alexandre Coutinho Mateus
Universidade Federal de Uberlândia
Examinador

Agradecimentos

Primeiramente, gostaria de agradecer a Deus pelo dom da vida.

Aos meus professores, em especial meu orientador, o professor André Luiz Aguiar da Costa pelos conhecimentos passados.

Aos meus pais, José e Ângela, e ao meu irmão Tiago por todo o apoio prestado durante o curso.

A minha companheira Camila, por todo o suporte durante a graduação.

Ao engenheiro Antônio Pereira por todo o apoio e ensinamentos durante o trabalho, sem sua ajuda não seria possível concluí-lo.

Ao professor Waldir Sabino da UFAM, por todo o apoio técnico durante o trabalho.

A todos os meus amigos feitos durante o curso (Adriel, Alex, Beatriz, Gabriel, Lukas e Neviton) por todos os momentos de estudos e descontrações.

Por último agradeço também a Universidade Federal de Uberlândia pelo ensino de qualidade e gratuito, além de todos os programas de bolsas, como o PET e bolsas de monitoria.

Resumo

Com o surgimento de novas tecnologias de redes móveis nos últimos anos, cada vez mais as redes de Telecomunicações precisam ser mais autônomas, identificando todos os parâmetros do usuário, incluindo o tipo de modulação. Dessa forma, esse trabalho tem como objetivo gerar uma base de dados com os principais tipos de modulação digital e construir um sistema que classifica automaticamente modulações digitais de um sinal que passou pelos canais AWGN e Rayleigh, além de selecionar e otimizar as melhores características do sinal para que a classificação seja feita de forma mais rápida e eficiente. Feita a seleção de características, utilizando o método de correlação somado ao algoritmo RFE, obteve-se um resultado muito promissor, onde as características foram diminuídas em cerca de 80% das características propostas inicialmente, além de ter atingido uma alta acurácia de acerto na classificação.

Palavras-chave: Aprendizado de Máquina. Classificação Automática de Modulações. Inteligência Artificial. Canal de Rayleigh. Correlação. Seleção de características.

Abstract

In the last years, lot of new mobile network technologies have emerged. So, Telecommunications networks need to be more autonomous, identifying all user parameters, including the type of modulation.

The objective of this work is to generate a database with the main types of digital modulation and build a system that automatically classifies digital modulations of a signal that passed through the AWGN and Rayleigh channels, select and optimize the best characteristics of the signal so that the classification be done more quickly and efficiently. The selection of features using the correlation and RFE methods, proved to be very promising, decreasing by about 80% of the initially proposed features, in addition to having reached a high accuracy of correct classification.

Keywords: Machine Learning. Automatic Modulation Classifications. Artificial Intelligence. Rayleigh Channel. Correlation. Feature Selection.

Sumário

Capítulo 1	9
Introdução	10
1.1 Justificativa	14
1.2 Objetivo geral	15
1.3 Objetivos específicos	15
1.4 Organização do trabalho	15
Capítulo 2	16
Aprendizado de máquina	16
2.1 Aprendizado supervisionado	17
2.1.1 Algoritmos do aprendizado supervisionado	19
2.2 Aprendizado não supervisionado	21
2.3 Aprendizado por reforço	22
Capítulo 3	23
Classificação de modulações	23
3.1 Introdução à Classificação de modulações.	23
3.2 Tipos de modulações utilizadas no trabalho	24
3.2.1 Sinais modulados em fase	24
3.3 Características do sinal	29
3.3.1 Características espectrais do sinal	29
3.3.2 Desvio padrão da fase instantânea absoluta não linear centralizada	29
3.3.3 Desvio padrão da fase não linear direta centralizada	30
3.3.4 Desvio padrão do valor absoluto da amplitude do sinal normalizada e centralizada	30
3.3.5 Desvio padrão da frequência instantânea normalizada e centralizada	30
3.3.6 Desvio padrão absoluto da amplitude instantânea normalizada e centralizada	31
3.3.7 Características baseadas em filtros	31
3.3.8 Características baseadas em estatísticas de séries	31
Capítulo 4	33
Simulações e resultados	33
4.1 Esquemático de simulações	33
4.2 Extração das características.	37
4.3 Projeto do classificador	37
4.4 Seleção das características	40
Capítulo 5	47
Conclusões e estudos futuros	47
5.1 Conclusões	47
5.2 Estudos futuros	49
Referências	50

Lista de figuras

Figura 1: Métodos de aprendizado de ML.	17
Figura 2: Processo de aprendizado supervisionado.	18
Figura 3: Árvore de Decisão.	20
Figura 4: Árvore de Decisão.	21
Figura 5: Constelação BPSK.	24
Figura 6: Modulador BPSK.	25
Figura 7: Constelação QPSK.	26
Figura 8: Modulador QPSK.	26
Figura 9: Constelação QAM.	28
Figura 10: Modulador QAM.	28
Figura 11: Setup de simulação.	33
Figura 12: Sinais BPSK e QPSK com 30dB.	34
Figura 13: Sinais 8PSK e 16QAM com 30dB.	34
Figura 14: Sinal 32QAM com 30dB.	34
Figura 15: Sinal BPSK E QPSK após o canal AWGN.	35
Figura 16: Sinal 8PSK E 16QAM após o canal AWGN.	35
Figura 17: Sinal 32QAM após o canal AWGN.	35
Figura 18: Sinal BPSK E QPSK após os canais AWGN e Rayleigh.	36
Figura 19: Sinal 8PSK E 16QAM após os canais AWGN e Rayleigh.	36
Figura 20: Sinal 32QAM após os canais AWGN e Rayleigh.	36
Figura 21: Matriz confusão para o algoritmo <i>Random Forest</i>	39
Figura 22: Matriz confusão para o algoritmo <i>Adaboost</i>	39
Figura 23: Matriz confusão para o algoritmo SVM.	39
Figura 24: Matriz de Correlação.	43
Figura 25: Matriz confusão para o algoritmo <i>Random Forest</i>	45
Figura 26: Matriz confusão para o algoritmo <i>Adaboost</i>	45
Figura 27: Matriz confusão para o algoritmo SVM.	45

Lista de Tabelas

Tabela 1: Características dos sinais	37
Tabela 2: <i>Random Forest</i>	38
Tabela 3: SVM	38
Tabela 4: <i>Adaboost</i>	38
Tabela 5: RFE	41
Tabela 6: Características selecionadas pelo RFE	41
Tabela 7: Correlação	42
Tabela 8: Características com mais de 0,5 de correlação	44
Tabela 9: Características selecionadas nos dois métodos	44

Lista de abreviações

2G	<i>Segunda Geração</i>
3G	<i>Terceira Geração</i>
4G	<i>Quarta Geração</i>
5G	<i>Quinta Geração</i>
AI	<i>Artificial Intelligence</i>
AMPS	<i>Sistema Avançado de Telefonia Móvel</i>
AMC	<i>Automatic Modulation Classification</i>
AWGN	<i>Additive white Gaussian noise</i>
BPSK	<i>Binary Phase Shift Keying Modulation</i>
CDMA	<i>Code-Division Multiple Access</i>
GSM	<i>Global System for Mobile communication</i>
LTE	<i>Long Term Evolution</i>
ML	<i>Machine Learning</i>
OFDM	<i>Orthogonal frequency-division multiplexing</i>
PCA	<i>Principal Component Analysis</i>
PCM	<i>Pulse Code Modulation</i>
QPSK	<i>Quadrature Phase Shift Keying</i>
QAM	<i>Quadrature Amplitude Modulation</i>
RFE	<i>Recursive Feature Extraction</i>
SVM	<i>Support Vector Machines</i>
TDMA	<i>Code-Division Multiple Access</i>
VOIP	<i>Voice Over Ip</i>
WCDMA	<i>Wideband Code Division Multiple Access</i>

Capítulo 1

Introdução

As telecomunicações são um conjunto de tecnologias capazes de eliminar distâncias entre cidades, países e continentes. Por séculos, informações eram transportados por mensageiros ou correspondências, em cavalos, carruagens, barcos e, até mesmo, por sinais de som ou de fogo. Com o advento das telecomunicações, uma mensagem não precisa mais de um mensageiro, e ao invés de demorar meses, elas chegam ao destino em alguns segundos [1].

A palavra telecomunicações tem sua origem no Latim: *communicatio*, que era o processo de troca de informações, a esse termo foi adicionado o prefixo *tele*, que vem de distância. Ela foi criada por Edouard Estaunie (1862-1942) no seu livro *Pratique de télécommunication électrique (télégraphie-téléphonie)* [1].

É considerado que as telecomunicações modernas têm seu início com a criação dos telégrafos ópticos, que também eram chamados de comunicação visual ou semafórica. Tais telégrafos eram compostos por um ou mais aparelhos que ficavam no topo de edifícios em cadeia que repetiriam a mensagem da anterior, fazendo com que ela chegasse de uma maneira muito mais rápida do que através de um mensageiro. O primeiro telégrafo óptico foi criado em 1792 pelo francês Claude Chappe (1763-1805), e se espalhou rapidamente pela França [2].

O telégrafo óptico por algum tempo dominou as comunicações na Europa. Porém, com o advento de descobertas relacionadas ao eletromagnetismo no século XIX, tais como as descobertas sobre a corrente elétrica de André-Marie Ampère (1775-1836), sobre o campo eletromagnético, por Christian Oerstedt (1777–1851), a Lei de Ohm e algumas invenções como o Galvâmetro [3], tornaram possível a criação do telégrafo elétrico [4].

Apesar de existirem várias propostas de telégrafos elétricos na época, o modelo mais aceito foi o proposto por Carl Friederich Gauss e Wilhelm Weber. Tal modelo utilizava apenas um circuito que gerava pulsos elétricos positivos ou negativos, os quais poderiam ser detectados por galvanômetros a mais de um quilômetro de distância do aparelho [1].

Na Europa e na América surgiram várias outras propostas de telégrafos elétricos, como os propostos por Charles Wheatstone e por Carl August von Steinheil. Porém, todos tinham algum problema, fosse a codificação antiquada do telégrafo de Steinheil ou a quantidade de agulhas do de Wheatstone [1].

Foi patenteado em 1837, pelo físico e pintor Samuel Finley Breese Morse e pelo seu assistente Alfred Vail, um novo conceito de telégrafo, no qual, surgiu o famoso código Morse. Outra inovação importante desse telégrafo foi a sua automação. É importante ressaltar que todos os telégrafos da época precisavam de uma pessoa escrevendo as mensagens quando necessário, enquanto no de Morse a mensagem escrita automaticamente em uma fita de papel, podendo ser decodificada a qualquer momento [5]

Em 1861, cerca de 20 anos depois, foi construído o primeiro sistema que transmitia mensagens, da costa Leste à costa Oeste dos Estados Unidos. Com isso, tornou-se possível enviar mensagens por todo o continente através de postes, o que anteriormente era feito a cavalo, podendo levar até 10 dias. Porém, mensagens entre continentes ainda não eram possíveis [1].

Após alguns anos, o inventor Werner Von Siemens, patenteou o 12º telégrafo registrado, mas o seu telégrafo tinha o diferencial de ser patenteado junto com cabos que poderiam ser instalados e isolados em ambientes subterrâneos, incluindo na água. Com a ajuda de Siemens e de outros intelectuais da época, como Michael Faraday e o Lorde Kelvin, foi possível passar cabos submarinos sob oceanos para a transmissão de telégrafos. No início do século XX já existiam vários cabos submarinos na costa mundial e interligando continentes [1].

Enquanto vários cientistas estavam com sua atenção voltada para os telégrafos, alguns poucos tinham seus estudos voltados para a transmissão de sinais sonoros. O inglês Robert Hooke foi o primeiro a sugerir que seria possível a transmissão de sinais de som através de longas distâncias.

A principal invenção para a transmissão de voz ocorreu em 1876, pelo americano Alexander Graham Bell. Uma curiosidade sobre a patente de Bell, é que ele descreve seu aparelho como uma evolução da telegrafia, não citando as palavras voz e telefone em nenhum momento do seu registro. A telefonia evoluiu muito no fim do século XIX, tendo registrados mais de 2 milhões de aparelhos no ano de 1900 [1], [6].

Na mesma época, vários cientistas estudaram a descoberta do campo Eletromagnético de Oerstedt, em 1820. Michael Faraday, 12 anos depois, confirmou a descoberta de Oerstedt, e previu a existência de ondas eletromagnéticas, incluindo as linhas de força terminadas em cargas elétricas ou em polos magnéticos. O americano Joseph Henry, também em 1832, descobriu o fenômeno da indução eletromagnética enquanto construía eletroímãs, porém, seu trabalho só foi publicado em 1843, e foi usado para gerar altas tensões pelo alemão Heinrich Daniel Ruhmkorf [1].

O primeiro registro de comunicações elétricas sem fio ocorreu em 1866, com o americano Mahlon Loomis, que chamou sua invenção de telégrafo aéreo. Vários outros cientistas tentaram desenvolver aparelhos utilizando a Lei de Faraday, mas, só com as descobertas de Heinrich Rudolf Hertz que foram possíveis maiores avanços na comunicação sem fio [7].

Hertz nasceu em Hamburgo no ano de 1857 e foi um grande estudioso das ciências naturais; estudou em Berlin com dois físicos muito famosos a época: Helmholtz e Kirchoff. Em 1888, ele publicou um artigo chamado *Ondas Eletromagnéticas no Ar e sua reflexão*, descobrindo, então, as ondas eletromagnéticas que comprovavam as teorias de Maxwell. Com essa descoberta, Hertz alcançou reconhecimento internacional, além de ter sido o responsável pela origem de várias inovações com tecnologias sem fio [7].

Após o período da descoberta das ondas eletromagnéticas, por alguns anos houve comprovações de suas descobertas, mas nenhuma construção relevante as utilizando [1].

O primeiro rádio prático é datado de 1896, pela patente do italiano Guglielmo Marconi, em um sistema que operava em aproximadamente 2,5MHz e transmitia em uma potência entre 10 W e 20 W [8]. No início do Século XX, com o advento da eletrônica, foram permitidos vários avanços nas tecnologias existentes, como: telefonia de longa distância, multiplexação na telefonia e canais de telégrafos, retransmissão de rádio e radiotelefonia com transmissores de ondas contínuas [1].

Então, depois da Segunda Guerra Mundial, quando grande parte da infraestrutura de telecomunicações da Europa e da Ásia foi destruída, o mundo precisou se reinventar e, conseqüentemente, foram desenvolvidas novas tecnologias dentro das redes de telecomunicações, sendo que estas podem ser divididas, a partir daqui, em dois tipos: as redes com fio e as redes sem fio. Neste trabalho daremos ênfase aos sistemas sem fio.

A primeira descoberta importante na época foram os semicondutores, materiais normalmente sólidos em forma cristalizada, cuja condutividade elétrica se encontra no meio da curva, não sendo um condutor como o cobre [9].

A grande evolução das telecomunicações da segunda metade do Século XX se deve a outra descoberta: o início do uso das transmissões utilizando códigos, sejam eles sem fio, por cabos de cobre ou por Fibras Ópticas. A primeira patente registrada é a patente da Modulação por Código de Pulso (PCM), datada da década de 60.

Também na década de 60, as redes de transporte começaram a ser transportadas por cabos coaxiais através de codificações PCM e “*switchs*” eletromecânicos. Existiam dois principais tipos de canais que utilizavam a hierarquia Digital Plesiócrona, com dois tipos de PCM: o americano, com 24 canais por par de fios e o europeu/asiático, com 30.

Quanto às comunicações sem fio, também houve importantes contribuições nos últimos 50 anos, começando com 1965, quando foi lançado o primeiro satélite comercial, o Intelsat I, que contava com 240 canais de telefonia, um canal de televisão e cobria áreas do Oceano Atlântico.

Também no final da década de 60 foram criadas, no Japão, as primeiras estações de rádio digital, operando em 2GHz e com capacidade equivalente a 240 canais de telefone [1].

Já na década de 70, foram criadas as primeiras redes em pequena escala de telefonia celular analógica, utilizando o AMPS (Sistema de Telefonia Móvel Avançada), desenvolvida pelo Bell Labs, em que as estações bases enviariam e receberiam chamadas de voz, e quaisquer células adjacentes operariam em frequências diferentes, evitando interferências; esse sistema começou a ser comercializado em 1983 [10].

Durante os anos posteriores, vários países desenvolveram tecnologias analógicas de telefonia celular. Assim, no ano de 1990 havia mais de 500 mil assinantes no mundo todo. Porém, já na época, existiam várias pesquisas através da segunda geração de redes móveis, também conhecida como 2G [11].

Na tecnologia 2G, os sistemas de rádio analógico passaram por uma migração para os sistemas digitais, também foi possível transmitir voz e mensagens de texto. Os principais sistemas do 2G foram: IS-136 (TDMA, *Time Division Multiple Acces*), o IS-95 (CDMA, *Code Division Multiple Acces*) e o GSM (*Global System Mobile Communications*, sistema global para comunicações móveis).

Todas essas evoluções resultaram no crescimento mais rápido de aparelhos de telecomunicações, com mais de 80 milhões de usuários com aparelhos celulares registrados no ano de 1995. [11].

No final do século XX, foram iniciados os estudos da terceira geração de redes móveis, ou 3G, em que foram separadas a voz dos dados de rede, além de utilizadas novas técnicas de transmissão de dados como: WCDMA, OFDM, TDMA e CDMA2000. Com o 3G, foi possível se comunicar utilizando o VoIP (Voz sobre IP), foi possível, também, o início do *streaming* de músicas e o acesso remoto a redes através de celulares e notebooks [11].

Com a evolução dos smartphones e de outros aparelhos móveis eletrônicos, foi preciso que os dados fossem priorizados em relação à voz. Com isso, começaram os estudos para a quarta geração de redes móveis, o 4G. Ele utiliza da LTE (Evolução de Longo Termo) e, inicialmente, alcançava valores de *downlink* e *uplink* muito altos, porém, precisava das redes 2G e 3G para utilizar as funções de voz, que passaram a ser integradas à rede com as evoluções VoLTE e o LTE *Advanced* [12].

Para o futuro começa a implementação da quinta geração, o 5G, que passará a ter velocidades muito altas, com latências extremamente baixas e aplicações de Inteligência Artificial, Aprendizado de Máquina e Internet das Coisas, sendo possível, em alguns anos, a criação de casas totalmente automatizadas, com sensores ligados à rede através do 5G [13].

1.1 Justificativa

Com o aumento de usuários nas redes móveis nos últimos anos, uma das maiores preocupações dos governos e da academia é em relação ao espectro de frequência, já que esse é um recurso limitado, e a maioria das frequências utilizáveis já estão em uso em múltiplas aplicações. Assim, a atribuição fixa, na qual uma faixa de frequência é reservada para um usuário, pode limitar o uso de múltiplos usuários, subutilizando um recurso muito importante.

Portanto, é importante identificar técnicas para facilitar a alocação dinâmica de frequência, para que a rede se torne autônoma e seja possível utilizar todos os recursos do espectro. Um dos desafios para essa alocação, é identificar o tipo de modulação que está vindo de uma transmissão. Nisso, a Classificação Automática de Modulações se mostra um agente relevante para a evolução das redes móveis.

1.2 Objetivo geral

O objetivo geral deste trabalho é estudar a classificação automática de modulações digitais por meio do aprendizado de máquinas, que sofreram algum tipo de degradação, após propagarem pelos canais AWGN e Rayleigh.

1.3 Objetivos específicos

- Revisar a literatura sobre aplicação do aprendizado de máquina, seu histórico, tipos e principais algoritmos.
- Estudar os tipos de modulações utilizadas no trabalho, além de suas características ou *features* e equações.
- Gerar uma base de dados composta por modulações digitais sob os efeitos degradantes dos canais de AWGN e Rayleigh, para o treino e teste.
- Selecionar as melhores características das modulações digitais, com objetivo de otimizar o poder de predição da máquina.

1.4 Organização do trabalho

Este trabalho será dividido em cinco capítulos. No Capítulo 1 foram apresentadas as considerações iniciais, a justificativa e o objetivo; já no Capítulo 2 será realizada uma revisão da literatura, com os principais conceitos do aprendizado de máquina, os tipos de classificação e os principais algoritmos; o Capítulo 3 contém informações sobre a classificação automática de modulações, as principais características dos sinais, além dos tipos de modulação utilizados no trabalho; o Capítulo 4 descreverá como foram realizadas as simulações e os resultados; por fim, o Capítulo 5 apresentará as discussões sobre os resultados e as propostas de estudos futuros.

Capítulo 2

Aprendizado de máquina

De acordo com [14], o aprendizado de máquina (*Machine Learning*, ou ML) é um tipo de técnica de Inteligência Artificial (AI) e de ciência da computação, focado em usar algoritmos e dados para ensinar as máquinas a aprenderem, imitando técnicas humanas, melhorando gradualmente sua precisão.

O principal objetivo é fazer com que sistemas aprendam automaticamente, eliminando o fator humano dessa aprendizagem e utilizando padrões presentes nos dados, através de metodologias estatísticas e de algoritmos feitos especificamente para classificações ou previsões [15].

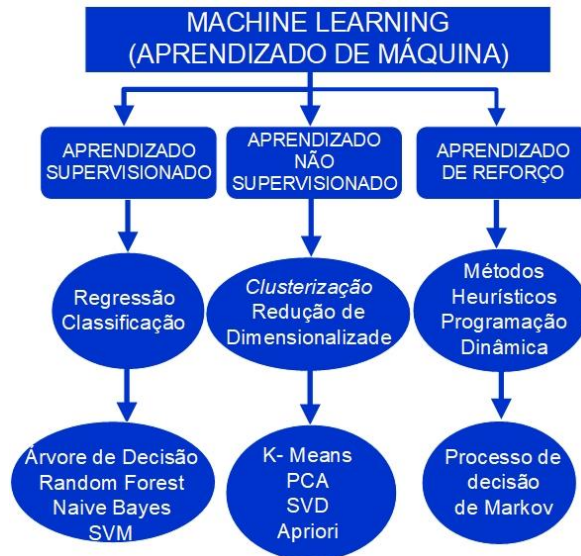
O termo ML surgiu do conceito matemático de modelagem de redes neurais. Um artigo de Walter Pitts e Warren McCulloch [16] tentou mapear matematicamente os processos de pensamento e de decisão humana. Já em 1950, surgiu o famoso teste de Turing, implementado pelo matemático Alan Turing, que media se um computador poderia ou não ter inteligência [17]. Após isso, surgiram vários algoritmos de ML e programas de computador para várias aplicações, desde jogos como damas, até algoritmos para planejar viagens durante campeonatos longos de esportes [18].

Um caso emblemático de evolução do aprendizado de máquina foi o jogo de xadrez entre o supercomputador da IBM, *Deep Blue*, e o campeão mundial de xadrez da época, Gary Kasparov, que ocorreu em duas séries diferentes, em 1996 e 1997 [19]. Na primeira o enxadrista foi o vencedor; já na segunda, o computador venceu por 3,5 a 2,5. Hoje em dia, não existem mais partidas de homem contra máquina, pois os algoritmos são tão evoluídos que mesmo os Grandes Mestres usam recursos computacionais em seus estudos para evolução [20].

O ML pode ser classificado de duas diferentes formas: pelos tipos de modelos e pelos métodos de aprendizagem, sendo a segunda forma a mais comum.

Os algoritmos de aprendizagem de máquina podem ser classificados em: algoritmo de aprendizado supervisionado, não supervisionado e por reforço. A Figura 1 ilustra um fluxograma que descreve os processos de aprendizado de máquina e seus respectivos algoritmos de classificação.

Figura 1: Métodos de aprendizado de ML.



Fonte: elaborado pelo autor

2.1 Aprendizado supervisionado

O aprendizado de máquina supervisionado é definido pela utilização de grupos de dados já rotulados que auxiliarão na classificação de grupos de dados semelhantes [21]. Antes de tal classificação, faz-se necessário um processo de treinamento supervisionado, que segue alguns padrões e ordens bem conhecidos, os quais podem ser resumidos em quatro passos.

O primeiro passo é a definição do problema, no qual deve-se olhar para os requisitos e as definições do que deve ser solucionado por meio do aprendizado de máquina, incluindo as etapas do problema, os resultados esperados e a resolução do problema aceitável; o segundo passo é a identificação dos dados por meio da utilização métodos estatísticos, para entender quais características podem ser utilizadas para composição da base de dados (*dataset*), a qual será utilizada para o treinamento e teste. Esse processo também é conhecido como *discovery* [22].

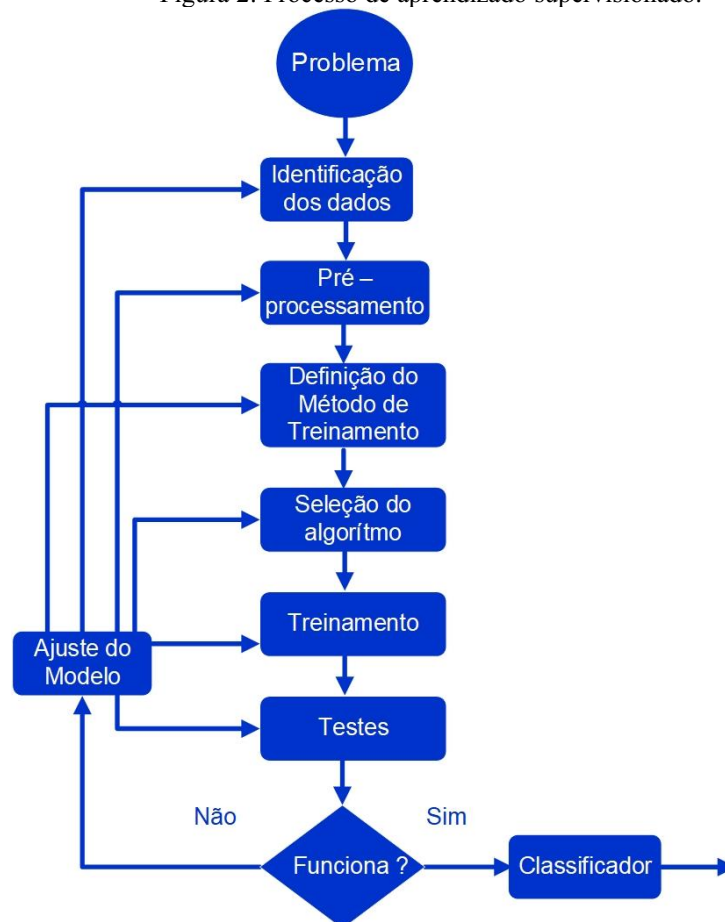
Em um terceiro passo os dados serão preparados e pré-processados. Esse processo pode incluir a construção de um armazenamento de características ou uma base de dados com as características (ou *feature store*) que irão conter possíveis características importantes, reduzindo substancialmente a quantidade de dados a serem processados, fazendo com que os algoritmos operem de forma melhor e mais rápida. Dessa forma, as *features* corretas devem ser selecionadas para que as irrelevantes e redundantes sejam removidas da classificação, eliminando a possibilidade de que os resultados do algoritmo sejam inflacionados, por meio de características dependentes entre si.

Finalmente, no quarto passo deve ser realizado já com a base de dados definida, deve-se definir qual o algoritmo será utilizado para a criação do modelo. A avaliação do algoritmo pode ser feita através de testes preliminares que, em muitas vezes, leva em conta a porcentagem de predições corretas.

Existem duas técnicas principais para calcular a acurácia de um classificador: separar a base de dados em duas partes diferentes, uma parte para teste e a outra para treino; outra técnica é dividir a base de dados em subseções (*subsets*), para cada um é feito um treinamento da união de todos os outros.

Portanto, com a média de erros, é possível estimar a média do classificador. Nessa etapa, pode-se entender se existe algum algoritmo que satisfaz as expectativas do projeto ou não. Em caso negativo é possível voltar nas outras etapas, como o pré-processamento dos dados ou a seleção das características e refazer todas as etapas seguintes. Todos os passos citados anteriormente são ilustrados pela Figura 2.

Figura 2: Processo de aprendizado supervisionado.



Fonte: adaptado de [21].

É importante ressaltar que o aprendizado supervisionado é um dos mais utilizados nos sistemas de ML, logo, existem vários algoritmos que podem ser utilizados como método de aprendizagem, podendo ser técnicas de classificação ou de regressão.

2.1.1 Algoritmos do aprendizado supervisionado

Em uma base de dados utilizada para classificação e, por consequência, para a solução de um problema, existem várias possibilidades de respostas. Dessa forma, deverão ser utilizados algoritmos de classificação capazes de selecionar a melhor solução para tal problema. A seguir serão apresentados os principais algoritmos de classificação utilizados no aprendizado supervisionado.

O algoritmo de *Naive Bayes* é um dos modelos mais populares de ML, foi desenvolvido pelo estatístico Thomas Bayes (1701-1761) e define-se como:

$$P(E)P(H) = P(H)P(E), \quad (2.1)$$

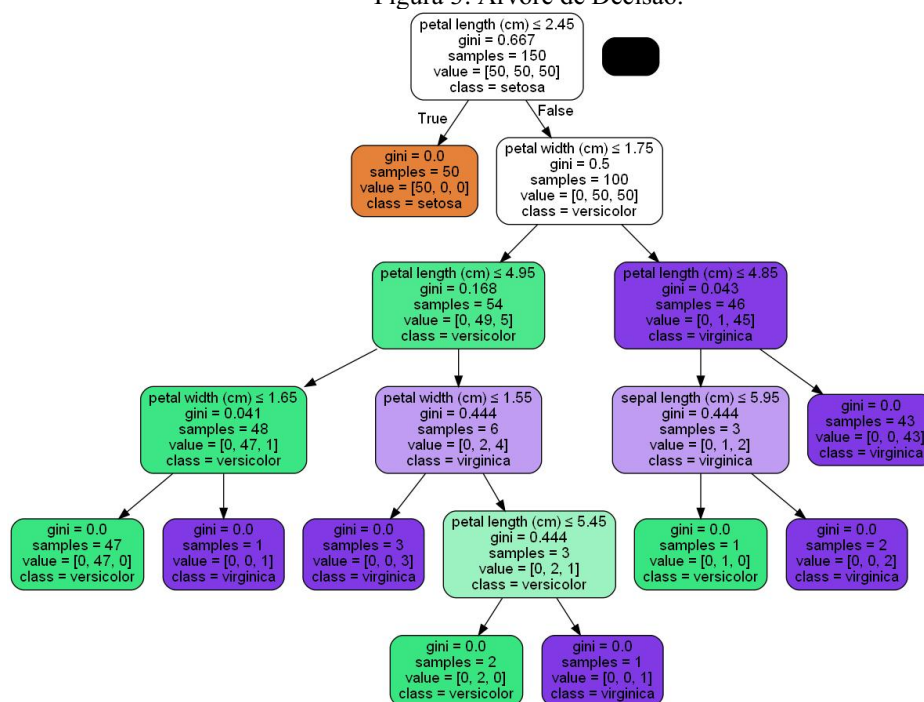
sendo que $P(E)$ $P(H)$ são, respectivamente, as probabilidades de evidência (E) e a hipótese (H) verdadeira.

Como é um teorema conhecido por calcular a probabilidade posterior de acordo com cada hipótese dada pelos dados de treino, a teoria bayesiana é muito utilizada para aplicações simples de aprendizado de máquina, como separar e-mails entre SPAM ou não SPAM [23]. É muito utilizado por não ter um custo computacional alto, além de precisar de poucos dados para ser executado. Entretanto, caso seja necessária a correlação de muitos fatores, o *Naive Bayes* provavelmente irá falhar na sua predição.

Os algoritmos de Árvore de Decisão são construídos para particionar os dados recursivamente até que o resultado esperado seja o da classificação ou da regressão escolhida. Ele é muito utilizado em aplicações não computacionais, como gráficos, fluxogramas, entre outros [24]. Porém, quando se trata de aprendizado de máquina, é um algoritmo muito poderoso, que utiliza de várias características dos dados escolhidos, para que possa fazer seu próprio fluxograma de classificação para o sinal [25].

Para exemplificar melhor o uso do algoritmo, foi utilizado um conjunto de dados da biblioteca *SciKitLearn* do *Python* [26]. Esse conjunto contém informações de largura e comprimento de pétalas e sépalas de três flores diferentes: *Iris setosa*, *Iris virginica* e *Iris versicolor*. Ao combinar essas quatro características, tem-se processo de decisão da árvore, como ilustra a Figura 3.

Figura 3: Árvore de Decisão.



Fonte: elaborada pelo autor.

Na figura acima, pode ser visto como é aplicado o método de classificação da árvore de decisão. Basicamente, de acordo com as características da base de dados, a árvore chega a três grupos diferentes: o de cor laranja, o verde e o roxo. Cada um deles representa um tipo de flor. Quanto mais escura a cor, mais certo o algoritmo está da classificação.

Baseado nas árvores de decisão, um algoritmo muito utilizado em problemas de Classificação é a Random Forest(Floresta Aleatória). Nele, são criadas diversas árvores de decisão aleatórias, e com elas, o algoritmo faz a decisão da classificação. [27]

Além disso, também é possível utilizar algoritmos como o Adaboost, que é um algoritmo que faz a classificação utilizando outros algoritmos, mas vai se adaptando a cada vez que é utilizado, aumentando assim, a acurácia. [28]

Outro algoritmo muito importante no aprendizado supervisionado são as máquinas vetores de suporte (SVM- *Support Vector Machine*), as quais buscam uma linha de separação entre duas diferentes classes, traçando uma linha entre dois pontos de grupos iguais e, com isso, dividindo-os entre diferentes grupos de acordo com suas características. Cada linha gerada pelo SVM é uma divisão diferente por agrupamento de classes, retornando à classificação dos dados [29].

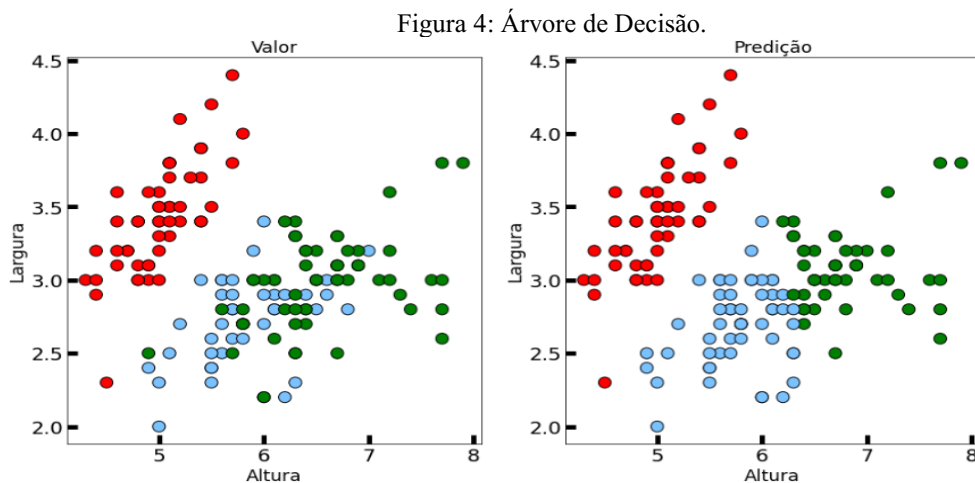
2.2 Aprendizado não supervisionado

O aprendizado não supervisionado consiste em receber dados não classificados, e utilizar métodos de divisão por grupos (*cluster*) para que estes possam ser classificados de acordo com suas características.

Nesse caso, não existe um processo de treinamento e teste. São algoritmos muito utilizados para a classificação automática de modulação, pois muitas vezes o processo para definição do algoritmo requer técnicas e conceitos mais simples do que o aprendizado supervisionado.

Um dos principais algoritmos não supervisionados é o *K-Means*, sendo um algoritmo bastante utilizado para separação de dados. Ele consiste em dividir os dados em K grupos, baseados na distância entre K inicial (que é aleatório) e suas classes. Esse processo é repetido inúmeras vezes até que restem apenas algumas pequenas médias a serem consideradas.

Apesar de ser um algoritmo muito utilizado em aplicações, ele tem uma certa complexidade computacional, dependendo da dimensionalidade dos dados e do número de interações necessárias para que o algoritmo tenha respostas aceitáveis [30]. Assim como foi feito no exemplo do algoritmo da Árvore de Decisão, a Figura 4 ilustra a divisão em clusters dos três tipos de pétalas, em que o algoritmo sem treinamento classificou as flores em seus três tipos, sendo a taxa de acerto foi muito satisfatória.



Fonte: elaborada pelo autor

Como pode ser visto na Figura 4, os tipos de pétalas são automaticamente divididos em três grupos, que dependem de características específicas de cada uma, e existe uma taxa de acerto muito alta. Cada cor representa um *cluster*, ou seja, um grupo, que seria um tipo de pétala.

Um outro algoritmo utilizado é o de análise do principal componente (PCA-*Principal Component*), sendo um método que reduz a dimensionalidade de dados muito grandes, transformando-os em diferentes variáveis que contenham a maior parte da informação no grupo [31]. Com a redução dos dados, sempre ocorre a redução da acurácia do modelo, ou seja, o ideal é sempre conseguir reduzir seu grupo de maneira que ele não perca muitas informações importantes. O aprendizado não supervisionado por meio dos algoritmos de PC e Apriori pode ser utilizado em aplicações de vendas e mineração de dados.

De acordo com [32], o algoritmo Apriori foi proposto em 1994 para minerar diferentes base dados utilizando associações booleanas. Tal algoritmo é baseado no uso do conhecimento das propriedades da base de dados, utilizando um método iterativo para agrupar todos os itens, e, então, o algoritmo retorna a melhor ordem dos dados, de acordo com as regras de associação utilizadas no método.

2.3 Aprendizado por reforço

O aprendizado por reforço é o tipo de aprendizado em que a máquina utiliza todos os resultados para alavancar o aprendizado, algumas vezes por métodos de tentativa e erro para encontrar um encaixe ideal [33]. Aqui será retratado o algoritmo do processo de decisão de Markov.

Esse algoritmo é utilizado em situações em que várias decisões devem ser tomadas sequencialmente, e seu resultado nem sempre é claro. Utilizando o resultado estatístico das decisões anteriores, é possível determinar algum tipo de algoritmo ou ordem que faça com que o resultado seja maximizado. Pode ser utilizado em sistemas de empresas que precisam decidir com frequência o preço de mercadorias e quando podem ser implementadas promoções [34].

Capítulo 3

Classificação de modulações

3.1 Introdução.

De acordo com [35], por muitas décadas a classificação automática de modulação (AMC- *Automatic Modulation Classification*), foi uma preliminar de várias aplicações, como sensoriamento de espectro, recuperação de sinais e seleção da modulação [36].

A AMC tem aplicações em todos os âmbitos da sociedade, servindo em importantes tecnologias já existentes. De acordo com [37], durante períodos de guerra, as aplicações militares de telecomunicações têm três componentes principais: o suporte, os ataques e a proteção. Para ataques, como citado em [38], é possível gerar uma rede semi-supervisionada para falsificar sinais modulados, e atacar outras aplicações que utilizam modelos de AMC, inserindo um sinal parecido na mesma frequência para que ele não consiga fazer a transmissão corretamente. É possível usar a AMC para proteção, pois com ela se torna mais fácil mudar várias características do sinal para evitar a interceptação deste. No dia a dia também é possível encontrar aplicações de AMC, como nas redes de telecomunicações.

As redes de comunicações móveis, LTE e 5G, também são aplicações da AMC em nosso cotidiano. Tais sistemas necessitam de uma gerência (*control software*) responsável pelo controle, configuração, análise de desempenho e qualidade dos serviços. Tais parâmetros podem ser obtidos pelo indicador de qualidade do canal (CQI - *Channel Quality Indicator*), que indica, de acordo com a distância, qual a melhor opção de modulação, taxa de transmissão e código corretor erro devem ser utilizados. Como citado em [39], com o lançamento do 5G para o mundo todo, já se iniciam discussões para o 6G. As cidades inteligentes e sinal inteligente continuam sendo alguns principais requisitos, nos quais, a AMC terá um papel muito importante para o sucesso do futuro dos sistemas de comunicações móveis.

Por ser um trabalho de AMC, nos próximos tópicos serão abordados os tipos de modulações utilizadas neste trabalho, além das características do sinal utilizadas para sua classificação automática.

3.2 Tipos de modulações utilizadas no trabalho

Vale ressaltar que, neste trabalho, foram utilizadas várias modulações digitais, em sua maioria as de comunicações móveis, podendo o processo de modulação ocorrer em fase ou em amplitude.

3.2.1 Sinais modulados em fase

A modulação chaveada em fase (*PSK: phase shift keying*) descreve vários tipos de sinais modulados e utilizados em larga escala nas telecomunicações. Neste trabalho foram abordadas as duas principais constelações do PSK: o binário (BPSK) e o em quadratura (QPSK).

O sinal BPSK é a representação de um dado binário por dois sinais em diferentes fases, que normalmente são 0 e π . Esses sinais podem ser representados pelas equações:

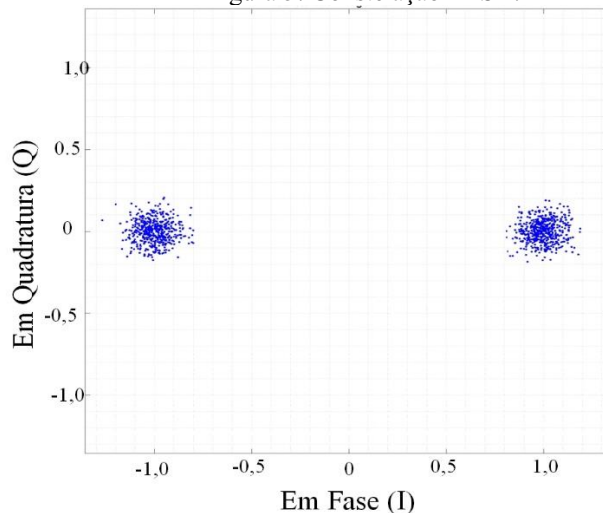
$$s_1(t) = A \cos 2\pi f_c t, 0 \leq t \leq T, \text{ para o bit } 0, \quad (3.1)$$

e

$$s_2(t) = -A \cos 2\pi f_c t, 0 \leq t \leq T, \text{ para o bit } 1, \quad (3.2)$$

sendo A e f_c a amplitude e a frequência do sinal, respectivamente. Esses sinais são opostos e chamados de antipodais. O sinal PSK pode ser representado graficamente por uma constelação em um sistema cartesiano de duas coordenadas, que é ilustrado na Figura 5.

Figura 5: Constelação BPSK.



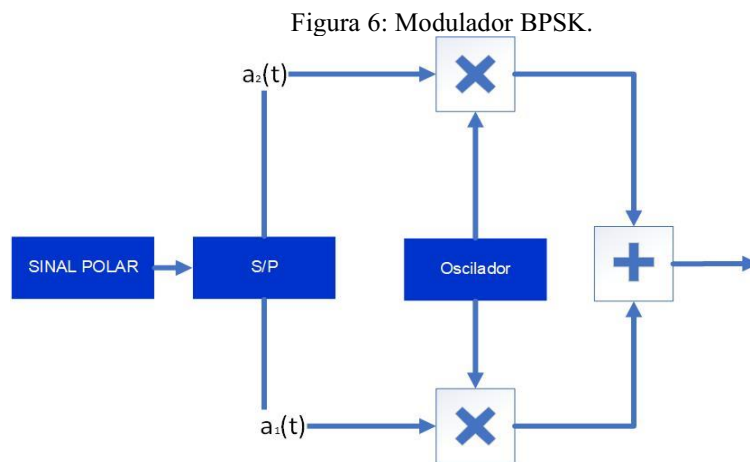
Fonte: elaborada pelo autor

Como foi ilustrado na Figura 5, a constelação é um sistema de coordenadas de duas dimensões, que pode ser mostrado por dois ângulos φ , cada um deles representando um bit transmitido, sendo definidos por:

$$\varphi_1(t) = \sqrt{\frac{2}{T}} \cos(2\pi f_c t), 0 \leq t \leq T \quad (3.3)$$

$$\varphi_2(t) = \sqrt{\frac{2}{T}} \cos(2\pi f_c t), 0 \leq t \leq T \quad (3.4)$$

Dito isso, vale ressaltar que um sinal PSK é formado pelo produto de um sinal codificado em linha e por um sinal conhecido como portadora, gerado com base em um oscilador local, como ilustra a Figura 6.



Fonte: elaborada pelo autor

Considerando um sinal digital, portanto um trem de pulsos, representado por:

$$a(t) = \sum_{k=-\infty}^{\infty} a_k p(t - kT), \quad (3.5)$$

sendo $p(t)$ o pulso retangular definido entre 0 e T . Tem-se, então, o sinal resultante BPSK:

$$s(t) = Aa(t) \cos 2\pi f_c t, -\infty < t < \infty. \quad (3.6)$$

O outro sinal modulado utilizado no trabalho é uma derivação dos sinais modulados em fase ou PSK. Nesses tipos de sinais, a modulação ocorrerá por desvio de fase e em quadratura (QPSK - *Quadrature Phase Shift Keying*). Tais sinais são muito utilizados por não sofrerem tanta degradação de taxa de erro de bit, já que sua banda de eficiência é maior e são definidos por (3.7)

$$s_i(t) = A \cos(2\pi f_c t + \theta_i), 0 \leq t \leq T, \text{ para } i = 1, 2, 3 \text{ e } 4, \quad (3.7)$$

sendo cada fase representada em um local da constelação, definido por:

$$\theta_i = \frac{(2i - 1)\pi}{4}. \quad (3.8)$$

Os valores de $\frac{\pi}{4}$, $\frac{3\pi}{4}$, $\frac{5\pi}{4}$, e $\frac{7\pi}{4}$ representarão um dabit. Dessa forma, o sinal resultante da modulação QPSK é definido por:

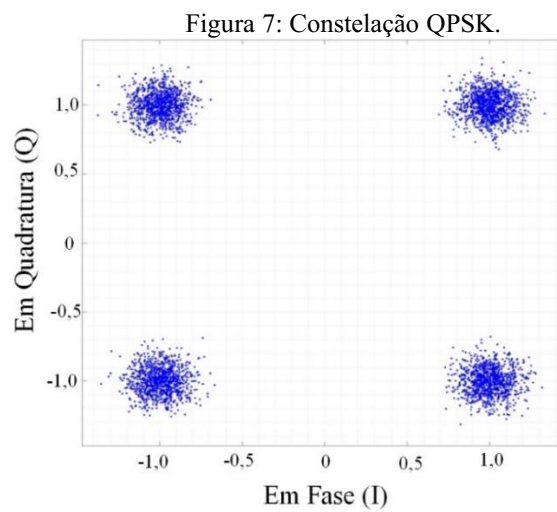
$$s_i(t) = s_{i1}\varphi_1(t) + s_{i2}\varphi_2(t) \quad (3.9)$$

Com s_{i1} e s_{i2} definidos pelas equações abaixo:

$$s_{i2} = \sqrt{E} \text{sen } \theta_i \quad (3.10)$$

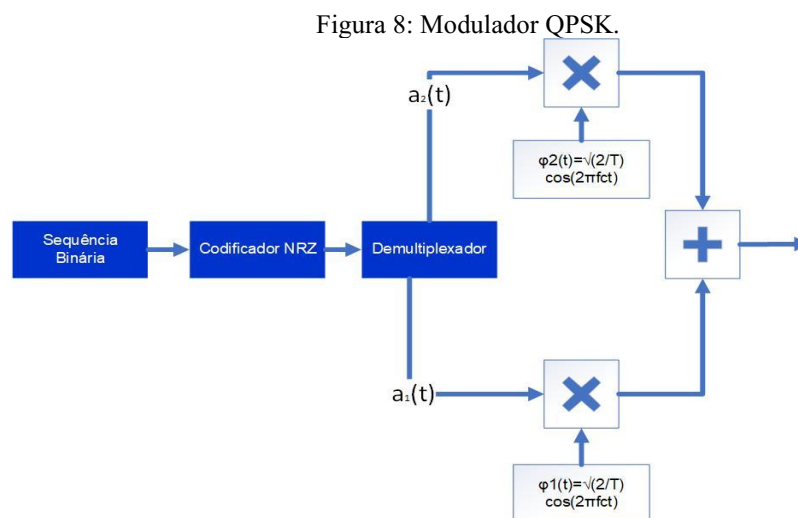
$$s_{i1} = \sqrt{E} \text{cos } \theta_i \quad (3.11)$$

O QPSK é representado por uma constelação em fase e uma quadratura, como ilustra a Figura 7.



Fonte: elaborada pelo autor

Assim como o BPSK, o sinal em quadratura também tem um esquema de modulação bem característico, ilustrado na Figura 8 :



Fonte: elaborada pelo autor

As modulações digitais também podem apresentar desvios de amplitude, dando origem aos famosos sinais de desvios de amplitude em quadratura (QAM - *Quadrature Amplitude Modulation*). Trata-se de uma variação da modulação em Amplitude, a qual é dividida em amplitude e quadratura. Sendo classificado de acordo com a quantidade de símbolos M, podendo ser 4, 16, 32, 64, 128 etc. Neste trabalho foram utilizados os sinais com 8, 16, 32 e 64 símbolos. O sinal modulado em QAM é descrito em:

$$s_i(t) = A_i \cos(2\pi f_c t + \theta_i), i=1, 2, \dots, M. \quad (3.12)$$

sendo A_i e θ_i , respectivamente, a amplitude e a fase do sinal para a modulação de ordem M. O sinal QAM também é muito utilizado com modelagem de pulso, em que ele é multiplicado por um pulso $p(t)$, o que pode ser visto na equação abaixo:

$$s(t) = A_i p(t) \cos(2\pi X f_c t + \theta_i), i=1, 2, \dots, M, \quad (3.13)$$

$$A_{i1} = A_i \cos \theta_i, \quad (3.14)$$

$$A_{i2} = A_i \sin \theta_i, \quad (3.15)$$

$$A_i = \sqrt{A_{i1}^2 + A_{i2}^2}. \quad (3.16)$$

O sinal modulado QAM também pode ser expresso pela Equação (3.9) do QPSK, porém o φ é calculado por:

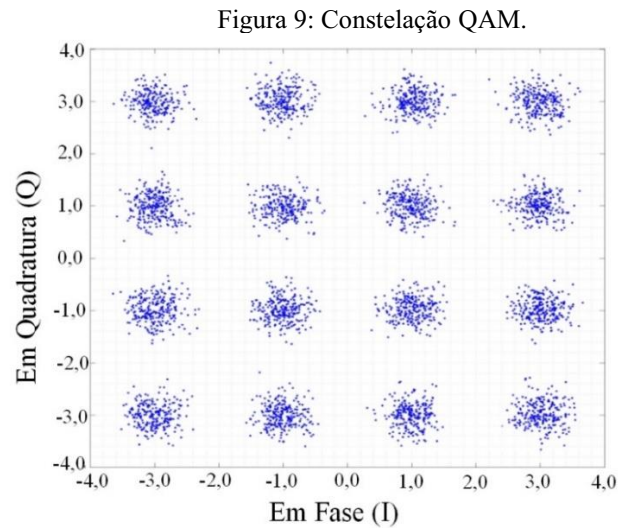
$$\varphi_i(t) = \sqrt{\frac{2}{E_p}} p(t) \cos(2\pi f_c t), 0 < t < T, \quad (3.17)$$

com E_p sendo a energia do pulso $p(t)$. Assim como no QPSK o QAM também pode ser descrito através de uma constelação, como ilustra a

Figura 9, que representa uma constelação de um sinal 16 QAM.

Também é possível calcular a distância entre os fasores das constelações de acordo com o M da modulação, sendo descrito por

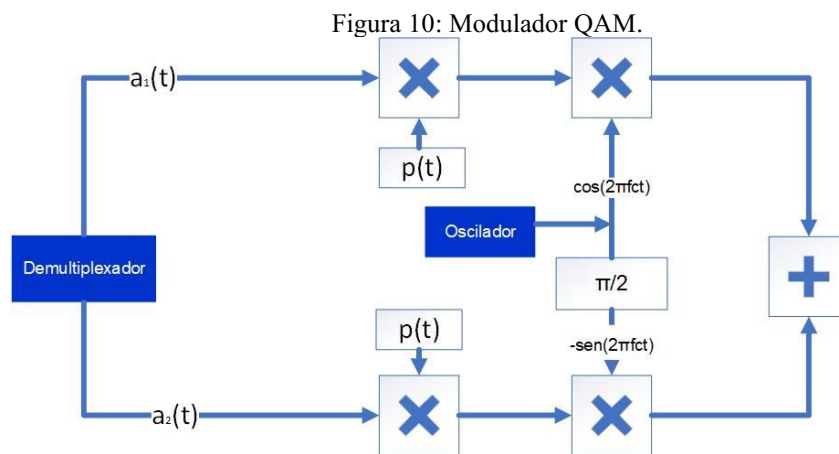
$$d_{ij} = \sqrt{|s_i - s_j|^2}, i, j = 1, 2, \dots, M \quad (3.18)$$



Fonte: elaborada pelo autor

Após a apresentação dos sinais modulados utilizados para o processo de classificação, na próxima seção serão apresentadas as *features* utilizadas no processo de aprendizado de máquina neste trabalho.

O sinal QAM também pode ser representado por um esquema de modulação, ilustrado na Figura 10: Modulador QAM.:



Fonte: elaborada pelo autor

3.3 Características do sinal

Neste trabalho, durante o período de revisão bibliográfica, foram estudadas por volta de 30 características de sinais, as quais são classificadas em estatísticas ou em variadas de séries. A seguir será apresentado um resumo de tais características.

3.3.1 Características espectrais do sinal

Estatisticamente, o desvio padrão de uma distribuição de dados é definido como raiz da média da diferença dos quadrados dos valores do conjunto de dados em relação à média.

$$S^2 = \sqrt{\frac{\sum_i^n (X_i - \bar{X})^2}{n - 1}} \quad (3.19)$$

Neste trabalho foram calculadas várias características com base no desvio padrão, sendo estas: desvio padrão na banda base, no qual, é o cálculo simples do desvio padrão do sinal na banda base, ou seja, o sinal codificado em linha. O desvio padrão na banda passante, sendo o mesmo cálculo do desvio padrão banda base, entretanto, nesse caso, o sinal será deslocado em outra frequência, tratando-se, portanto, do sinal modulado. A seguir serão abordados outros tipos de desvios padrões específicos utilizados como características para a classificação desses sinais.

3.3.2 Desvio padrão da fase instantânea absoluta não linear centralizada

O desvio padrão de fase absoluta instantânea pode ser definido por:

$$\sigma_{ap} = \sqrt{\frac{1}{N_c} \left(\sum_{A_n[n] > A_t} \phi_{NL}^2[n] \right) - \left(\frac{1}{N_c} \sum_{A_n[n] > A_t} |\phi_{NL}^2[n]| \right)^2} \quad (3.20)$$

sendo N_c , A_t , $A_n[n]$ e ϕ_{NL}^2 , respectivamente, o número de amostras, o limite de valores, o número de amostras de amplitude, a amplitude em relação ao tempo e a variação de fase não linear do sinal.

3.3.3 Desvio padrão da fase não linear direta centralizada

Uma outra característica do sinal utilizada neste trabalho, e relacionada ao desvio padrão, é o desvio padrão de fase não linear direta, que pode ser definido por:

$$\sigma_{dp} = \sqrt{\frac{1}{N} \left(\sum_{n=1}^N (A_{cn}^2[n]) \right) - \frac{1}{N_c} \left(\sum_{n=1}^N (A_{cn}^2[n]) \right)^2}, \quad (3.21)$$

sendo $A_{cn}[n]$ o número de amostras centralizadas.

3.3.4 Desvio padrão do valor absoluto da amplitude do sinal normalizada e centralizada

Além dos desvios padrões citados anteriormente, utilizou-se também, neste trabalho, o desvio padrão do valor absoluto da amplitude do sinal normalizado e centralizado.

$$\sigma_a = \sqrt{\frac{1}{N_c} \left(\sum_{A_n[n] > A_t} a_{cn}^2[n] \right) - \left(\frac{1}{N_c} \sum_{A_n[n] > A_t} a_{cn}[n] \right)^2}. \quad (3.22)$$

3.3.5 Desvio padrão da frequência instantânea normalizada e centralizada

Ainda com relação ao estudo de classificação automática de modulações abordado neste trabalho, utilizou-se o desvio padrão em frequência instantânea normalizada e centralizada, definida por:

$$\sigma_{af} = \sqrt{\frac{1}{N} \left(\sum_{A_n[n] > A_t} f_N^2[n] \right) - \left(\frac{1}{N_c} \sum_{A_n[n] > A_t} |f_N[n]| \right)^2}, \quad (3.23)$$

sendo f_N a frequência instantânea normalizada pela frequência de amostragem.

3.3.6 Desvio padrão absoluto da amplitude instantânea normalizada e centralizada

Por fim, no que diz respeito às características dos sinais relacionadas aos desvios padrões, utilizou-se também o desvio padrão absoluto da amplitude instantânea normalizada e centralizada definido por:

$$\sigma_{aa} = \sqrt{\frac{1}{N} \left(\sum_{A_n[n] > A_t} f_N^2[n] \right) - \left(\frac{1}{N_c} \sum_{A_n[n] > A_t} |f_N[n]| \right)^2} \quad (3.24)$$

3.3.7 Características baseadas em filtros

Algumas características foram baseadas em filtros passa baixa de *Nyquist* basicamente no modelo do Cosseno Levantado. Tal equação do filtro é descrita por [40].

$$g(t) = \text{sinc} \left(\frac{t}{T} \right) \sqrt{\cos \left(\frac{\left(\frac{\pi \alpha t}{T_s} \right)}{\left(\frac{1 - 4\alpha^2 t^2}{T^2} \right)} \right)}. \quad (3.25)$$

Com $g(t)$ sendo representado pela função *sinc* de t/T multiplicado pelo cosseno de T_s , e α o fator de *roll-off*. Usando o filtro, foram feitas variações nos seus parâmetros, gerando, com isso, diferentes características do sinal para que este possa ser analisado futuramente.

Após a resposta ao degrau, o sinal era dividido em forma real e imaginária, e sua média calculada, gerando, assim, várias características resultantes do filtro.

3.3.8 Características baseadas em estatísticas de séries

Muitas características do sinal foram retiradas da expansão da Série de Taylor, mais especificamente na cumulante, em que X é uma variável aleatória e E o operador esperança. As cumulantes podem ser descritas como em [41]:

$$K(t) = \sum_{n=1}^{\infty} k_n \frac{t^n}{n!} = k_1 t + k_2 \frac{t^2}{2} \quad (3.26)$$

sendo k_n as cumulantes de ordem n da função $k(t)$. Foram utilizadas as de segunda, terceira e quarta ordem, com filtro de cosseno levantado e sem filtro.

Assim, foram geradas as características das modulações BPSK, QPSK, 8QAM, 16QAM e 32QAM. No próximo capítulo, será discutida a metodologia do trabalho.

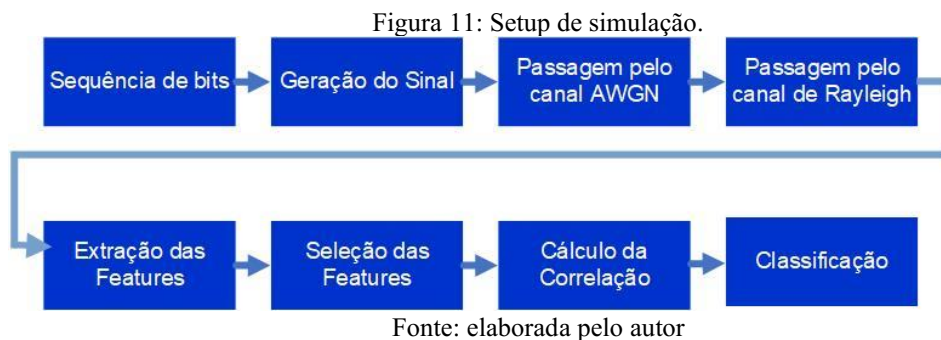
Capítulo 4

Simulações e resultados

4.1 Esquemático de simulações

Neste capítulo será apresentada a metodologia adotada para a realização das simulações e para os resultados do trabalho. Vale ressaltar que, à princípio, utilizou-se o software Matlab, com a licença de número 41013326, para o processo de geração e propagação dos sinais nos canais a AWGN e Rayleigh.

Já a linguagem de programação Python foi utilizada no processo de classificação e extração das características dos sinais, por meio das bibliotecas: *matplotlib*, *pandas*, *sklearn*, *scipy*, *numpy*, *mlxtend* e *seaborn*. Dessa forma, a Figura 11 ilustra o diagrama em blocos do *setup* no processo.

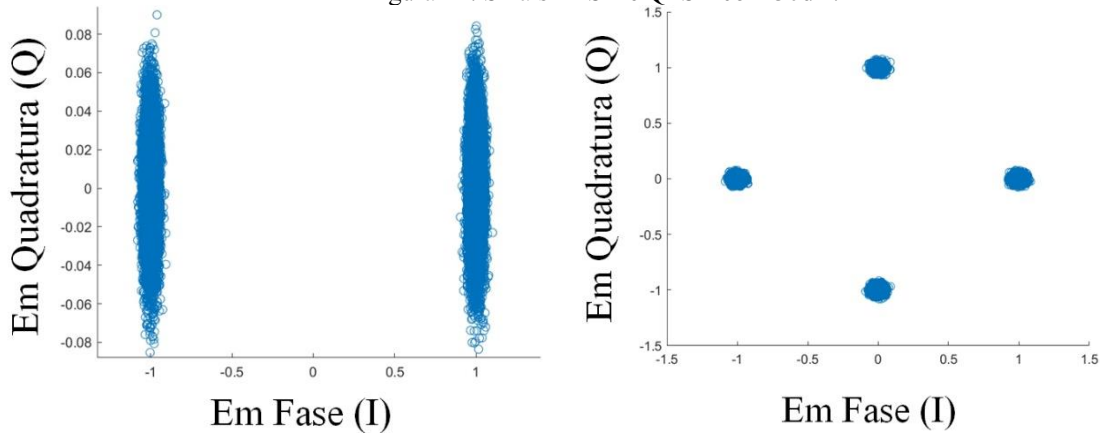


Inicialmente, gerou-se uma sequência de bits aleatórios, que alimentaram as entradas dos moduladores m-PSK ou m-QAM. Após geração de tais sinais, foi adicionado o ruído gaussiano branco, por meio da propagação em um canal AWGN, variando-se a relação sinal ruído de 0 a 30 dB, com intervalos de 2.5 dB.

Posteriormente, os sinais foram propagados por um canal de Rayleigh com atrasos, devido aos percursos iguais a $[0, 200, 800, e 1200] * 10e-9$ segundos, ganhos médios iguais a $[0, -0,9 -4,9, -8]$ dB e o deslocamento de frequência devido ao efeito *Doppler* igual 50 Hz.

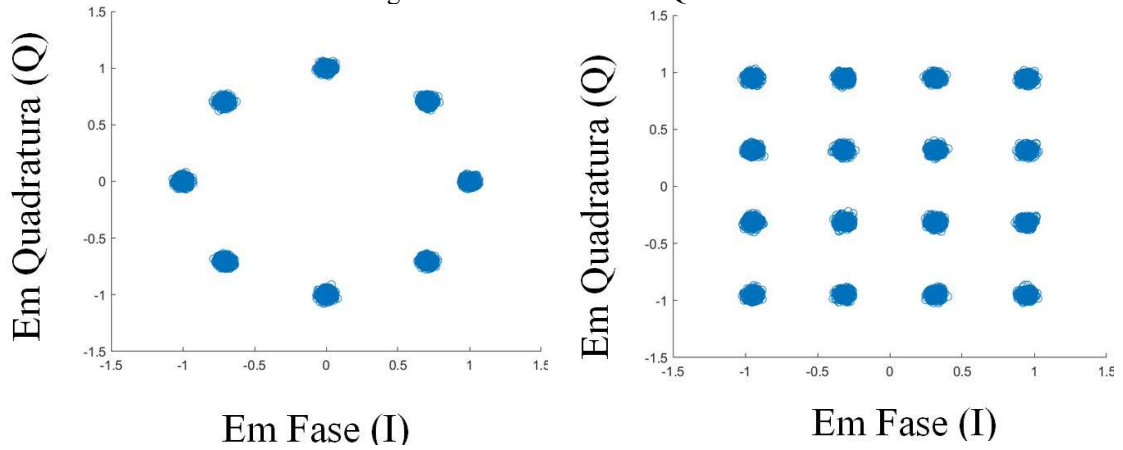
As figuras 12, 13 e 14 ilustram os sinais m-PSK e m-QAM utilizados no processo de classificação. Já as figuras 15, 16 e 17 ilustram tais sinais com a adição do ruído gaussiano branco, oriundo do canal AWGN. E, por fim, as figuras 18, 19 e 20 ilustram os mesmos sinais sobre o efeito da propagação por meio do canal de Rayleigh.

Figura 12: Sinais BPSK e QPSK com 30dB.



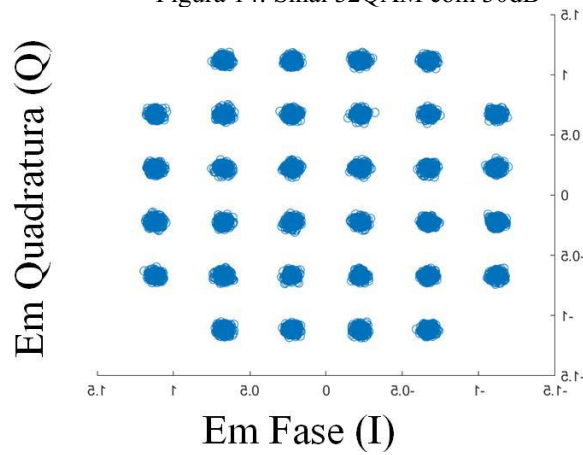
Fonte: elaborada pelo autor

Figura 13: Sinais 8PSK e 16QAM com 30dB



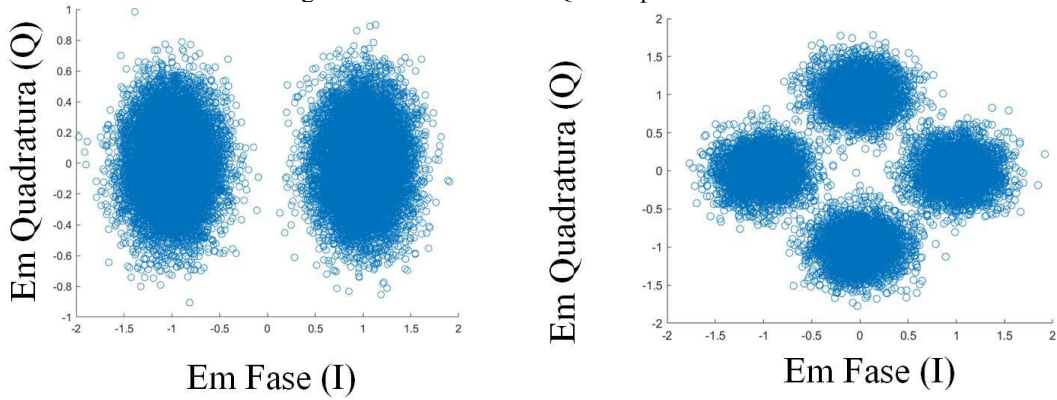
Fonte: elaborada pelo autor

Figura 14: Sinal 32QAM com 30dB



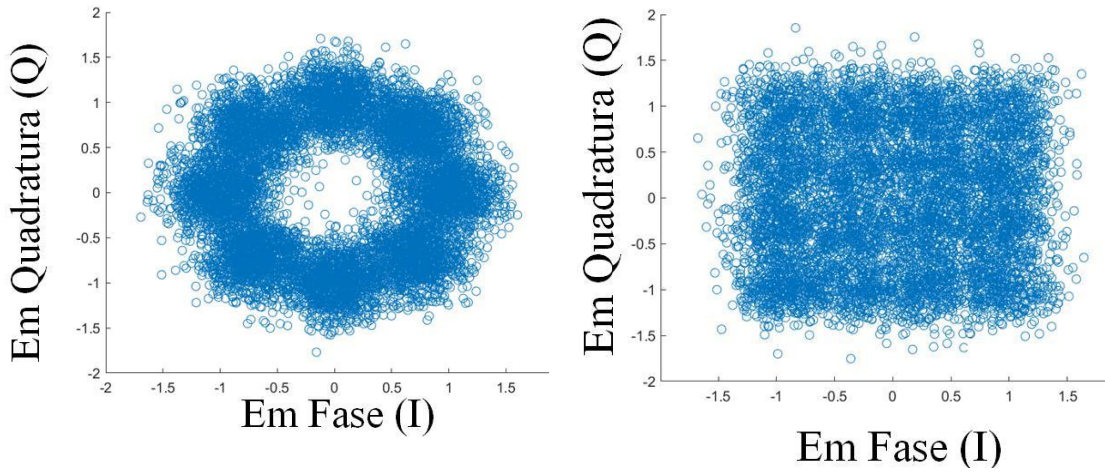
Fonte: elaborada pelo autor

Figura 15: Sinal BPSK E QPSK após o canal AWGN.



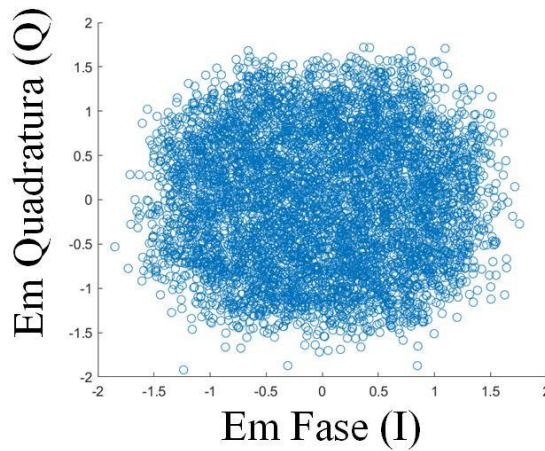
Fonte: elaborada pelo autor

Figura 16: Sinal 8PSK E 16QAM após o canal AWGN.



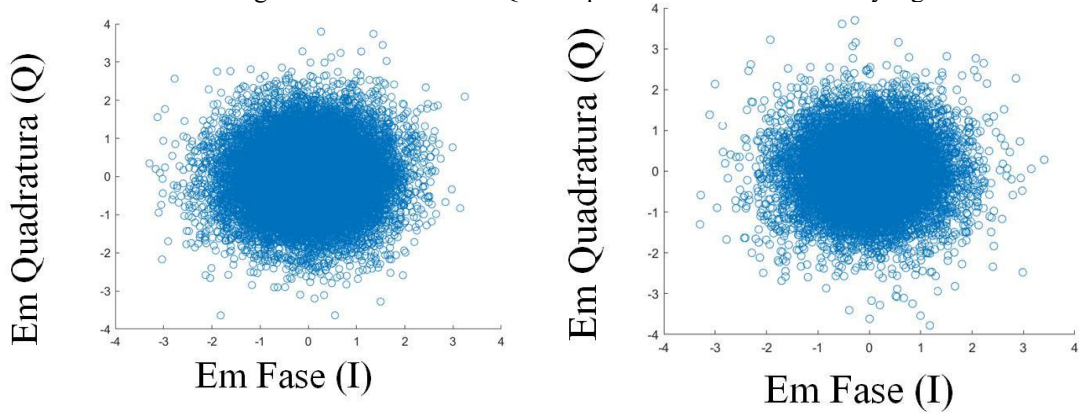
Fonte: elaborada pelo autor

Figura 17: Sinal 32QAM após o canal AWGN



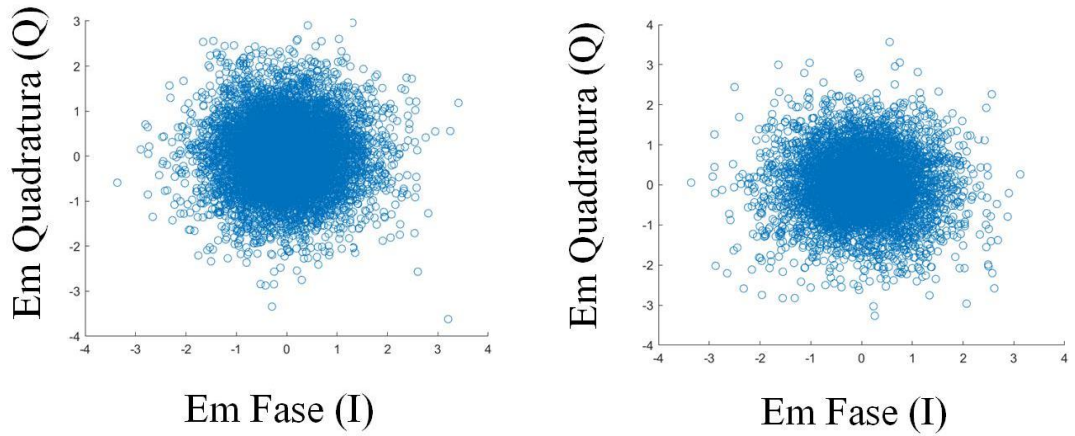
Fonte: elaborada pelo autor

Figura 18: Sinal BPSK E QPSK após os canais AWGN e Rayleigh.



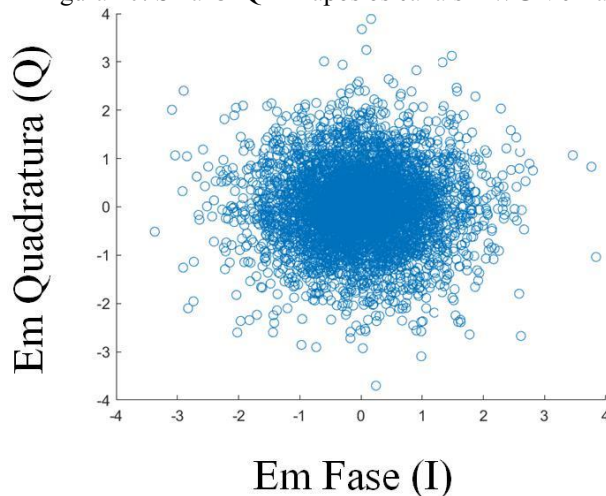
Fonte: elaborada pelo autor

Figura 19: Sinal 8PSK E 16QAM após os canais AWGN e Rayleigh.



Fonte: elaborada pelo autor

Figura 20: Sinal 32QAM após os canais AWGN e Rayleigh.



Fonte: elaborada pelo autor

4.2 Extração das características.

Ainda seguindo o *setup* da Figura 11, após serem gerados, os sinais foram salvos em arquivos de dados, suas *features* foram extraídas ainda utilizando o *matlab*. A Tabela 1 contém as 30 características utilizadas para o processo inicial de classificação, descritas no capítulo 3 deste trabalho [37].

Tabela 1: Características dos sinais

Valor máximo da densidade espectral de potência da amplitude instantânea do sinal.
Desvio padrão da fase instantânea absoluta não linear centralizada
Desvio padrão da fase não linear direta centralizada
Desvio padrão do valor absoluto da amplitude do sinal normalizada e centralizada
Desvio padrão da frequência instantânea normalizada e centralizada
Desvio padrão absoluto da amplitude instantânea normalizada e centralizada
Valor máximo da densidade espectral de potência da amplitude da frequência instantânea do sinal.
Desvio padrão da transformada DCT do sinal
Desvio padrão da transformada rápida de Walsh-Hadamard do sinal
Cumulante de Primeira ordem da parte real da envoltória complexa
Cumulante de Segunda ordem da parte real da envoltória complexa
Cumulante de Terceira ordem da parte real da envoltória complexa
Cumulante de Quarta ordem da parte real da envoltória complexa
Cumulante de Primeira ordem da parte imaginária da envoltória complexa
Cumulante de Segunda ordem da parte imaginária da envoltória complexa
Cumulante de Terceira ordem da parte imaginária da envoltória complexa
Cumulante de Quarta ordem da parte imaginária da envoltória complexa
Cruzamento das Cumulantes de Parte Real e imaginária
Teste da Gaussianidade
Cálculo da Variância e do Desvio Padrão do sinal após o filtro Cosseno Levantado
Média da transformada de Hilbert
Média do sinal após a envoltória complexa
Cumulantes Cíclicas de Segunda, Terceira e Quarta Ordem
Cálculo da Curtose do sinal após o filtro Cosseno Levantado
Soma dos componentes da resposta ao degrau do sinal após o filtro Cosseno Levantado
Resposta ao degrau do sinal após o filtro Cosseno Levantado

4.3 Projeto do classificador

Para o projeto do classificador foram utilizados 280 mil sinais com suas características extraídas, separados em dois grupos: 252 mil (90% dos sinais) foram destinados para treino e 28 mil para testes (10% dos sinais). Assim, foram gerados 40 mil

sinais para cada diferente relação entre sinal e ruído, propagados também pelo canal de Rayleigh.

Após a separação entre treino e teste, os sinais foram separados novamente em dois grupos, nos quais, o primeiro contém os identificadores dos sinais e o tipo de modulação; e o segundo o identificador do sinal e as suas características. Escolheu-se, então, três algoritmos para classificação dos sinais: o algoritmo SVM, o *RandomForest* e o *Adaboost* com um estimador de Árvore de Decisão. As tabelas 2, 3 e 4 mostram, respectivamente, as configurações utilizadas para o *Random Forest*, para o SVM e para o *Adaboost*.

Tabela 2: *Random Forest*

Parâmetros	
Nome	Valor
Estimadores	10
Método de aleatorização	Bootstrap
Número de amostras	280.000
Máximo de Features	30
Profundidade	5
Mínimo de Folhas	1
Critério de qualidade	Gini
Máximo de Folhas	∞

Tabela 3: SVM

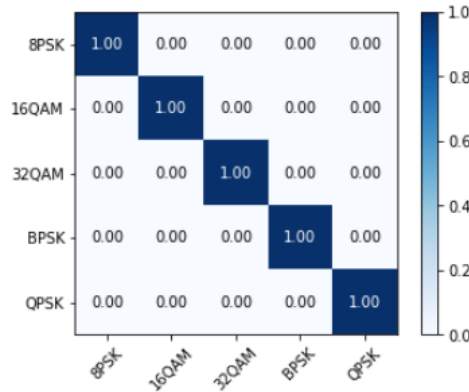
Parâmetros	
Nome	Valor
Grau	3
Gamma	Scala
Função de decisão	Ovo
Kernel	Rbf
Random State	0
Desempate	Falso
Tolerância	0,001

Tabela 4: *Adaboost*

Parâmetros	
Nome	Valor
Estimadores	50
Ritmo de Aprendizado	1
Algoritmo	SAMME
Estimador	Árvore de Decisão

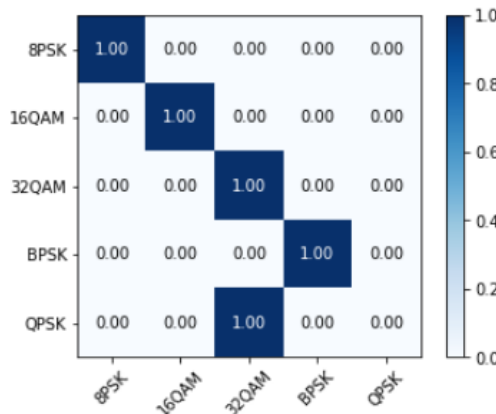
Para a primeira classificação, foram utilizadas todas as 30 características calculadas no trabalho, junto a todos os sinais gerados. As figuras 21, 22 e 23 ilustram, respectivamente, as matrizes confusões referentes aos algoritmos utilizados neste trabalho.

Figura 21: Matriz confusão para o algoritmo *Random Forest*



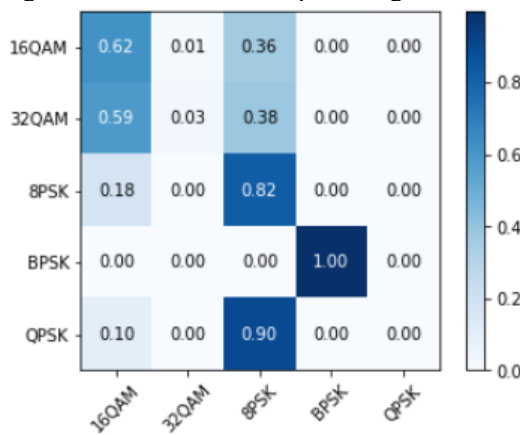
Fonte: elaborada pelo autor

Figura 22: Matriz confusão para o algoritmo *Adaboost*



Fonte: elaborada pelo autor

Figura 23: Matriz confusão para o algoritmo *SVM*



Fonte: elaborada pelo autor

Como pode ser visto, com o algoritmo *Random Forest* a classificação obteve excelentes resultados, com acertos de 100% na classificação das modulações. Entretanto,

quando se verifica a classificação obtida pelos algoritmos *Adabost* e SVM, os resultados não são tão bons. No caso do *Adabost*, a matriz confusão mostra um erro de 100% na classificação entre as modulações QPSK e 32 QAM. Já para a matriz do SVM, pode-se verificar inúmeros erros.

Os erros podem ser atribuídos ao excesso de características extraídas dos sinais, uma vez que nem sempre muitas característica ou *features* garantirão uma performance otimizada na classificação das modulações [42]. Dessa forma, na próxima secção, será executado o processo de seleção das características dos sinais, de maneira a utilizar somente características importantes, com o objetivo de otimizar o processo de classificação de tais modulações melhorando, portanto, o desempenho classificador.

4.4 Seleção das características

Os cálculos das 30 características dos sinais foram realizados no software Matlab, em um computador com processador i7 e 32GB de memória RAM (*Random Access Memory* ou Memória de Acesso Aleatório). O tempo médio de tais cálculos foi de 12 minutos para cada sinal, resultando em, aproximadamente, quatro dias [43].

Para que um modelo de classificação automática de modulação possa ser utilizado em sistemas de telecomunicações, tal classificação, deve ocorrer em tempo real e em pequenos dispositivos, o que demanda um baixo custo computacional e pouco tempo para calcular as características e classificar o sinal, e então passar pelos processos de modulação e de decodificação.

Dessa forma, mesmo que os resultados da classificação automática de modulações referente ao algoritmo *Random Forest* tenham apresentado excelentes resultados, o custo computacional por se utilizar todas as características deve ser levado em consideração. Assim, torna-se necessária uma seleção criteriosa das características dos sinais, com o objetivo de otimizar e melhorar os resultados obtidos dos algoritmos, além do custo computacional. Para tanto, foram utilizados dois métodos para a seleção das características: o algoritmo Recursivo de Eliminação de Características (RFE-*Recursive Feature Elimination*) e a correlação entre as características do sinal e o tipo de modulação.

O algoritmo RFE trabalha em conjunto com a máquina durante o treinamento, por meio de um modelo que procura por atributos da base de treinamento, ranqueando por importância cada característica, e selecionando as mais importantes para que o

modelo possa aprender a classificar os sinais. Além disso, remove as características que afetam o modelo com informações que podem levar a erro, aumentando portanto o nível da qualidade da predição do modelo [44].

Para esse método, escolheu-se utilizar o treinamento do *Random Forest*, por ter apresentado o melhor desempenho no processo de classificação, algo crucial e que deve ser levado em consideração. Além disso, o algoritmo *Adaboost*, por utilizar uma árvore de decisão, pode se beneficiar, devido à origem dos dois algoritmos ser a mesma [45]. Assim utilizou-se a função RFE da biblioteca do *SKLearn* do Python; suas configurações podem ser vistas na Tabela 5.

Tabela 5: RFE

Parâmetros	
Nome	Valor
Estimador	<i>Random Forest</i>
Verbosidade	0
Importância das Características	Automática

Para o *setup* do algoritmo, foram selecionadas as 12 características (ou 40%) consideradas pelo algoritmo mais relevantes durante o processo de treinamento. As outras características foram ranqueadas entre os números dois e sete, porém, neste trabalho serão utilizadas apenas as 15 ranqueadas como o número 1, que podem ser vistas na Tabela 6.

Tabela 6: Características selecionadas pelo RFE

Desvio padrão da fase instantânea absoluta não linear centralizada
Desvio padrão do valor absoluto da amplitude do sinal normalizada e centralizada
Desvio padrão da frequência instantânea normalizada e centralizada
Desvio padrão da transformada DCT do sinal
Desvio padrão da transformada rápida de Walsh-Hadamard do sinal
Teste da Gaussianidade
Desvio Padrão do sinal após o filtro Cosseno Levantado
Média da Frequência instantânea do sinal
Cumulante de Terceira ordem da parte imaginária da envoltória complexa
Média
Média da transformada de Hilbert

Além da seleção pelo método RFE, realizou-se também uma seleção de características por meio da utilização da correlação estatística entre as características, comparando-as com o tipo de modulação. De acordo com [46], o sistema que combina tais características com o maior grau de correlação com as de maior ranking na seleção feita pelo método RFE, apresentará um custo computacional menor, aumentando a performance da predição [47].

Incluiu-se, então, em uma tabela todas as informações do sinal, suas características e tipos, com o objetivo de se calcular a correlação por meio da função corr do pandas, com os parâmetros na Tabela 7.

Tabela 7: Correlação

Parâmetros	
Nome	Valor
Método	Pearson
Períodos Mínimos	∞

Pode-se notar na tabela acima que é utilizado o método de Pearson para o cálculo da correlação. O seu valor pode variar entre -1 e 1, em que um valor negativo indica uma relação inversa, e um valor positivo uma correlação, enquanto números próximos a 0 não representam nenhum tipo de relação. [48].

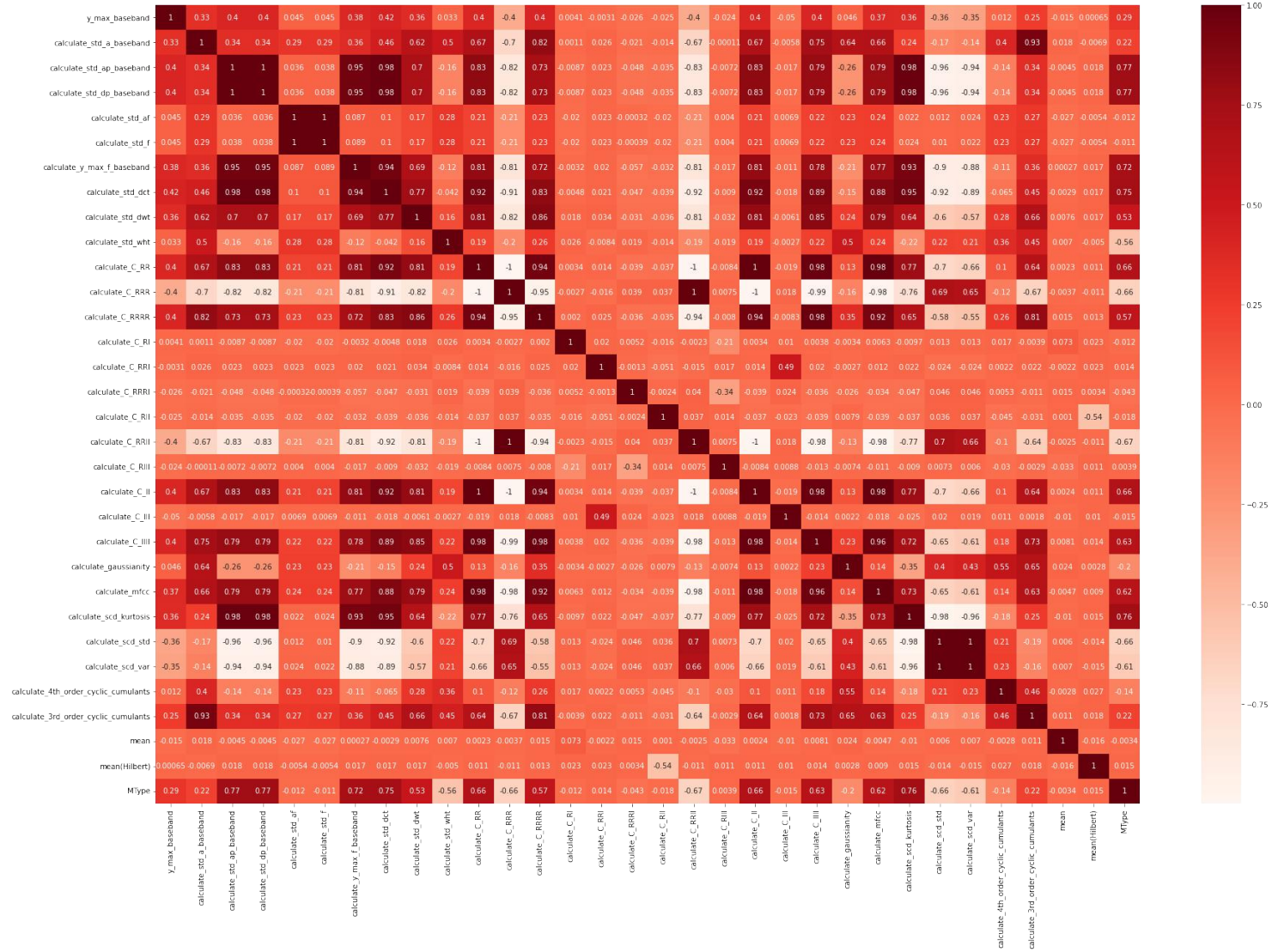
Esse método pode ser definido pela equação:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2] [n \sum y^2 - (\sum y)^2]}}, \quad (4.1)$$

na qual r é o coeficiente de Pearson, n o tamanho das amostras, e x e y as amostras, no caso deste trabalho, os sinais.

A Figura 24 ilustra a Matriz de Correlação entre as 30 características e o tipo de modulação e entre si.

Figura 24: Matriz de Correlação



Fonte: elaborada pelo autor

Por fim, considerou-se as características com correlação de módulo maior que 0,5 em relação ao tipo de modulação. A Tabela 8 mostra todas as características selecionadas pelo método de correlação, junto ao valor do coeficiente.

Tabela 8: Características com mais de 0,5 de correlação

Nome	Correlação
Desvio padrão da fase instantânea absoluta não linear centralizada	0,76
Desvio padrão do valor absoluto da amplitude do sinal normalizada e centralizada	0,76
Desvio padrão da frequência instantânea normalizada e centralizada	0,72
Desvio padrão absoluto da amplitude instantânea normalizada e centralizada	0,65
Desvio padrão do valor absoluto da amplitude do sinal normalizada e centralizada	0,76
Média da frequência instantânea do sinal	0,66
Soma dos componentes da resposta ao degrau do sinal após o filtro Cosseno Levantado	0,57
Resposta ao degrau do sinal após o filtro Cosseno Levantado	0,66
Desvio padrão da transformada DCT do sinal	0,74
Cálculo da Curtose do sinal após o filtro Cosseno Levantado	0,75
Desvio padrão da transformada rápida de Walsh-Hadamard do sinal	0,55
Variância do sinal após o filtro Cosseno Levantado	0,60
Desvio Padrão do sinal após o filtro Cosseno Levantado	0,65

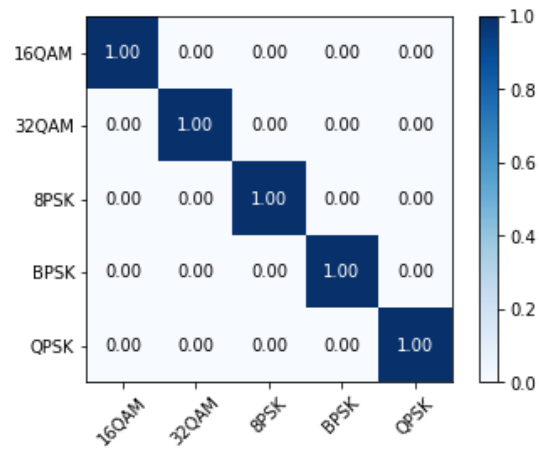
Dessa forma, foram selecionadas apenas seis características presentes em ambos os métodos que podem ser vistas na Tabela 9.

Tabela 9: Características selecionadas nos dois métodos

Nome
Desvio padrão do valor absoluto da amplitude do sinal normalizada e centralizada
Desvio padrão da frequência instantânea normalizada e centralizada
Desvio padrão absoluto da amplitude instantânea normalizada e centralizada
Desvio padrão da transformada DCT do sinal
Média da frequência instantânea do sinal
Desvio padrão da transformada rápida de Walsh-Hadamard do sinal

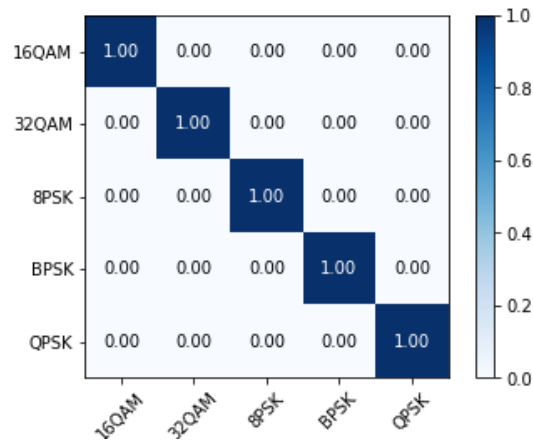
Foi possível, então, diminuir o número de características para classificação de 30 para apenas 6 (uma redução de 80%). Utilizando apenas as características da Tabela 9 foi possível treinar e classificar novamente o sinal com os mesmos três algoritmos da seção 4.3. Os resultados podem ser observados pelas figuras 25, 26 e 27.

Figura 25: Matriz confusão para o algoritmo *Random Forest*



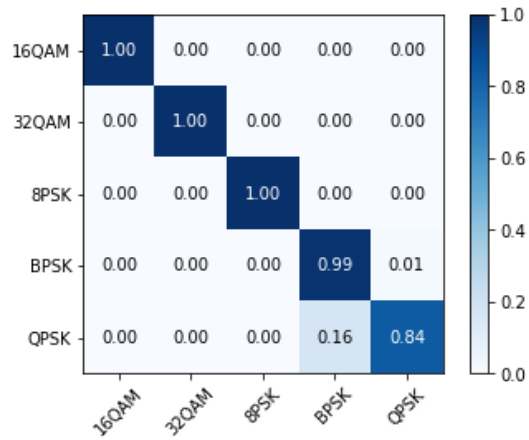
Fonte: elaborada pelo autor

Figura 26: Matriz confusão para o algoritmo *Adaboost*



Fonte: elaborada pelo autor

Figura 27: Matriz confusão para o algoritmo SVM



Fonte: elaborada pelo autor

Como pode ser visto nas figuras acima, o algoritmo *Random Forest* manteve seu alto padrão de acertos na predição e com um menor tempo de processamento para todos os algoritmos, onde os dados foram executados em um tempo menor que 80% do inicial. Para o algoritmo de SVM, a melhora foi significativa com características que prediziam melhor o tipo do sinal, o que nos mostra que a seleção pode ajudar na acurácia do modelo. Para o *Adaboost*, usando Árvore de Decisão, o resultado também foi melhor utilizando a seleção de características.

Capítulo 5

Conclusões e estudos futuros

5.1 Conclusões

Após a geração, propagação e extração de 30 características específicas dos sinais, com o objetivo aumentar o poder de predição das modulações para cada um dos algoritmos de classificação utilizados neste trabalho, notou-se que o algoritmo *Random Forest* obteve excelentes resultados, ao contrário dos algoritmos de Árvore de Decisão e o *Adaboost*. Isso se deve ao fato de o algoritmo ser composto por múltiplas árvores de decisões, facilitando, portanto, a seleção das características do sinal por si só. Esse fato pode ser notado devido aos valores de correção entre características e o tipo de modulação, utilizados no modelo, apresentarem valores maiores que 0,5 [49].

Já para o algoritmo SVM foi possível uma distribuição maior entre as características, com pesos muito parecidos e correlações baixas, o que pode ter levado a um erro de interpretação das características pelo algoritmo.

Para o algoritmo *Adaboost* as características foram muito parecidas com as *do Random Forest*, o que levou a resultados melhores que o do SVM, porém, por se tratar de uma Árvore de Decisão de pouca profundidade, o algoritmo não foi capaz de selecionar um número suficiente de características para que pudesse classificar os sinais BPSK e QPSK, levando a uma confusão entre os dois sinais.

Vale ressaltar que o *Adaboost* aprende a cada classificação, então se houvessem mais dados, o algoritmo provavelmente evoluiria para um resultado melhor sem a seleção de características, porém, por ser susceptível a ruído, talvez enfrentasse dificuldade em sinais que chegassem com relações sinal ruído menores [50].

Apesar de o *Random Forest* ter apresentado um resultado excelente em termos de classificação, o uso de todas as características demanda um custo computacional que não pode ser colocado em prática em um sistema de telecomunicações. Por exemplo, em sistemas de telefonia móvel de sexta geração (6G) a latência deve ser da ordem de 1ms, dessa forma, o tempo de processamento para a classificação de um sinal não deverá ultrapassar esse intervalo de tempo. Além disso, os equipamentos que irão realizar os cálculos de características e

classificação, quanto ao tipo modulações, terão um poder de processamento menor que um computador de mesa.

Com isso, mostrou-se necessária a seleção de algumas características dos sinais, para diminuir o tempo de processamento para os algoritmos utilizados neste trabalho. Para isso foi utilizado o método de Recursão de Características (RFE) aliado ao cálculo da correlação de *Pearson* entre as características e o tipo de modulação dos sinais.

Para o método de RFE foram selecionadas as principais características utilizadas no algoritmo *Random Forest*, que apresentou maior precisão. Tal método trabalha junto à classificação e faz um ranqueamento das características. Com isso, foram separadas 12 para serem utilizadas em futuras classificações.

Já para a correlação de *Pearson* foram consideradas apenas aquelas que têm um módulo maior que 0,5, o que deixou 13 características a serem utilizadas pelo método de correlação. Assim, com os dois métodos feitos e selecionados, foi construída uma intersecção entre eles, restando apenas seis características, uma redução de 80% do total, diminuindo, portanto, o tempo de processamento.

Para a segunda classificação o resultado foi muito satisfatório. O algoritmo *Random Forest* continuou com o mesmo resultado, o que é excelente, mas o custo computacional do cálculo de características seria bem mais rápido; no mais, por não apresentar tantas variáveis no algoritmo, ele executa o treinamento e a classificação em um tempo bem menor. Assim como na primeira classificação, as características utilizadas no treinamento e teste foram praticamente as mesmas, com pesos muito parecidos.

Já o *Adaboost*, obteve um resultado melhor, não havendo mais a confusão entre sinais que ocorreu na primeira classificação, e sua seleção de características se manteve parecida, porém com menos características.

O algoritmo SVM mostrou uma melhora significativa na classificação devido ao número reduzido de características, porém nos sinais BPSK e QPSK, o algoritmo não conseguiu classificar o sinal, o que poderia ser resolvido na escolha e configuração dos classificadores, contudo esse não era o foco deste trabalho.

Por fim, vale ressaltar que os resultados proeminentes da classificação de características podem ser considerados um sucesso, já que foi possível fazer com que apenas seis características classificassem todo um conjunto de sinais.

5.2 Estudos futuros

Para uma primeira sugestão de trabalho futuro seria interessante usar métodos reais de geração, transmissão e recepção de sinais através de rádios definidos por software para que fosse possível fazer o *setup* em uma transmissão real em alta frequência.

É possível, também, utilizar de modelos de Aprendizado de Máquina para monitorar outras questões importantes quando se trata de redes sem fio, como a análise de dados de qualidade do sinal na rede, por exemplo, a relação sinal ruído e o uso da rede, sendo possível definir parâmetros como o tipo de modulação, a banda utilizada, e outras características para melhorar a recepção do sinal [51].

Por último, pode-se propor também utilizar outros métodos de classificação para fazer a comparação entre suas diferentes características e resultados.

Referências

- [1] Huurdeman, A. A., *The WorldWide History of Telecommunications*, Hooken, New Jersey: A JOHN WILEY & SONS, 2003.
- [2] Selleri, S., “Claude Chappe and the first telecommunication network (without electricity),” *URSI Radio Science Bulletin*, pp. 96-101, 03 2017.
- [3] Dibner, B., “History of Electrical Engineering Andre Marie Ampere,” *IEEE Power Engineering Review*, pp. 15-16, 02 1984.
- [4] Schulz, Peter, “Unicamp,” 28 06 2019. [Online]. Available: <https://www.unicamp.br/unicamp/ju/artigos/peter-schulz/quem-inventou-o-telegrafo-esquerda-direita-direita-direita-esquerda-direita#4>. [Acesso em 04 2021].
- [5] Andrews, F. T., “The heritage of telegraphy,” *IEEE Communications Magazine*, pp. 12-18, 08 1989.
- [6] Flood, J. E., “Alexander Graham Bell and the invention of the telephone,” *Electronics and Power*, vol 22, pp. 159-162, 03 1976.
- [7] Susskind, C., “Heinrich Hertz: a short life,” *IEEE Transactions on Microwave Theory and Techniques*, pp. 802-805, 05 1988.
- [8] Falciassecca G. and Valotti B., “Guglielmo Marconi: The pioneer of wireless communication,” *2009 European Microwave Conference (EuMC)*, pp. 544-546, 2009.
- [9] Hitachi, “History of semiconductors,” Hitachi, [Online]. Available: <https://www.hitachi-hightech.com/global/en/knowledge/semiconductor/room/about/history.html>. [Acesso em 01 05 2022].
- [10] “THE INVENTION OF MOBILE PHONES,” Science Museum, 12 11 2018. [Online]. Available: <https://www.sciencemuseum.org.uk/objects-and-stories/invention-mobile-phones#:~:text=Martin%20Cooper%2C%20the%20engineer%20from,launched%20the%20DynaTAC%20in%201984..> [Acesso em 09 05 2021].

- [11] Mshvidobadze, T., “Evolution mobile wireless communication and LTE networks,” *2012 6th International Conference on Application of Information and Communication Technologies (AICT)*, pp. 1-7, 2012.
- [12] Sauter, Martin, em *Evolution from 2G over 3G to 4G*, Wiley, 2013, pp. 1-7.
- [13] Sector, ITU’s Telecommunication Standardization, “ITU G.652,” 24 12 2019. [Online]. Available: <https://www.itu.int/rec/T-REC-G.652/en>. [Acesso em 7 05 2021].
- [14] Education, IBM Cloud, “<https://www.ibm.com>,” A International Business Machines Corporation, 15 07 2020. [Online]. Available: <https://www.ibm.com/cloud/learn/machine-learning>. [Acesso em 28 04 2022].
- [15] Sujatha, R. Saravanan and P., “A State of Art Techniques on Machine Learning Algorithms: A Perspective of Supervised Learning Approaches in Data Classification,” *Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 345-349, 2018.
- [16] Mcculloch, W. & Pitts, W., “A Logical Calculus of the Ideas Immanent in Nervous Activity,” *Bulletin of Mathematical Biophysics*, n° 5, pp. 127-147, 1943.
- [17] Turing, A.M, “Computing Machinery and Intelligence.,” *Mind*, pp. 433-460, 1950.
- [18] IDZHAM, K. K. et al., “Study of Artificial Intelligence into Checkers Game using HTML and JavaScript,” em *IOP Publishing, 2020. p. 012091.*, 2020.
- [19] GOODMAN, David; KEENE, Raymond., “Man versus machine: Kasparov versus deep blue.,” *ICGA Journal.*, pp. 186-187, 1997.
- [20] Leitão, Rafael, “Rafael Leitão,” 2015 10 10. [Online]. Available: <https://rafaelleitao.com/o-homem-e-a-mquina-o-match-kasparov-x-deep-blue/>. [Acesso em 2022 04 24].
- [21] Kotsiantis, S. B., “Supervised Machine Learning: A Review of Classification,” *Informatica*, n° 31, pp. 249-268, 2007.
- [22] BOGATU, Alex et al., “Dataset discovery in data lakes,” *2020 IEEE 36th International Conference on Data Engineering (ICDE)*, pp. 709-720, 2020.

- [23] SCHNEIDER, Karl-Michael, "A comparison of event models for naive bayes anti-spam e-mail filtering.," em *0th Conference of the European Chapter of the Association for Computational Linguistics*, 2003.
- [24] QUINLAN, J. Ross, "Learning decision tree classifiers.," *ACM Computing Surveys (CSUR)*, pp. 71-72, 1996.
- [25] P. Argentero, R. Chin and P. Beaudet, "An Automated Approach to the Design of Decision Tree Classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 4, n° 1, pp. 51-57, 1982.
- [26] Pedregosa, F. et al., "Scikit-learn," *Journal of Machine Learning Research*, vol. 12, pp. 285-2830, 2011.
- [27] Reddy P. D. and Parvathy L. R. , "Prediction Analysis using Random Forest Algorithms to Forecast the Air Pollution Level in a Particular Location.," *2022 3rd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India*, pp. 1585-1589, 2022.
- [28] Yadahalli S. and Nighot M. K. , "Adaboost based parameterized methods for wireless sensor networks," *International Conference On Smart Technologies For Smart Nation (SmartTechCon)*, pp. 1370-1374, 2017.
- [29] Xue, Hui, Qiang Yang, and Songcan Chen., "SVM: Support vector machines," em *The top ten algorithms in data mining.*, Chapman and Hall/CRC, 2009, pp. 51-74.
- [30] Rube, I. El, "Image Color Reduction Using Progressive Histogram Quantization and Kmeans Clustering," *2019 International Conference on Mechatronics, Remote Sensing, Information Systems and Industrial Information Technologies (ICMRSISIT)*, pp. 1-5, 2019.
- [31] Nachar, I. Dagher and R., "Face recognition using IPCA-ICA algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, n° 6, pp. 996-1000, 2006.

- [32] Aparna S Varde, Makiko Takakshi, Elke A Rundensteiner, Mathew O ward, “Apriori algorithm and game-of-life for predictive analysis in materials science,” *International journal of knowledge based and intelligent engineering systems* 8, 2004.
- [33] Lyu L., Shen Y. Zhang S, “The Advance of Reinforcement Learning and Deep Reinforcement Learning,” *IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA)*, pp. 644-648, 2022.
- [34] STROOCK, Daniel W, “An introduction to Markov processes.,” *Springer Science & Business Media*, 2013.
- [35] Ghunaim S.A. , Nasir Q. Talib M. A. , “Deep Learning Techniques for Automatic Modulation Classification: A Systematic Literature Review,” *14th International Conference on Innovations in Information Technology*, n° 14, pp. 108-113, 2020.
- [36] Valadão Myke D. M. , Amoedo Diego, Costa André , Carvalho Celso , Sabino Waldir , “Deep Cooperative Spectrum Sensing Based on Residual Neural Network Using Feature Extraction and Random Forest Classifier,” em *Sensors*, 2021.
- [37] Zhu, Zhechen, *Automatic Modulation Classification: Principles, Algorithms and Applications*, Wiley, 2015.
- [38] Sabino Waldir , L. A. Costa André , Pereira Antonio , Tavares Samuel , Barbosa Carvalho Celso , A. Amoedo Diego , Valadão Myke D. M. , Alves dos Santos Beatriz, “Rede Adversária Generativa Semi-Supervisionada para Falsificação de Sinais Modulados Utilizados em Simulação de Ataque a Modelos de Reconhecimento Automático de Modulações,” em *XL Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT2022)*, 2022.
- [39] Xuan Qui C. P., Hong Quang D. , Duy P. T. , Thi Thu Hien D. and V. -H., “Strengthening IDS against Evasion Attacks with GAN-based Adversarial Samples in SDN-enabled network,” *2021 RIVF International Conference on Computing and Communication Technologies (RIVF)*, pp. 1-6, 2021.

- [40] MathWorks, “MathWorks,” 2022. [Online]. Available: <https://www.mathworks.com/help/comm/ug/raised-cosine-filtering.html>. [Acesso em 09 09 2022].
- [41] Nandi, Zhechen Zhu and Asoke K., Automatic Modulation Classification: Principles, Algorithms and Applications, John Wiley & Sons, Ltd, 2015.
- [42] Lee, Sang Hoon, Kwang-Yul Kim, and Yoan Shin, “Effective Feature Selection Method for Deep Learning-Based Automatic Modulation Classification Scheme Using Higher-Order Statistics,” *Applied Sciences* 10.2, p. 588, 2020.
- [43] Lee, Sang Hoon, Kwang-Yul Kim, and Yoan Shin, “Effective feature selection method for deep learning-based automatic modulation classification scheme using higher-order statistics,” *Applied Sciences*, p. 588, 2020.
- [44] Srichavengsup, T. Visetbunditkun and W., “DDoS Attack Detection Using Ensemble Machine Learning Models with RFE Algorithm,” *2022 7th International Conference on Business and Industrial Research*, pp. 269-273, 2022.
- [45] Yildirim, Soner, “Decision Trees and Random Forests — Explained,” *Towards Data Science*, 8 Fevereiro 2020. [Online]. Available: <https://towardsdatascience.com/decision-tree-and-random-forest-explained-8d20ddabc9dd#:~:text=Random%20forest%20is%20an%20ensemble,received%20from%20each%20decision%20tree..> [Acesso em 01 12 2022].
- [46] Jacob, Shomona, and Geetha Raju., “Software defect prediction in large space systems through hybrid feature selection and classification.,” *Int. Arab J. Inf. Technol.* 14.2, pp. 208-214, 2017.
- [47] Gopika M.E, N. and Kowshalaya A. M, “Correlation Based Feature Selection Algorithm for Machine Learning,” *3rd International Conference on Communication and Electronics Systems (ICCES)*, pp. 692-695, 2018.
- [48] Iranmanesh V. , Ahmad S. , Wan Adnan W. A. , Layth Malallah F. , S. Yussof, “Online signature verification using neural network and pearson correlation features,” *2013 IEEE Conference on Open Systems (ICOS)*, pp. 18-21, 2013.

- [49] Dubey, Akash, “Towards Data Science,” Towards Data Science, 14 Dezembro 2018. [Online]. Available: <https://towardsdatascience.com/feature-selection-using-random-forest-26d7b747597f#:~:text=Feature%20selection%20using%20Random%20forest%20comes%20under%20the%20category%20of,built%2Din%20feature%20selection%20methods..> [Acesso em 29 12 2022].
- [50] Sun B. ,Luo J. , Shu S.and Yu N. , “Introduce randomness into AdaBoost for robust performance on noisy data,” *2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery*, pp. 1858-1861, 2010.
- [51] Kwon J,Jung D, Park H,, “Traffic Data Classification using Machine Learning Algorithms in SDN Networks,” *International Conference on Information and Communication Technology Convergence (ICTC)*,, pp. 1031-1033, 2020 Jeju, Korea .