



DETECÇÃO DE DEFEITOS EM PLANTAS INDUSTRIAIS ATRAVÉS DE CARACTERÍSTICAS ESPECTRAIS DO SOM

Luana Gantert Reis

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Elétrica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Miguel Elias Mitre Campista

Rio de Janeiro

Maio de 2021

DETECÇÃO DE DEFEITOS EM PLANTAS INDUSTRIAIS ATRAVÉS DE
CARACTERÍSTICAS ESPECTRAIS DO SOM

Luana Gantert Reis

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE
ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO
PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU
DE MESTRE EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Orientador: Miguel Elias Mitre Campista

Aprovada por: Prof. Miguel Elias Mitre Campista
Prof. Rodrigo de Souza Couto
Prof. Daniel Sadoc Menasche

RIO DE JANEIRO, RJ – BRASIL
MAIO DE 2021

Reis, Luana Gantert

Detecção de Defeitos em Plantas Industriais através de Características Espectrais do Som/Luana Gantert Reis. – Rio de Janeiro: UFRJ/COPPE, 2021.

XI, 48 p.: il.; 29,7cm.

Orientador: Miguel Elias Mitre Campista

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2021.

Referências Bibliográficas: p. 46 – 48.

1. Aprendizado de Máquina. 2. Redes Industriais.
3. Internet das Coisas. I. Campista, Miguel Elias Mitre. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

À minha família e amigos.

Agradecimentos

Agradeço aos meus pais, Iris Gantert e Everaldo Reis por acreditarem em mim e serem fundamentais na minha jornada. A Renan Martinez por ser meu principal apoiador desde a minha escolha em iniciar a graduação e por me ensinar a não desistir jamais. A Beatriz Parente pela amizade acima dos laços de sangue e por vibrar com cada conquista como se fosse sua. Agradeço também aos familiares que sempre estiveram ao meu lado nesta jornada.

Agradeço aos colegas do Grupo de Teleinformática e Automação pelo acolhimento e contribuição. Em especial, a Mariana Maciel, Daiane Pereira, Roberto Pacheco e Lucas Gomes pela amizade e companheirismo.

Agradeço ao meu orientador Miguel Elias Mitre Campista pelos ensinamentos, paciência e apoio. Obrigada pelo exemplo e por sempre me motivar com novos desafios. Agradeço aos professores Rodrigo de Souza Couto e Daniel Sadoc Menasche por fazerem parte da banca examinadora.

Por fim, agradeço ao apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

DETECÇÃO DE DEFEITOS EM PLANTAS INDUSTRIAIS ATRAVÉS DE CARACTERÍSTICAS ESPECTRAIS DO SOM

Luana Gantert Reis

Maio/2021

Orientador: Miguel Elias Mitre Campista

Programa: Engenharia Elétrica

Esta dissertação propõe a adoção de características do espectro sonoro no reconhecimento de falhas do maquinário industrial analisando quatro tipos distintos de componentes: ventiladores industriais, bombas hidráulicas, trilhos deslizantes e válvulas. Enquanto as abordagens baseadas em espectrogramas adotam redes com alto número de parâmetros, uma vez que podem ser classificados como imagens, as características adotadas neste trabalho permitem que as ondas de som sejam traduzidas de forma concisa. Logo, a classificação é realizada com algoritmos que exijam menor poder computacional. Avalia-se o desempenho de cinco algoritmos de aprendizado supervisionado comumente utilizados na classificação binária: Naïve Bayes, SVM, MLP, XGBoost e Floresta Aleatória. Como resultado, a AUC obtida é até 106,68% superior em comparação com a abordagem baseada em Mel-espectrogramas. A classificação binária, porém, permite apenas a comparação entre estados de um único tipo de componente. Dessa forma, na segunda parte dessa dissertação, é realizada a classificação multiclasse, que considera métricas adicionais e o tempo necessário para o treinamento e teste do modelo. O objetivo é avaliar se as características adotadas permitem que não apenas o estado do equipamento seja identificado, como também o tipo da máquina. A avaliação é realizada através da matriz de confusão, onde percebe-se que as amostras normais de sons são classificadas corretamente na maior parte dos casos. As amostras anômalas, exceto pelas válvulas, tendem a ser classificadas corretamente na maior parte das vezes, mas em menor quantidade que as amostras normais. Quando ocorre a classificação incorreta, a confusão ocorre entre os estados, mas acerta-se o tipo de máquina. Este comportamento é, em parte, explicado pela natureza desbalanceada da base de dados. Nas válvulas não é possível classificar corretamente as amostras anômalas.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

A SOUND-BASED CORRECTIVE MAINTENANCE AT INDUSTRIAL PLANTS

Luana Gantert Reis

May/2021

Advisor: Miguel Elias Mitre Campista

Department: Electrical Engineering

The fourth industrial revolution is described as the factory processes modernization and use the Internet of Things, Artificial Intelligence, and sensors to process automation, improve workers' safety, and monitoring industrial assets. In this work, we propose to use the spectral features for fault recognition on industrial machinery analyzing four types of components: fans, pumps, slide rails, and valves. While the approaches spectrograms-based adopts networks with a high number of parameters, once spectrograms can be treated as images, the features selected allow representing sound waveforms concisely. Thus, the classification can utilize algorithms that require lower computational power. We evaluate the performance of five supervised algorithms usually applied on classification tasks in binary classification: Naive Bayes, Support Vector Machine, Multilayer Perceptron, XGBoost, and Random Forrest. As result, the AUC value on our approach is 106,68% higher than the work based on Mel-spectrograms. However, the binary classification only allows the comparison between the status of a single type of component. Then, in the second part of the experiments, we conduct a multiclass classification. The algorithm is selected considering additional metrics and the time required to train and test the model. The goal is to evaluate if the features enable identify also the type of machine in addition to the equipment status. We evaluate the results by confusion matrix, where we realize the normal samples are classified correctly on major of times. Without consider the valves, the abnormal samples also tend to be classified correctly, but less than the normal samples. When the classification is wrong the confusion occurs between status, but the machine type is right. This behavior is, in part, explained by the imbalanced nature of the dataset. On valves is not possible to classify the abnormal samples.

Sumário

Lista de Figuras	ix
Lista de Tabelas	x
1 Introdução	1
2 Trabalhos Relacionados	5
2.1 Aprendizado de Máquina para Reconhecimento de Falhas e Eventos .	5
2.2 Desafio DCASE 2020	7
3 Aprendizado Supervisionado para Classificação	8
3.1 Algoritmos de Aprendizado de Máquinas	9
3.2 Métricas para avaliação	10
4 Metodologia	14
4.1 Características Espectrais do Som	14
4.2 Etapas do experimento	16
5 Resultados e Discussões	20
5.1 Classificação Binária	20
5.2 Classificação Multiclasse	35
5.2.1 Tempo de Conversão do Conjunto de Dados	40
6 Conclusão e Trabalhos Futuros	43
Referências Bibliográficas	46

Lista de Figuras

3.1	Separação de duas classes pelo algoritmo SVM.	10
3.2	Exemplo de rede neural com camadas totalmente conectadas.	10
3.3	Matriz de Confusão para classificação binária.	11
3.4	Curva ROC.	13
3.5	Matriz de Confusão para classificação multiclasse.	13
4.1	Etapas realizadas no experimento.	16
5.1	Comparação da acurácia obtida pelos cinco modelos de aprendizado supervisionado para classificação binária.	24
5.2	Comparação da Precisão obtida pelos cinco modelos de aprendizado supervisionado para classificação binária.	28
5.3	Comparação da <i>recall</i> obtida pelos cinco modelos de aprendizado supervisionado para classificação binária.	31
5.4	Comparação da F1-Score obtida pelos cinco modelos de aprendizado supervisionado para classificação binária.	34
5.5	Comparação da AUC obtida pelos cinco modelos de aprendizado supervisionado e o trabalho Purtohit et al. e para classificação binária.	36
5.6	Número total de amostras por equipamento e estado.	38
5.7	Matriz de Confusão para classificação Multiclasse para utilizando SVM sem <i>oversampling</i>	40
5.8	Matriz de Confusão para classificação Multiclasse para utilizando SVM com <i>oversampling</i>	41
5.9	Tempo total de conversão das amostras de som do conjunto de dados em Mel-espectrogramas e para extração de características espectrais	42

Lista de Tabelas

4.1	Conjunto de dados MIMII.	17
5.1	Tempo em segundos para treinar e testar os modelos.	20
5.2	Acurácia para os modelos na classificação binária de amostras anômalas dos ventiladores industriais.	21
5.3	Acurácia para os modelos na classificação binária de amostras anômalas das bombas hidráulicas.	22
5.4	Acurácia para os modelos na classificação binária de amostras anômalas dos trilhos deslizantes.	22
5.5	Acurácia para os modelos na classificação binária de amostras anômalas das válvulas.	23
5.6	Precisão para os modelos na classificação binária de amostras anômalas dos ventiladores.	25
5.7	Precisão para os modelos na classificação binária de amostras anômalas das bombas hidráulicas.	25
5.8	Precisão para os modelos na classificação binária de amostras anômalas dos trilhos deslizantes.	26
5.9	Precisão para os modelos na classificação binária de amostras anômalas das válvulas.	27
5.10	<i>Recall</i> para os modelos na classificação binária de amostras anômalas dos ventiladores.	27
5.11	<i>Recall</i> para os modelos na classificação binária de amostras anômalas das bombas hidráulicas.	29
5.12	<i>Recall</i> para os modelos na classificação binária de amostras anômalas dos trilhos deslizantes.	29
5.13	<i>Recall</i> para os modelos na classificação binária de amostras anômalas das válvulas.	30
5.14	F1-score para os modelos na classificação binária de amostras anômalas dos ventiladores.	32
5.15	F1-score para os modelos na classificação binária de amostras anômalas das bombas hidráulicas.	33

5.16	F1-score para os modelos na classificação binária de amostras anômalas dos trilhos deslizantes.	33
5.17	F1-score para os modelos na classificação binária de amostras anômalas das válvulas.	35
5.18	AUCs para os modelos na classificação binária de amostras anômalas dos ventiladores industriais.	35
5.19	AUCs para os modelos na classificação binária de amostras anômalas das bombas hidráulicas.	37
5.20	AUCs para os modelos na classificação binária de amostras anômalas dos trilhos deslizante.	37
5.21	AUCs para os modelos na classificação binária de amostras anômalas das válvulas.	37

Capítulo 1

Introdução

A Indústria 4.0 ou I4.0, também conhecida como a quarta revolução industrial, caracteriza o processo de modernização das fábricas através de tecnologias digitais. A partir desta modernização é possível monitorar as máquinas, sistemas e processos nas plantas industriais, permitindo que a configuração ótima dos processos seja obtida, considerando sua natureza dinâmica e como se adequar eficientemente ao mercado global [1]. A I4.0 está fortemente relacionada com tópicos como: sistemas ciberfísicos, inteligência artificial, Internet das Coisas Industrial (IIoT), Computação em nuvem e na borda [2].

As fábricas inteligentes (*smart factories*) fazem parte da quarta revolução industrial e utilizam soluções automatizadas para reduzir trabalhos desnecessários e reduzir o desperdício de recursos [3]. Um dos desafios de uma fábrica inteligente é evitar que a operação e acesso aos dados de seus componentes seja interrompido, devido à sua natureza crítica. Esta característica significa que as interrupções não planejadas não apenas podem impactar financeiramente, como colocar em risco a segurança dos trabalhadores. Ao adotar tecnologias para integrar e monitorar seus sistemas, é possível aumentar a eficiência dos equipamentos, otimizar a linha de produção e, conseqüentemente, reduzir os custos e criar um ambiente sustentável [4].

Uma das maneiras de evitar a interrupção do trabalho industrial é a partir da adoção de sistemas de monitoramento, adotados para detecção e diagnóstico inteligente de falhas, monitoramento da condição das máquinas e outras tarefas relacionadas ao acompanhamento e gerenciamento da saúde dos componentes da indústria [5]. Desta maneira, pode-se agendar a manutenção ou adiar a troca de uma peça de um equipamento até que o comportamento deixe de estar dentro de um padrão considerado adequado para seu funcionamento.

Entre as técnicas de inteligência artificial estão as de aprendizado de máquina que podem ser escolhidas a partir do grau de complexidade da tarefa. As técnicas de aprendizado de máquinas buscam modelar o comportamento de uma função através de uma representação matemática. Para isso, é realizado o treinamento do modelo,

onde seus parâmetros serão ajustados utilizando atributos de amostras selecionadas. O objetivo é que o modelo seja capaz de prever o comportamento de novas amostras e de forma autônoma realizar tarefas como classificação ou predição de comportamentos. Essas tarefas estão relacionadas com a saída do modelo, onde o formato dos dados de entrada na etapa de treinamento ditam a maneira que o modelo trabalhará. Desta forma, Bochie et al. apresentam aplicações de aprendizado profundo para resolução de problemas em Redes Desafiadoras, isto é, redes que geram alto volume de dados dada a alta complexidade de operação [6]. As soluções baseadas em aprendizado profundo para redes industriais selecionadas pelos autores evidenciam a emergente necessidade de soluções inteligentes em aplicações como a inspeção de produtos com defeitos na linha de produção.

Os algoritmos de aprendizado de máquina surgem como uma das tecnologias que possibilitam os sistemas inteligentes de monitoramento que, por sua vez, auxiliam as manutenções corretivas e/ou preventivas. As abordagens de manutenção buscam encontrar o ponto ótimo entre custos e confiabilidade considerando os prós e contras de cada uma [7]. Enquanto a manutenção corretiva é baseada na troca ou reparo de um componente apenas após a sua falha, a preventiva utiliza as manutenções agendadas justamente para evitar que a falha ocorra. Desta maneira, a manutenção corretiva maximiza a vida útil das peças e reduz o desperdício, mas impede que o sistema trabalhe dentro do padrão esperado até que ocorra a manutenção. A manutenção preventiva funciona da forma oposta.

A manutenção corretiva inteligente pode ser realizada com o auxílio de um sistema responsável pela classificação do estado das máquinas. Então, os operadores podem ser avisados que o comportamento não corresponde ao esperado e que a manutenção é necessária. As fontes de dados de entrada de um classificador podem ser as mais diversas possíveis. Em abordagens recentes, amostras de sons são adotadas em modelos baseados em aprendizado supervisionado. As amostras são normalmente transformadas em espectrogramas ou Mel-espectrogramas e então podem ser consideradas como imagens. Desta maneira, classificadores populares em tarefas de classificação de imagens, como Redes Neurais Convolucionais (*Convolutional Neural Networks* – (CNNs)) podem ser adotados.

Entre as opções de dados de entrada estão as amostras sonoras captadas de múltiplos equipamentos industriais. Nesse contexto, os trabalhos apresentados no desafio de detecção e classificação de eventos acústicos - no desafio *Detection and Classification of Acoustic Scenes and Events 2020* (DCASE 2020) - apresentam soluções baseadas em aprendizado de máquina para solucionar tal tarefa. O conjunto de som adotado no desafio foi o conjunto de dados MIMII (*Malfunctioning Industrial Machine Investigation and Inspection*) [8]. Neste trabalho, utiliza-se nos experimentos o mesmo conjunto de dados, que contém amostras produzidas por quatro diferentes

tipos de máquinas: ventiladores industriais, bombas hidráulicas, trilhos deslizantes e válvulas. Por sua vez, cada tipo é composto por quatro máquinas diferentes, totalizando 16 máquinas individuais. Todas as máquinas possuem amostras de som normais e anômalas. Além disso, o conjunto de dados apresenta três diferentes relações de sinal-ruído, obtidas mixando às amostras originais os ruídos de fundo de uma fábrica.

Nesta dissertação adotam-se características extraídas dos espectros sonoros ao invés dos Mel-espectrogramas. O objetivo é ser capaz de utilizar algoritmos supervisionados que necessitem de menos poder computacional que a proposta baseada em Mel-espectrogramas e Redes Autoassociativas (*Autoencoder*). Este trabalho utilizado para comparação é proposto por Purohit et al., autores do conjunto de dados adotado, e utiliza a classificação binária para identificar o mau funcionamento dos equipamentos. Na classificação binária busca-se identificar se determinado modelo individual está ou não operando conforme o esperado. A avaliação da classificação binária é realizada através das métricas acurácia, AUC, F1-score, *recall* e precisão. A AUC permite que os modelos da proposta sejam comparados com o trabalho de Purohit et al., além da comparação entre os modelos realizada através das demais métricas [9]. A adoção das características selecionadas nesta dissertação possibilitaram a melhoria dos resultados de Purohit et al., mesmo utilizando uma rede neural com menor número de neurônios e camadas. Além disso, os resultados promissores da classificação binária possibilitaram a próxima etapa experimental, onde o algoritmo SVM foi escolhido pelos resultados obtidos para realizar a classificação multiclasse.

Na classificação multiclasse os modelos individuais são agrupados para representar seu tipo de máquina. O objetivo é que um único modelo seja capaz de identificar a partir da amostra de som o tipo de máquina e seu estado de funcionamento. Os resultados da classificação multiclasse são realizados com e sem o método de *oversampling*. Este método foi aplicado em todos os modelos de classificação binária desta dissertação e é utilizado para equilibrar o número de amostras anômalas e normais do conjunto de dados. Esta análise mostra altas taxas de acertos na detecção de anomalia ao utilizar *oversampling*. As válvulas, que possuem a maior diferença entre a quantidade de amostras, não conseguem ter seu defeito detectado em nenhuma das SNR.

Este trabalho está organizado da seguinte maneira: no Capítulo 2 são revisados os trabalhos relacionados a sistemas de monitoramento e classificação utilizando amostras sonoras como entrada; no Capítulo 3 são descritos alguns conceitos teóricos sobre os algoritmos de aprendizado supervisionado selecionados, assim como as métricas para avaliação de desempenho dos modelos; no Capítulo 4 é apresentada a metodologia adotada para os experimentos e as características do espectro sonoro

selecionadas; Capítulo 5 são apresentados os resultados dos experimentos de classificação binária e multiclases; o Capítulo 6 apresenta a conclusão deste trabalho.

Capítulo 2

Trabalhos Relacionados

O monitoramento das plantas industriais de maneira inteligente pode ser realizado através de algoritmos de aprendizado de máquina e adotando-se como entrada os dados coletados de diversos tipos de sensores. Os trabalhos relacionados a seguir foram divididos em dois tópicos de interesse: o reconhecimento de falhas ou eventos perigosos através de outros tipos de dados e os trabalhos apresentados no desafio DCASE Challenge 2020.

2.1 Aprendizado de Máquina para Reconhecimento de Falhas e Eventos

Adotando como dados de entrada os coletados através de sensores e atuadores de uma turbina eólica, Zhang et al. combinam os algoritmos XGBoost e Floresta Aleatória (Random Forest – RF) para detecção de falhas [10]. Apesar de sua aplicação em tarefas de classificação e regressão, RF pode ser adotado na seleção de atributos. Neste trabalho, os dados dos sensores compõem as características de entrada que terão seus graus de impacto na classificação avaliados com RF. Então, os mais importantes em cada falha são ranqueados e os três mais importantes são selecionados para treinamento do modelo baseado em XGBoost. O sistema pode detectar dez tipos de falhas diferentes a partir de dados como a velocidade da turbina, por exemplo.

Para detectar os defeitos dos produtos na linha de produção, Li et al. desenvolvem um sistema que utiliza as imagens das câmeras [11]. O sistema é baseado em CNNs e não apenas detecta se o produto não está dentro de um padrão pré-estabelecido, como também mensura o grau do defeito. Para realizar tal diagnóstico o modelo é treinado com um conjunto de dados contendo dez tipos diferentes de defeitos. A solução é baseada na colaboração dos nós na borda e computação em nuvem, adotando a estratégia de parada antecipada (Early-stopping). Desta ma-

neira, quando o poder computacional dos nós locais for capaz de treinar o modelo dentro do limite determinado pelos autores para a métrica de avaliação, a nuvem não é utilizada. Assim, o tempo de inferência também pode ser reduzido.

Cheng et al. utilizam os sinais de vibração dos rolamentos para detecção de anomalias e prognóstico de falhas [12]. Três tipos de características podem ser extraídas dos sinais de vibração: as no domínio do tempo, no domínio tempo-frequência. Essas características são utilizadas em um algoritmo de agrupamento para detectar as anomalias. Uma vez detectada, o comportamento da máquina é monitorado a partir de uma Long Short-Term Memory Recurrent Neural Network (LSTM-RNN). Zhang et al. também propuseram um sistema para monitoramento do estado de rolamentos a partir de sinais de vibração, mas em sua proposta utilizam um modelo baseado em uma Deep Belief Network (DBN) [13].

Liu et al. e Diez-Olivan et al. realizam uma revisão de trabalhos baseados em inteligência artificial ligados à indústria [5, 14]. No trabalho de Li et al., são selecionados e comparados os trabalhos para diagnóstico de falhas especificamente em máquinas rotativas. São comparados algoritmos como K-Nearest Neighbors (K-NN), classificados Naive Bayes (NB), Máquina de Vetores de Suporte (Support Vector Machine – SVM), Redes Neurais e abordagens baseadas em aprendizado profundo. Nestes casos, a comparação entre os desempenhos demonstrou que SVMs e Redes Neurais podem alcançar uma acurácia satisfatória, mas possuem baixa robustez a ruído. O algoritmo NB pode ser realizar a inferência mais rápido que os demais, mas não alcança alta acurácia. As abordagens baseadas em aprendizado profundo é robusto a ruídos e alcança alta acurácia, mas demora mais tempo que os demais para realizar a classificação. Isto significa que cada problema a ser resolvido possui uma abordagem mais vantajosa ou adequada. Os trabalhos analisados por Li et al. utilizam como dados os sinais de sensores de vibração, temperatura, pressão, dados acústicos, entre outros.

Diez-Olivan et al. comparam trabalhos baseados em aprendizado de máquina para manutenção nos setores industriais. Os modelos analisados são separados em três categorias principais: Descritivos, Preditivos e Prescritivos. Desta maneira, as propostas de classificação e reconhecimento de padrões para detecção de danos e falhas se enquadram na categoria descritiva.

Por último, Sammarco and Detyniecki compara modelos para detecção de eventos perigosos na estrada a partir de dados no domínio do tempo e da frequência em conjunto com Mel-espectrogramas na melhoria da segurança dos motoristas [15].

2.2 Desafio DCASE 2020

O desafio DCASE 2020 é uma competição composta por 6 tarefas de detecção e classificação de eventos baseados em amostras de som. Em uma das tarefas o objetivo é realizar a detecção de falhas do maquinário industrial a partir do conjunto de dados MIMII de Purohit et al. [8]. Os autores do conjunto de dados realizam uma análise inicial baseada em uma Rede Autoassociativa Densa (*Dense Autoencoder Network*) de cinco camadas. A arquitetura da rede conta com 8 neurônios na camada central e 64 neurônios em todas as outras camadas. As amostras sonoras são convertidas em Mel-espectrogramas. Este trabalho é utilizado na comparação de resultados com a proposta desta dissertação, sendo apresentada no Capítulo 5.

A tarefa original era baseada em abordagens baseadas em aprendizado não supervisionado. Desta forma, o melhor desempenho no reconhecimento de falhas foi o trabalho proposto por Giri et al [16], onde os Mel-espectrogramas foram utilizados como entrada de uma Rede Residual Profunda pré-treinada conhecida como ResNet50. A rede consiste em mais de 23 milhões de parâmetros treináveis. A alta quantidade de parâmetros torna necessário uma quantidade de dados para treinamento do modelo para evitar sobreajuste (*overfitting*) e, conseqüentemente, a baixa generalização do modelo. Além dos problemas computacionais e de treinamento, a quantidade de amostras produzidas durante o comportamento normal do maquinário é múltiplas vezes maior que as produzidas nas situações de falha. A característica desbalanceada dos dados disponíveis impõem um problema a mais a ser enfrentado pelos classificadores.

Daniluk et. al. realizam a tarefa utilizando Redes Autoassociativas Variacionais e Condicionadas [17]. Além disso, uma rede baseada em *Deep Complex U-Net* é utilizada na etapa de pré-processamento para redução de ruído. O terceiro colocado no mesmo ranking foi o modelo proposto por Primus [18], também baseado em uma Rede Neural Residual. Apesar dos trabalhos alcançarem resultados promissores, ainda possuem os requisitos e desafios citados anteriormente para dos modelos baseados em aprendizado profundo na classificação de eventos através de Mel-espectrogramas.

Capítulo 3

Aprendizado Supervisionado para Classificação

O aprendizado de máquinas pode ser dividido em supervisionado e não supervisionado. Diferente do aprendizado não supervisionado, que dispensa a presença de dados rotulados, o aprendizado supervisionado é caracterizado pela presença de algoritmos que se servem dos rótulos para o treinamento. Desta maneira, seus parâmetros se ajustam para identificar uma nova entrada considerando entradas semelhantes que foram apresentadas na fase de treinamento. Os algoritmos de aprendizado supervisionado podem ser aplicados em tarefas de classificação ou regressão. Enquanto nas tarefas de classificação o objetivo é lidar com valores discretos, nas tarefas de regressão busca-se prever valores contínuos.

Neste capítulo são apresentados os algoritmos de classificação selecionados para detecção de defeitos a partir do som, assim como as métricas de avaliação adotadas para analisar os resultados da classificação binária e multiclasse. Cada resultado obtido representa direta ou indiretamente um comportamento que se deseja incentivar ou inibir, uma vez que os classificadores não apenas precisam aumentar a taxa de acertos, mas também reduzir o tanto quanto for possível os alarmes falsos ou situações onde o mau funcionamento não é detectado. As métricas acurácia, precisão, *recall*, F1-score e AUC (*Area Under the Curve*) foram utilizadas para captar essas informações e comparar os modelos entre si. A métrica AUC permitiu que os modelos adotados nesta proposta também fossem comparados com o trabalho baseado em redes autoassociativas e Mel-espectrogramas, a fim de escolher o modelo mais vantajoso para a etapa de classificação multiclasse. Por último, a matriz de confusão permitiu avaliar os resultados do segundo tipo de classificação.

3.1 Algoritmos de Aprendizado de Máquinas

Os algoritmos de aprendizado supervisionado apresentados foram selecionados entre os frequentemente aplicados nas tarefas de classificação [19]. A seção descreve a ideia central de funcionamento de cada um, assim como a configuração adotada em modelos onde o padrão pré-configurado não foi utilizado. As bibliotecas Scikit-learn [20] e Keras possibilitaram a implementação dos modelos e sua avaliação.

Naive Bayes

Os classificadores do tipo Naive Bayes são baseados no Teorema de Bayes e assumem que as características das amostras são condicionalmente independentes entre si, dada a classe. Considerando que existem n classes e cada amostra do conjunto de treino possui m atributos, as probabilidades de uma amostra inédita x pertencer a cada uma das n classes é calculada assumindo que os m atributos são independentes entre si. Desta maneira, o rótulo de saída é definido da maior probabilidade calculada.

Floresta Aleatória

O método de floresta aleatória é comumente aplicado em tarefas de regressão e classificação, apesar de também existirem trabalhos que o adotam para eleição de atributos. São utilizadas múltiplas árvores de decisão e, no caso de trabalhos de classificação, o rótulo de saída é a moda entre as saídas das árvores individuais.

Máquinas de Vetores de Suporte

A técnica de Máquinas de Vetores de Suporte (*Support Vector Machines* – SVM) tem sua ideia principal apresentada na Figura 3.1, onde é ilustrada a separação de duas classes distintas a partir de um hiperplano. Em SVM a distância mínima entre as classes é denominada margem e o objetivo do algoritmo é encontrar o hiperplano capaz de maximizá-la.

XGBoost

O XGBoost (*eXtreme Gradient Boosting*) é uma versão escalável do algoritmo *gradient boosting* [21]. De forma similar ao método da floresta aleatória, o algoritmo utiliza árvores de decisão para obter o rótulo da amostra. Entretanto, ao invés de decidir a saída por maioria de votos, combina os classificadores considerados fracos iterativamente. A função de custo do modelo é minimizada adotando o método do gradiente cada vez que uma nova árvore é adicionada.

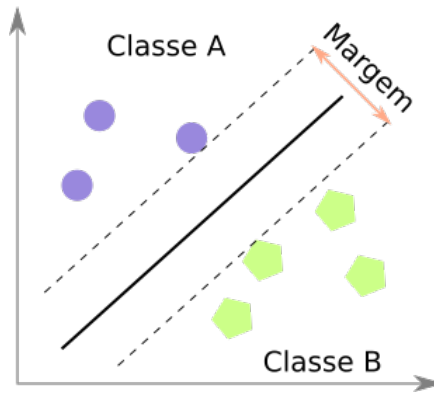


Figura 3.1: Separação de duas classes pelo algoritmo SVM.

Perceptron de Múltiplas Camadas

As redes neurais do tipo Perceptron de Múltiplas Camadas (*Multilayer Perceptron* — MLP) possuem uma ou mais camadas ocultas de neurônios, além das camadas de entrada e saída [22].

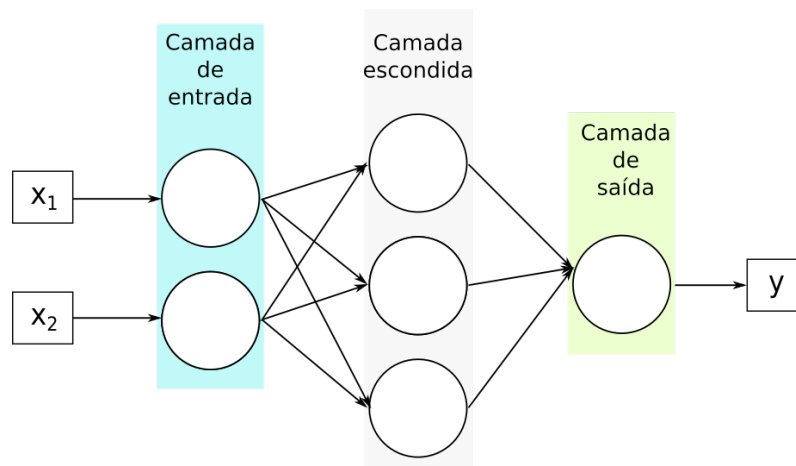


Figura 3.2: Exemplo de rede neural com camadas totalmente conectadas.

Na Figura 3.2 é apresentada uma rede neural totalmente conectada com dois neurônios na camada de entrada, três na chamada camada escondida e um neurônio na camada de saída. x_1 e x_2 representam os atributos de entrada da rede e y sua saída.

3.2 Métricas para avaliação

As métricas adotadas para avaliação dos resultados na classificação binária utilizam os conceitos apresentados na Figura 3.3. Os quadrados em verde representam as classificações realizadas corretamente e os em laranja a situação inversa. Na tarefa de reconhecimento de falha realizada na etapa experimental, as amostras que

descrevem o comportamento anômalo recebem o rótulo positivo e as amostras normais o rótulo negativo. Diferentes métricas foram adotadas para avaliar problemas possíveis como, por exemplo, a tendência dos modelos a rotular uma nova entrada como pertencente à classe com maior número de amostras. Apesar da prioridade ser aumentar a taxa de verdadeiros positivos (VP), é importante reduzir as taxas de falso positivo (FP) e falso negativo (FN). Enquanto altas taxas de FP revelam que o sistema considera que muitas amostras normais descrevem o mau funcionamento do sistema, altas taxas de FN deixam de avisar quando um equipamento não está funcionando conforme o esperado. Para avaliar o desempenho dos modelos na classificação binária, foram escolhidas as métricas *recall*, precisão, F1-score, acurácia e AUC.

		Rótulo verdadeiro	
		Positivo	Negativo
Rótulo predito	Positivo	Verdadeiro Positivo	Falso Positivo
	Negativo	Falso Negativo	Verdadeiro Negativo

Figura 3.3: Matriz de Confusão para classificação binária.

Acurácia

A acurácia, vista na Equação 3.1, calcula a relação das classificações corretas, verdadeiro positivo (VP) e verdadeiro negativo (VN), e todas as classificações realizadas, que incluem também os Falsos Negativos (FN) e Falsos Positivos (FP). Logo, essa métrica pode ser interpretada como a taxa de acertos do modelo. No caso da classificação de estados anômalos, porém, a acurácia pode mascarar um sistema com taxas significativas de FP e FN.

$$acurácia = \frac{VP + VN}{VP + FN + VN + FP}. \quad (3.1)$$

Recall

A métrica *recall* (ou sensibilidade) descrita na Equação 3.2 revela a taxa de VP entre todas as classes positivas do conjunto de dados. Ou seja, avalia quantas das amostras anômalas foram classificadas corretamente. Considerando que a *recall*

varia entre 0 e 1, quanto menor seu valor, maior o número de amostras anômalas sendo classificadas como normais.

$$recall = \frac{VP}{VP + FN}. \quad (3.2)$$

Precisão

A precisão avalia a taxa de VP entre todos os valores positivos preditos, sendo descrita pela Equação 3.3. Enquanto a *recall* reduz em situações onde o classificador não interpreta corretamente um cenário de falha, a precisão diminui com o aumento de falsos alarmes.

$$precisão = \frac{VP}{VP + FP}. \quad (3.3)$$

F1-score

As métricas *recall* e precisão sinalizam, indiretamente, se a taxa de falsos positivos e falsos negativos estão comprometendo a classificação. A F1-score é responsável por combinar ambas as métricas da maneira descrita na Equação 3.4. Dessa forma, busca-se maximizar o valor da métrica assim como a precisão e a *recall*. Quanto mais próximo de 1 o valor da F1-score estiver, mais êxito em identificar a classe positiva.

$$F1 = \frac{2 \cdot recall \cdot precisão}{precisão + recall}. \quad (3.4)$$

AUC

Para compreender a AUC, é necessário entender a Curva Característica de Operação do Receptor (*Receiver Operating Characteristic* – ROC) representada na Figura 5.5. A curva descreve a relação entre a taxa de verdadeiros positivos e a taxa de falsos positivos. A taxa de verdadeiros positivos é a relação entre VP e o total de amostras positivas no conjunto de dados. Logo, a taxa de verdadeiros positivos é o valor de *recall*. Por outro lado, a taxa de falsos positivos é a relação entre FP e o total de amostras positivas.

A AUC é a área sob a curva ROC e facilita a comparação entre diferentes modelos. Assim como as outras métricas adotadas, seu valor pode variar entre 0 e 1, onde quanto mais alto o valor mais correta foi realizada a classificação. Essa foi a única métrica adotada pelo trabalho base e foi escolhida para comparação entre os modelos adotados nesta dissertação e a abordagem baseada em Mel-espectrogramas.

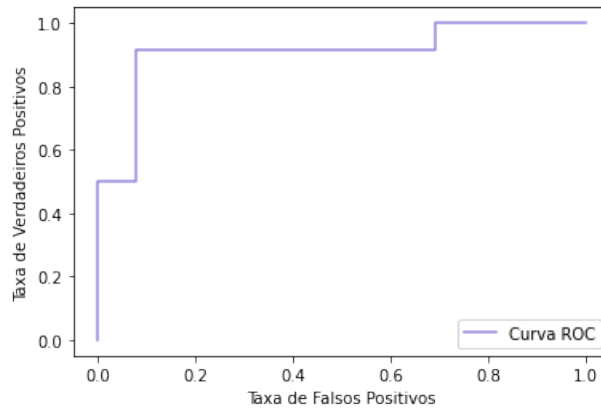


Figura 3.4: Curva ROC.

Matriz de Confusão Multiclasse

Uma forma de avaliar o desempenho de um classificação com mais de uma classe é através de uma matriz de confusão, como a descrita na Figura 3.5. Os quadrados em laranja representam as situações onde o classificador assumiu que determinada amostra pertencia a uma classe diferente da verdadeira. A diagonal de uma matriz de confusão representa as classificações realizadas corretamente. É possível apresentar a matriz tanto com a quantidade total de amostras, quanto na forma normalizada. Neste trabalho adota-se esta última opção, onde os valores das matrizes estão entre 0 e 1.

		Rótulo verdadeiro		
		Classe A	Classe B	Classe C
Rótulo predito	Classe A	Classificação correta	Classificação incorreta	Classificação incorreta
	Classe B	Classificação incorreta	Classificação correta	Classificação incorreta
	Classe C	Classificação incorreta	Classificação incorreta	Classificação correta

Figura 3.5: Matriz de Confusão para classificação multiclasse.

Capítulo 4

Metodologia

Neste capítulo são apresentados os resultados obtidos pelos classificadores binários e na classificação de múltiplas classes adotando os algoritmos selecionados e descritos no Capítulo 3. A primeira classificação é apenas capaz de identificar, analisando um modelo de equipamento específico, se o funcionamento é normal ou não. Apesar da informação poder ser adotada para auxiliar a manutenção corretiva, necessita que para cada tipo de máquina um modelo diferente seja treinado. A classificação multiclasse busca possibilitar que um único modelo seja capaz de identificar o tipo de equipamento e seu estado.

4.1 Características Espectrais do Som

As características extraídas das ondas sonoras foram selecionadas buscando representar de forma resumida as amostras de som. O objetivo é aumentar o desempenho da classificação, tornando a adoção de imagens e seus classificadores com alto número de parâmetros desnecessários. As características adotadas e seus significados na descrição dos espectros são apresentadas nesta seção.

- Croma (Chroma): a característica de croma é capaz de representar concisamente os harmônicos e sinais musicais das doze notas musicais através de um vetor com doze elementos. A magnitude do espectro é obtida através da Transformada de Fourier de Tempo Curto (*Short Time Fourier Transform – STFT*).
- Coeficientes Mel-cepstrais (Mel-frequency Cepstral Coefficients – MFCCs): os coeficientes são capazes de descrever foneticamente o envelope do som. O envelope é a amplitude instantânea de um sinal oscilatório. Essa característica é amplamente adotada em reconhecimento de voz por sua representação compacta da energia do espectro, como expressa na escala de

frequência Mel [23]. A conversão entre a frequência f e a escala Mel é descrita na Equação 4.1:

$$Mel(f) = 2595 \log_{10} \left(1 + \frac{f}{100} \right). \quad (4.1)$$

É possível obter os MFCCs através da STFT de cada frequência convertida na escala Mel. Essa etapa é descrita na Equação 4.2, onde M_{mfcc} é o número de bandas de frequência na escala Mel, $\tilde{X}_m(t)$ é a energia da banda da m -ésima frequência na escala Mel e c em $[1 - M_{mfcc}]$ é o índice do coeficiente cepstral.

$$MFCC = \sqrt{\frac{2}{M_{mfcc}} \sum_{m=1}^{M_{mfcc}} \log(\tilde{X}_m(t)) \cos\left(\frac{c(n - \frac{1}{2})}{M_{mfcc}}\right)} \quad (4.2)$$

- Centroide Espectral (Spectral Centroid – SC): representa a magnitude do centro de massa do espectro. Essa característica está associada à ideia subjetiva de clareza do som. Os sinais com frequências mais altas tendem a ser percebidos como sinais mais claros. Na Equação 4.3, o centroide espectral é a média das frequências presentes no i -ésimo quadro presente no sinal ponderadas por suas amplitudes $X_i(k)$.

$$SC_i = \frac{\sum_{k=1}^M k X_i(k)}{\sum_{k=1}^M X_i(k)}; \quad (4.3)$$

- Largura Espectral (Spectral Band – SB): a Largura Espectral é calculada utilizando o Centroide Espectral SC e é definida como descrita na Equação 4.4. Nessa equação, a amplitude $X_i(k)$ do i -ésimo quadro do sinal é multiplicada pela diferença da frequência e do SC ao quadrado. A largura espectral é a raiz quadrada do somatório realizado para toda a amostra.

$$SB = \sqrt{\sum_{k=1}^M X_i(k) (f_k - SC)^2}; \quad (4.4)$$

- Decaimento Espectral (Spectral roll-off – SR): o Decaimento Espectral é a frequência abaixo da qual está 85% da energia do espectro. O SR é obtido através da Equação 4.5:

$$\sum_{k=1}^{SR} X_i(k) = 0.85 \sum_{k=1}^M X_i(k). \quad (4.5)$$

- Taxa de Cruzamento por Zero (Zero Crossing Rate – ZCR): é o

número de vezes que um sinal de onda de áudio cruza o zero e é definido da seguinte forma.:

$$ZCR = \frac{1}{2N} \sum_{t=0}^{N-2} |sgn(x(t+1)) - sgn(x(t))|, \quad (4.6)$$

$$\text{onde } sgn(x(t)) = \begin{cases} 1, & \text{se } x(t) \geq 0, \\ -1, & \text{se } x(t) < 0 \end{cases}.$$

4.2 Etapas do experimento

A Figura 4.1 descreve todas as etapas utilizadas nos dois tipos de classificação. As amostras de sons originais são convertidas para a extração dos atributos selecionados. Estes passos são realizados apenas uma vez. O conjunto de dados convertido é pré-processado em cada uma das rodadas, garantindo que os dados apresentam o formato adequado para o treinamento e avaliação dos modelos. A média das métricas é realizada após 10 rodadas de treinamento e teste e este valor final é utilizado na avaliação dos resultados. A escolha do algoritmo adotado na classificação multiclasse baseia-se no resultados das métricas acurácia, precisão, *recall* e F1-score para os cinco algoritmos, no resultado da métrica AUC comparando-se os cinco algoritmos ao trabalho de Purohit et al. e o tempo necessário para treinamento e teste dos modelos. Nesta segunda classificação os resultados são avaliados através da matriz de confusão multiclasse.

A seguir, as diferentes etapas abordadas na Figura 4.1 serão detalhadas:

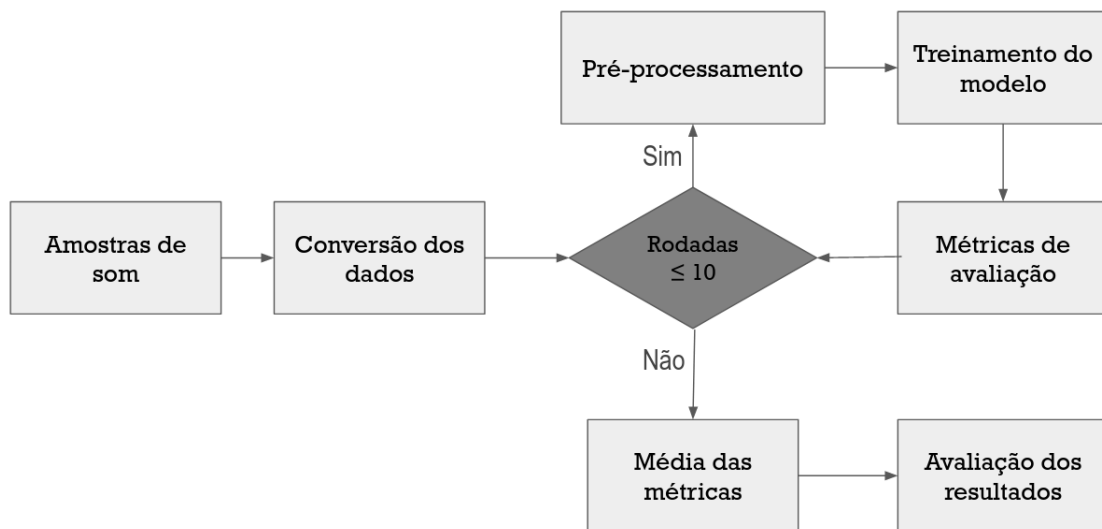


Figura 4.1: Etapas realizadas no experimento.

- Amostras de som: O conjunto de dados MIMII (*Malfunctioning Industrial Machine Investigation and Inspection*) é composto por quatro tipos de máquinas diferentes, ventiladores industriais, bombas hidráulicas, trilhos deslizantes e válvulas solenóides. Além disso, este conjunto de dados possui as quantidades de amostras normais e anômalas conforme apresentadas na Tabela 4.1. Cada amostra tem duração igual a 10 segundos e cada tipo de máquina contém quatro modelos individuais, identificados pelos IDs 00, 02, 04 e 06. Os dados são desbalanceados, tendo a classe de estados normais com um maior número de itens que a classe dos estados de falha, dificultando o ajuste dos modelos. Como exemplo, é possível observar que a quantidade de amostras anômalas da bomba hidráulica de ID 06 é aproximadamente igual a 10% da quantidade de amostras normais. Além disso, foram adicionados ruídos de fábrica aos sons originais resultando em três valores distintos de SNR: 6dB, 0dB e -6dB. Quanto maior o valor da relação sinal-ruído, mais facilmente a tarefa é realizada, pois é maior a potência do sinal de interesse em relação aos sinais de ruído.

Tabela 4.1: Conjunto de dados MIMII.

Tipo de Máquina	ID da Máquina	Amostras normais	Amostras anômalas
Ventilador Industrial	00	1011	407
	02	1016	359
	04	1033	348
	06	1015	361
Bomba Hidráulica	00	1006	143
	02	1005	111
	04	702	100
	06	1036	102
Trilho Deslizante	00	1068	356
	02	1068	267
	04	534	178
	06	534	89
Válvula	00	991	119
	02	708	120
	04	1000	120
	06	992	120

- Conversão dos dados: As características dos espectros de som são extraídas e salvas no formato CSV na etapa de conversão de dados. O novo conjunto de dados será utilizado nas etapas de classificação. As funções da biblioteca python librosa [24] são utilizadas para extração das características selecionadas.

- Pré-processamento: A etapa de pré-processamento pode englobar quatro ações distintas: divisão dos conjuntos de treino e teste; *oversampling*, padronização e transformação dos rótulos. O conjunto de treino de cada um dos modelos apresenta 60% das amostras normais e 60% das amostras anômalas do conjunto de dados original selecionadas aleatoriamente em cada uma das rodadas. As amostras restantes compõem o conjunto de teste. Em seguida as amostras são padronizadas, isto é, todas as amostras são transformadas para terem média igual a 0 e desvio padrão igual a 1. Uma vez divididos, utiliza-se o método *oversampling* para igualar o número de amostras anômalas ao número de amostras normais no conjunto de treino, buscando facilitar o ajuste de parâmetros na etapa de treinamento. Para isto, amostras anômalas do conjunto de treino são selecionadas aleatoriamente e duplicadas até que a quantidade de amostras de ambas as categorias se iguale. O novo valor é calculado como descrito na Equação 4.7, onde x é o valor da amostra, μ o valor médio das amostras e σ o desvio padrão. Ao colocar as variáveis na mesma escala, é possível comparar os valores entre todas as características selecionadas.

$$z = \frac{(x - \mu)}{\sigma} \quad (4.7)$$

Por último, os rótulos das amostras são convertidos em valores binários. Para a classificação binária, o estado normal de operação assume o valor 0 e o estado anômalo o valor 1. Para a classificação multiclasse, considera-se apenas o tipo da máquina, unindo todos os IDs e apenas preocupando-se com o estado e o tipo da máquina. Dessa forma, totaliza-se 8 classes: ventilador normal, ventilador anômalo, bomba normal, bomba anômala, trilho deslizante normal, trilho deslizante anômalo, válvula normal e válvula anômala. Espera-se que o mau funcionamento possa não apenas ser detectado confrontando as amostras de um mesmo equipamento, mas também identificado junto com o tipo de máquina a qual pertence.

- Treinamento do modelo: a etapa de treinamento dos modelos consiste no fornecimento de dados aos algoritmos supervisionados para que ocorra o processo de aprendizagem. Os modelos devem ser capazes de identificar os padrões nos dados de treinamento e mapear as características de entrada com o rótulo final.
- Avaliação dos resultados: A avaliação de desempenho na classificação binária é realizada a partir das métricas AUC, acurácia e F1-score. As métricas permitem que os modelos sejam comparados entre si e a AUC viabiliza a comparação com o modelo de base proposto por Purohit et al. Após a classifi-

cação binária, o algoritmo com melhor desempenho, considerando as métricas e o tempo necessário para treinar e testar os modelos, é selecionado para a classificação multiclasse. Nesta classificação, a avaliação das métricas é realizada a partir da matriz de confusão.

Capítulo 5

Resultados e Discussões

Neste capítulo são apresentados os resultados obtidos na classificação binária e na classificação multiclasse. O objetivo é analisar se a adoção das características extraídas das amostras de som tornam possível a identificação dos cenários e, no caso da classificação binária, se o desempenho obtido por Purohit et al. é superado. O algoritmo selecionado para a classificação multiclasse é adotado com e sem o método de *oversampling* para que o impacto do número de amostras desbalanceado do conjunto de dados seja demonstrado. A seguir, os resultados obtidos a partir dos dois tipos de classificação adotados, binária e multiclasse, são apresentados.

5.1 Classificação Binária

O objetivo da classificação binária é treinar um modelo capaz de identificar amostras sonoras fora do padrão esperado para cada máquina individual. Dessa forma, para cada SNR, 16 modelos por algoritmo são treinados. Os modelos são analisados a partir das métricas acurácia, *recall*, precisão, F1-score e AUC.

Tabela 5.1: Tempo em segundos para treinar e testar os modelos.

Rodada	NB	RF	SVM	XGBoost	MLP
1 ^a Rodada	2.06	11.75	3.23	6.32	136.00
2 ^a Rodada	2.40	11.26	3.20	7.22	138.51
3 ^a Rodada	2.37	11.20	3.26	6.83	142.32
4 ^a Rodada	1.89	11.35	3.21	6.43	145.01
5 ^a Rodada	2.16	11.20	3.43	5.60	145.24
6 ^a Rodada	1.89	11.25	3.38	6.02	148.45
7 ^a Rodada	1.90	11.28	3.47	5.12	141.59
8 ^a Rodada	1.88	11.36	3.19	5.10	157.65
9 ^a Rodada	1.88	11.56	3.21	5.34	146.26
10 ^a Rodada	1.89	11.43	3.45	4.31	143.94
Média das Rodadas	2.03	11.36	3.30	5.83	144.50

A acurácia simboliza a taxa de sucesso da classificação. As métricas *recall* e precisão evidenciam possíveis valores de falsos positivos e falsos negativos que a acurácia pode mascarar. A F1-score combina as métricas *recall* e precisão, uma vez que o objetivo é encontrar a melhor combinação de ambos. Por último, a AUC permite a comparação dos modelos desta dissertação com o proposto por Purohit et al. [25] que adota redes autoassociativas em conjunto com Mel-espectrogramas. Os modelos são treinados por 10 rodadas independentes e a média de valores obtidos nas métricas foi o adotado para compará-los. A Tabela 5.1 contém os tempos em segundos necessários pelos algoritmos para o treinamento e teste de todos os 48 modelos. Os modelos podem ser organizados da seguinte forma por ordem crescente de tempo: Naive Bayes (NB), Máquina de Vetores de Suporte (SVM), XGBoost, Floresta Aleatória (RF) e Perceptron de Múltiplas Camadas (MLP). O modelo baseado em redes neurais é mais de 12 vezes mais lento que RF e aproximadamente 71 vezes mais lento que NB.

Acurácia

As acurácias médias atingidas após as 10 rodadas são apresentadas nas Tabelas 5.2, 5.3, 5.4 e 5.5. Os valores mais altos são evidenciados em negrito. As acurácias

Tabela 5.2: Acurácia para os modelos na classificação binária de amostras anômalas dos ventiladores industriais.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Ventilador Industrial	00	0 dB	0.6051	0.7446	0.8501	0.7601	0.8143
		6 dB	0.7778	0.8242	0.9695	0.8702	0.9314
		- 6dB	0.5198	0.6593	0.7420	0.6584	0.7033
	02	0 dB	0.7935	0.8729	0.978	0.9231	0.9603
		6 dB	0.9224	0.9422	0.9945	0.9398	0.9850
		- 6dB	0.7113	0.7615	0.9013	0.7916	0.8818
	04	0 dB	0.8301	0.8585	0.9746	0.8772	0.9428
		6 dB	0.9295	0.7719	0.9987	0.8784	0.9746
		- 6dB	0.5042	0.7571	0.8500	0.7884	0.8396
	06	0 dB	0.8367	0.9076	0.9889	0.8225	0.9806
		6 dB	0.6813	0.9709	0.9985	0.8409	0.9954
		- 6dB	0.8111	0.8607	0.9735	0.9385	0.9709

obtidas para os ventiladores industriais são apresentadas na Tabela 5.2. Em todas as taxas de SNR os melhores resultados são alcançados pelo algoritmo SVM, que alcança taxas de acerto próximas ao valor máximo. As bombas hidráulicas e trilhos deslizantes também apresentam bons resultados de acurácia e são apresentados respectivamente nas Tabelas 5.3, 5.4. Nesses casos SVM possui o melhor resultado na maioria dos resultados, perdendo para MLP apenas na relação SNR mais baixa.

Tabela 5.3: Acurácia para os modelos na classificação binária de amostras anômalas das bombas hidráulicas.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Bomba Hidráulica	00	0 dB	0.6479	0.7479	0.8983	0.7978	0.7220
		6 dB	0.6351	0.7664	0.8924	0.7800	0.7325
		- 6dB	0.6137	0.7614	0.8913	0.7858	0.7997
	02	0 dB	0.6670	0.6957	0.8417	0.7516	0.7827
		6 dB	0.6446	0.7630	0.8984	0.7926	0.8253
		- 6dB	0.6702	0.6572	0.7222	0.7148	0.7265
	04	0 dB	0.7134	0.8134	0.9103	0.7866	0.7759
		6 dB	0.8069	0.9106	0.9679	0.8389	0.7900
		- 6dB	0.6523	0.7293	0.8879	0.7598	0.7963
	06	0 dB	0.6400	0.6820	0.9035	0.7927	0.8094
		6 dB	0.6422	0.6901	0.9180	0.7890	0.7757
		- 6dB	0.5578	0.6547	0.7760	0.7396	0.7845

Uma das possibilidades é o fato do ruído dificultar que o hiperplano de separação seja encontrado por SVM. As acurácias das válvulas estão na Tabela 5.5. O equipamento é o que se comporta de maneira menos previsível, tendo melhores resultados divididos entre MLP, SVM e XGBoost.

Tabela 5.4: Acurácia para os modelos na classificação binária de amostras anômalas dos trilhos deslizantes.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Trilho Deslizante	00	0 dB	0.8105	0.8081	0.9859	0.8756	0.9621
		6 dB	0.9011	0.8977	0.9909	0.9295	0.9687
		- 6dB	0.6910	0.7506	0.9671	0.8492	0.9409
	02	0 dB	0.7318	0.8539	0.9273	0.8507	0.9227
		6 dB	0.7588	0.8419	0.9783	0.8916	0.9436
		- 6dB	0.6346	0.7895	0.9028	0.8320	0.8641
	04	0 dB	0.7677	0.7463	0.9547	0.7870	0.9178
		6 dB	0.8491	0.8151	0.9782	0.8358	0.9318
		- 6dB	0.6639	0.7028	0.9116	0.7586	0.9103
	06	0 dB	0.5896	0.7156	0.8364	0.7440	0.8274
		6 dB	0.6164	0.7440	0.8772	0.7528	0.8167
		- 6dB	0.4688	0.6452	0.7736	0.7128	0.7817

Os resultados podem indicar que existem dois fatores determinantes para o diagnóstico de falhas à partir do som: a quantidade de amostras da classe de interesse e a diferença entre os sons de interesse e os demais.

Apesar de aplicado o *oversampling* para contornar o desbalanceamento dos subconjuntos de cada modelo de equipamento, o método apenas repete amostras já presentes no conjunto de treino original. Logo, em subconjuntos com grande diferença

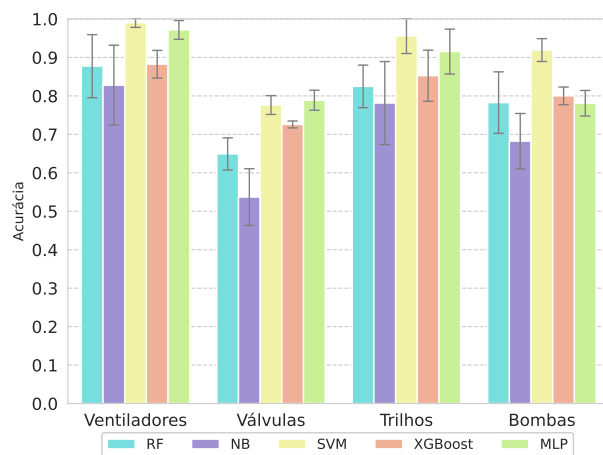
Tabela 5.5: Acurácia para os modelos na classificação binária de amostras anômalas das válvulas.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Válvula	00	0 dB	0.5097	0.6520	0.6718	0.7182	0.6817
		6 dB	0.4797	0.6131	0.7381	0.7209	0.7453
		- 6dB	0.4414	0.6581	0.6473	0.6991	0.6694
	02	0 dB	0.6033	0.6012	0.7175	0.6746	0.6907
		6 dB	0.4483	0.7118	0.7967	0.7396	0.8101
		- 6dB	0.4779	0.6163	0.6659	0.7178	0.6843
	04	0 dB	0.6022	0.6011	0.7190	0.7060	0.7580
		6 dB	0.6078	0.6089	0.7987	0.7154	0.8074
		- 6dB	0.5739	0.5949	0.6855	0.7134	0.6854
	06	0 dB	0.4838	0.6034	0.6978	0.7196	0.7500
		6 dB	0.6117	0.6622	0.7715	0.7270	0.7922
		- 6dB	0.4106	0.6126	0.6751	0.7337	0.7392

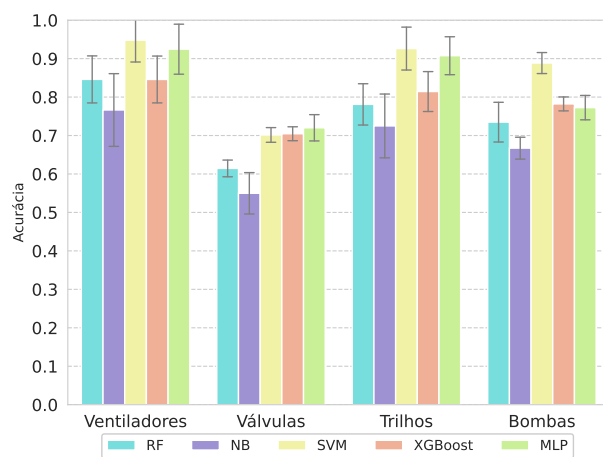
entre as amostras normais e anômalas a probabilidade do método repetir diversas vezes a mesma amostra é alta. Os ventiladores industriais são os que possuem a menor diferença proporcional entre as classes. Entretanto, este não é o único fator para explicar os resultados, uma vez que as bombas hidráulicas e as válvulas possuem números próximos de amostras normais e anômalas. A segunda possibilidade para o desempenho abaixo dos demais equipamentos é a diferença pouco perceptível do padrão do som entre as classes.

Analisando separadamente as SNR e realizando a média dos valores obtidos nas Tabelas é possível comparar visualmente o desempenho dos modelos para cada tipo de componente. Estes valores são representados através de gráficos de barras na Figura 5.1. A barra azul claro refere-se ao algoritmo RF; a barra roxa representa o NB; a barra amarela o SVM; a barra laranja o XGBoost; e por último, a barra verde representa MLP. As barras em cinza demarcam os valores mínimos e máximos obtidos. Os resultados das diferentes SNRs são apresentados em ordem decrescente pelas Figuras 5.1a, 5.1b e 5.1c, respectivamente.

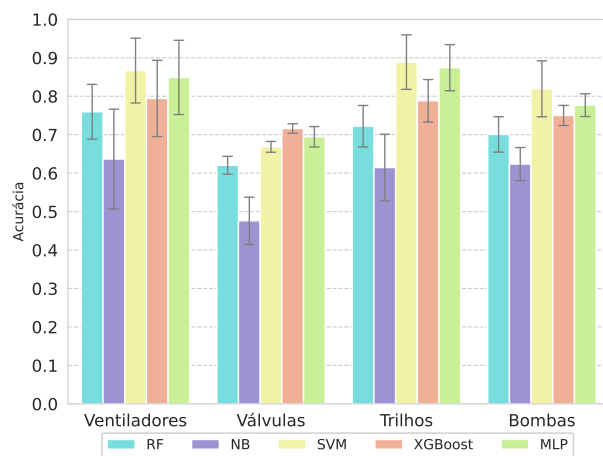
Começando pela mais alta relação sinal-ruído presente na Figura 5.1a, o algoritmo SVM alcança maior acurácia média para os ventiladores. O segundo melhor resultado é alcançado pelo MLP. O NB é o que apresenta a acurácia mais baixa entre todos os outros modelos. Em quase todos os tipos de máquinas, a tendência se mantém, exceto para as válvulas, onde a acurácia do MLP supera a do SVM. O MLP e o SVM se mantêm como os melhores nas SNRs iguais a 0dB e -6dB, nas Figuras 5.1b e 5.1c, respectivamente.



(a) Acurácia dos modelos para SNR = 6 dB.



(b) Acurácia dos modelos para SNR = 0 dB.



(c) Acurácia dos modelos para SNR = -6 dB.

Figura 5.1: Comparação da acurácia obtida pelos cinco modelos de aprendizado supervisionado para classificação binária.

Tabela 5.6: Precisão para os modelos na classificação binária de amostras anômalas dos ventiladores.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Ventiladores industriais	00	0 dB	0.3954	0.5652	0.6951	0.5743	0.6280
		6 dB	0.5954	0.6557	0.9192	0.7142	0.8232
		- 6 dB	0.3252	0.4246	0.5484	0.4261	0.4963
	02	0 dB	0.6151	0.7238	0.9312	0.8051	0.8874
		6 dB	0.7939	0.8616	0.9798	0.8200	0.9465
		- 6 dB	0.4580	0.5414	0.7737	0.5766	0.7206
	04	0 dB	0.6223	0.6629	0.9152	0.6903	0.8182
		6 dB	0.7854	0.5488	0.9950	0.6808	0.9093
		- 6 dB	0.3043	0.5166	0.6467	0.5610	0.6237
	06	0 dB	0.6219	0.7820	0.9690	0.6241	0.9413
		6 dB	0.4720	0.9560	0.9945	0.7944	0.9829
		- 6 dB	0.5858	0.7078	0.9376	0.8519	0.9262

Precisão

A precisão obtida por todos os modelos é apresentada nas Tabelas 5.6, 5.7, 5.8 e 5.9.

Os valores baixos de precisão indicam que os valores de falsos alarmes podem estar altos. Os resultados de precisão para os ventiladores industriais são apresentados na Tabela 5.6. Assim como na acurácia SVM alcança os melhores resultados para o equipamento, tendo na maior SNR valores com acima de 0.91 para todos os modelos individuais. O aumento da potência do ruído reduz a precisão, mas a classificação ainda é realizada de maneira satisfatória.

Tabela 5.7: Precisão para os modelos na classificação binária de amostras anômalas das bombas hidráulicas.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Bomba Hidráulica	00	0 dB	0.2519	0.3139	0.5696	0.3667	0.3122
		6 dB	0.2537	0.3423	0.5424	0.3672	0.3207
		- 6 dB	0.2233	0.3145	0.5497	0.3407	0.3837
	02	0 dB	0.1905	0.2302	0.3798	0.2768	0.3159
		6 dB	0.2179	0.2945	0.4965	0.3241	0.3741
		- 6 dB	0.1748	0.1912	0.2541	0.2308	0.2650
	04	0 dB	0.3038	0.4034	0.5833	0.3702	0.3585
		6 dB	0.3935	0.5855	0.8018	0.4404	0.3757
		- 6 dB	0.2635	0.3145	0.5297	0.3378	0.3810
	06	0 dB	0.1884	0.2124	0.4851	0.2925	0.3237
		6 dB	0.1998	0.2241	0.5331	0.3015	0.2956
		- 6 dB	0.1493	0.1862	0.2719	0.2241	0.2888

Para as bombas hidráulicas e trilhos deslizantes os resultados são apresentados

respectivamente nas Tabelas 5.7 e 5.8. Assim como observado na acurácia, SVM também é a melhor opção na maioria dos casos. Entretanto, a diferença entre as amostras dos modelos individuais de cada tipo de máquina é evidenciada nesta métrica.

Tabela 5.8: Precisão para os modelos na classificação binária de amostras anômalas dos trilhos deslizantes.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Trilho deslizante	00	0 dB	0.5692	0.566	0.9469	0.6689	0.8702
		6 dB	0.7166	0.7097	0.965	0.7804	0.8895
		- 6 dB	0.4462	0.5013	0.8893	0.6248	0.8147
	02	0 dB	0.4174	0.6056	0.748	0.5903	0.7289
		6 dB	0.4454	0.5792	0.9052	0.6630	0.7879
		- 6 dB	0.3302	0.4862	0.7171	0.5613	0.6103
	04	0 dB	0.5210	0.5002	0.8681	0.5470	0.7631
		6 dB	0.6353	0.5822	0.9384	0.6171	0.7941
		- 6 dB	0.4154	0.4503	0.7715	0.5137	0.7593
	06	0 dB	0.2412	0.3196	0.4741	0.3434	0.4509
		6 dB	0.2630	0.3582	0.5433	0.3600	0.4392
		- 6 dB	0.1761	0.2548	0.3829	0.2929	0.3875

Como exemplo, enquanto a bomba hidráulica de ID 04 obtém 0.80 de precisão para SNR igual a 6dB, a bomba de ID 02 alcança aproximadamente 0.5. Ou seja, a qualidade ou tipo de mal-funcionamento apresentado por máquinas distintas pertencentes ao mesmo tipo influenciam na qualidade da classificação realizada. Para cenários onde a potência do sinal de interesse é maior que os ruídos do ambiente, a classificação pode ser realizada com alta precisão para equipamentos como trilhos deslizantes e ventiladores. O resultado das válvulas são apresentados na Tabela 5.9.

A abordagem que mais demonstra reduzir os falsos alarmes é a MLP. Porém, o baixo valor da métrica demonstra que mesmo no melhor dos casos a quantidade de falsos positivos do equipamento é alta.

Este comportamento pode ser explicado novamente pela proximidade do padrão dos sons normais e anômalos em conjunto com a grande diferença entre a quantidade de amostras das classes.

A Figura 5.2 resume os resultados obtidos em cada equipamento, separando os cenários por SNR. A média das métricas representada pelas barras coloridas revela que as abordagens MLP e SVM alcançam resultados próximos. O método NB é apontado como a pior opção em todos os casos, sendo este um resultado esperado por basear-se na independência dos atributos entre si. Os métodos RF e XGBoost, alcançam resultados próximos e intermediários para precisão. Esta tendência indica que propostas baseadas em árvores de decisão podem não ser as mais indicadas para captar as diferenças de padrão para amostras que representam defeitos.

Tabela 5.9: Precisão para os modelos na classificação binária de amostras anômalas das válvulas.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Válvula	00	0 dB	0.1252	0.1706	0.2102	0.2036	0.2218
		6 dB	0.1373	0.1878	0.2686	0.2152	0.2817
		- 6 dB	0.1210	0.1603	0.1825	0.1788	0.2045
	02	0 dB	0.1920	0.2148	0.3073	0.2237	0.3032
		6 dB	0.1735	0.3138	0.4116	0.3354	0.4322
		- 6 dB	0.1646	0.1904	0.2500	0.2382	0.2703
	04	0 dB	0.1401	0.1716	0.2584	0.2049	0.2873
		6 dB	0.1882	0.1969	0.3343	0.2458	0.3589
		- 6 dB	0.1187	0.1526	0.2114	0.1790	0.2175
	06	0 dB	0.1369	0.1625	0.2364	0.2101	0.2845
		6 dB	0.1606	0.2019	0.3101	0.2225	0.3359
		- 6 dB	0.1195	0.1455	0.2105	0.1760	0.2633

Tabela 5.10: *Recall* para os modelos na classificação binária de amostras anômalas dos ventiladores.

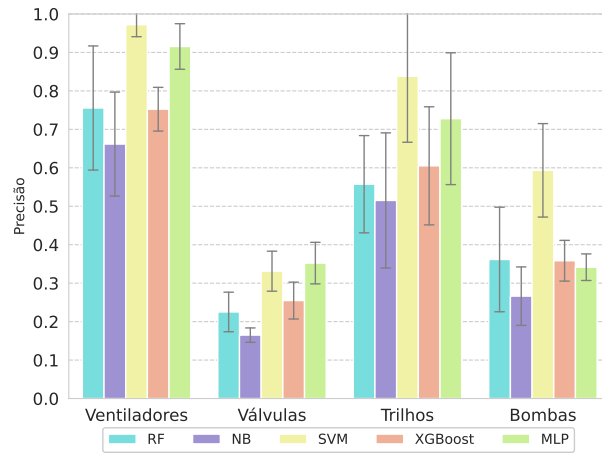
Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Ventilador Industrial	00	0 dB	0.6822	0.5344	0.8644	0.6638	0.8821
		6 dB	0.7196	0.8325	0.9804	0.9227	0.9772
		- 6dB	0.6190	0.5104	0.6117	0.5356	0.7008
	02	0 dB	0.5806	0.8333	0.9896	0.9382	0.9741
		6 dB	0.9528	0.9458	1.0000	0.9910	1.0000
		- 6dB	0.5375	0.6535	0.8840	0.7979	0.9028
	04	0 dB	0.8683	0.9007	0.9921	0.9439	0.9952
		6 dB	0.9928	0.9957	1.0000	0.9906	1.0000
		- 6dB	0.7532	0.7309	0.9043	0.7849	0.9308
	06	0 dB	0.9653	0.9007	0.9896	0.9618	0.9880
		6 dB	1.0000	0.9333	1.0000	0.7118	1.0000
		- 6dB	0.9590	0.7993	0.9632	0.9285	0.9667

Recall

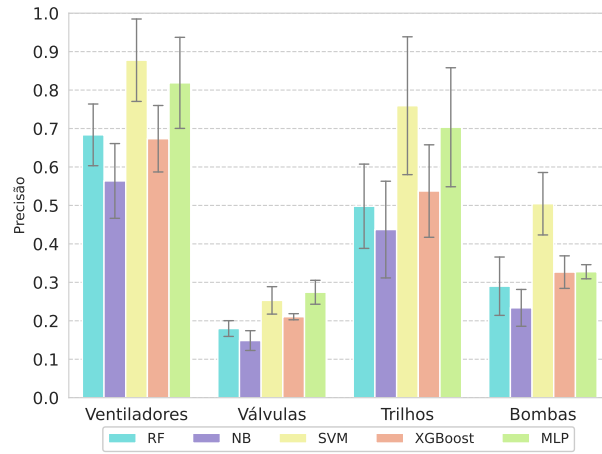
As Tabelas 5.10, 5.11, 5.12 e 5.13 apresentam os resultados da métrica *recall* para, respectivamente, os ventiladores industriais, bombas hidráulicas, trilhos deslizantes e válvulas.

Os melhores resultados obtidos para a classificação de mal-funcionamento em ventiladores industriais apresentados na Tabela 5.10 são alcançados SVM e MLP na maior parte dos modelos.

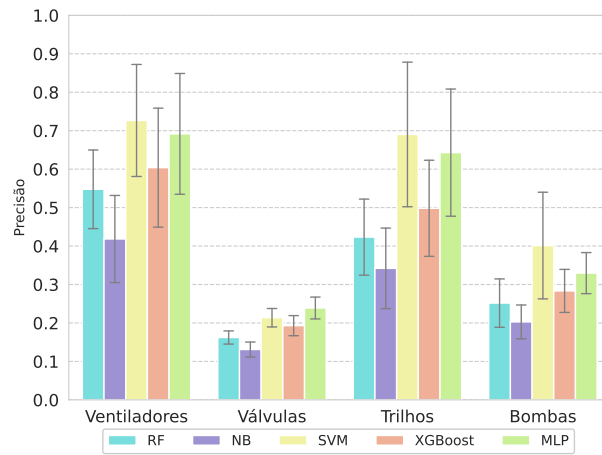
O valor 1 obtido em 3 dos 4 modelos para a mais alta SNR indica ausência de falsos negativos. Isto significa que em todas as situações onde o equipamento é determinado como em seu funcionamento normal, isto é verdade. O resultado



(a) Precisão para SNR = 6 dB.



(b) Precisão para SNR = 0 dB.



(c) Precisão para SNR = -6 dB.

Figura 5.2: Comparação da Precisão obtida pelos cinco modelos de aprendizado supervisionado para classificação binária.

Tabela 5.11: *Recall* para os modelos na classificação binária de amostras anômalas das bombas hidráulicas.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Bomba Hidráulica	00	0 dB	0.9298	0.8491	0.9263	0.8456	0.9186
		6 dB	0.9912	0.9175	1.0000	0.9667	0.9837
		- 6dB	0.8298	0.7737	0.8298	0.7596	0.8419
	02	0 dB	0.7250	0.8909	0.9341	0.9295	0.9706
		6 dB	1.0000	0.9955	0.8182	0.9864	1.0000
		- 6dB	0.6273	0.7614	0.9318	0.7909	0.9529
	04	0 dB	1.0000	1.0000	0.9950	0.9925	1.0000
		6 dB	1.0000	1.0000	1.0000	1.0000	1.0000
		- 6dB	0.9950	0.9825	0.9725	0.9550	1.0000
	06	0 dB	0.9000	0.9293	0.9488	0.9073	0.9645
		6 dB	0.9854	0.9829	0.9488	0.9707	0.9903
		- 6dB	0.8268	0.8220	0.8707	0.7463	0.9161

da *recall* para as bombas e trilhos deslizantes nas Tabelas 5.11 e 5.12 apresentam modelos alcançando o valor máximo possível para a métrica mais vezes ainda. Para bombas e trilhos é possível observar cenários onde todos os modelos conseguem zerar o número de falsos negativos. Por último, as válvulas novamente apresentam desempenho abaixo dos demais componentes. Entretanto, o valor da métrica não fica abaixo de 0.7 em nenhum dos modelos conforme apresentado na Tabela, 5.13.

Tabela 5.12: *Recall* para os modelos na classificação binária de amostras anômalas dos trilhos deslizantes.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Trilho Deslizante	00	0 dB	0.9986	1.0000	1.0000	0.9993	1.0000
		6 dB	1.0000	1.0000	1.0000	1.0000	1.0000
		- 6dB	0.9817	0.9923	0.9930	0.9923	0.9907
	02	0 dB	0.8421	0.7860	0.9701	0.8514	0.9813
		6 dB	0.8215	0.8047	0.9991	0.9617	1.0000
		- 6dB	0.7963	0.7271	0.8561	0.7458	0.9113
	04	0 dB	0.8592	0.8761	0.9704	0.9070	0.9870
		6 dB	0.9338	0.9577	0.9775	0.9451	0.9926
		- 6dB	0.7423	0.8028	0.9183	0.7296	0.9519
	06	0 dB	0.8611	0.8639	0.9611	0.8194	0.9885
		6 dB	0.9194	0.9722	0.9944	0.9139	1.0000
		- 6dB	0.7250	0.7583	0.8917	0.6972	0.9308

A média dos valores obtidos para os diferentes tipo de máquina é realizada e apresentada pela Figura 5.3. As barras próximas ao limite superior do gráfico indicam que na maior parte dos tipos de máquinas os modelos alcançam ou se aproximam do valor nulo de falsos negativos.

Tabela 5.13: *Recall* para os modelos na classificação binária de amostras anômalas das válvulas.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Válvula	00	0 dB	0.5875	0.5708	0.7333	0.5417	0.7722
		6 dB	0.7167	0.7708	0.8104	0.5938	0.8528
		- 6dB	0.6646	0.4979	0.6458	0.4875	0.7083
	02	0 dB	0.5396	0.6479	0.7375	0.4979	0.8083
		6 dB	0.7458	0.8188	0.9062	0.7750	0.9417
		- 6dB	0.6188	0.4979	0.6354	0.4208	0.6528
	04	0 dB	0.5292	0.7042	0.8604	0.5896	0.8222
		6 dB	0.7833	0.8604	0.875	0.7958	0.9528
		- 6dB	0.4646	0.6062	0.6958	0.4625	0.7250
	06	0 dB	0.6833	0.6312	0.7979	0.5667	0.8389
		6 dB	0.6104	0.7146	0.9021	0.6125	0.9028
		- 6dB	0.7021	0.5146	0.7250	0.3937	0.7528

A análise de ambas as métricas simultaneamente revela que em equipamentos como válvulas e bombas hidráulicas, o sistema tem altas taxas de falso positivo e baixas situações de falso negativo. Isso significa que os modelos tendem a considerar que uma amostra inédita de som é anômala. Apesar da redução da taxa de falsos positivos ser um dos objetivos, no contexto de detecção de falhas é menos preocupante o alarme falso que o não alarme em casos de componente com defeito.

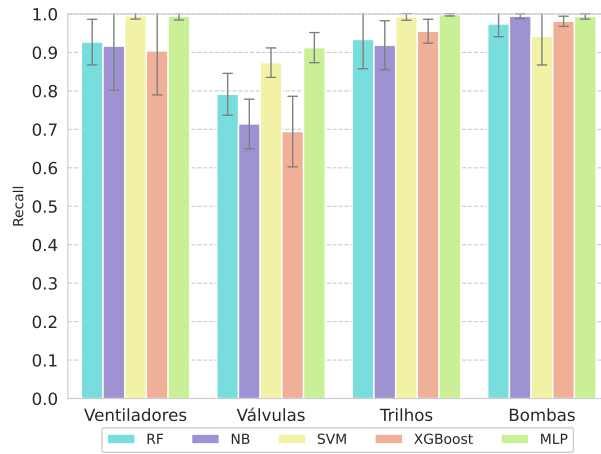
Na Figura 5.3, as barra de média de *recall* para a maioria dos equipamentos alcança valores próximos ao limite superior.

A válvula mantém os valores baixos para a precisão evidenciando que a classificação incorreta se mantém alta, tanto para FP quanto para FN. Assim como mencionado anteriormente, este problema pode ter duas origens: a pouca quantidade de amostras anômalas para o equipamento que torna a alta a diferença entre o número de amostras das classes e a possível semelhança entre o padrão do som para os comportamentos para os estados de falha e em operação normal.

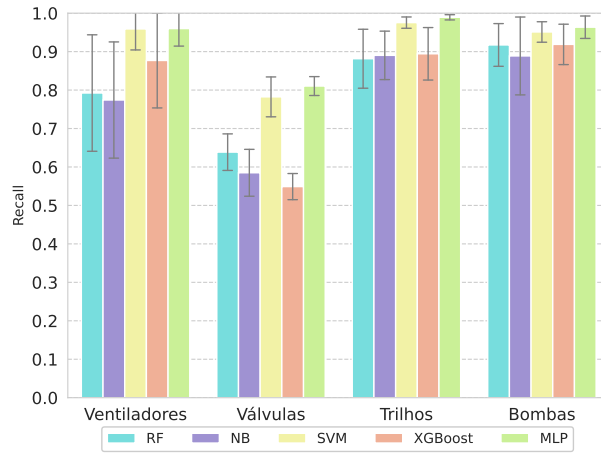
F1-score

A F1-score tem seu valor médio para cada modelo apresentado nas Tabelas 5.14, 5.15, 5.16 e 5.17 e é uma forma de obter o equilíbrio entre as métricas *recall* e precisão.

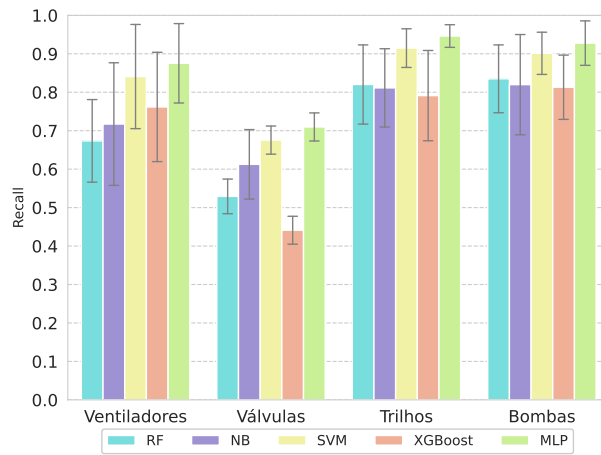
A métrica segue o comportamento da precisão, uma vez que os resultados de *recall* são próximos entre os modelos. O SVM obtém os resultados mais altos na maioria dos equipamentos, tendo novamente as válvulas como exceção. Para as válvulas, o melhor resultado é obtido através da MLP. Apesar dos resultados para as válvulas estar abaixo de 0.5 na maior parte dos modelos, os demais equipamentos



(a) *Recall* para SNR = 6 dB.



(b) *Recall* para SNR = 0 dB.



(c) *Recall* para SNR = -6 dB.

Figura 5.3: Comparação da *recall* obtida pelos cinco modelos de aprendizado supervisionado para classificação binária.

apresentam na maior parte dos casos um valor satisfatório para a métrica.

Na Figura 5.4 é possível observar a média obtida pela métrica observando separadamente as taxas de sinal-ruído. Os equipamentos mais comprometidos em todas as SNR foram os que apresentavam maior diferença entre a quantidade de amostras normais e anômalas, reforçando que este fator pode ser determinante no sucesso da detecção de eventos de interesse.

AUC

A última métrica de avaliação de desempenho é a AUC. Além dos cinco modelos adotados na proposta, as Tabelas 5.18, 5.19, 5.20 e 5.21 comparam os resultados com os de Purohit et al., apresentados na última coluna da tabela.

O modelo utilizado para comparação foi disponibilizado em um repositório no Github, possibilitando que também fosse treinado por 10 rodadas independentes e a média dos valores obtidos.

A Figura 5.5 apresenta a média dos resultados de AUC. É acrescentada uma barra cinza para simbolizar o trabalho de Purohit et al. às barras descritas nos gráficos que simbolizam a média das métricas anteriores. Os autores adotam redes neurais autoassociativas para detectar os defeitos, utilizando os Mel-espectrogramas como entrada da abordagem baseada em aprendizado não-supervisionado. Em todos os equipamentos e SNR os algoritmos SVM e MLP obtêm AUC mais alta que a proposta de Purohit et al., sendo MLP visualmente superior aos demais em todas as situações. Este resultado demonstra que as características do espectro são capazes de não apenas descrever eventos de falha, como melhorar os resultados de outras abordagens.

Tabela 5.14: F1-score para os modelos na classificação binária de amostras anômalas dos ventiladores.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Ventilador industrial	00	0 dB	0.4996	0.5471	0.7690	0.6147	0.7329
		6 dB	0.6510	0.7327	0.9487	0.8043	0.8928
		- 6dB	0.4259	0.4627	0.5763	0.4738	0.5780
	02	0 dB	0.5966	0.7736	0.9594	0.8655	0.9281
		6 dB	0.8659	0.8994	0.9897	0.8968	0.9723
		- 6dB	0.4942	0.5867	0.8244	0.6679	0.8006
	04	0 dB	0.7232	0.7630	0.9519	0.7962	0.8976
		6 dB	0.8768	0.6990	0.9975	0.8058	0.9522
		- 6dB	0.4331	0.6041	0.7533	0.6532	0.7461
	06	0 dB	0.7562	0.8365	0.9791	0.7513	0.9640
		6 dB	0.6347	0.9442	0.9972	0.7317	0.9913
		- 6dB	0.7270	0.7502	0.9501	0.8881	0.9458

Tabela 5.15: F1-score para os modelos na classificação binária de amostras anômalas das bombas hidráulicas.

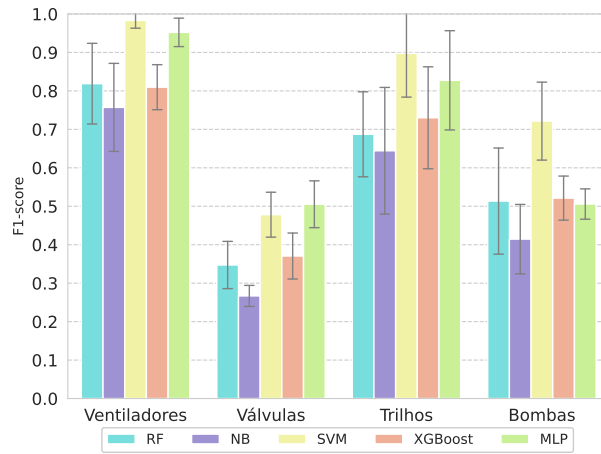
Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Bomba Hidráulica	00	0 dB	0.3963	0.4576	0.7016	0.5106	0.4618
		6 dB	0.4037	0.4972	0.7013	0.5294	0.4822
		- 6dB	0.3511	0.4471	0.6584	0.4698	0.5223
	02	0 dB	0.3016	0.3658	0.5391	0.4260	0.4761
		6 dB	0.3576	0.4542	0.6159	0.4869	0.5424
		- 6dB	0.2733	0.3046	0.3990	0.3557	0.4145
	04	0 dB	0.4658	0.5740	0.7350	0.5387	0.5274
		6 dB	0.5644	0.7376	0.8883	0.6102	0.5452
		- 6dB	0.4166	0.4760	0.6851	0.4986	0.5513
	06	0 dB	0.3114	0.3456	0.6411	0.4419	0.4833
		6 dB	0.3322	0.3648	0.6801	0.4582	0.4532
		- 6dB	0.2528	0.3028	0.4138	0.3438	0.4384

Tabela 5.16: F1-score para os modelos na classificação binária de amostras anômalas dos trilhos deslizantes.

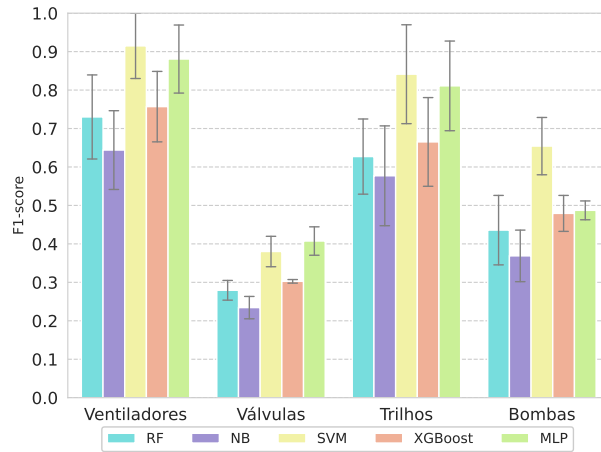
Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Trilho deslizante	00	0 dB	0.7249	0.7227	0.9727	0.8010	0.9302
		6 dB	0.8348	0.8301	0.9821	0.8765	0.9413
		- 6dB	0.6135	0.6657	0.9381	0.7667	0.8938
	02	0 dB	0.5580	0.6837	0.8437	0.6961	0.8359
		6 dB	0.5775	0.6723	0.9492	0.7832	0.8791
		- 6dB	0.4666	0.5819	0.7793	0.6402	0.7295
	04	0 dB	0.6485	0.6355	0.9154	0.6814	0.8594
		6 dB	0.7558	0.7230	0.9574	0.7445	0.8813
		- 6dB	0.5289	0.5757	0.8380	0.6014	0.8436
	06	0 dB	0.3768	0.4664	0.6330	0.4820	0.6181
		6 dB	0.4089	0.5232	0.7017	0.5161	0.6083
		- 6dB	0.2832	0.3811	0.5345	0.4116	0.5459

A Tabela 5.18 apresenta o resultado para os ventiladores industriais. MLP alcança o melhor resultado para todas as diferentes taxas de sinal-ruído, superando inclusive a proposta de Purohit et al. O comportamento se repete novamente para as bombas hidráulicas na Tabela 5.19, trilhos deslizantes na Tabela 5.20 e para as válvulas na Tabela 5.21.

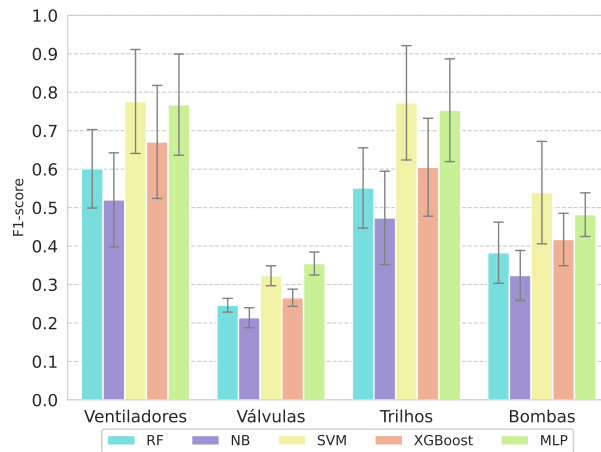
Com base nos resultados das métricas e no tempo necessário para treinamento e teste de cada modelo conforme apresentado na Tabela 5.1 o algoritmo SVM é escolhido para a realização da classificação multiclasse.



(a) F1-Score para SNR = 6 dB.



(b) F1-Score para SNR = 0 dB.



(c) F1-Score para SNR = -6 dB.

Figura 5.4: Comparação da F1-Score obtida pelos cinco modelos de aprendizado supervisionado para classificação binária.

Tabela 5.17: F1-score para os modelos na classificação binária de amostras anômalas das válvulas.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP
Válvula	00	0 dB	0.2061	0.2616	0.3263	0.2950	0.3442
		6 dB	0.2303	0.3015	0.4025	0.3151	0.4223
		- 6dB	0.2046	0.2413	0.2843	0.2595	0.3172
	02	0 dB	0.2829	0.3222	0.4332	0.3076	0.4376
		6 dB	0.2815	0.4530	0.5653	0.4663	0.5916
		- 6dB	0.2572	0.2746	0.3574	0.3025	0.3789
	04	0 dB	0.2214	0.2757	0.3970	0.3028	0.4245
		6 dB	0.3023	0.3203	0.4833	0.3752	0.5190
		- 6dB	0.1885	0.2431	0.3236	0.2577	0.3337
	06	0 dB	0.2267	0.2578	0.3641	0.3050	0.4234
		6 dB	0.2531	0.3141	0.4613	0.3257	0.4879
		- 6dB	0.2041	0.2254	0.3259	0.2424	0.3883

Tabela 5.18: AUCs para os modelos na classificação binária de amostras anômalas dos ventiladores industriais.

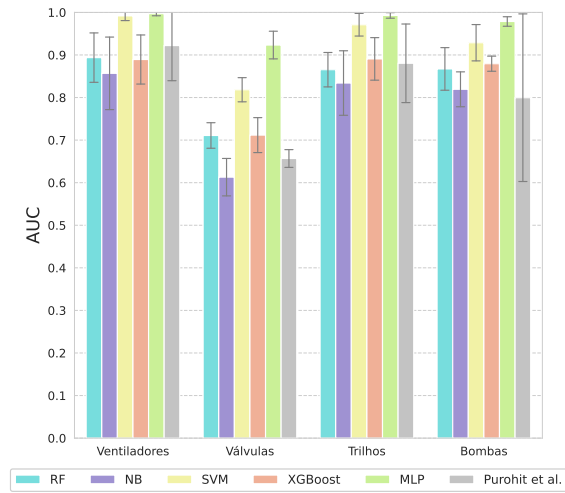
Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP	Purohit et al.
Ventilador Industrial	00	0 dB	0.6281	0.6819	0.8544	0.7314	0.9223	0.6236
		6 dB	0.7604	0.8267	0.9727	0.8859	0.9884	0.7875
		- 6 dB	0.5494	0.6149	0.7031	0.6218	0.7742	0.5534
	02	0 dB	0.7248	0.8601	0.9817	0.9280	0.9915	0.8539
		6 dB	0.9322	0.9434	0.9963	0.9563	0.9998	0.9809
		- 6 dB	0.6552	0.7266	0.8957	0.7937	0.9587	0.6850
	04	0 dB	0.8428	0.8725	0.9804	0.8993	0.9950	0.7560
		6 dB	0.9505	0.8461	0.9992	0.9157	0.9998	0.9224
		- 6 dB	0.5868	0.7484	0.868	0.7872	0.9509	0.5753
	06	0 dB	0.8782	0.9054	0.9891	0.8675	0.9953	0.9785
		6 dB	0.7841	0.9588	0.9990	0.7993	1.0000	0.9974
		- 6 dB	0.8588	0.8409	0.9701	0.9353	0.9894	0.8703

5.2 Classificação Multiclasse

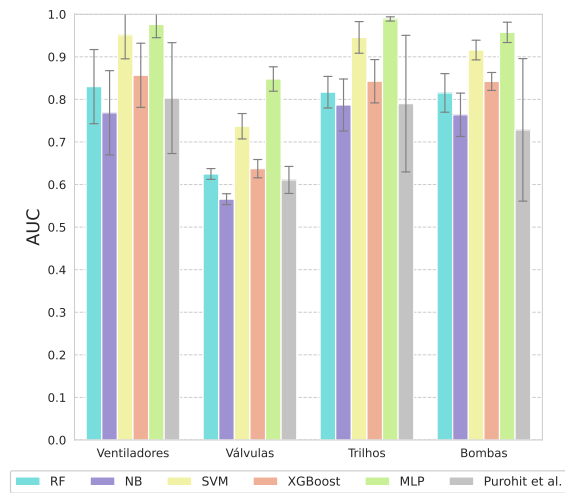
A classificação multiclasse consiste em identificar uma classe entre múltiplas classes possíveis. Um dos objetivos ao aplicar a classificação multiclasse ao invés da binária é identificar não apenas se um modelo de equipamento específico está em falha, mas também identificar qual equipamento não está operando conforme o padrão esperado.

Logo, um único modelo pode ser adotado para classificação de diversos equipamentos. Uma vez que a classificação binária para a maior parte dos equipamentos alcança altos valores na maioria das métricas utilizadas, espera-se que a classificação multiclasse também alcance êxito na detecção de falhas.

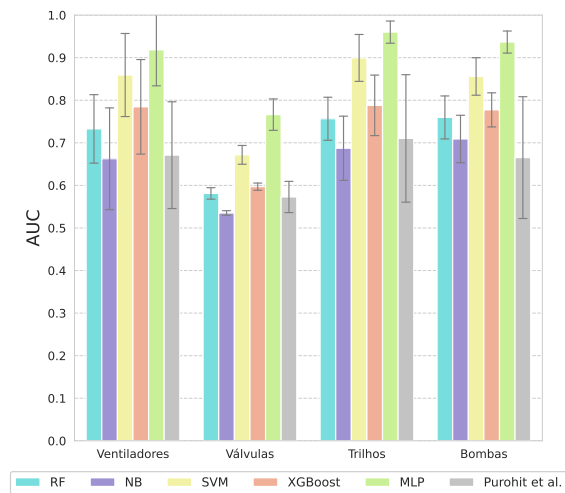
A Figura 5.6 descreve o número de amostras totais para cada um dos tipos de



(a) AUC para SNR = 6 dB.



(b) AUC para SNR = 0 dB.



(c) AUC para SNR = -6 dB.

Figura 5.5: Comparação da AUC obtida pelos cinco modelos de aprendizado supervisionado e o trabalho Purohit et al. e para classificação binária.

Tabela 5.19: AUCs para os modelos na classificação binária de amostras anômalas das bombas hidráulicas.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP	Purohit et al.
Bomba Hidráulica	00	0 dB	0.7689	0.7914	0.9103	0.8183	0.9180	0.6720
		6 dB	0.7879	0.8313	0.9386	0.8601	0.9688	0.8199
		- 6 dB	0.7065	0.7667	0.8649	0.7746	0.9050	0.5918
	02	0 dB	0.6928	0.7826	0.8828	0.8308	0.9632	0.4919
		6 dB	0.8029	0.8665	0.8627	0.8789	0.9803	0.4743
		- 6 dB	0.6511	0.7036	0.8155	0.7487	0.9390	0.5388
	04	0 dB	0.8363	0.8934	0.9466	0.8749	0.9822	0.9459
		6 dB	0.8897	0.9489	0.9817	0.9080	0.9962	0.9858
		- 6 dB	0.7993	0.8379	0.9242	0.8435	0.9763	0.9074
	06	0 dB	0.7571	0.7934	0.9239	0.8444	0.9662	0.8036
		6 dB	0.7968	0.8220	0.9319	0.8709	0.9691	0.9184
		- 6 dB	0.6790	0.7301	0.8187	0.7426	0.9263	0.6229

Tabela 5.20: AUCs para os modelos na classificação binária de amostras anômalas dos trilhos deslizante.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP	Purohit et al.
Trilho Deslizante	00	0 dB	0.8733	0.8721	0.9906	0.9169	0.9959	0.9846
		6 dB	0.9341	0.9319	0.9939	0.9530	0.9994	0.9915
		- 6 dB	0.7880	0.8313	0.9758	0.8969	0.9950	0.9369
	02	0 dB	0.7731	0.8285	0.9434	0.8510	0.9882	0.8263
		6 dB	0.7823	0.8280	0.9861	0.9178	0.9966	0.9223
		- 6 dB	0.6952	0.7661	0.8853	0.7997	0.9499	0.7283
	04	0 dB	0.7983	0.7897	0.9600	0.8271	0.9898	0.8107
		6 dB	0.8774	0.8628	0.9780	0.8723	0.9933	0.8682
		- 6 dB	0.6901	0.7362	0.9138	0.7489	0.9707	0.6513
	06	0 dB	0.7025	0.7773	0.8883	0.7754	0.9820	0.5384
		6 dB	0.7424	0.8389	0.9260	0.8198	0.9822	0.7396
		- 6 dB	0.5754	0.6923	0.8227	0.7063	0.9244	0.5246

Tabela 5.21: AUCs para os modelos na classificação binária de amostras anômalas das válvulas.

Tipo de Máquina	ID da Máquina	SNR	NB	RF	SVM	XGBoost	MLP	Purohit et al.
Válvula	00	0 dB	0.5439	0.6164	0.6989	0.6407	0.8100	0.5762
		6 dB	0.5838	0.6824	0.7699	0.6651	0.8685	0.6627
		- 6 dB	0.5395	0.5877	0.6467	0.6061	0.7419	0.6315
	02	0 dB	0.5769	0.6206	0.7258	0.6013	0.8326	0.5827
		6 dB	0.5719	0.7562	0.8422	0.7543	0.9416	0.6613
		- 6 dB	0.5364	0.5672	0.6532	0.5945	0.7277	0.5715
	04	0 dB	0.5701	0.6465	0.7812	0.6548	0.8643	0.6357
		6 dB	0.685	0.7196	0.8322	0.7508	0.9521	0.6230
		- 6 dB	0.5258	0.5999	0.6900	0.6030	0.7712	0.5324
	06	0 dB	0.5715	0.6156	0.7418	0.6524	0.8846	0.6485
		6 dB	0.6111	0.6853	0.8289	0.6767	0.9309	0.6801
		- 6 dB	0.5387	0.5695	0.6970	0.5843	0.8240	0.5557

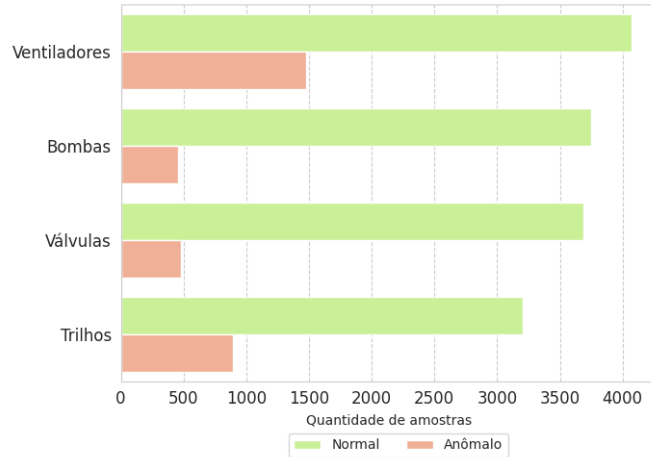


Figura 5.6: Número total de amostras por equipamento e estado.

máquinas após a união dos IDs. A imagem ressalta a natureza desbalanceada do conjunto de dados, uma vez que as barras vermelhas indicam o estado anormal do equipamento e as barras verdes o padrão esperado. Os ventiladores industriais possuem aproximadamente 4000 amostras normais contra pouco menos que 1500 amostras anômalas. As bombas hidráulicas e as válvulas ultrapassam as 3500 amostras normais e contam com menos de 500 amostras de mal-funcionamento. Por último, os trilhos deslizantes contêm mais de 3000 amostras de funcionamento normal e menos que 1000 amostras anômalas.

A classificação multiclasse pode ser realizada de duas maneiras distintas, um-vs-todos ou um-vs-um, como pode ser visto abaixo:

- Um-vs-todos: a estratégia consiste em ajustar um modelo por classe como uma espécie de classificação binária: ou a amostra pertence à classe ou não pertence. São necessários n classificadores em uma tarefa com n classes, sendo computacionalmente menos custosa que a estratégia um-vs-um. Desta forma, essa abordagem foi a escolhida para a etapa de classificação multiclasse.
- Um-vs-um: a estratégia consiste em treinar um classificador para cada par de rótulos, sendo selecionado aquele que obtiver o maior número de votos. O número de classificadores necessário é descrito na Equação 5.1, resultando em 28 classificadores necessários para a tarefa. O número é 3,5 vezes superior ao utilizado pela estratégia adotada.

$$N_{classificadores} = \frac{n_{classes} \cdot (n_{classes} - 1)}{2} \quad (5.1)$$

Matriz de Confusão Normalizada

Os resultados da classificação adotando SVM e a estratégia um-vs-todos podem ser analisados através das matrizes de confusão das Figuras 5.7 e 5.8. Os valores estão dentro do intervalo entre 0 e 1, onde 1 representa 100% das classificações. Quanto mais escuro o tom da cor, maior a quantidade de classificações pertencentes ao rótulo. Em um cenário ideal, espera-se que a diagonal principal apresente o tom mais escuro e os demais valores sejam iguais a zero, simbolizando 100% das classificações corretas. Para avaliar se a característica desbalanceada do conjunto de dados influencia nos resultados da classificação multiclasse, foram realizadas duas análises: com e sem *oversampling* das amostras.

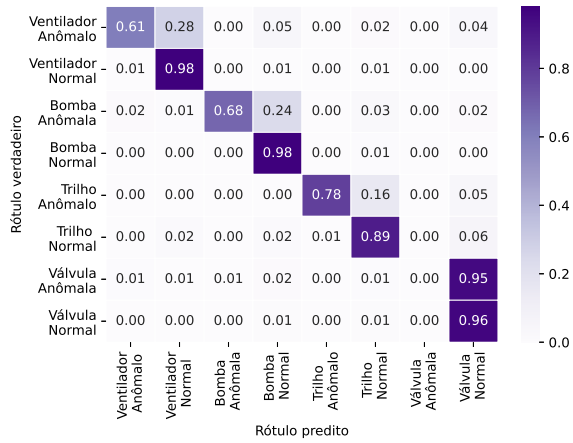
Sem método *oversampling*

Para a SNR igual a 6dB, na Figura 5.7a, o estado normal é identificado corretamente entre 89% e 98% das vezes, tendo o ventilador e a bomba apresentando os melhores resultados e o trilho deslizante apresentando o pior. A classificação das amostras anômalas do ventilador, bomba e trilho deslizante acontece de forma incorreta, sendo confundidas com as amostras normais do mesmo equipamento entre 16% e 28% das vezes. Entretanto, para a classificação de mal-funcionamento das válvulas, o algoritmo erra em todas as vezes e classifica 95% das amostras como pertencentes à classe normal. A mesma tendência de resultados é encontrada para as demais relações sinal-ruído nas Figuras 5.7b e 5.7c.

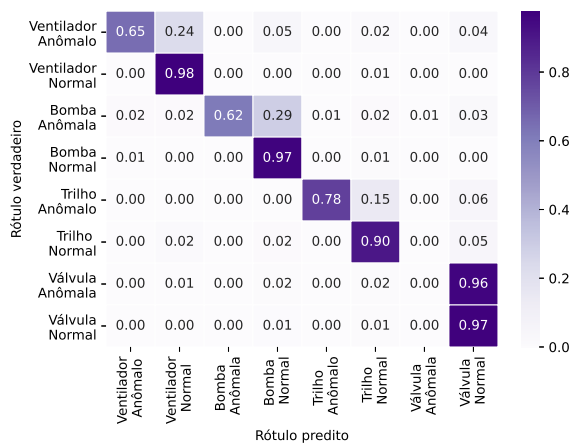
Com método *oversampling*

A Figura 5.8 apresenta a matriz de confusão normalizada pra o método de multi-classificação aplicando *oversampling*. É possível perceber que a classificação de sinais anômalos melhora significativamente quando comparada com os valores da classificação sem *oversampling*. Os estados anômalos das válvulas podem ser identificados ao realizar cópias das amostras de defeito. A classificação passa a ser realizada corretamente em mais de 60% dos casos, independente da relação sinal-ruído analisada.

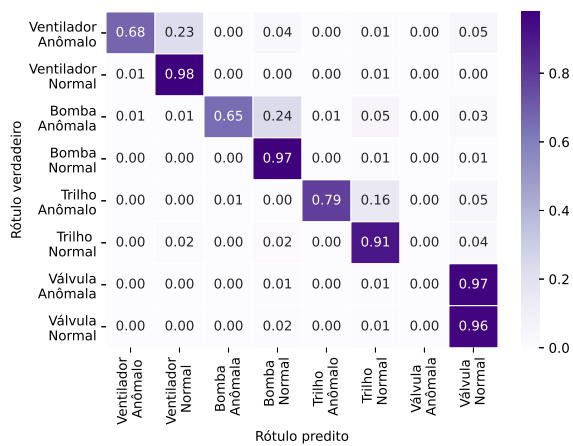
Os resultados demonstram que as características selecionadas podem ser adotadas para manutenção corretiva dos equipamentos industriais. O sucesso dos modelos está relacionado aos dados disponíveis para treinamento dos modelos, tendo para a maior parte dos equipamentos taxa de acerto superior à 85% nas diversas SNR analisadas. O único equipamento que possui comportamento diferente dos demais pode ter seu desempenho melhorado a partir da aquisição de novas amostras de defeitos.



(a) Matriz de Confusão para SNR = 6 dB.



(b) Matriz de Confusão para SNR = 0 dB.

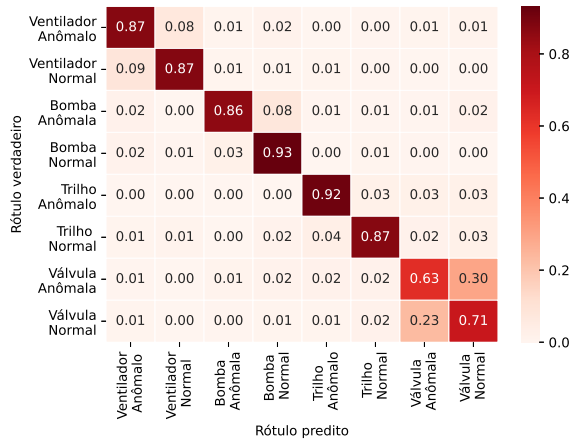


(c) Matriz de Confusão para SNR = -6 dB.

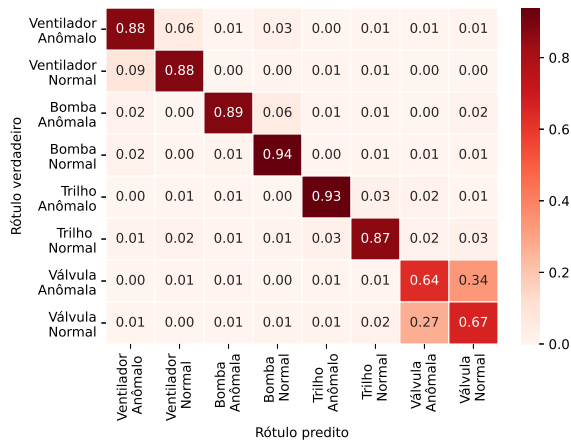
Figura 5.7: Matriz de Confusão para classificação Multiclasse para utilizando SVM sem *oversampling*.

5.2.1 Tempo de Conversão do Conjunto de Dados

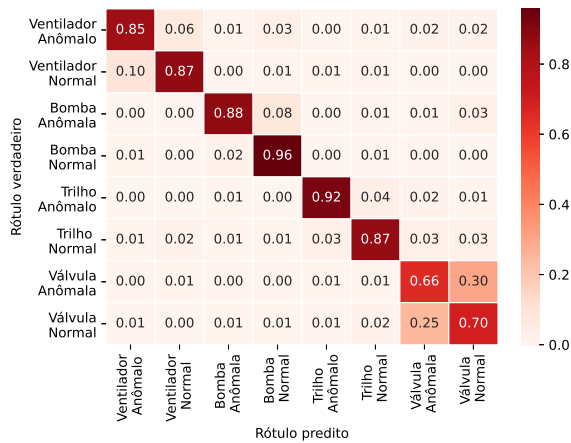
A melhora dos resultados com a adoção das características espectrais possuem como preço o aumento do tempo necessário para conversão do conjunto de dados.



(a) Matriz de Confusão para SNR = 6 dB.



(b) Matriz de Confusão para SNR = 0 dB.



(c) Matriz de Confusão para SNR = -6 dB.

Figura 5.8: Matriz de Confusão para classificação Multiclasse para utilizando SVM com *oversampling*.

A Figura 5.9 mostra o tempo em segundos considerando que as amostras convertidas em Mel-espectrogramas estão em formato de imagens png características espectrais são salvas em arquivos CSVs. As barras azul e laranja representam a mé-

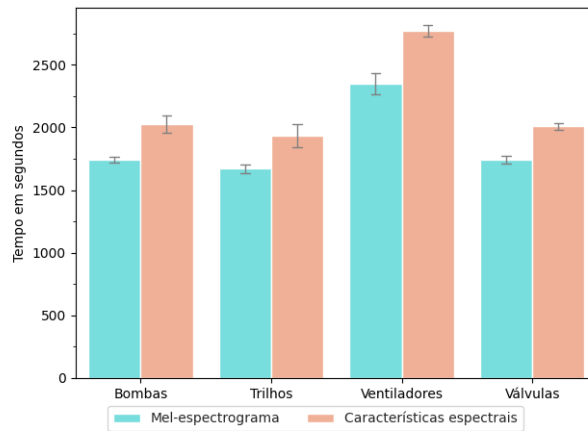


Figura 5.9: Tempo total de conversão das amostras de som do conjunto de dados em Mel-espectrogramas e para extração de características espectrais .

dia de tempo necessária para o conjunto de dados considerando as diferentes SNR. Os ventiladores possuem a maior diferença entre os tempos para conversão, sendo a proposta baseada em características espectrais aproximadamente 20% mais lenta.

Para manutenção corretiva, que não possui requisitos de tempo real e sim que as ações sejam realizadas o mais rápido possível, a diferença entre os tempos pode ser compensada adotando-se algoritmos mais simples. A vantagem é reduzir os requisitos computacionais necessários e tornar a classificação de defeitos possíveis para mais pontos de uma fábrica.

Capítulo 6

Conclusão e Trabalhos Futuros

Esta dissertação teve como proposta estudar o diagnóstico de mau funcionamento no maquinário industrial a partir da análise de amostras sonoras. Para classificação a partir de amostras de som, outras abordagens adotaram a conversão das amostras em Mel-espectrogramas. Após a conversão, as amostras foram tratadas como imagens e algoritmos comumente adotados para tarefas de classificação de imagens puderam ser utilizados, como por exemplo, os baseados em redes convolucionais.

Porém, o poder computacional necessário para este tipo de tarefa e a quantidade de dados disponíveis para treinamento dos modelos podem representar um obstáculo na classificação dos eventos.

Os algoritmos que foram escolhidos para classificação de defeitos utilizando as características do espectro ao invés dos Mel-espectrogramas podem ser mais simples quando comparados com redes utilizadas para classificação de imagens. Entretanto, a conversão das amostras sonoras para extração das características precisaram de mais tempo que a conversão do som em Mel-espectrogramas. Este fator não necessariamente é um problema, uma vez que a tarefa analisada não possui requisitos de tempo real.

Este trabalho utilizou o conjunto de dados MIMII que contém amostras do funcionamento esperado e do mau funcionamento de quatro tipos diferentes de máquinas. As amostras anômalas são as de maior interesse para o reconhecimento de falhas, mas também possuem quantidade reduzida se comparadas ao número de amostras normais. Esta característica adiciona um desafio no treinamento dos modelos. Para a classificação binária o método de *oversampling* foi aplicado nas amostras de mau funcionamento. O método selecionou aleatoriamente amostras anômalas e as repetiu até que sua quantidade se igualasse ao número de amostras normais.

Para teste da abordagem sugerida foram adotados cinco classificadores baseados em aprendizado supervisionado para a classificação binária dos estados. Os resultados dos modelos foram comparados entre si, uma vez que trabalhos relacionados demonstraram que um modelo pode ser adequado para determinada tarefa e menos

vantajoso em outras. A avaliação é realizada a partir das métricas acurácia, precisão, *recall*, F1-score e AUC. A AUC possibilitou a comparação com o trabalho de Purohit et al. baseado em redes autoassociativas e Mel-espectrogramas. O trabalho adotado nesta comparação foi proposto pelos autores do conjunto de dados. O SVM e a MLP foram os classificadores com melhores desempenhos, tendo a MLP superado o trabalho de Purohit et al. em todos os equipamentos testados. Entretanto, o tempo necessário para treinamento da rede neural é múltiplas vezes maior que o do SVM. Sendo assim, este último modelo foi o selecionado para a segunda etapa experimental: a classificação multiclasse. O modelo com menor tempo de treinamento foi selecionado pensando em possíveis aplicações futuras baseadas em treinamento online.

Na classificação multiclasse foi possível com um único modelo detectar as amostras anômalas e seus respectivos tipos de máquinas correspondentes. Para analisar o impacto do método *oversampling* para contornar o desbalanceamento do conjunto de dados foram realizados experimentos de multiclasse com e sem a utilização do método. Sem igualar o número das amostras, a classificação da anomalia é realizada corretamente em pelo menos 60% das vezes para os equipamentos ventilador, bomba hidráulica e trilho deslizante em todas SNR. Entretanto, o classificador falhou na classificação do estado anômalo da válvula hidráulica em todas as vezes, assumindo que tratava-se da válvula no estado normal. Por outro lado, o estado normal dos equipamentos foi classificado corretamente pelo menos 89% das vezes em todos os equipamentos, alcançando 98% de classificações corretas para os ventiladores em todas as SNRs. Um dos fatores desta discrepância é o número de amostras disponível para treinamento dos modelos devido à natureza desbalanceada dos dados sendo comprovado ao aplicar o método de *oversampling*.

As classificações corretas de anomalias foram aumentadas ao equilibrar o número de amostras entre as classes. Apesar do aumento de falsos positivos para a válvulas, apenas com a aplicação de *oversampling* o comportamento fora do padrão pode ser identificado. Dessa forma, os experimentos comprovaram que é possível utilizar as características do espectro de som para reconhecimento de padrões e automação do reconhecimento de falhas.

Entre os futuros trabalhos pretende-se avaliar a utilização de Mel-espectrogramas através das métricas que não foram originalmente adotadas por Purohit et al. e, se em equipamentos com comportamento semelhante ao apresentado pelas válvulas a combinação das características baseadas nos espectros sonoros pode ser combinada aos Mel-espectrogramas para aumentar os verdadeiros positivos e reduzir os alarmes falsos. Além disso, pretende-se avaliar o sistema implementado para detecção de defeitos online e explorar uma nova versão do conjunto de dados disponibilizada [26]. Outra direção futura é a combinação de modelos de predição de falhas aos modelos de

classificação de defeitos. Dessa maneira, será possível encontrar o ponto ótimo entre manutenções corretivas e preditivas para os componentes industriais. Além disso, a proximidade do padrão sonoro entre comportamentos normais e anômalos pode ser melhor investigada e quantificada para que melhores soluções sejam exploradas. Por fim, outras métricas podem ser adotadas para melhor avaliação de resultados como a AUC-PR.

Referências Bibliográficas

- [1] ZHONG, R. Y., XU, X., KLOTZ, E., et al. “Intelligent manufacturing in the context of industry 4.0: a review”, *Engineering*, v. 3, n. 5, pp. 616–630, 2017.
- [2] ACETO, G., PERSICO, V., PESCAPE, A. “A survey on information and communication technologies for Industry 4.0: state-of-the-art, taxonomies, perspectives, and challenges”, *IEEE Communications Surveys & Tutorials*, v. 21, n. 4, pp. 3467–3501, 2019.
- [3] RADZIWON, A., BILBERG, A., BOGERS, M., et al. “The smart factory: exploring adaptive and flexible manufacturing solutions”, *Procedia engineering*, v. 69, pp. 1184–1190, 2014.
- [4] SHROUF, F., ORDIERES, J., MIRAGLIOTTA, G. “Smart factories in Industry 4.0: A review of the concept and of energy management approached in production based on the Internet of Things paradigm”. In: *2014 IEEE international conference on industrial engineering and engineering management*, pp. 697–701. IEEE, 2014.
- [5] DIEZ-OLIVAN, A., DEL SER, J., GALAR, D., et al. “Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0”, *Information Fusion*, v. 50, n. 1, pp. 92–111, 2019.
- [6] BOCHIE, K., DA SILVA GILBERT, M., GANTERT, L., et al. “Aprendizado Profundo em Redes Desafiadoras: Conceitos e Aplicações”, *Sociedade Brasileira de Computação*, 2020.
- [7] ERKOYUNCU, J. A., KHAN, S., EIROA, A. L., et al. “Perspectives on trading cost and availability for corrective maintenance at the equipment type level”, *Reliability Engineering & System Safety*, v. 168, n. 1, pp. 53–69, 2017.
- [8] PUROHIT, H., TANABE, R., ICHIGE, T., et al. “MIMII Dataset: Sound Dataset for Malfunctioning Industrial Machine Investigation and Inspection”.

In: *4th Workshop on Detection and Classification of Acoustic Scenes and Events (DCASE 2019)*, pp. 209–213, out. 2019.

- [9] GANTERT, L., SAMMARCO, M., DETYNIECKI, M., et al. “A Supervised Approach for Corrective Maintenance Using Spectral Features from Industrial Sounds”. In: *2021 IEEE World Forum on Internet of Things (WF-IoT)*. IEEE, 2021.
- [10] ZHANG, D., QIAN, L., MAO, B., et al. “A data-driven design for fault detection of wind turbines using random forests and XGboost”, *IEEE Access*, v. 6, pp. 21020–21031, 2018.
- [11] LI, L., OTA, K., DONG, M. “Deep learning for smart industry: Efficient manufacture inspection system with fog computing”, *IEEE Transactions on Industrial Informatics*, v. 14, n. 10, pp. 4665–4673, 2018.
- [12] CHENG, Y., ZHU, H., WU, J., et al. “Machine health monitoring using adaptive kernel spectral clustering and deep long short-term memory recurrent neural networks”, *IEEE Transactions on Industrial Informatics*, v. 15, n. 2, pp. 987–997, 2018.
- [13] ZHANG, C., ZHANG, Y., HU, C., et al. “A Novel Intelligent Fault Diagnosis Method Based on Variational Mode Decomposition and Ensemble Deep Belief Network”, *IEEE Access*, v. 8, n. 1, pp. 36293–36312, 2020.
- [14] LIU, R., YANG, B., ZIO, E., et al. “Artificial intelligence for fault diagnosis of rotating machinery: A review”, *Mechanical Systems and Signal Processing*, v. 108, pp. 33–47, 2018.
- [15] SAMMARCO, M., DETYNIECKI, M. “Crashzam: Sound-based Car Crash Detection”. In: Helfert, M., Gusikhin, O. (Eds.), *Proceedings of the 4th International Conference on Vehicle Technology and Intelligent Transport Systems, VEHITS 2018, Funchal, Madeira, Portugal, March 16-18, 2018*, pp. 27–35. SciTePress, 2018. doi: 10.5220/0006629200270035. Disponível em: <<https://doi.org/10.5220/0006629200270035>>.
- [16] GIRI, R., TENNETI, S. V., CHENG, F., et al. *Unsupervised Anomalous Sound Detection Using Self-supervised Classification and Group Masked Autoencoder for Density Estimation*. Relatório técnico, Challenge on Detection and Classification of Acoustic Scenes and Events (DCASE 2020 Challenge), 2020.

- [17] DANILUK, P., GOŹDZIEWSKI, M., KAPKA, S., et al. *Ensemble of Auto-encoder based and Wavenet like Systems for Unsupervised Anomaly Detection*. Relatório técnico, Challenge on Detection and Classification of Acoustic Scenes and Events (DCASE 2020 Challenge), 2020.
- [18] PRIMUS, P. *Reframing Unsupervised Machine Condition Monitoring as a Supervised Classification Task with Outlier-Exposed Classifiers*. Relatório técnico, Challenge on Detection and Classification of Acoustic Scenes and Events (DCASE 2020 Challenge), 2020.
- [19] ZHOU, Z.-H. *Ensemble methods: foundations and algorithms*. EUA, CRC press, 2012.
- [20] PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., et al. “Scikit-learn: Machine learning in Python”, *the Journal of machine Learning research*, v. 12, pp. 2825–2830, 2011.
- [21] CHEN, T., GUESTRIN, C. “Xgboost: A scalable tree boosting system”. In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794, 2016.
- [22] HAYKIN, S. *Redes neurais: princípios e prática*. Porto Alegre, Bookman Editora, 2001.
- [23] ZHENG, F., ZHANG, G., SONG, Z. “Comparison of different implementations of MFCC”, *Journal of Computer Science and Technology*, v. 16, n. 6, pp. 582–589, 2001.
- [24] MCFEE, B., RAFFEL, C., LIANG, D., et al. “librosa: Audio and music signal analysis in python”. In: *Proceedings of the 14th python in science conference*, v. 8, pp. 18–25. Citeseer, 2015.
- [25] PUROHIT, H., TANABE, R., ICHIGE, K., et al. “MIMII Dataset: Sound dataset for malfunctioning industrial machine investigation and inspection”, *arXiv preprint arXiv:1909.09347*, 2019.
- [26] TANABE, R., PUROHIT, H., DOHI, K., et al. “MIMII DUE: Sound Dataset for Malfunctioning Industrial Machine Investigation and Inspection with Domain Shifts due to Changes in Operational and Environmental Conditions”, *arXiv preprint arXiv:2105.02702*, 2021.