



INSTITUTO TECNOLÓGICO VALE



**Programa de Pós-Graduação em Instrumentação, Controle e
Automação de Processos de Mineração (PROFICAM)
Escola de Minas, Universidade Federal de Ouro Preto (UFOP)
Associação Instituto Tecnológico Vale (ITV)**

Dissertação

**DETECTOR DE FALHAS DE ROLOS DE TRANSPORTADORES DE CORREIA
BASEADO EM APRENDIZADO DE MÁQUINA**

Emanuel Balduino da Cruz

**Ouro Preto
Minas Gerais, Brasil
2020**

Emanuel Balduino da Cruz

**DETECTOR DE FALHAS DE ROLOS DE TRANSPORTADORES DE CORREIA
BASEADO EM APRENDIZADO DE MÁQUINA**

Dissertação apresentada ao Programa de Pós-Graduação em Instrumentação, Controle e Automação de Processos de Mineração da Universidade Federal de Ouro Preto e do Instituto Tecnológico Vale, como parte dos requisitos para obtenção do título de Mestre em Engenharia de Controle e Automação.

Orientador: Prof. Agnaldo José da Rocha Reis, D.Sc.

Coorientador: Prof. Frederico Gadelha Guimarães, D.Sc.

Coorientador: Profa. Silvia Grasiella Moreira Almeida, D.Sc.

Ouro Preto
2020

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

C957d Cruz, Emanuel Balduino da .
Detector de falhas de rolos de transportadores de correia baseado em
aprendizado de máquina. [manuscrito] / Emanuel Balduino da Cruz. -
2020.

53 f.: il.: color., gráf., tab..

Orientador: Prof. Dr. Agnaldo José da Rocha Reis.

Coorientadores: Prof. Dr. Frederico Gadelha Guimarães, Profa. Dra.
Silvia Grasiella Moreira Almeida.

Dissertação (Mestrado Profissional). Universidade Federal de Ouro
Preto. Programa de Mestrado Profissional em Instrumentação, Controle e
Automação de Processos de Mineração. Programa de Pós-Graduação em
Instrumentação, Controle e Automação de Processos de Mineração.

Área de Concentração: Engenharia de Controle e Automação de
Processos Mineraiis.

1. Aprendizado do computador . 2. Correias transportadoras - Rolos.
3. Correias transportadoras. I. Almeida, Silvia Grasiella Moreira. II.
Guimarães, Frederico Gadelha. III. Reis, Agnaldo José da Rocha. IV.
Universidade Federal de Ouro Preto. V. Título. CDU 681.5:622.2

Bibliotecário(a) Responsável: Maristela Sanches Lima Mesquita - CRB-1716



FOLHA DE APROVAÇÃO

Emanuel Balduino da Cruz

Detector de Falhas de Rolos de Transportadores de Correia Baseado em Aprendizado de Máquina

Membros da Banca

Agnaldo José da Rocha Reis – Doutor - Universidade Federal de Ouro Preto
Frederico Gadelha Guimarães – Doutor – Universidade Federal de Minas Gerais
Sílvia Grasiella Moreira Almeida - Doutora – Instituto Federal de Minas Gerais (Ouro Preto)
Rodrigo Cesar Pedrosa Silva – Ph.D. – Universidade Federal de Ouro Preto
Gustavo Pessin – Doutor – Instituto Tecnológico Vale
Paulo Raimundo Pinto – Doutor - Instituto Federal de Minas Gerais (Ouro Preto)

Versão final

Aprovada em 11 de Setembro de 2020

De acordo,

Agnaldo José da Rocha Reis.



Documento assinado eletronicamente por **Agnaldo Jose da Rocha Reis**, **COORDENADOR(A) DO CURSO DE PÓS-GRADUAÇÃO EM INSTRUMENTAÇÃO, CONTROLE E AUTOMAÇÃO DE PROC DE MINERAÇÃO**, em 11/12/2020, às 01:20, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0113884** e o código CRC **FA5010A5**.

*Quando as raízes são profundas
não há razão para temer o
vento.(Provérbio Chinês).*

Agradecimentos

Gostaria de primeiramente agradecer às minhas famílias Balduino e Cruz por sempre estarem do meu lado e apoiado em diferentes etapas da minha vida. Agradeço aos meus pais e irmãos pelo apoio incondicional e pelo suporte, nunca serei grato o suficiente. Sem esquecer das minhas sobrinhas, vocês sempre me dão força para seguir em frente.

Agradeço à UFOP e ao ITV, incluindo professores e demais funcionários, pela estrutura e oportunidade de fazer parte e contribuir de alguma forma na busca de conhecimentos neste porte que o curso de mestrado profissional proporcionou. Agradeço ao Felipe Rocha e Gustavo Pessin por terem acreditado em mim e deixado contribuir com o projeto do ROSI. Em especial, agradeço aos professores Frederico Gadelha e Silvia Almeida por toda ajuda e conhecimento compartilhado neste processo e ao professor Agnaldo Reis por toda orientação, suporte e paciência.

Agradeço à todos colegas de turma e amigos, em especial aos da República Badalação, vocês tornaram esta jornada muito mais proveitosa. Sou muito agradecido pela oportunidade de ter estudado com Julio, você sempre será lembrado como membro da terceira turma. Por fim, agradeço e dedico a conclusão desta etapa ao meu pai. Não foi fácil encerrar este projeto depois de você ter nos deixado, mas sei que você não gostaria que eu desistisse e espero ter te deixado orgulhoso.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES), Código de Financiamento 001; do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG); e da Vale SA.

Resumo

Resumo da Dissertação apresentada ao Programa de Pós Graduação em Instrumentação, Controle e Automação de Processos de Mineração como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

DETECTOR DE FALHAS DE ROLOS DE TRANSPORTADORES DE CORREIA BASEADO EM APRENDIZADO DE MÁQUINA

Emanuel Balduino da Cruz

Setembro/2020

Orientadores: Agnaldo José da Rocha Reis
Frederico Gadelha Guimarães
Silvia Grasiella Moreira Almeida

O Transportador de Correia (TC) é o meio mais versátil e difundido para movimentação de cargas em instalação industrial. Os roletes são equipamentos que dão suporte e guiam as correias. Como o rolo fica no apoio da correia, quando ele apresenta falha, o funcionamento do transportador é prejudicado. A detecção de falhas nesses equipamentos é feita por funcionários da manutenção que se baseiam, dentre outros fatores, nos ruídos produzidos por rolos defeituosos. Como uma correia pode ter centenas de metros e estar em locais de difícil acesso, determinar qual rolo está defeituoso se torna uma tarefa bastante difícil e custosa. Apresenta-se neste trabalho uma proposta para detecção de rolos defeituosos através de sons por ele produzido, por meio de aprendizado de máquina. Para tanto, dados reais de rolos bons e defeituosos foram coletados em campo e utilizados no desenvolvimento do sistema supracitado. Os resultados obtidos são discutidos em detalhes e as principais conclusões são apresentadas.

Palavras-chave: Transportador de correia, Rolo, Aprendizado de Máquina.

Macrotema: Mina; **Linha de Pesquisa:** Robótica aplicada à mineração; **Tema:** Melhoria de Procedimentos de Manutenção; **Área Relacionada da Vale:** Manutenção e inspeção.

Abstract

Abstract of Dissertation presented to the Graduate Program on Instrumentation, Control and Automation of Mining Process as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

BELT CONVEYOR ROLLER FAILURE DETECTION SYSTEM BASED ON MACHINE LEARNING

Emanuel Balduino da Cruz

September/2020

Advisors: Agnaldo José da Rocha Reis
Frederico Gadelha Guimarães
Silvia Grasiella Moreira Almeida

The Conveyor Belt (CB) is the most versatile and widespread means of transporting loads in industrial installations. In this apparatus, the rollers are equipment which provide support and guide the belt. As the roller sits on the belt support, when it fails, the operation of the conveyor belt is impaired. Failure detection on these equipment is done by maintenance personnel who rely on the noise produced by defective rollers, amongst other factors, to make a diagnosis. Moreover, given that a belt can be hundreds of meters long and stretching across places of difficult access, determining which roll is defective becomes a rather difficult and costly task. In this paper, a proposal for detection of defective rolls through sounds by means of machine learning is presented. For this purpose, real data of good and defective rolls were collected in the field and used in the development of the above-mentioned system. The obtained results are fully discussed and the main conclusions are presented.

Keywords: Conveyor belt, Roll, Machine Learning.

Macrotheme: Mine; **Research Line:** Robotics applied to mining; **Theme:** Improvement of Maintenance Procedures; **Related Area of Vale:** Maintenance and inspection.

Lista de Figuras

Figura 2.1	Principais componentes de um TC.	15
Figura 2.2	Nomenclatura de alguns dos elementos do TC.	17
Figura 2.3	Partes principais de um rolo.	18
Figura 2.4	Sinal digital vs sinal analógico.	20
Figura 2.5	Dois hiperplanos separadores do espaço de atributos.	23
Figura 2.6	Coleta de dados de um TC usando drone.	25
Figura 2.7	Robô de inspeção infravermelho.	26
Figura 2.8	Instalação do cabo de fibra-óptica.	27
Figura 2.9	Comunicação direta entre os rolos.	27
Figura 2.10	Correia e sensor de vibração instalado.	28
Figura 2.11	Gravação de ultrassom de um rolo em um transportador.	28
Figura 3.1	Equipamentos e programa para coleta de áudios.	33
Figura 3.2	Porto de Tubarão, Vitória - Es	33
Figura 3.3	Gravação dos áudios em campo.	34
Figura 3.4	Gráfico do som de um rolo bom no domínio da frequência.	35
Figura 3.5	Gráfico do som de um rolo defeituoso no domínio da frequência.	35
Figura 3.6	Espectrograma de um rolo bom.	36
Figura 3.7	Espectrograma de um rolo ruim.	36
Figura 4.1	Matriz confusão.	41
Figura 4.2	Matrizes de confusão dos algoritmos a serem comparados.	42
Figura 4.3	Gráfico ROC básico.	44
Figura 4.4	Gráfico ROC para os algoritmos testados.	45

Lista de Tabelas

Tabela 3.1	Base de dados dos áudios dos rolos.	38
Tabela 4.1	Taxa de acertos de cada kernel do algoritmo SVM para o conjunto de dados de áudios de rolos.	40
Tabela 4.2	Taxa de acertos de cada configuração do algoritmo RF para o conjunto de dados de áudios de rolos.	40
Tabela 4.3	Média da taxa de acertos dos algoritmos escolhidos.	41
Tabela 4.4	Média das precisões de cada algoritmo.	43
Tabela 4.5	Média da taxa de acertos dos algoritmos para o novo conjunto de dados. . .	45
Tabela 4.6	Média da taxa de acertos de cada base de dados de amostras individuais dos rolos.	46

Lista de Siglas e Abreviaturas

ABT *Analytical Base Table*

ACF *Canais de Características Agregadas*

Auto-ML *Automatic Machine Learning*

AUC *Area Under Curve*

CQT *Constant-Q-Transform*

CWT *Continuous Wavelets Transform*

DNN *Deep Neural Network*

DTS *Distributed Temperature Sensing*

GMM *Gaussian Mixture Models*

GRNN *Gated Recurrent Neural Network*

HPSS *Harmonic Percussive Source Separation*

IoT *Internet of Things*

KNN *K-Nearest Neighbor*

LSTM *Long Short-Term Memory*

MFCC *Mel-Frequency Cepstral Coefficient*

MLP *Multilayer Perceptron*

RF *Random Forest*

RFID *Radio Frequency Identification*

RNN *Recurrent Neural Network*

ROC *Receiver Operating Characteristic*

SIF *Spectrogram Image File*

STFT *Short-Time Fourier Transform*

SVM *Support Vector Machine*

TC *Transportador de Correias*

TPOT *Tree-based Pipeline Optimization Toll*

WPD *Wavelet Packet Decomposition*

Sumário

1	Introdução	13
1.1	Motivação	13
1.2	Objetivos	14
1.2.1	Objetivos Específicos	14
1.3	Estrutura do trabalho	14
2	Fundamentação científica e referencial teórico	15
2.1	Fundamentação científica	15
2.1.1	Transportador de correia e rolos	15
2.1.2	Sons e gravação digital	20
2.1.3	Aprendizado de máquina	21
2.2	Referencial Teórico	25
2.2.1	Análises de rolos de Transportadores de Correia	25
2.2.2	Aprendizado de máquina em classificação e identificação de sons ambientes	29
3	Metodologia	32
3.1	Materiais e métodos	32
3.2	Coleta de dados	33
3.3	Processamento de dados	34
3.4	Algoritmos de aprendizado de máquina para a base de dados de áudios de rolos	37
4	Resultados e Discussão	39
4.1	Definições iniciais	39
4.2	Resultados da precisão dos testes de cada algoritmo de aprendizado	39
4.3	Matriz de confusão e precisão	41
4.4	ROC e AUC	43
4.5	Teste dos algoritmos para amostras restantes dos rolos	45
4.6	Classificação para base de dados com apenas uma amostra de cada rolo	46
5	Conclusão e trabalhos futuros	47

Referências Bibliográficas	49
Apêndices	52

1. Introdução

1.1. Motivação

O Transportador de Correias (TC) se tornou o equipamento mais usado no transporte de materiais devido a sua confiabilidade, versatilidade e habilidade de conseguir transportar uma larga variedade de materiais. O material é carregado e descarregado enquanto a correia está em movimento, não havendo perda de tempo em todo processo. De todo sistema de manipulação de materiais, o TC normalmente opera com o menor custo de transporte, manutenção, energia e mão de obra por tonelada e maior capacidade (LUCAS *et al.*, 2007).

O uso de TC é extenso em processos da cadeia de mineração que envolvem mina, usina de beneficiamento, pelletização e porto. A Vale, uma das maiores empresas de mineração do mundo, conta com mais de 2000 TCs, totalizando 1000 km de extensão de correias (CARVALHO JÚNIOR, 2018). Apesar de todas vantagens em usar TCs, devido a sua grande extensão e componentes situados ao longo de todo o seu comprimento, conseguir monitorar e atuar quando esses apresentam falhas é uma tarefa muito custosa, sendo que um dos componentes que necessita de maior atenção é o rolo (NASCIMENTO *et al.*, 2017).

Uma vez que a esteira corre sobre os rolos é relativamente simples monitorá-la no início ou no fim do TC. Inspeccionar os rolos é uma situação diferente. O procedimento atual necessita que um funcionário tenha que andar ao longo do TC para inspecionar/monitorar a situação dos rolos. O problema é que um TC pode chegar a ter de 10000 a 100000 rolos (LODEWIJKS *et al.*, 2016). Com isso, a probabilidade de ocorrer um erro de manutenção é considerável e a ocorrência do evento de quebra pode causar danos extremos como incêndios que podem danificar outros rolos, correia, sistema elétrico e de automação, sem contar a perda do tempo de operação no momento em que ficou parado (CARVALHO JÚNIOR, 2018).

Quando o funcionário inspeciona os rolos, além de verificar o seu funcionamento, ele também fica atento aos sons por ele produzido. O som de operação do TC é diferente quando os rolos estão em perfeita condição ou com falha, e essa diferença varia de falha para falha. Os distintos sons emitidos possuem informações sobre as características das falhas, o que pode ser usado para sua detecção (CARVALHO JÚNIOR, 2018). Com essas informações, foi analisado a possibilidade de se usar inteligência artificial para diferenciar um rolo em perfeito funcionamento de um defeituoso através de seus sons.

A classificação por áudio tem como objetivo identificar e classificar um áudio descritivo com base em características semelhantes de um grupo de sons. Geralmente, o áudio pode ser dividido em 3 subdomínios principais: classificação de som ambiente, musical e fala. Classificação do som ambiente é muito importante para máquinas entenderem o ambiente que as cerca, mas isto tem sido um desafio e tem atraído muito interesse recentemente (ZHU *et al.*, 2018). Habilitar uma máquina a entender o ambiente através dos sons é o principal objetivo de pesquisa em *Machine Listening*. O sistema de *Machine Listening* atua processando tarefas similarmente

a um sistema auditivo humano, e são partes de um vasto tema de pesquisa que liga áreas como aprendizado de máquina, robótica e inteligência artificial (BARCHIESI *et al.*, 2015).

É nesse contexto que se insere o trabalho aqui proposto. Aqui, se objetiva fazer a análise do ruído sonoro dos rolos e classificá-los, a priori, entre rolos em bom estado e rolos defeituosos, utilizando aprendizado de máquina.

1.2. Objetivos

O objetivo geral é desenvolver um detector baseado em aprendizado de máquina que informe ao operador o estado de um rolo (estado bom ou com avaria).

1.2.1. Objetivos Específicos

- Fazer a coleta de sons de rolos em uma mineração com equipamentos e métodos apropriados.
- Preparar uma base de dados com os áudios coletados.
- Comparar diferentes tipos de algoritmos de aprendizado de máquina e definir qual o mais eficiente para este problema.

1.3. Estrutura do trabalho

Os próximos capítulos deste trabalho estão organizados da seguinte forma: No capítulo 2 - Fundamentação científica e referencial teórico, são apresentados trabalhos utilizados como referência desta dissertação. No capítulo 3 – Metodologia, são abordados os equipamentos, programas e métodos utilizados neste trabalho. No capítulo 3 - Resultado e discussão, os resultados são analisados e comparados. No capítulo 4 - Conclusão e trabalhos futuros, é feita uma conclusão geral das atividades e resultados deste trabalho e também é proposto opções de trabalhos para continuidade desta pesquisa.

2. Fundamentação científica e referencial teórico

2.1. Fundamentação científica

2.1.1. Transportador de correia e rolos

Por muitos anos o TC tem provado ser um equipamento com relativo baixo custo de mão de obra, energia e manutenção; aliando tudo com alta capacidade e confiabilidade (YANG, 2014). O TC apresenta um papel importante no transporte contínuo de materiais pesados em indústrias de mineração, terminais de granel, plantas de energia, produção química e outros. Comparado a outros meios de transporte que geralmente são usados para transporte de materiais sólidos pesados, como trens ou caminhões, TC é o meio de transporte mais confiável para transportar grandes volumes com rapidez e eficiência através dos processos de produção em áreas onde infraestrutura de estradas ou trilhos não existem ou são inviáveis (PANG, 2010).

Componentes de um transportador de correia

Independentemente do tipo de transportador, seu objetivo é formar um sistema de transporte que receba um material oriundo de um TC ou outro equipamento, transportá-lo ao longo de uma rota específica, por distâncias e trajetos adaptáveis, e descarregá-lo em uma pilha, navio ou transferi-lo para outro TC. Para garantir que esse objetivo seja atingido é necessário manter um plano de manutenção regular que garanta o bom funcionamento dos componentes do TC. Um TC possui múltiplos equipamentos. A lista abaixo resume e apresenta uma visão básica dos principais componentes dele (NASCIMENTO, 2018):

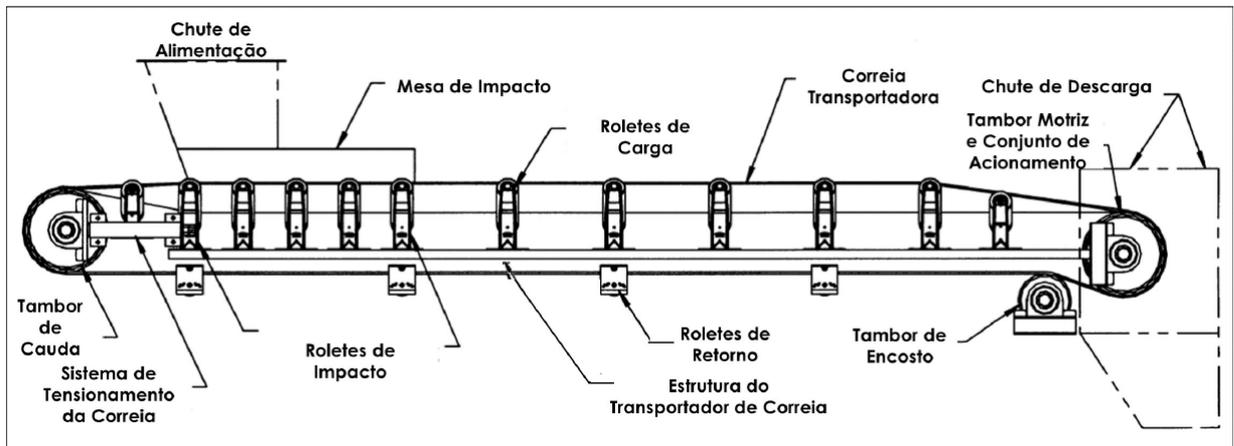


Figura 2.1: Principais componentes de um TC.
(NASCIMENTO, 2018)

Chute de Alimentação e Chute de Descarga: o primeiro é responsável por receber o material a partir de outro TC ou equipamento e posicioná-lo na correia para que possa ser transportado. Já o chute de descarga é usado para transferir o material transportado para um outro TC. Quando

a descarga do material é feita diretamente em pilhas ou mesmo em navios, geralmente ele não é usado. Além disso, em alguns casos, é necessário descarregar total ou parcialmente o material em um ou vários pontos intermediários do TC. Para isso, pode-se adotar desviadores ou trippers;

Mesa de Impacto: como o material a ser transportado vem de uma região mais elevada em relação ao TC atual, existe impacto quando o material é despejado pelo chute de alimentação. Logo, essa região é preparada para absorver e acomodar o material que será transportado;

Roletes de Impacto: são posicionados abaixo das mesas de transferência de modo a absorver o impacto do material despejado pelo chute de alimentação. Nessa região, os roletes são colocados próximos uns aos outros, em geral a cada 30 cm, e empregam rolos com revestimentos especiais para absorção de impacto;

Correia Transportadora: pode ser vista como o principal componente de um TC, já que é o elemento de contato com o material e que responde por entre 30% e 40% do custo total do TC. É composta pela carcaça e por uma cobertura. Esses dois elementos podem ter diferentes tipos de composições para fornecer à correia propriedades específicas às mais diferentes aplicações. Algumas das características resultantes da composição da correia são a resistência à tração, abrasão, alongamento, tipo de material suportado e ambiente de operação, visto que alguns compostos são específicos para operação em condições de temperaturas extremas e com exposição a elementos químicos;

Roletes de Carga: são responsáveis por suportar a correia e, conseqüentemente, o material transportado. Nos TCs tradicionalmente usados na mineração, os roletes de carga são espaçados a cada 1 m, mas essa distância varia de acordo com o projeto para condições específicas do TC ou da carga. Cada rolete é formado por um suporte e por três rolos, inclinados entre si em ângulos que normalmente variam de 0° a 45°. A inclinação varia em função do ângulo de acomodação do material, espessura da correia, velocidade de operação e tonelagem de transporte necessária, podendo mudar dentro de um mesmo TC;

Roletes de Retorno: suportam a parte inferior da correia, que está sem carga. Os roletes de retorno ficam geralmente espaçados entre 1,5 e 2,5 m, com variações de acordo com o projeto. Podem empregar 1 ou 2 rolos, podendo também desempenhar a função de alinhamento da correia em alguns projetos;

Conjunto de Acionamento: é geralmente composto por um motor elétrico, acoplamento hidráulico e um sistema de transmissão. É responsável por movimentar a correia ao fornecer tração para o tambor motriz, além de regular sua velocidade de operação;

Tambores (Motriz, de Encosto e Cauda): os tambores são geralmente feitos de aço e desempenham diferentes papéis em um TC. O tambor motriz é conectado ao conjunto de acionamento e fornece tração à correia. O tambor de encosto aumenta o ângulo de contato da correia com o tambor motriz, auxiliando na tração. Por fim, o tambor de cauda retorna a correia para sua posição inicial e também pode desempenhar funções relativas ao tensionamento da correia;

Sistema de Tensionamento da Correia: também conhecido como sistema de esticamento, é responsável por tensionar a correia durante o acionamento do TC e compensar mudanças na sua extensão, que ocorrem em função das condições de operação. Assim, esse sistema procura manter a tensão ideal de operação da correia em todos os momentos;

Estrutura do Transportador de Correia: é o conjunto de todos os elementos que suportam e dão sustentação ao TC, como estruturas metálicas, treliças, longarinas, torres, parafusos, conectores, soldas, etc.

As nomenclaturas entre os roletes e os rolos são muitas vezes confundidas. Segundo a norma NBR 6177 (Associação Brasileira de Normas Técnicas, 2016), denomina-se rolete o conjunto formado por um ou mais rolos montados em um suporte, cuja função é a de guiar e sustentar a correia transportadora. Por sua vez, ainda segundo a mesma norma, os rolos são elementos cilíndricos capazes de girar em seu próprio eixo com o objetivo de apoiar a correia (NASCIMENTO, 2018). Mais detalhes sobre rolos serão apresentados na próxima seção.

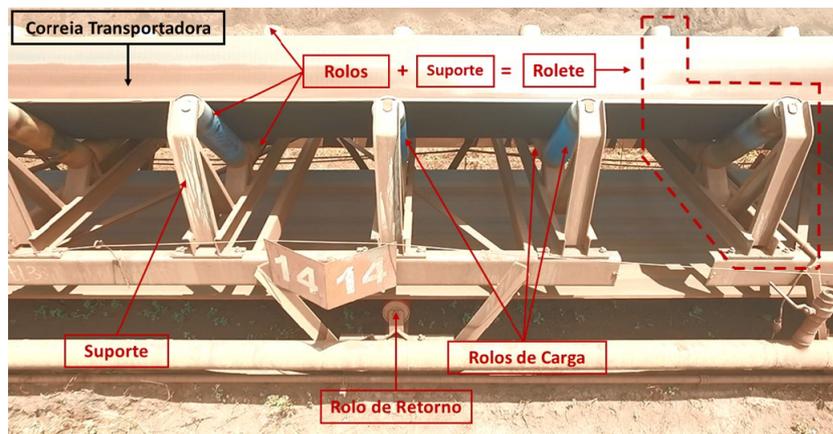


Figura 2.2: Nomenclatura de alguns dos elementos do TC.
(NASCIMENTO, 2018)

Rolos e suas falhas

A principal tarefa de um rolete de rolos é prover suporte de carga com pouca resistência em seu movimento. O principal requisito da maioria dos rolos é fazer o carregamento radial da correia e do material em cima de um rolete. Esta tarefa também inclui absorver impacto nos pontos de carregamento e tensões de correias quando se encontram em uma curva. O rolo também permite o movimento longitudinal, com apenas um pouco de aumento da tensão na correia. Parte disso é atribuído à resistência torcional à rotação do rolo. Minimizar essa resistência e melhorar a vida útil dos rolos e esteira são os objetivos primários do projeto de carregamento e rotação dos equipamentos de um TC (REICKS, 2008).

O rolo possui uma superfície cilíndrica externa rotativa acoplada a um eixo estacionário que liga dois rolamentos. Existem uma diversidade de tamanhos de eixo, posicionamento de rolamentos e diferentes materiais para o cilindro externo. O rolo normalmente pode

ser identificado com um diâmetro, largura e tipo de material, cada variação impacta de formas diferentes na performance geral do TC (REICKS, 2008).

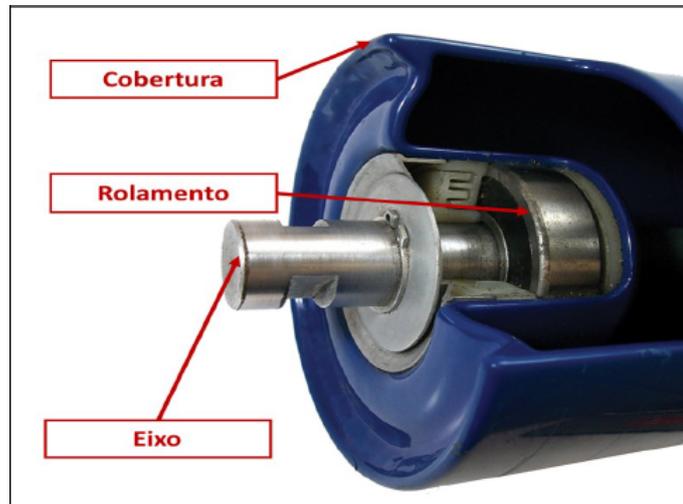


Figura 2.3: Partes principais de um rolo.
(NASCIMENTO, 2018)

Por ser o elemento de contato direto com a correia e em constante movimento, o rolo está suscetível ao desgaste natural e a diferentes tipos de falhas que reduzem seu ciclo de vida. O revestimento está sujeito a rasgos e fissuras causados pela degradação devido ao uso, atrito com um ponto específico da correia quando um rolamento trava ou pela queda de materiais perfurantes sobre ele. Contudo, as falhas nos rolamentos são as mais comuns e já é esperado que um número alto de rolos venha a falhar prematuramente. Os rolamentos podem travar, superaquecer ou até mesmo quebrar em virtude de deterioração ou de imperfeições em sua fabricação (NASCIMENTO, 2018).

Existem diversos fatores que podem levar os rolamentos a falhar. Segundo (CARVALHO JÚNIOR, 2018), são eles:

Carga excessiva: A carga excessiva do rolamento causa a fadiga do aço nas superfícies dos elementos rolantes ou pistas de rolamento. É necessário reduzir a carga operacional ou redimensionar os rolamentos;

Super-aquecimento: Temperaturas elevadas podem resultar na perda de dureza do material, na deformação dos elementos rolantes, na degradação do lubrificante e alterações dimensionais dos elementos do rolamento que poderá resultar na quebra do rolamento;

Esmagamento (falso brinell): O esmagamento ou falso brinell é causado por vibração excessiva do rolamento e deixa marcas elípticas na pista de rolamento. A vibração faz com que o elemento rolante não gire corretamente, e desta forma o filme de óleo que previne o desgaste não é formado;

Brinell: Essa falha ocorre quando a carga excede o limite elástico do anel do rolamento, gerando marcas chamadas de marcas de Brinell. Uma das causas pode ser o uso de marretas

durante a montagem do rolamento;

Fadiga normal: Ocorre quando parte superficial do material (descascamento) da pista de rolamento externo ou interno é perdida, e uma vez iniciado esse processo ele é progressivo e irá aumentar a vibração do rolamento e também haverá o espalhamento da perda de material. Esse tipo de falha é relacionada com o tempo de vida útil estimado do rolamento;

Carga reversa: Ocorre em rolamentos de carga axial que são desenhados para girar em somente uma direção;

Contaminação: A contaminação é uma das causas mais frequentes de quebras em rolamentos, ela ocorre pela infiltração de sujeira ou qualquer material abrasivo no rolamento. Para o presente projeto esse é um grande problema, pois o minério de ferro é um grande contaminante dos rolamentos;

Falha de lubrificante: O rolamento de esferas depende da presença contínua de um filme lubrificante muito fino (milionésimos de polegada de espessura) para evitar o desgaste. A lubrificação incorreta irá gerar o desgaste excessivo dos componentes causando sobre-aquecimento ou mesmo a quebra do rolamento;

Corrosão: Resulta da exposição do rolamento a materiais ou ambientes corrosivos. A lubrificação ajuda a evitar essa falha;

Desalinhamento: Quando os elementos rolantes giram, eles deixam marcas na pista de rolamento, se essas marcas não são paralelas à pista é um sinal de desalinhamento, podendo gerar sobre-temperatura e desgaste excessivo;

Erro de montagem: A montagem do rolamento obedece a limites de folgas pré definidos. Caso o rolamento é montado muito apertado ou muito folgado haverá falhas prematuras nele.

Como os rolamentos são os componentes dos rolos que apresentam os maiores desafios na sua inspeção, o monitoramento deve primariamente ser capaz de avaliar o estado deles. Os principais instrumentos disponíveis para esse fim baseiam-se nos mesmos princípios que permitem determinar a condição de qualquer equipamento com partes rotativas usando três tipos de sinais: acústico, térmico e vibração (NASCIMENTO, 2018).

Apesar de ser possível empregar esses três tipos de sinais para avaliar um rolamento, cada um deles possui características específicas que precisam ser avaliadas ao se construir a solução de monitoramento. Tanto o sinal acústico quanto a vibração conseguem detectar defeitos em fases iniciais, mas o sinal acústico, mesmo que em baixas velocidades de rotação, consegue fazê-lo de forma mais prematura, quando as frequências emitidas sequer são audíveis. No mais, ambos os sinais trazem consigo desafios na captura e processamento, visto que a operação ocorre em ambientes ruidosos e com fontes de vibração diversas (NASCIMENTO, 2018).

2.1.2. Sons e gravação digital

O som é um fenômeno físico causado pela vibração da matéria, sendo que esta vibração provoca alterações de pressão no ar que rodeia a matéria. Estas alterações de pressão propagam-se através do ar e quando atingem o ouvido humano, perturbam o tímpano onde o som é captado e posteriormente interpretado no cérebro. O áudio define-se como o processo de gravação, reprodução, transmissão ou recepção de som. O som é a base do áudio, que é a componente audível no intervalo de frequência de 20Hz a 20kHz (RIJO, 2018).

Quando a necessidade de gravar e reproduzir som surgiu, este era armazenado em suportes analógicos, como vinis e cassetes. Porém com a evolução tecnológica e o aparecimento do microprocessador, a digitalização do áudio tornou-se possível. Para que o computador seja capaz de interpretar o som é necessário discretizar a perturbação sonora através de Conversor Analógico Digital (Conversores A/D). Por sua vez, para a reprodução do áudio a partir de suporte digital são necessários Conversores Digital Analógico (Conversores D/A) (RIJO, 2018).

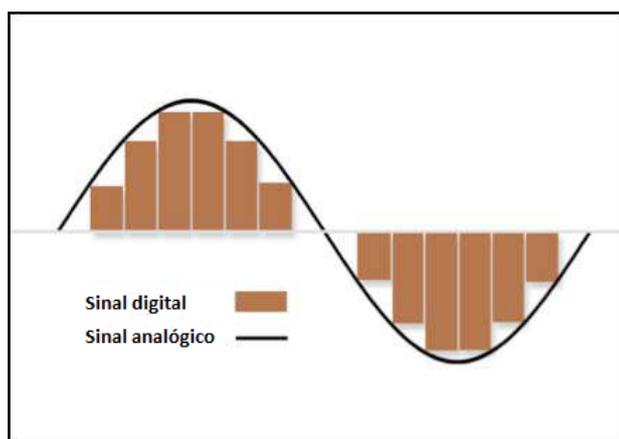


Figura 2.4: Sinal digital vs sinal analógico.
(RIJO, 2018)

Na gravação de um áudio digital é fundamental garantir que o arquivo de áudio gravado contenha o maior número possível de informações do sinal original. Essa regra vale tanto para gravações quanto para digitalizações de uma gravação analógica existente. Na construção de um áudio digital é recomendável que os gravadores registrem o áudio em formato não comprimido, sendo WAV o mais recomendado. Quanto ao equipamento adequado é importante considerar quais são as suas taxas de amostragem e resolução, as entradas para microfone, o sistema de alimentação, a resposta em frequência e o alcance dinâmico (FREITAG, 2014).

As definições de alguns desses conceitos são:

Taxa de amostragem: É o número de amostras (pontos) existentes em um segundo de um sinal gravado. Por exemplo, uma taxa de amostragem de 44,1 kHz significa que, por segundo, existem 44100 amostras (igualmente espaçadas no tempo), o que é equivalente a existir uma amostra a cada 0,0227 ms. Quanto maior a taxa de amostragem, mais próximo o sinal é da sua representação original, sendo que o desejável é ter a melhor representação possível

do sinal. No entanto há que considerar que altas taxas de amostragem resultam em arquivos significativamente maiores (RIJO, 2018).

Resolução: A resolução é o número de bits disponíveis para representar o valor de cada amostra de áudio. Tal como a taxa de amostragem, quanto maior o número de bits, melhor é a exatidão do sinal digitalizado em relação ao sinal original, tornando o arquivo maior em tamanho. Existem diversos padrões como 16, 20 e 24 bits. Uma amostra de áudio com uma resolução de 16 bits permite 65536 valores de amplitude diferentes (RIJO, 2018).

Forma de Onda: É representável através de um gráfico que demonstra o comportamento do som, em termos da sua amplitude (eixo Y), no domínio do tempo (eixo X).

Microfone

O microfone é parte fundamental na aquisição de ondas sonoras, senão o mais importante. Escolher o microfone que melhor se adequa aos propósitos da aquisição e às especificações do equipamento de gravação exige algum conhecimento básico sobre a estrutura e o mecanismo de operação desses equipamentos. Os microfones são descritos pelo seu princípio transdutor (dinâmico, condensador), por sua direcionalidade (omnidirecional, bidirecional, cardióide) e pelo tipo de uso a que se destinam (*shotguns*, de lapela, de mão, *headsets* etc) (FREITAG, 2014).

Os microfones dinâmicos são bastante resistentes, não precisam de baterias ou fontes de alimentação externas, mas não vêm equipados com pré-amplificadores. Os microfones condensadores respondem com clareza a sons transientes, o que resulta em sons mais naturais, limpos e ricos em detalhes. Além disso, pesam muito menos e podem ser muito menores. No entanto, porque requerem uma fonte de energia e têm cápsulas muito mais delicadas, podem apresentar algumas dificuldades de uso, maior cuidado de manipulação e maior fragilidade a ruídos (FREITAG, 2014).

Os microfones também podem ser identificados por suas propriedades direcionais, isto é, pela forma como foram planejados para captar o som de diversas direções. Assim, os microfones podem ser classificados geralmente entre um de dois grupos principais: omnidirecional e direcional. Os microfones omnidirecionais captam igualmente sons de todas as direções. Os direcionais podem ter várias configurações, sendo a mais comum a dos cardioides, que captam sons provenientes da frente, e as bidirecionais, que registram sons que vêm de direções opostas (FREITAG, 2014).

2.1.3. Aprendizado de máquina

Para resolver um problema em um computador é necessário um algoritmo. Um algoritmo é uma sequência de instruções que deve ser executado para transformar a entrada em uma saída. Contudo, para algumas tarefas não se sabe o algoritmo correto que faz esta conversão, mesmo sabendo qual é a entrada e como deve ser a saída o conteúdo que faz esta transformação

é desconhecido. A falta de conhecimento desse algoritmo, pode ser compensado com dados. Com uma grande quantidade de dados é possível “aprender” o que faz uma entrada convergir para uma saída. Com isso, seria ideal que uma máquina pudesse extrair automaticamente o algoritmo nestas situações, (ALPAYDIN, 2009).

Alpaydin (2009) afirma que existe um processo que explica os dados que estão sendo observados. Pode não se saber os detalhes do processo subjacente a geração de dados, mas se sabe que não é completamente aleatório. Existem certos padrões nos dados. Pode não ser possível identificar completamente o processo, mas pode ser possível construir uma boa e útil aproximação. Esta aproximação pode ser capaz de explicar algumas partes dos dados, identificando certos padrões e regularidades. Este é o nicho do aprendizado de máquina. Tais padrões ajudam a entender o processo, podendo ser usado para fazer previsões. Assumindo que este futuro, ao menos futuro próximo, não será muito diferente do passado de quando as amostras de dados foram coletadas, a previsão do futuro tem uma expectativa de estar certa.

Mesmo sabendo que o aprendizado de máquina é uma poderosa ferramenta para aquisição de conhecimento, deve ser observado que não existe um único algoritmo que seja eficiente para diferentes tipos de problemas. Logo, é importante ter conhecimento das características do problema e quais são as limitações que cada algoritmo tem para determinar qual o melhor para cada situação, (MONARD e BARANAUSKAS, 2003).

Random Forest

O algoritmo *Random Forest* (RF), proposto por Breiman (2001), tem sido bem sucedido como método de classificação e regressão de uso geral. A abordagem, que combina várias árvores de decisão randomizadas e agrupa suas previsões por votação, tem mostrado excelente desempenho em ambientes onde o número de variáveis é muito maior que o número de amostras. Além disso é versátil o suficiente para ser aplicado a problemas de larga escala, é facilmente adaptável e retorna os valores de variáveis importantes (BIAU e SCORNET, 2016).

Melhorias significativas na precisão da classificação foram resultados do crescimento dos conjuntos de árvores e da votação da classe mais popular. No intuito de cultivar estes conjuntos, geralmente vetores randômicos são gerados de tal modo que administrem o crescimento de cada árvore do conjunto. Existem diversos procedimentos, mas o elemento comum em todos eles é que, para a k-ésima árvore, um vetor aleatório é gerado, independente dos vetores aleatórios anteriores, mas com a mesma distribuição. Cada árvore é cultivada usando o conjunto de treinamento e o vetor aleatório, resultando em um classificador que leva em consideração o vetor de entrada e o vetor aleatório. Depois que um grande número de árvores é gerado, eles votam na classe mais popular. Esse procedimento é chamado de *Random Forest* (BREIMAN, 2001).

Support Vector Machine (SVM)

Introduzido por Cortes e Vapnik (1995), *Support Vector Machine* (SVM) é uma ferramenta para classificação e regressão que usa a teoria de aprendizado de máquina para maximizar a precisão preditiva enquanto, automaticamente, evita o sobre ajuste dos dados. SVM pode ser definido como um sistema que usa espaço de hipóteses de uma função linear em um espaço de características de alta dimensão, treinados com um algoritmo de aprendizado de máquina da teoria de otimização, que implementa uma via de aprendizado derivado de teorias estatísticas (JAKKULA, 2006).

Segundo Jakkula (2006), SVM é uma parte ativa das pesquisas de aprendizado de máquina por todo mundo. O SVM ficou famoso usando mapas de pixels como entrada, isto resultou em uma precisão comparável a sofisticadas redes neurais com características elaboradas em uma tarefa de reconhecimento de escrita à mão. Também tem sido usado para muitas aplicações, como análises de escritas a mão, análise facial entre outros, especialmente para as aplicações padrões de classificação e regressão.

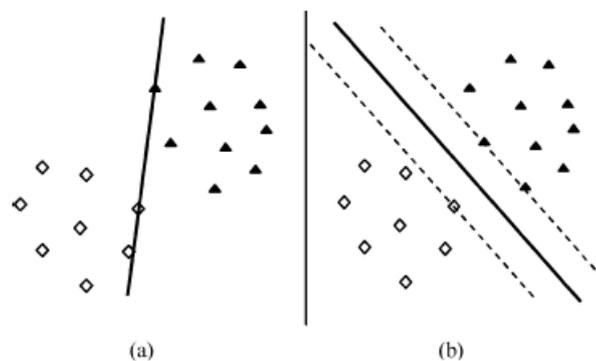


Figura 2.5: Dois hiperplanos separadores do espaço de atributos. (MAVROFORAKIS e THEODORIDIS, 2006)

O método utilizado pelo SVM para classificação consiste em encontrar o melhor hiperplano separador entre duas classes de amostras de treinamento no espaço de características, maximizando a margem de distância entre eles. Na figura 2.5 são vistas duas situações, em ambas os grupos são separados, mas em (a) o hiperplano separador não apresenta um distanciamento dos dados enquanto em (b) a reta apresenta o maior distanciamento entre os conjuntos de dados (MAVROFORAKIS e THEODORIDIS, 2006).

O SVM pertence à classe de métodos de Kernel e tem origem na teoria de aprendizado estatístico. Como todo algoritmo de aprendizado baseado em kernel eles são compostos de um propósito geral de aprendizado de máquina (no caso do SVM uma máquina linear) e um problema específico da função kernel. Desde que uma máquina linear pode apenas classificar dados em um espaço de atributos linearmente separáveis, a função do Kernel é a de induzir o espaço de atributos implicando o mapa de treinamento de dados a um alto espaço dimensional onde os dados são linearmente separáveis. Como o propósito geral do aprendizado de máquina

e a função kernel podem ser usados modularmente, é possível construir diferentes aprendizados de máquina caracterizados por diferentes superfícies de decisão não lineares (HOFMANN, 2006).

K-Nearest Neighbors (KNN)

K-Nearest Neighbor (KNN) é uma das técnicas de aprendizado e classificação mais populares. Muitas pesquisas tem mostrado que o KNN alcança bom desempenho em seus experimentos em diferentes conjuntos de dados. Introduzido por Fix (1951), o KNN tem provado ser um simples e poderoso algoritmo de reconhecimento (SUGUNA e THANUSHKODI, 2010).

KNN é um método de classificação, apesar de simples, muito efetivo. Seu apelo deriva do fato que sua superfície de decisões é não linear, existe apenas um único parâmetro inteiro (que é facilmente sintonizado com validação cruzada), e a qualidade esperada de precisão melhora automaticamente enquanto a quantidade de dados de treinamento aumenta. Estas vantagens, compartilhadas por muitos métodos não paramétricos, reflete o fato que apesar da máquina de classificação final ter grande capacidade, uma vez que acessa todo reservatório dos dados de treinamento durante o tempo de teste, o procedimento de aprendizado raramente causa sobrecarga (GOLDBERGER *et al.*, 2005).

No campo de reconhecimento de padrões, KNN é um dos algoritmos não paramétricos mais importantes e é um algoritmo de aprendizado supervisionado. A regra de classificação é gerada pelas próprias amostras de treinamento sem qualquer dado adicional. O algoritmo classificador KNN prevê as categorias das amostras de teste de acordo com a quantidade K de amostra de treinamento, que são os vizinhos próximos da amostra de teste, e escolhe qual das categorias tem maior probabilidade de ser a mais próxima (SUGUNA e THANUSHKODI, 2010).

Tree-based Pipeline Optimization Toll (TPOT)

O aprendizado de máquina automático é uma área crescente do aprendizado de máquina. O seu objetivo é similar a área de hiper-heurística, recomendar automaticamente pipelines otimizados, algoritmos ou parâmetros apropriados para uma tarefa específica sem depender muito do conhecimento do usuário. Apesar de estarmos longe de produzir programas completos sem um considerável nível de intervenção humana, o espaço entre programas automáticos e a habilidade necessária para produzir estes programas tem sido reduzido. Muitas iniciativas tem surgido nesta direção denominadas *Automatic Machine Learning* (Auto-ML) (DE SÁ *et al.*, 2017).

Desenvolvido por Olson e Moore (2016), *Tree-based Pipeline Optimization Toll* (TPOT) é um método de Auto-ML que de forma automática cria e otimiza pipelines de aprendizado de máquina para o domínio de um dado problema sem qualquer necessidade de intervenção humana. TPOT faz isso otimizando tarefas do aprendizado de máquina usando uma versão

de programação genética, uma técnica de evolução computacional bem conhecida que constrói automaticamente programas de computador.

2.2. Referencial Teórico

Nesta seção serão apresentadas pesquisas similares aos interesses desta dissertação. Pesquisas relativas a manutenção de transportadores de correia serão apresentadas primeiro e, em sequência, pesquisas de técnicas em aprendizado de máquina para classificação e identificação de sons ambientes.

2.2.1. Análises de rolos de Transportadores de Correia

Em Carvalho Júnior (2018), figura 2.6, é proposto um método para identificação de falhas em rolos usando um drone com uma câmera térmica embarcada. O uso do drone possibilita o acesso rápido à grande quantidade de rolos e a câmera térmica coleta os dados necessários para o processamento de imagens e identificação de defeitos. Foram utilizados dois algoritmos para a identificação da região dos rolos, o Viola & Jones e o Canais de Características Agregadas (ACF).



Figura 2.6: Coleta de dados de um TC usando drone.
(CARVALHO JÚNIOR, 2018)

A análise feita por imagens térmicas também está presente em Yang *et al.* (2016). Nesse projeto, figura 2.7, é apresentada uma inspeção dos equipamentos que compõem o TC. Um robô percorre trilhas ao longo da estrutura e captura em tempo real imagens térmicas do motor, rolos e outros componentes de transmissão. Estas imagens são classificadas em uma escala térmica de falha, sendo elas: comum, sério e urgente. As imagens são analisadas usando processamento de imagem e tecnologia de reconhecimento de padrões. Foi treinado um classificador SVM para

realizar a classificação automática da escala térmica de falha das imagens e obteve uma precisão de 96,7%.

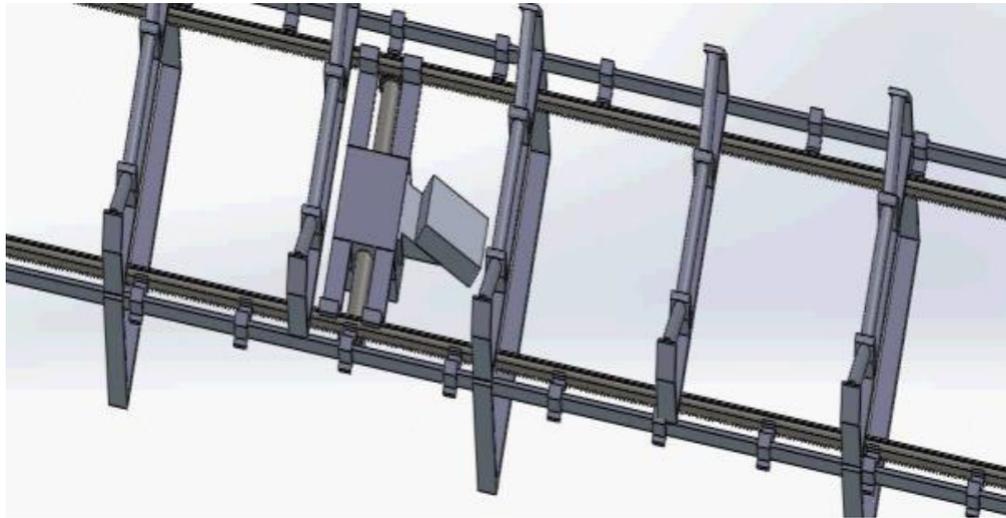


Figura 2.7: Robô de inspeção infravermelho.
(YANG *et al.*, 2016)

No projeto de Yang (2014) é desenvolvido um *Distributed Temperature Sensing* (DTS), sistema de monitoramento de temperatura, usando fibra óptica. O principal objetivo do projeto foi investigar várias opções de instalações dos cabos infravermelhos, ao longo do transportador de correia, para identificar a distribuição mais efetiva, figura 2.8. O sistema gera contínuas medidas e monitora a variação de temperatura dos componentes rolantes da esteira para identificar qualquer falha mecânica. A fibra óptica não é usada apenas para transmissão de dados, mas também para medir a temperatura. O sistema DTS fornece um perfil de distribuição de temperatura contínua, ao longo de todo comprimento do cabo da fibra óptica, com extrema precisão. Logo, o sistema gerencia as temperaturas do transportador de correia e avisa sempre que houver alguma temperatura acima da normal.

Em Lodewijks *et al.* (2016) foi feito um estudo onde foi analisada a implementação da *Internet of Things* (IoT) em um sistema de monitoramento de um TC. A pesquisa propõe colocar sensores em equipamentos que compõem o TC para análise contínua, mesmo nos rolos que são equipamentos numerosos e girantes. Neste caso, é proposto instalar, em cada rolo, um sensor *Radio Frequency Identification* (RFID) sem fio para verificar seu funcionamento, figura 2.9. No projeto é citado que o conceito de rolo inteligente tem sido bastante testado em laboratórios, em diferentes situações e configurações, e o sistema funciona. A barreira para implementação real do sistema é o seu alto custo.

O artigo de Li *et al.* (2013) aborda uma solução para detecção de falhas em rolos de TC usando a vibração por ele produzida quando esses apresentam falhas. Sensores de aceleração são colocados em lugares estratégicos para coletar sinais de vibração dos rolos do TC, figura 2.10. Os sinais são decompostos usando o método *Wavelet Packet Decomposition* (WPD), e então a energia de diferentes bandas de frequências do sinal original é obtida. O sinal de



Figura 2.8: Instalação do cabo de fibra-óptica.
(YANG, 2014)

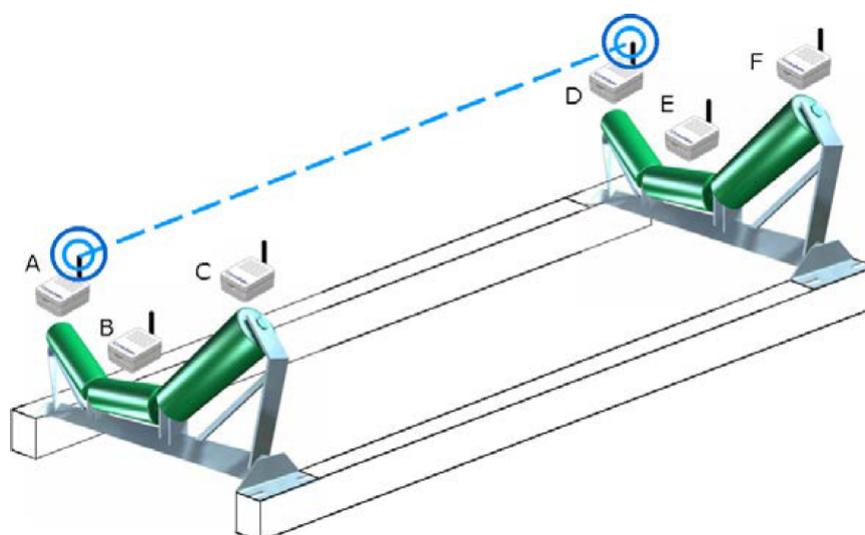


Figura 2.9: Comunicação direta entre os rolos.
(LODEWIJKS *et al.*, 2016)

vibração muda quando alguma falha acontece no rolo e a energia de cada banda de frequência é diferente para cada diagnóstico do rolo. Estes dados são usados para treinar uma rede SVM para monitorar e classificar as vibrações e determinar a situação dos rolos na proximidade dos sensores.

Em Ericeira (2019) é apresentada uma solução para detecção de defeitos em rolos de transportadores utilizando sensoriamento ultrassônico. Foram realizadas gravações de ultrassom em rolos sem ruídos perceptíveis, classificados como não-defeituosos e em rolos que apresentavam ruídos característicos de falhas já perceptíveis, classificados como defeituosos, figura 2.11. A base de dados foi utilizada para treinamento e teste de desempenho de algoritmos de aprendizado de máquina do tipo RF e *Multilayer Perceptron* (MLP). Foram realizados alguns experimentos variando o domínio dos dados e atributos estatísticos. Os resultados dos testes

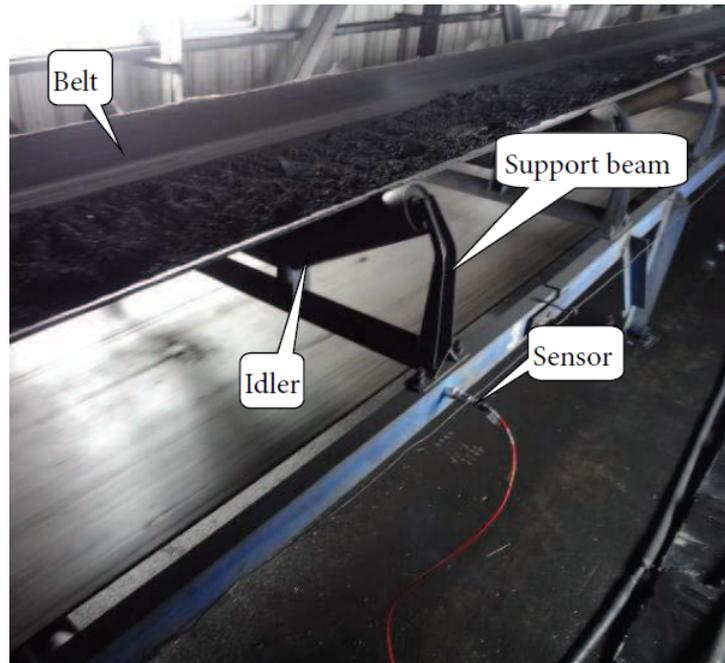


Figura 2.10: Correia e sensor de vibração instalado.
(LI *et al.*, 2013)

comprovaram que os rolos podem ser diferenciados pela faixa de ultrassom e dentro dos experimentos teve um que teve média de classificação correta de rolos de 83,68%, tendo o melhor resultado com taxa de acerto de 90%.



Figura 2.11: Gravação de ultrassom de um rolo em um transportador.
(ERICEIRA, 2019)

2.2.2. Aprendizado de máquina em classificação e identificação de sons ambientes

No artigo de Chu *et al.* (2006) foi apresentado um robô móvel que utiliza baixo nível do sinal de áudio em um sistema de reconhecimento para determinar a diferença de 5 ambientes baseado nos sons de fundo desses locais. Os lugares considerados estão nos arredores do prédio de engenharia de uma universidade e são: área de café, corredor, proximidade e dentro de um elevador, saguão e uma rua próxima do prédio. Para avaliar a performance do sistema de reconhecimento foram examinados 3 classificadores: KNN, *Gaussian Mixture Models* (GMM) e SVM. Todos os métodos obtiveram uma acurácia próxima de 90% em identificar o ambiente correto.

O artigo de Qian *et al.* (2017) contribuiu com uma análise da efetividade da característica de oscilação para classificação de cenas acústicas para a sub-tarefa do IEEE AASP Desafio em Detectar e Classificar Cenas Acústicas e Eventos (DCASE2017). Foi analisada a base de dados do DCASE2017 que consiste em 312 segmentos de 10 segundos de 15 classes: praia, ônibus, restaurante, carro, centro de cidade, floresta, mercado, casa, biblioteca, estação de metrô, escritório, parque, área residencial, trem e bonde. Foram comparados dois algoritmos de aprendizado SVM e o *Gated Recurrent Neural Network* (GRNN). Neste caso, SVM teve uma performance um pouco melhor que o GRNN, mas ambos com porcentagem de acerto maior que 80%.

Em Rijo (2018) é desenvolvido um sistema de classificação baseado apenas na forma de onda do áudio, sem recorrer a qualquer tipo de extração de características. Os valores que definiram o sinal foram a entrada direta da rede neural artificial. As classes de eventos selecionados para distinção através desta rede foram: voz, música, gargalhadas, aplausos, gritos e silêncio. As redes usadas para treinamento foram *Deep Neural Network* (DNN) e *Recurrent Neural Network* (RNN). O sistema de classificação foi eficiente, tanto pela rapidez do processo como pela sua precisão, que atingiu taxas de reconhecimento dos sons na ordem de 90%.

Em Zhu *et al.* (2018) é proposta uma arquitetura de rede CNN que usa a forma pura da onda como entrada. Essa abordagem utiliza uma série de CNNs, separadas paralelamente com diferentes tamanhos de filtro de convolução, com o intuito de aprender características com resolução multi-temporal. A arquitetura proposta também agrega características hierárquicas de camadas CNN usando conexões diretas entre camadas convolucionais. Os resultados mostraram que a combinação de características multi-nível e multi-escala melhoraram a performance do classificador. Foram usadas duas bases de dados, a ESC-50, que consiste em 50 classes que podem ser divididas em 5 grupos maiores: animais, paisagens sonoras naturais e sons de água, sons humanos sem falas, sons domésticos e sons urbanos, e a DCASE2017.

Já em Abdoli *et al.* (2019) é apresentada uma aproximação *end-to-end*, sinal sem pré-processamento, de uma rede CNN 1D. A proposta pode lidar com sinais de áudios de qualquer largura pois o sinal é dividido em quadros sobrepostos usando uma janela deslizante. A arqui-

tutura consiste de 3 a 5 camadas convolucionais, dependendo do tamanho do sinal de áudio. Ao invés de usar *filterbank* ou extrator de características *Mel-Frequency Cepstral Coefficient* (MFCC), a proposta CNN 1D aprende os filtros diretamente da forma de onda do áudio. A performance da proposta end-to-end em classificar sons ambientes foi avaliada na base de dados UrbanSound8k, com mais de 8000 amostras de áudio e o resultado experimental obteve uma média de acertos de 89%.

Em Huzaiifah (2017) é afirmado que a visualização de um sinal de áudio através de diferentes representações tempo-frequência, como o espectrograma, oferece uma rica representação da estrutura temporal e espectral do sinal original. Neste artigo são comparados vários métodos populares de processamento de sinal. São eles: *Short-Time Fourier Transform* (STFT) com escala linear e Mel, *Constant-Q-Transform* (CQT) e *Continuous Wavelets Transform* (CWT). Estes foram avaliados quanto a capacidade de classificação em 2 bases de dados de sons, ESC-50 e UrbanSound8k, usando uma rede CNN.

No artigo de Han e Lee (2016) é demonstrado como foi aplicado uma rede CNN para a base de dados DCASE2017. Neste projeto é proposta uma variedade de métodos de pré-processamento que enfatizam diferentes características acústicas. São eles: representação bi-neural, *Harmonic Percussive Source Separation* (HPSS) e *background subtraction*. Os resultados experimentais mostraram que a estrutura de rede proposta e os métodos de pré-processamento efetivamente aprenderam características acústicas e este modelo reduziu bastante a margem de erro tendo uma taxa de acerto de 91,7%.

É proposto no artigo de Bae *et al.* (2016) combinar as redes *Long Short-Term Memory* (LSTM) e CNN em paralelo como redes menores com objetivo de explorar correlações sequenciais e o local de informações espectro-temporal. Na camada LSTM, sequências de qualidades MFCCs são utilizadas como entradas para extrair informações sequenciais. A camada CNN aprende a localidade espectro-temporal do *Spectrogram Image File* (SIF). As saídas das duas camadas são combinadas por uma camada que consegue aprender características, de LSTM e CNN, que se complementam. Para comparar a performance do método proposto foi conduzido um número de experimentos no banco de dados TUT Acoustic Scenes 2016, 1170 gravações que compõem 15 classes de cenas acústicas diferentes. Os resultados revelaram que a combinação de LSTM e CNN foi mais eficiente, para este caso, do que redes convencionais como DNN, CNN e LSTM.

Em Weiping *et al.* (2017) também é utilizada uma rede CNN, mas neste caso múltiplos espectrogramas dos arquivos de áudios são produzidos e usados pra treinar uma rede CNN. O artigo explora duas diferentes produções de espectrogramas: espectrograma padrão e CQT. De acordo com a largura da janela deslizante e o espaço do salto, múltiplos padrões de espectrogramas com diferentes resoluções são geradas. Foram testadas as performances da rede CNN com o espectrograma de múltiplos padrões e CQT. Também foi testado um modelo unindo os métodos. Os experimentos concluíram que a fusão dos dois métodos teve melhor performance.

Em Cunningham *et al.* (2020) é apresentado um estudo que faz detecção de estímulos

emocionais resultantes de sons não musicais. O objetivo é entender melhor como sons provocam respostas emocionais em um público, isto possivelmente melhoraria o trabalho de designers de sons. Este estudo utiliza a base de dados International Affective Digitized Sound (IADS), um conjunto de 167 diferentes sons e suas emoções associadas. Um total de 76 características foram extraídas dos sons, abrangendo o domínio do tempo e frequência. As características foram usadas em dois modelos de aprendizado de máquina: modelos de regressão e redes neurais artificiais. As redes neurais rasas tiveram resultados melhores que os modelos de regressão e os melhores resultados foram próximos de 65%. Estes resultados apresentaram uma melhora significativa em comparação com outros estudos encontrados na literatura.

Em Akiyama *et al.* (2020) é proposto um método para detecção de eventos de sons em que o modelo de treinamento foca em eventos de sons de curta duração. O método proposto foca na relação entre a duração do evento de som e a facilidade/dificuldade do modelo de treinamento. Em particular, muitos eventos de sons de longa duração são estacionários, ou seja, tem pouca variação nas suas características e seu modelo de treinamento é simples. Entretanto, alguns eventos de curta duração têm mais que um padrão de áudio. Logo, foi aplicada uma classe que repondera a função de perda com entropia cruzada dependendo da facilidade/dificuldade do modelo de treinamento. Os experimentos de avaliação foram conduzidos usando os conjuntos de dados TUT Sound Events 2016/2017 e TUT Acoustic Scenes 2016 e os resultados mostraram que o método proposto melhorou a performance de detecção de eventos de som em 3,15 e 4,36 por cento em macro e micro-Fscores, respectivamente, em comparação com métodos convencionais que usam função de perda com entropia cruzada.

Analisando as demais propostas de identificação de falhas em rolo, a proposta por som apresenta uma solução menos invasiva, não tendo necessidade de contato com o rolo ou estrutura do transportador de correias. A detecção via som também não sofre com interferências que podem afetar a medição como horário do dia ou estação do ano, como é o caso de soluções usando temperatura. O grande desafio é isolar o som do entorno do rolo como o próprio movimento da correia, equipamentos e caminhões normais em uma mineradora. Os projetos citados para identificação e classificação de sons mostram uma variedade de métodos e técnicas de aprendizado de máquina. Os métodos obtiveram resultados significantivos para diferentes bases de dados de som. Conclui-se a partir desses trabalhos que é possível fazer a classificação de sinais sonoros de diversas formas.

3. Metodologia

Neste capítulo serão apresentados os procedimentos metodológicos que foram empregados para alcançar o objetivo de fazer um algoritmo de aprendizado conseguir diferenciar se um rolo de um TC está defeituoso ou não através de seu som.

3.1. Materiais e métodos

Para alcançar o objetivo proposto foi necessário coletar amostras de sons de rolos em diferentes condições, posições na esteira e áreas da mina em que se encontram. Estas amostras foram analisadas para ver se através das suas características seria possível distinguir rolos bons de rolos defeituosos. Mas antes disso é muito importante determinar os materiais, configuração dos arquivos gravados e programas adequados para esta tarefa.

Sabendo que o ambiente próximo a um transportador de correias é muito ruidoso, o instrumento de medida de som tem que ter características para reduzir ao máximo a interferência de sons indesejados na gravação. O microfone possui duas variações, condensador e dinâmico. O microfone condensador é muito sensível, mais usado em lugares controlados pois é suscetível a ruídos, além disso necessita de uma fonte de energia. Nessas condições, o microfone dinâmico é ideal para nossa aplicação pois ele é menos sensível, sendo apropriado para lugares mais ruidosos, e não precisa de uma alimentação de energia especial, podendo ser ligado direto em um computador portátil para coleta de áudios dentro de uma mineradora. Outro fator relevante a ser considerado é a direcionalidade do microfone, ele pode ser direcional ou omnidirecional. O microfone omnidirecional coleta o áudio de todas as direções, enquanto o direcional possui diferentes configurações, sendo a mais relevante para este projeto o microfone cardioide, que coleta o som de uma única direção. O microfone cardioide possui variações, podendo ser supercardióide e hipercardióide. Essas funções só aumentam a direcionalidade do microfone, reduzindo ao máximo sons indesejáveis de outras direções. Baseado nessas informações foi utilizado para coleta de áudios de rolos um microfone dinâmico hipercardióide.

Os arquivos dos áudios gravados precisam ter suas configurações definidas para garantir que o arquivo digitalizado apresente dados que representem bem o original. É importante ter uma quantidade grande de informações do som original para que sua representação digital seja o mais parecida possível. Foi definida uma taxa de amostragem de 96 KHz, representando 96000 amostras de som por segundo, e uma resolução de 24 bits, significando que cada amostra terá 24 bits para representar o seu valor. Na escolha das configurações do arquivo de áudio também foi levado em consideração qual será a duração da gravação e conseqüentemente qual será o tamanho de cada arquivo pensando em não sobrecarregar o computador em campo e ser de fácil manipulação para análises. A gravação foi feita por apenas um canal (mono) e no formato WAV, formatos compactados devem ser evitados pois dados podem ser perdidos na compactação. O programa definido para gravação de áudio foi o Audacity que é *open source* e foi eficiente nos

testes preparatórios.

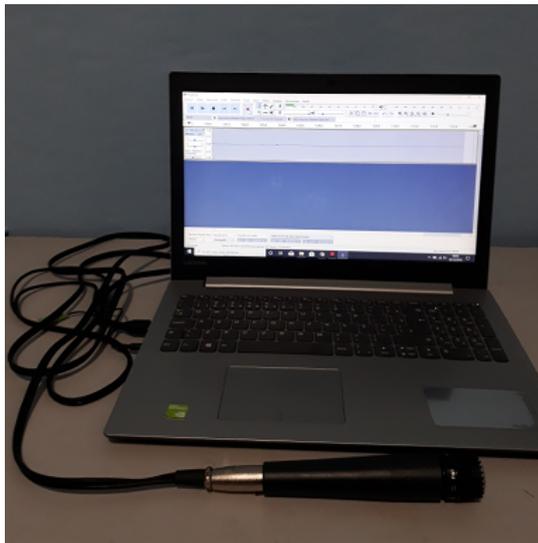


Figura 3.1: Equipamentos e programa para coleta de áudios.

Na figura 3.1 vê-se o microfone escolhido integrado com um computador portátil, material utilizado para fazer as gravações de áudios dos rolos. O som será capturado pelo microfone e gravado no computador portátil usando o programa Audacity.

3.2. Coleta de dados

A coleta de dados de áudio foi feita do dia 9 ao dia 11 de julho de 2019 no Porto de Tubarão no Espírito Santo, figura 3.2. O porto recebe o minério retirado de Minas Gerais, via trem, e esse minério é transportado dentro do porto, via TC, até os navios que seguem com ele para outros países. Com um grande conjunto de transportadores de correias, o Porto de Tubarão é um lugar ideal para a coleta de sons de rolos em diferentes condições.



Figura 3.2: Porto de Tubarão, Vitória - Es (WEBPORTOS)

Uma vez dentro do porto, foi traçado um trajeto ao lado de diferentes transportadores. Com o auxílio de um inspetor de manutenção de transportador de correias sempre que um rolo candidato foi identificado, áudios foram gravados. O inspetor identificou quais eram os rolos que seriam considerados para troca e estes foram gravados e identificados como áudio de rolo defeituoso. No processo de gravação o microfone foi posicionado próximo ao rolo, a uma distância segura entre 20 a 30 cm, capturando o seu som por volta de 30 segundos, figura 3.3.

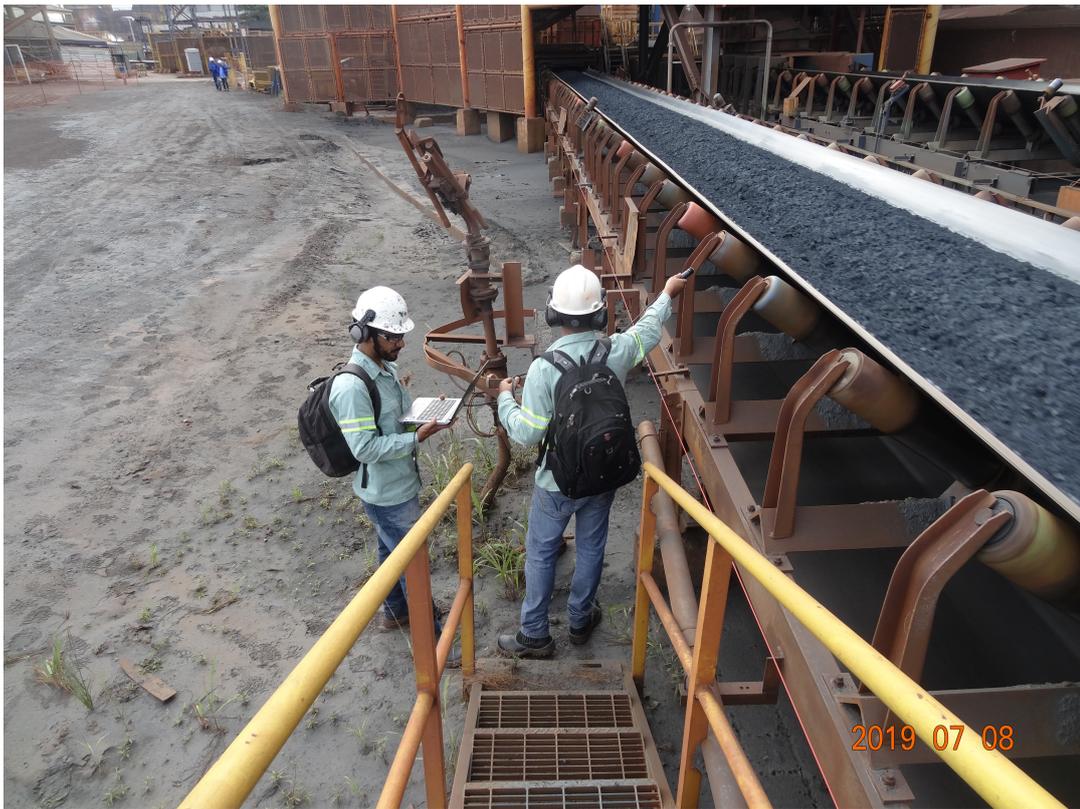


Figura 3.3: Gravação dos áudios em campo.

Os áudios foram gravados em diferentes etapas do dia-a-dia do porto e em diferentes locais pelos quais o transportador de correias passava, sendo alguns próximos de outras máquinas ruidosas ou com trânsito de automóveis de grande porte. Notou-se que o som gravado teve pouca interferência e que o microfone conseguiu gravar nitidamente os sons dos rolos. No final dos três dias foram coletados um total de 19 áudios de rolos bons e 38 áudios de rolos defeituosos, sendo cada áudio referente a um rolo diferente.

3.3. Processamento de dados

Com os arquivos gravados em mãos, os áudios tinham que ser interpretados pelo computador. Um áudio digitalizado é representado como um vetor único que contém todas as amostras coletadas do som original ao longo da duração do mesmo. Como a taxa de amostragem foi de 96KHz, as primeiras 96000 casas do vetor representam cada amostra feita dentro de 1 segundo.

Os sons dos rolos, independente da sua condição, possuem um padrão constante. Consequentemente, com o intervalo de 1 segundo de som já é possível distinguir a qualidade dos rolos. Este intervalo teria 96000 amostras, o suficiente para se ter uma boa representação do som original. Logo, para aumentar a quantidade de amostras, os áudios que tinham aproximadamente 30 segundos foram divididos em segundos e nestes seguiram a classificação do áudio original. Cada áudio gerou em média 30 amostras que representavam um determinado rolo.

Para comparar as diferenças das condições dos rolos através de seus áudios, se analisar a variação de intensidade do som no domínio do tempo percebe-se que a diferença entre eles não é constante e os sons ambientes interferem nestas condições. Logo, a proposta foi analisar as amostras no domínio da frequência. No domínio da frequência os sons do ambiente interferem de forma semelhante nas amostras, logo a variação da amplitude nas frequências de rolos defeituosos e bons representam a diferença entre eles e pode ser utilizada para distingui-los.

Para fazer o tratamento e análise dos dados foi utilizado o programa Matlab. Nele, inicialmente, foram selecionadas as amostras dentro de um segundo de áudio de um rolo defeituoso e de um bom para ver as diferenças dos sinais. Em cada conjunto foi aplicado a Transformada de Fourier, convertendo os dados no domínio do tempo em dados no domínio da frequência.

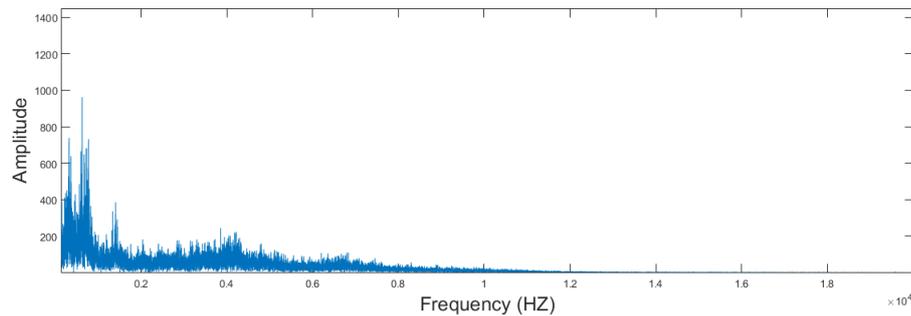


Figura 3.4: Gráfico do som de um rolo bom no domínio da frequência.

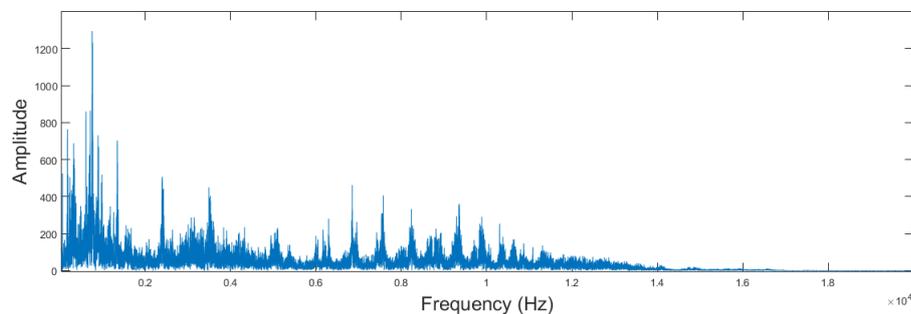


Figura 3.5: Gráfico do som de um rolo defeituoso no domínio da frequência.

Nas figuras 3.4 e 3.5 é possível ver o gráfico de cada conjunto de amostras. O gráfico mostra a amplitude de cada frequência contida nas amostras de cada conjunto. O intervalo analisado é o de 0 a 20KHz pois é o intervalo de frequência audível pelo ser humano, uma vez que o som pode ser distinguido pelo inspetor. Notou-se que algumas faixas de frequência

apresentaram grandes diferenças. Neste caso, mesmo se algumas destas frequências forem influenciadas por algum sinal externo, ao se analisar diferentes conjuntos de amostras o algoritmo de aprendizado irá detectar quais delas serão dominantes para distinguir os rolos.

O sinal inicial mostra a intensidade do som ao longo do tempo. Notou-se que em alguns casos está intensidade, para rolos de mesma condição, tinham amplitudes diferentes. Isso pode acontecer porque a intensidade de coleta de dados do microfone foi ajustada com níveis diferentes para distintos rolos, ou até mesmo interferências na alimentação do microfone. Essa variação na amplitude do sinal no domínio do tempo interfere na amplitude das frequências no domínio da frequência. Caso um algoritmo de aprendizado leve em consideração apenas as intensidades das frequências, sinais que foram gravados com intensidades diferentes podem interferir nos resultados finais. Baseado nessas informações, foi tomada a decisão de normalizar os dados iniciais no domínio do tempo em um intervalo entre 1 e -1, com isso todos os áudios ficaram com amplitudes semelhantes. Logo, a análise no domínio da frequência tem menos interferência por causa da variação de amplitudes diferentes no domínio do tempo. Assim, a análise pode ser feita pelas variações entre as frequências (formato da onda), fator que evidencia as diferenças entre os sons dos rolos.

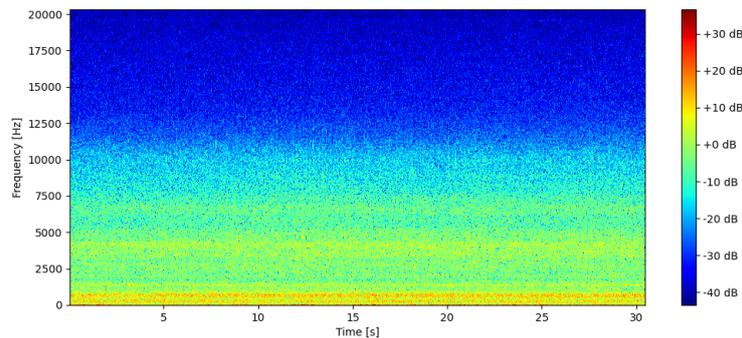


Figura 3.6: Espectrograma de um rolo bom.

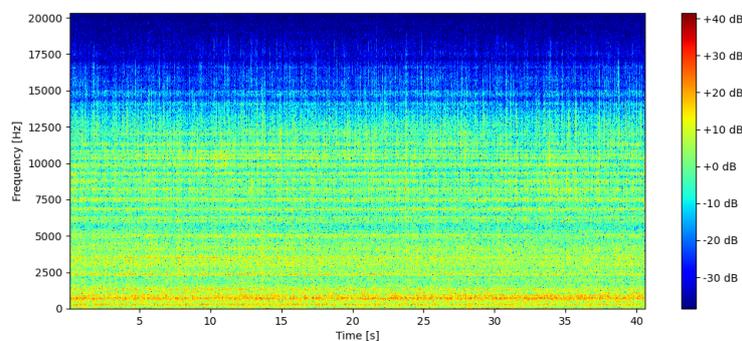


Figura 3.7: Espectrograma de um rolo ruim.

Outro modo de ver as diferenças entre as frequências de rolos bons e ruins pode ser visto nas figuras 3.6 e 3.7. As figuras ilustram o espectrograma da gravação de um rolo em

boas condições e um defeituoso. O espectrograma mostra ao longo do tempo as frequências que foram gravadas e as cores indicam a intensidade de cada frequência ao longo da gravação. Logo, analisando as figuras é possível ver que na gravação do rolo ruim existem diversas faixas de frequências, ao longo da gravação, com intensidades diferentes das apresentadas na gravação de um rolo bom.

Sabendo que os rolos defeituosos podem ser identificados pela diferença de intensidade de algumas faixas de frequência, o propósito agora é criar uma *Analytical Base Table* (ABT) com os áudios coletados. Esta tabela irá conter características selecionadas das amostras dos áudios dos rolos gravados em funcionamento na mina. As amostras desta tabela serão classificados como amostras de rolos bons ou rolos ruins.

No início do processamento de dados o sinal de áudio foi dividido em segundos, sendo que cada segundo continha a definição de qual classe de rolo ele pertencia. Com isso, foi formado uma matriz onde cada linha representava um segundo do áudio e as colunas representavam cada uma das 96000 amostras contida em cada segundo. Com a passagem dos dados para o domínio da frequência, passou a se ter 20000 colunas e estas continham a amplitude das frequências de 1 a 20000 Hz pertencente a cada segundo. Desta maneira, as linhas dessa tabela serviriam como amostras para um algoritmo de aprendizado e as colunas seriam as características definidas para este problema, ou seja, as entradas desse algoritmo. Como desse modo o algoritmo teria 20000 entradas, foi decidido reduzir o número de colunas agrupando por média da soma do conteúdo de 200 colunas, ou seja, a média da soma da amplitude de cada hertz num grupo de 200 em 200. Os picos que diferenciam os rolos ficaram divididos em diferentes grupos deste intervalo e a tabela passa a ter 100 colunas, um número de entradas mais aceitável de se manipular nos algoritmos de aprendizado. Com isso, a tabela manteve a organização de cada linha corresponder cada segundo e as colunas à média aritmética das intensidades de cada grupo de 200 Hz (200, 400, ..., 19800, 20000).

Na tabela 3.1 é mostrado uma representação da base de dados. Nela é possível ver as linhas que correspondem a cada amostra do algoritmo de aprendizado, já definida as saídas como rolo bom ou ruim (rb,rr), e as colunas que são as entradas que correspondem aos 100 atributos escolhidos para treiná-lo. A construção da tabela de dados finalizou com um total de 1665 amostras.

3.4. Algoritmos de aprendizado de máquina para a base de dados de áudios de rolos

Com a base de dados pronta, o próximo passo foi definir o algoritmo de aprendizado apropriado para este projeto. Foi proposto fazer um teste com 3 algoritmos diferentes e compará-los para ver qual deles seria mais eficiente para esta base de dados. A linguagem de programação Python foi utilizada para testar os algoritmos de aprendizado pois possui um am-

Tabela 3.1: Base de dados dos áudios dos rolos.

Amostras	Atr1	Atr2	Atr3	...	Atr98	Atr99	Atr100	Saída
1	202.47	214.05	303.87	...	1.09	1.04	0.98	rb
2	141.90	208.71	328.61	...	1.15	1.07	1.10	rb
3	201.81	210.22	280.04	...	1.43	1.14	1.18	rb
...
557	265.38	165.01	82.38	...	1.88	1.85	1.77	rb
558	293.50	177.69	86.91	...	2.04	1.87	1.75	rb
559	174.16	178.27	73.56	...	1.87	1.64	1.65	rb
560	180.76	118.11	140.99	...	0.49	0.59	0.53	rr
561	89.66	100.71	130.74	...	0.41	0.36	0.40	rr
562	61.89	100.56	148.66	...	0.40	0.42	0.37	rr
...
1663	155.21	181.70	192.88	...	1.43	1.22	1.29	rr
1664	189.30	178.68	190.41	...	1.32	1.34	1.26	rr
1665	222.23	189.13	205.08	...	1.35	1.30	1.22	rr

plo conjunto de ferramentas e bibliotecas. Foram escolhidos dois algoritmos já bem conhecidos e eficientes em projetos de classificação, RF e SVM. O terceiro algoritmo foi escolhido usando o TPOT, uma ferramenta de auto-ML que possui uma biblioteca no Python, que irá aplicar uma coleção de algoritmos na base de dados aqui considerada, retornando qual seria o melhor modelo classificador segundo alguns critérios.

Para utilizar o TPOT inicialmente é necessário definir os seus parâmetros. Existem dois deles que são bastante importantes, um é o tamanho da população, que define quantos conjuntos de algoritmos e seus parâmetros serão julgados, e o outro é a geração, que é o número de vezes que novas populações serão analisadas. No teste feito com a base de dados gerada com os áudios foi utilizada uma população de 100 métodos que serão analisados em 50 gerações.

Na primeira geração, o TPOT escolhe aleatoriamente diferentes métodos e avalia suas taxas de acertos com a base de dados escolhida. Nas demais gerações, o algoritmo escolhe os 20 melhores métodos da geração anterior e cada um desses métodos produz 5 cópias, variando seus parâmetros seguindo as definições internas do algoritmo de seleção do TPOT. O algoritmo repete esse processo até a última geração, alterando e sintonizando parâmetros dos métodos selecionados na população, melhorando a precisão da classificação. Ao final da última geração, o algoritmo selecionado terá a configuração com melhor precisão da população analisada e qual a sua seleção de parâmetros, sendo possível extrair a codificação em Python para a base de dados em questão (OLSON e MOORE, 2016).

4. Resultados e Discussão

4.1. Definições iniciais

Foram seguidos alguns procedimentos com o objetivo de escolher condições próximas das melhores em cada algoritmo de aprendizado. Para a base de dados dos áudios dos rolos o método que apresentou melhor precisão no TPOT foi o KNN com 5 vizinhos mais próximos de cada amostra de teste e o peso desses vizinhos foram definidos como uniforme, ou seja, todos os vizinhos terão o mesmo peso independente da distância.

Para definir os parâmetros do SVM foi feito um teste variando os seus parâmetros e analisando qual deles apresentou melhor taxa de acertos para a base de dados dos rolos. O parâmetro a ser definido no SVM é o kernel, que pode variar entre linear, radial, polinomial e sigmoide. As definições de qual desses kernels é mais eficiente depende do espalhamento do conjunto de dados. Logo, foi feita uma análise testando o SVM para cada kernel. O SVM foi aplicado e analisado 30 vezes para cada kernel e no final foi avaliada a média da precisão de cada kernel.

Quanto ao RF os parâmetros que foram alterados para analisar a melhor performance do algoritmo foram a quantidade de árvores na floresta e a profundidade de cada árvore. Normalmente quanto maior o número de árvores melhor o aprendizado. Entretanto, adicionar muitas árvores pode aumentar consideravelmente o tempo de processamento durante o treinamento. Em relação a profundidade, quanto maior o seu valor maior será o número de divisões da árvore, o que aumenta o número de informações retiradas do conjunto de dados, aumentando também a probabilidade do sistema sofrer sobreajuste (*overfitting*, em inglês). Logo, se definiu um valor padrão de 100 árvores e profundidade 5 e foram avaliados valores acima e abaixo destes parâmetros. Cada situação foi implementada no RF 30 vezes, usando o conjunto de dados dos áudios, e a média total da taxa de acertos no treinamento foi usada para comparação.

4.2. Resultados da precisão dos testes de cada algoritmo de aprendizado

Os testes foram feitos com a base de dados de áudios de rolos construída neste projeto. O total de amostras de áudios de rolos contavam com uma base de dados de 1665 amostras, dentre essas, 559 amostras são de rolos bons e 1106 de rolos defeituosos. Para esta fase de treinamento e teste dos algoritmos foi usada uma base balanceada, ou seja, foi usada a mesma quantidade de amostras de cada condição do rolo, 559 de cada.

Os resultados de cada modelo serão apresentados e as melhores configurações serão escolhidas para comparação entre os algoritmos de aprendizado definidos. Para o método KNN, que foi eleito usando um programa de auto-ML, sua configuração foi definida pelo TPOT. Com a configuração dos parâmetros, usando 5 vizinhos para análise de cada amostra e o peso destes

vizinhos uniforme, o algoritmo KNN foi testado 30 vezes com a base de dados dos rolos e apresentou uma média de acertos de 99,1%.

Os testes para o SVM apresentaram os resultados mostrados na tabela 4.1. Comparando-se as taxas de acertos dos kernels, observa-se que os SVMs com kernels lineares tiveram um desempenho superior àqueles com kernels não lineares. Uma provável justificativa para tal fato diz respeito à natureza dos dados do problema em questão e dos atributos escolhidos que sugerem que se está lidando com um problema linearmente separável ou muito próximo de linearmente separável. Os SVMs com kernels não lineares poderiam até apresentar resultados melhores que os lineares para este problema. Porém, além de ser necessário fazer uma sintonia mais apurada dos parâmetros, via de regra, os SVMs não lineares implicam em modelos mais complexos e, conseqüentemente, modelos mais suscetíveis a 'overfitting' (ou sobreajuste). Portanto, neste trabalho, decidiu-se utilizar o SVM com kernel linear para fins de comparação com as demais técnicas de Aprendizado de Máquina aqui avaliadas.

Tabela 4.1: Taxa de acertos de cada kernel do algoritmo SVM para o conjunto de dados de áudios de rolos.

Kernel	Taxa de acertos
Linear	95,57 % \pm 0,54%
Radial	84,37 % \pm 0,90%
Polinomial	81,73 % \pm 0,77%
Sigmoide	57,23 % \pm 1,30%

Na tabela 4.2 é apresentado o resultado final dos testes para as configurações do RF. O resultado sugere que alterando o número de árvores, a taxa de acertos teve pouca alteração, sendo que nesse caso, o aumento de árvores diminuiu a média de exatidão. A alteração da profundidade das árvores mostrou que a redução deste parâmetro diminuiu significativamente a precisão do sistema e o aumento teve uma melhora de mais de 4%. Com base nestas informações, optou-se por usar o RF mantendo o número de árvores em 80 e uma profundidade igual a 7, resultado que teve melhor taxa de acertos.

Tabela 4.2: Taxa de acertos de cada configuração do algoritmo RF para o conjunto de dados de áudios de rolos.

Número de árvores	Profundidade	Taxa de acertos
80	3	86,13% \pm 0,78%
100	3	86,03% \pm 0,77%
120	3	86,07% \pm 0,76%
80	5	92,63% \pm 0,94%
100	5	93,17% \pm 0,83%
120	5	93,07% \pm 0,67%
80	7	97,87% \pm 0,31%
100	7	97,57% \pm 0,40%
120	7	97,27% \pm 0,50%

Na tabela 4.3 é mostrado os resultados das médias da taxa de acertos das configurações escolhidas de cada algoritmo. Os 3 métodos chegaram a soluções acima de 95%, o KNN teve melhor desempenho seguido do RF. O SVM teve uma taxa de acertos alta mas teve desempenho pior que os outros dois algoritmos.

Tabela 4.3: Média da taxa de acertos dos algoritmos escolhidos.

Algoritmos	Taxa de acertos
KNN	99,1 % \pm 0,28%
SVM	95,57 % \pm 0,54%
RF	97,87 % \pm 0,31%

4.3. Matriz de confusão e precisão

Além da taxa de acertos, em sistemas de classificação binária é interessante analisar também outras medidas de desempenho. Estas medidas apontam fórmulas que analisam os acertos, tanto a quantidade de acertos para valores esperados de serem verdadeiros (VP, verdadeiro positivo) quanto valores esperados de serem falsos (VN, verdadeiro negativo), e as taxas de erros, para erros onde o resultado esperado era verdadeiro e foi encontrado falso (FN, verdadeiro negativo) e onde o resultado esperado era negativo e algoritmo classificou como positivo (FP, falso positivo). Estes valores são apresentados em uma matriz chamada de matriz de confusão.

		Detectada	
		Sim	Não
Real	Sim	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (VN)

Figura 4.1: Matriz confusão.
(RODRIGUES)

Em uma matriz de confusão as linhas apresentam os valores reais dos alvos dos algoritmos e as colunas representam os valores preditos por ele (figura 4.1). A primeira coluna da primeira linha representa o VP, quantidade de amostras onde o resultado previsto esperado era verdadeiro e assim foi classificado, e a segunda coluna representa o FN, quantidade de amostras esperadas de serem positiva e foram classificadas como negativas. A primeira coluna da segunda linha representa o FP, quantidade de amostras onde o esperado era negativo e foi classificado como positivo, e a segunda coluna representa o VN, quantidade de amostras esperadas de serem negativa e foram corretamente classificadas.

Na figura 4.2 é apresentada a matriz confusão dos 3 algoritmos escolhidos para a base de dados dos rolos. Nesta situação se nota que os resultados são bem semelhantes, mas esta é uma representação situacional pois estes resultados são de um único teste. Para resultados consistentes é necessário analisar a média dos resultados de mais de um teste.

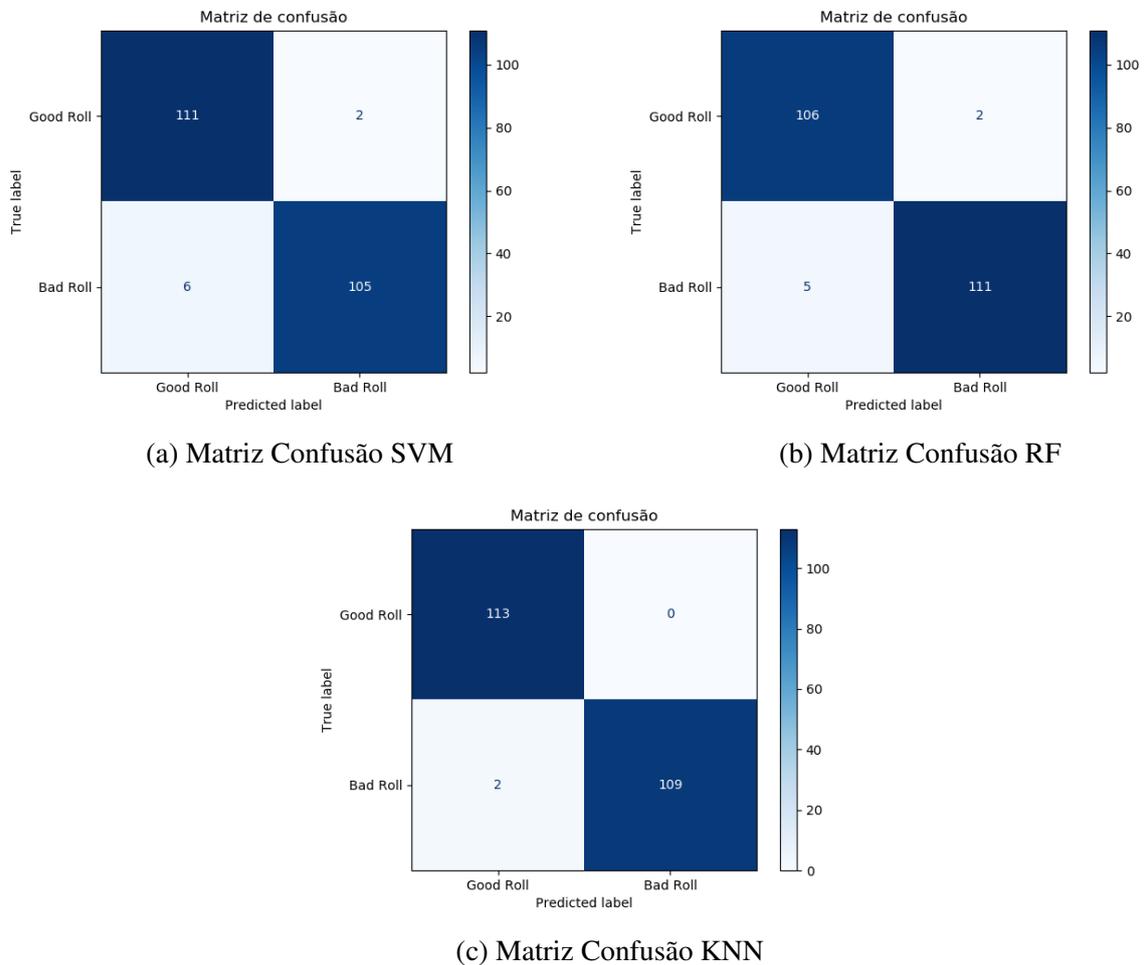


Figura 4.2: Matrizes de confusão dos algoritmos a serem comparados.

As medidas de desempenho são equações que usam os valores da matriz de confusão para fazer análises e comparações. São elas: precisão, revocação e acurácia. Dentre as medidas de desempenho as que apresentam propostas de análises relevantes para o nosso conjunto de dados são a acurácia, que já foi analisada, e a precisão. A precisão é representada por uma fórmula que leva em consideração os VP e os FP. No caso do sistema de classificação de rolos, os valores FP são os mais preocupantes. Ele representa a situação de análise onde rolos definidos como defeituosos foram classificados pelo sistema como rolos bons, podendo gerar avarias no transportar. Esta é uma situação de erro mais preocupante do que os FN, neste caso o rolo definido como bom foi classificado como defeituoso, desse modo quando a troca for ser realizada é possível perceber que foi uma classificação indevida, não gerando dano ao sistema.

$$Precisão = \frac{VP}{VP + FP}$$

Para comparar a precisão dos algoritmos escolhidos, foi seguido o procedimento de analisar a média da precisão de cada algoritmo 30 vezes para a base de dados dos rolos. Em cada teste foram guardados os valores de VP e FP e calculado a precisão de cada situação. Na tabela 4.4 é apresentado a média das precisões, sendo que os 3 métodos tiveram ótimos resultados. O KNN teve a melhor precisão, ou seja, apresentou um menor número de falsos positivos. O SVM apresentou uma precisão menor que o RF, mas a diferença entre os dois foi menor que a apresentada na comparação da taxa de acertos.

Tabela 4.4: Média das precisões de cada algoritmo.

Algoritmo	Precisão
KNN	98,30 % ±0,41%
SVM	94,60 % ±0,83%
RF	96,77 % ±0,31%

4.4. ROC e AUC

Outro modo de avaliação dos métodos de classificação é a curva *Receiver Operating Characteristic* (ROC) e a *Area Under Curve* (AUC). O gráfico ROC é uma técnica para visualização, organização e seleção de classificadores baseados na sua performance. O ROC é um gráfico bi-dimensional que permite uma melhor visualização da multidimensionalidade do problema a ser avaliado. O gráfico ROC é baseado na probabilidade de detecção, ou taxa VP (eixo Y), e na probabilidade de falsos alarmes, ou taxa FP (eixo X) (FAWCETT, 2006).

$$taxaVP = \frac{VerdadeirosPositivos}{TotaldePositivos}$$

$$taxaFP = \frac{FalsosNegativos}{TotaldeNegativos}$$

No gráfico ROC, figura 4.3, alguns pontos são importantes para o seu entendimento. O ponto mais baixo do lado esquerdo (0,0) representa a estratégia de não assumir nenhuma classificação como positiva. Assim o classificador não comete nenhum erro de FP, mas também não apresenta nenhum VP. A estratégia oposta, de assumir todas classificações como positiva, é representada pelo ponto (1,1). O ponto (0,1) representa uma classificação perfeita (FAWCETT, 2006).

Para avaliar os pontos dentro do gráfico ROC pode-se dizer que pontos que estão acima e à esquerda são melhores, ou seja, maior taxa VP e menor taxa FP. Classificadores que aparecem abaixo e à esquerda do gráfico ROC podem ser considerados "conservadores", eles são

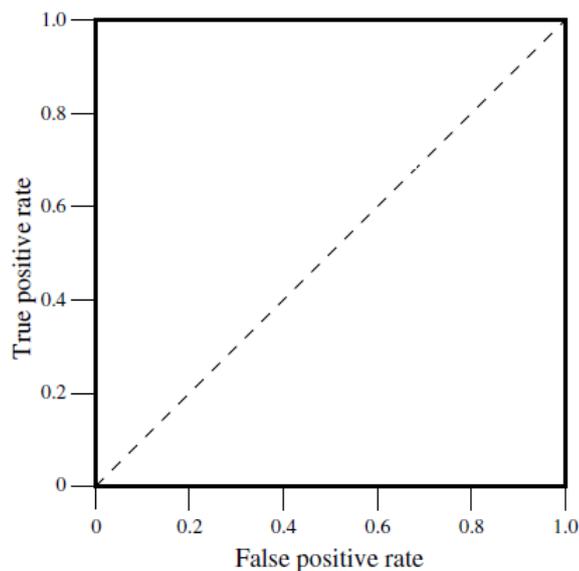


Figura 4.3: Gráfico ROC básico.
(FAWCETT, 2006)

mais rigorosos para fazer uma classificação positiva, com isso a taxa FP é baixa, porém a taxa VP tende a abaixar também. Classificadores que aparecem acima e à direita do gráfico são considerados "liberais"; eles são menos rigorosos para fazer uma classificação positiva, logo a taxa VP é alta, porém as taxas de FP poderão aumentar também (FAWCETT, 2006).

No caso deste problema procura-se classificadores com altas taxas de VP e baixas de FP. Mas para casos semelhantes os classificadores que forem mais conservadores serão melhores pois diminuem a chance de algum rolo defeituoso ser classificado como um rolo bom e continuar na planta.

AUC é uma medida de performance para problemas de classificação em várias configurações de limites. ROC é a curva de probabilidade e AUC representa o grau ou medida de separabilidade. Isto diz o quanto o modelo é capaz de distinguir entre classes. Quanto maior o valor de AUC, melhor será o modelo em prever os VP e FP. A AUC varia de 0 a 1. Um modelo excelente tem AUC próximo de 1, que implica que o modelo tem boa separabilidade. Um modelo com AUC próximo de 0,5 significa que ele tem nenhuma separabilidade. Quando o valor de AUC se aproxima de 0 o modelo está prevendo todos valores invertidos (NARKHEDE, 2018).

Na figura 4.4 é ilustrado o gráfico ROC dos 3 algoritmos testados para a base de dados dos áudios de rolos. A posição da curva ROC mostra que o modelo KNN está à esquerda e acima dos demais, assim como RF está do SVM. Isto mostra que ele teve uma taxa de VP maior e FP menor que os outros. A AUC dos 3 algoritmos apresentou valores próximos de 1 mostrando que todos tem um grau de separabilidade alto, mas o KNN foi melhor para distinguir a qualidade dos rolos.

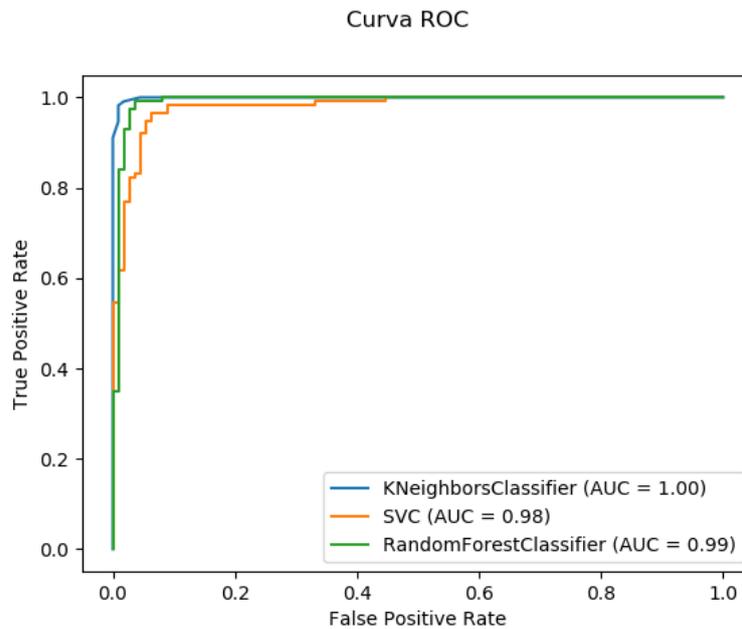


Figura 4.4: Gráfico ROC para os algoritmos testados.

4.5. Teste dos algoritmos para amostras restantes dos rolos

O total de amostras de áudios de rolos contavam com uma base de dados de 559 amostras de rolos bons e 1106 de rolos defeituosos. Para fazer os treinamentos e testes dos algoritmos foram usadas todas as amostras de rolos bons e a mesma quantidade foi usada dos rolos ruins para garantir o balanceamento da base de dados. Com isso, sobraram quase 500 amostras de rolos que não participaram do processo. Foi proposto utilizar estes dados restantes para avaliar os mesmos 3 algoritmos.

Tabela 4.5: Média da taxa de acertos dos algoritmos para o novo conjunto de dados.

Algoritmos	Taxa de acertos
KNN	97,80 % $\pm 0,30\%$
SVM	90,67 % $\pm 0,43\%$
RF	94,37 % $\pm 0,30\%$

Na tabela 4.5 é visto que o método KNN se manteve com a melhor acurácia dos 3 modelos, mesmo com sua taxa de acertos diminuindo. Todos os modelos tiveram queda em seu aproveitamento, mas mantiveram a taxa acertos acima de 90%. Como só foram avaliadas amostras reais de rolos ruins a matriz confusão teve apenas valores de VN e FP. Com isso, não faz sentido analisar a precisão dos sistemas. A taxa de acertos mostra os rolos defeituosos que foram corretamente classificados e todos os erros correspondem a rolos que foram erroneamente classificados como rolos bons, fato indesejado no nosso sistema. Também não é possível fazer uma análise da curva ROC porque o eixo Y representa a taxa de VP e este teste só gerou valores de VN e FP.

4.6. Classificação para base de dados com apenas uma amostra de cada rolo

Esta foi outra análise feita com os áudios coletados. Nesta análise apenas uma amostra por rolo foi selecionada para treinar um algoritmo para fazer classificação dos rolos. Neste caso, a base de dados ficou com um total de 19 amostras de rolos bons e 38 de rolos ruins. Mantendo a ideia de utilizar uma base de dados balanceada, a base de dados final ficou com um total de 38 amostras, 19 de rolos bons e 19 de rolos defeituosos.

Para fazer a classificação foi utilizado o algoritmo Random Forest com 100 árvores e profundidade 7. Foi separado 80% da base de dados para treinamento do algoritmo e 20% para teste do mesmo. Para uma análise mais completa desta situação, foram separados 3 conjuntos de diferentes amostras de cada rolo, ou seja, 3 bases de dados. O algoritmo foi treinado e testado 30 vezes e a média da taxa de acertos de cada base de dados é apresentada na tabela 4.6.

Tabela 4.6: Média da taxa de acertos de cada base de dados de amostras individuais dos rolos.

Base de dados	Taxa de acertos
BD1	79,93 % \pm 3,79%
BD2	82,87 % \pm 4,96%
BD3	84,37 % \pm 4,56%

A taxa de acertos para esta situação apresentou resultados próximos de 80%. Com um número menor de amostras a taxa de acertos do algoritmo ficou menor em comparação a análise com a base de dados com várias amostras de cada rolo, mas mesmo assim apresentou uma boa acurácia.

5. Conclusão e trabalhos futuros

O estudo nos mostrou a importância que a correia transportadora tem em empresas, principalmente nas de grande porte. A identificação precoce dos rolos defeituosos evita danos nos componentes do TC. Um TC em manutenção pode parar e atrasar a produção de uma mina inteira. A proposta apresentada obteve resultados expressivos na detecção dos rolos defeituosos. O uso da análise de sons dos rolos, mesmo num ambiente tão ruidoso, se mostrou bastante eficiente.

Os 3 algoritmos foram analisados de diferentes formas e em todas o algoritmo KNN se mostrou melhor para a base de dados dos áudios de rolos. Apresentou uma melhor acurácia, precisão e apresentou melhor grau de separabilidade do que os algoritmos SVM e RF. Apesar disso, os 3 algoritmos apresentaram bons resultados, com taxas acima de 90% em todos os critérios de comparação.

O teste feito com as amostras de rolos defeituosos que não foram utilizadas no treinamento apresentou resultado compatível ao do treinamento. O KNN se manteve na frente dos demais, contudo todos tiveram aproveitamento maior que 90%.

Além do artigo aceito e em fase de correção no automining, apêndice A, este trabalho contribuiu com informações sobre a escolha do microfone ideal para coleta de sinais sonoros em ambientes ruidosos. O uso do microfone ideal foi de grande importância para os resultados positivos deste trabalho. Conseguir bons áudios de rolos em um ambiente tão ruidoso como o próximo de um TC não é uma tarefa simples. Outra contribuição foi o uso da normalização dos sinais sonoros. Foi notado que a variação da intensidade de gravação interfere bastante quando se analisa diferenças de sinais no domínio da frequência. Normalizando o sinal inicial do áudio garante-se que todas as amostras terão a mesma variação de amplitude no domínio tempo. Com isso, ao se analisar o sinal no domínio da frequência será levado em consideração apenas as diferenças das frequências de cada rolo. A normalização evita que diferenças causadas por gravações feitas em lugares mais ruidosos que outros ou áudios gravados com diferentes intensidades de captura do microfone interfiram na análise no domínio da frequência.

O estudo do som se mostrou uma ferramenta muito poderosa para detecção de falhas. Como opção de trabalhos futuros pode-se também montar uma base de dados com sons das diferentes falhas que um rolo pode apresentar. Assim, além de detectar a falha será possível classificá-la. Esta informação pode ser útil para controle das falhas.

Outra opção de trabalho futuro seria a possibilidade de identificar um rolo defeituoso com antecedência, ou até estipular o tempo de vida deles. Para isto o estudo teria que analisar os sons de rolos por todo o seu período de vida, desde instalação do rolo novo até o período de falha. Nas análises desses dados, poderiam ser identificadas quais frequências foram alteradas e o período que elas aconteceram. Com isso, poderia ser montado um banco de dados da vida do rolo. Assim, quando um rolo fosse analisado poderia se saber, pelas frequências dele, em qual período da vida ele se encontra.

Como opção de trabalho futuro também se encaixa a integração deste sistema ao projeto ROSI, (GARCIA *et al.*, 2019). Um dispositivo robótico terrestre com a função de inspecionar as estruturas dos transportadores de correias. O robô se encontra em fase de testes e a integração desse sistema de detecção de falhas através do som auxiliaria na inspeção de rolos defeituosos.

Referências Bibliográficas

- ABDOLI, S., CARDINAL, P., KOERICH, A. L. “End-to-end environmental sound classification using a 1D convolutional neural network”, *Expert Systems with Applications*, v. 136, pp. 252–263, 2019.
- AKIYAMA, D., IMOTO, K., TONAMI, N., et al.. “Sound Event Detection Using Duration Robust Loss Function”, *arXiv preprint arXiv:2006.15253*, 2020.
- ALPAYDIN, E. *Introduction to machine learning*. MIT press, 2009.
- AUDACITY. “Audacity”. Disponível em: <<https://www.audacityteam.org/>>.
- BAE, S. H., CHOI, I., KIM, N. S. “Acoustic scene classification using parallel combination of LSTM and CNN”. Em: *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2016 Workshop (DCASE2016)*, pp. 11–15, 2016.
- BARCHIESI, D., GIANNOULIS, D., STOWELL, D., et al.. “Acoustic scene classification: Classifying environments from the sounds they produce”, *IEEE Signal Processing Magazine*, v. 32, n. 3, pp. 16–34, 2015.
- BIAU, G., SCORNET, E. “A random forest guided tour”, *Test*, v. 25, n. 2, pp. 197–227, 2016.
- BREIMAN, L. “Random forests”, *Machine learning*, v. 45, n. 1, pp. 5–32, 2001.
- CARVALHO JÚNIOR, J. R. D. *Processamento digital de imagens para a identificação automática de falhas em rolos dos transportadores de correias*. Tese de Mestrado, Universidade Federal de Ouro Preto, 2018.
- CHU, S., NARAYANAN, S., KUO, C.-C. J., et al.. “Where am I? Scene recognition for mobile robots using audio features”. Em: *2006 IEEE International conference on multimedia and expo*, pp. 885–888. IEEE, 2006.
- CORTES, C., VAPNIK, V. “Support-vector networks”, *Machine learning*, v. 20, n. 3, pp. 273–297, 1995.

- CUNNINGHAM, S., RIDLEY, H., WEINEL, J., et al.. “Supervised machine learning for audio emotion recognition: Enhancing film sound design using audio features, regression models and artificial neural networks”, *Personal and Ubiquitous Computing*, pp. 1–14, 2020.
- DE SÁ, A. G., PINTO, W. J. G., OLIVEIRA, L. O. V., et al.. “RECIPE: a grammar-based framework for automatically evolving classification pipelines”. Em: *European Conference on Genetic Programming*, pp. 246–261. Springer, 2017.
- ERICEIRA, D. R. *Detecção automática de defeitos em rolos de transportadores de correia utilizando sensoriamento Ultrassônico*. Tese de Mestrado, Universidade Federal de Ouro Preto, Ouro preto, 2019.
- FAWCETT, T. “An introduction to ROC analysis”, *Pattern recognition letters*, v. 27, n. 8, pp. 861–874, 2006.
- FIX, E. *Discriminatory analysis: nonparametric discrimination, consistency properties*. USAF School of Aviation Medicine, 1951.
- FREITAG, R. M. K. *Metodologia de coleta e manipulação de dados em sociolinguística*. Editora Blucher, 2014.
- GARCIA, G., ROCHA, F., TORRE, M., et al.. “ROSI: A Novel Robotic Method for Belt Conveyor Structures Inspection”. Em: *2019 19th International Conference on Advanced Robotics (ICAR)*, pp. 326–331. IEEE, 2019.
- GOLDBERGER, J., HINTON, G. E., ROWEIS, S. T., et al.. “Neighbourhood components analysis”. Em: *Advances in neural information processing systems*, pp. 513–520, 2005.
- HAN, Y., LEE, K. “Convolutional neural network with multiple-width frequency-delta data augmentation for acoustic scene classification”, *IEEE AASP Challenge on Detection and Classification of Acoustic Scenes and Events*, 2016.
- HOFMANN, M. “Support vector machines-kernels and the kernel trick”, *Notes*, v. 26, n. 3, 2006.
- HUZAIFAH, M. “Comparison of Time-Frequency Representations for Environmental Sound Classification using Convolutional Neural Networks”, *arXiv*, pp. arXiv–1706, 2017.
- JAKKULA, V. “Tutorial on support vector machine (svm)”, *School of EECS, Washington State University*, v. 37, 2006.

- LI, W., WANG, Z., ZHU, Z., et al.. “Design of online monitoring and fault diagnosis system for belt conveyors based on wavelet packet decomposition and support vector machine”, *Advances in Mechanical Engineering*, v. 5, pp. 797183, 2013.
- LODEWIJKS, G., LI, W., PANG, Y., et al.. “An application of the IoT in belt conveyor systems”. Em: *International Conference on Internet and Distributed Computing Systems*, pp. 340–351. Springer, 2016.
- LUCAS, J., THABET, W., WORLIKAR, P. “Using virtual reality (VR) to improve conveyor belt safety in surface mining”. Em: *24th W78 Conference Maribor 2007 & 5th IT-CEDU Workshop & 14th EG-ICE Workshop: Bringing ITC knowledge to work*, pp. 431–438, 2007.
- MATLAB. “Matlab”. Disponível em: <<https://www.mathworks.com/products/matlab.html>>.
- MAVROFORAKIS, M. E., THEODORIDIS, S. “A geometric approach to support vector machine (SVM) classification”, *IEEE transactions on neural networks*, v. 17, n. 3, pp. 671–682, 2006.
- MONARD, M. C., BARANAUSKAS, J. A. “Conceitos sobre aprendizado de máquina”, *Sistemas inteligentes-Fundamentos e aplicações*, v. 1, n. 1, pp. 32, 2003.
- NARKHEDE, S. “Understanding AUC-ROC Curve”, *Towards Data Science*, v. 26, 2018.
- NASCIMENTO, R., CARVALHO, R., DELABRIDA, S. E., et al.. “An Integrated Inspection System for Belt Conveyor Rollers-Advancing in an Enterprise Architecture.” Em: *ICEIS (2)*, pp. 190–200, 2017.
- NASCIMENTO, R. S. D. *Inspeção de transportadores de correia: arquitetura integrada para uma plataforma de inspeção com uso de VANTs*. Tese de Mestrado, Universidade Federal de Ouro Preto, 2018.
- OLSON, R. S., MOORE, J. H. “TPOT: A tree-based pipeline optimization tool for automating machine learning”. Em: *Workshop on automatic machine learning*, pp. 66–74, 2016.
- PANG, Y. *Intelligent belt conveyor monitoring and control*. Tese de Mestrado, Delft University of Technology, 2010.
- PYTHON. “Python”. Disponível em: <<https://www.python.org/>>.
- QIAN, K., REN, Z., PANDIT, V., et al.. “Wavelets revisited for the classification of acoustic scenes”. Em: *Proc. DCASE Workshop, Munich, Germany*, pp. 108–112, 2017.

- REICKS, A. V. “Belt conveyor idler roll behaviors”, *Bulk material handling by conveyor belt*, v. 7, n. 11, pp. 35–40, 2008.
- RIJO, S. M. A. *Técnicas de deep learning para detecção de eventos em áudio: treino de modelos acústicos a partir de sinais puros*. Tese de Mestrado, Universidade de Évora, 2018.
- RODRIGUES, V. “Métricas de Avaliação: acurácia, precisão, recall... quais as diferenças?” Disponível em: <<https://medium.com/@vitorborbarodrigues/m%C3%A9tricas-de-avalia%C3%A7%C3%A3o-acur%C3%A1cia-precis%C3%A3o-recall-quais-as-diferen%C3%A7as-c8f05e0a513c>>.
- SUGUNA, N., THANUSHKODI, K. “An improved k-nearest neighbor classification using genetic algorithm”, *International Journal of Computer Science Issues*, v. 7, n. 2, pp. 18–21, 2010.
- WEBPORTOS. “Porto Tubarão”. Disponível em: <<https://webportos.labtrans.ufsc.br/Tup/Index/105>>.
- WEIPING, Z., JIANTAO, Y., XIAOTAO, X., et al.. “Acoustic scene classification using deep convolutional neural network and multiple spectrograms fusion”. Em: *Detection and Classification of Acoustic Scenes and Events 2017 Workshop (DCASE2017)*, 2017.
- YANG, B. Y. *Fibre optic conveyor monitoring system*. Tese de Mestrado, The University of Queensland, 2014.
- YANG, W., ZHANG, X., MA, H. “An inspection robot using infrared thermography for belt conveyor”. Em: *2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pp. 400–404. IEEE, 2016.
- ZHU, B., XU, K., WANG, D., et al.. “Environmental sound classification based on multi-temporal resolution convolutional neural network combining with multi-level features”. Em: *Pacific Rim Conference on Multimedia*, pp. 528–537. Springer, 2018.

Apêndice A: Trabalhos Gerados

A seguinte produção foi gerada ao longo desta dissertação:

- Cruz, E.; Reis, A.; Guimarães, F.; Almeida, S. Conveyor Belt Roller Failure Detection System based on Machine Learning. In: 7th International Congress on Automation in Mining - Automining, Santiago, 2020.