



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CAMPUS ARARANGUÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIAS DA INFORMAÇÃO E
COMUNICAÇÃO

Heverton Marcos Costa

**Utilização de técnicas de Learning Analytics para Classificação e Detecção de Erros
Experimentais no Laboratório remoto VISIR.**

Araranguá
2022

Heverton Marcos Costa

**Utilização de técnicas de Learning Analytics para Classificação e Detecção de Erros
Experimentais no Laboratório remoto VISIR.**

Dissertação submetida ao Programa de Pós-Graduação em Tecnologias da Informação e Comunicação da Universidade Federal de Santa Catarina para a obtenção do título de mestre em Tecnologias da Informação e Comunicação.

Orientador: Prof. João Bosco da Mota Alves, Dr.

Coorientador: Prof. Juarez Bento da Silva, Dr.

Coorientador: Prof. Gustavo Ribeiro da Costa Alves, Dr.

Araranguá

2022

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Costa, Heverton Marcos

Utilização de técnicas de Learning Analytics para Classificação e Detecção de Erros Experimentais no Laboratório remoto VISIR. / Heverton Marcos Costa ; orientador, João Bosco da Mota Alves, coorientador, Juarez Bento da Silva, coorientador, Gustavo Ribeiro da Costa Alves, 2022.

74 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Campus Araranguá, Programa de Pós-Graduação em Tecnologias da Informação e Comunicação, Araranguá, 2022.

Inclui referências.

1. Tecnologias da Informação e Comunicação. 2. Learning Analytics. 3. Algoritmo de classificação. 4. VISIR. 5. Laboratório remoto. I. Alves, João Bosco da Mota. II. Silva, Juarez Bento da. III. Alves, Gustavo Ribeiro da Costa IV. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Tecnologias da Informação e Comunicação. V. Título.

Heverton Marcos Costa

Utilização de técnicas de Learning Analytics para Classificação e Detecção de Erros Experimentais no Laboratório remoto VISIR.

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof.^a. Simone Meister Sommer Bilessimo, Dr.^a.

Universidade Federal de Santa Catarina

Prof. Max Pereira, Dr.

Universidade Federal de Santa Catarina

Prof. Paulo Manoel Mafra, Dr.

Instituto Federal de Santa Catarina

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de mestre em Tecnologias da Informação e Comunicação.

Coordenação do Programa de Pós-Graduação

Prof. João Bosco da Mota Alves, Dr.

Orientador

Araranguá, 2022.

Este trabalho é dedicado à minha esposa Jaqueline, minha filha Luna, aos meus pais Francisco e Sandra, ao meu irmão Lyon e meu irmão Peterson.

AGRADECIMENTOS

Agradeço primeiramente a Deus pelo dom da vida. Agradeço aos meus orientadores, Bosco, Juarez e Gustavo, sou muito grato por todos os ensinamentos adquirido com eles. Agradeço a minha família por todo apoio durante o tempo de dedicação para conquistar meu objetivo.

“A mente que se abre a uma nova ideia jamais voltará ao seu tamanho original”.

(Albert Einstein)

RESUMO

A disseminação da tecnologia na educação cresceu consideravelmente nos últimos anos. Com ela, o ensino consegue atingir um maior alcance de usuários para o setor educacional. Ao mesmo tempo, surgem algumas dificuldades do acompanhamento do desempenho dos alunos. Assim, o processo que sistematicamente coleta e analisa grandes volumes de dados, com o objetivo de melhorar o processo de ensino-aprendizado é chamado de *Learning Analytics*. Baseada em processamento de dados, mineração de dados educacionais e visualização, assim consegue-se acompanhar os progressos acadêmicos, potencializando ações de como o professor deverá conduzir a disciplina. O objetivo dessa pesquisa está em minerar dados de experimentos realizados no laboratório remoto denominado *Virtual Instrument Systems in Reality* (VISIR). De acordo com algumas literaturas com tema de classificação de erros experimentais no VISIR, apresenta-se muito a questão teórica dos principais pontos onde erros podem ser cometidos pelo usuário. A literatura apresenta alguns modelos como circuitos denominados 1:1, que consistem em uma fonte de energia para um componente consumidor de energia. A pesquisa utiliza o arquivo de log dos experimentos, denominado na pesquisa como base de dados do Instituto Superior de Engenharia do Porto (ISEP), com duzentos e cinquenta e cinco mil, duzentos e trinta experimentos. A escolha para obter os dados do ISEP, está vinculado a toda sua expertise no assunto da pesquisa. Estuda-se o modelo de circuitos de 1:1, 0:1 e 1:0 adquiridos na base de dados (BD), esses modelos correspondem a 17,24% da base. Realiza-se um pré-processamento das informações (constituídas no arquivo de log dos experimentos) antes da execução dos algoritmos com objetivo de criar grupos de classificação conforme análise teórica estudada em análise de circuitos (modelos 1:1, 0:1 e 1:0). Utilizou-se o modelo de classificação *k-NN* e *tree* para realização do estudo. Os resultados obtidos pelos algoritmos resultaram em amostras de acurácia maior que 0.9. Considerando uma excelente forma de análise de classificação para os três modelos de circuito elétricos simples propostos.

Palavras-chave: *Learning Analytics*. Algoritmo de classificação. VISIR. Laboratório remoto. Circuito elétricos simples.

ABSTRACT

The spread of technology in education has grown considerably in recent years. With it, teaching can reach a greater reach of users for the educational sector. At the same time, some difficulties arise in monitoring student performance. Thus, the process that systematically collects and analyzes large volumes of data to improve the teaching-learning process is called Learning Analytics. Based on data processing, educational data mining and visualization, it is possible to monitor academic progress, enhancing actions on how the teacher should conduct the discipline. The objective of this research is to mine data from experiments carried out in the remote laboratory called Virtual Instrument Systems in Reality (VISIR). According to some literature on the classification of experimental errors in VISIR, the theoretical question of the main points where errors can be made by the user is presented. The literature presents some models as 1:1 circuit, which consist of a power source for an energy-consuming component. The research uses the log file of the experiments, named in the research as the Polytechnic of Porto School of Engineering (ISEP) database, with two hundred and fifty-five thousand, two hundred and thirty experiments. The choice to obtain ISEP data is linked to all its expertise in the research subject. The model of 1:1, 0:1 and 1:0 circuits acquired in the database (BD) is studied, these models correspond to 17.24% of the base. A pre-processing of the information (consisting of the log file of the experiments) is carried out before the execution of the algorithms in order to create classification groups according to theoretical analysis studied in circuit analysis (models 1:1, 0:1 and 1:0). The k-NN and tree classification model was used to carry out the study. The results obtained by the algorithms resulted in samples of accuracy greater than 0.9. Considering an excellent form of classification analysis for the three proposed simple electrical circuit models.

Keywords: Learning Analytics. Classification algorithm. VISIR. Remote lab. Simple electrical circuit.

LISTA DE FIGURAS

Figura 1 - Resultados da pesquisa no banco de dados do programa.	20
Figura 2 - Visão geral da arquitetura de um laboratório remoto.	22
Figura 3 – Laboratório on-line.....	23
Figura 4 – Protoboard e instrumentos disponíveis no VISIR.....	24
Figura 5 – Exemplo de erro manual.	26
Figura 6 – Exemplo de erro conceitual.....	26
Figura 7 - Plataforma RELLE com o acesso ao VISIR.....	28
Figura 8 - Modelo de referência de <i>Learning Analytics</i>	29
Figura 9 - Ciclos de <i>Learning Analytics</i>	30
Figura 10 – Processamento de descoberta de conhecimento em bases de dados.	31
Figura 11 – Impacto do valor de k no algoritmo k-NN.....	37
Figura 12 – Árvore de decisão para definir se vamos ou não esperar por uma mesa.....	39
Figura 13 - Fluxo da classificação dos experimentos.....	42
Figura 14 - Fluxograma das etapas da pesquisa.	43
Figura 15 – Ciclo LA com objetivo da pesquisa.	44
Figura 16 – Ligação da matriz utilizando a protoboard.	45
Figura 17 – Explicação das etapas de otimização de circuito (circuito medindo resistência)..	47
Figura 18 – Circuito equivalente entre laboratório remoto VISIR e laboratório convencional.	47
Figura 19 – Explicação das etapas de otimização de circuito (circuito medindo corrente).	48
Figura 20 – Circuito equivalente entre laboratório remoto VISIR e laboratório convencional.	49
Figura 21 – Modelo de execução do k-NN.....	50
Figura 22 - Classificação do experimento do k-NN.....	51
Figura 23 - Modelo de execução da <i>tree</i>	52
Figura 24 - Classificação do experimento da <i>tree</i>	53
Figura 25 – Apresentação da análise superficial na base de dados.	54
Figura 26 – Análise da amostra de dados entre informações trabalhadas, informações a serem trabalhadas (etapa futura), e dados desconsiderados no momento.....	54
Figura 27 – Tabela de classificação de experimentos 1:1, 1:0 e 0:1.	55
Figura 28 – Apresentação da classificação em forma de gráfico dos experimentos.	56
Figura 29 – Análise da proporção dos experimentos corretos.....	56

Figura 30 – Análise da classificação dos experimentos corretos e não corretos.....	57
Figura 31 - Gráfico demonstrativo de 100% de classificação.	58
Figura 32 - Dados classificados corretamente pelo modelo 19 do experimento quantitativo do algoritmo k-NN.	61
Figura 33 - Dados classificados incorretamente pelo modelo 19 do experimento quantitativo do algoritmo k-NN.	61
Figura 34 - Resultado da classificação geral pelo modelo 19 do experimento quantitativo do algoritmo k-NN	62
Figura 35 - Classificação correta dos dados do algoritmo <i>Tree</i>	63
Figura 36 - Classificação incorreta dos dados do algoritmo <i>Tree</i>	63
Figura 37 - Classificação mista dos dados do algoritmo <i>Tree</i>	64
Figura 38 - Resultados de análise do modelo qualitativo com algoritmo k-NN.	66
Figura 39 - Resultados de análise do modelo quantitativo com algoritmo k-NN.	66

LISTA DE QUADROS

Quadro 1 - Pseudocódigo do algoritmo.....	38
Quadro 2 - Algoritmo para construção de uma árvore de decisão.	40
Quadro 3 - Classificação da pesquisa.	41
Quadro 4 - Dados do circuito com <i>tag circuitlist</i>	46
Quadro 5 - Dados do circuito com <i>tag circuitlist</i>	48
Quadro 6 - Pseudocódigo para análises.	49

LISTA DE TABELAS

Tabela 1 - Possíveis erros ao medir a tensão, corrente elétrica e resistência em cenário de circuito 1:1	27
Tabela 2 - Nomenclatura dos componentes do circuito.	46
Tabela 3 - Nomenclatura dos componentes do circuito.	48
Tabela 4 - Análise dos testes do algoritmo k-NN com dados qualitativos.....	59
Tabela 5 - Análise dos testes do algoritmo k-NN com dados quantitativos.....	60
Tabela 6 - Análise do desempenho de classificação do algoritmo <i>Tree</i>	62
Tabela 7 - Análise de nós e folhas pelo algoritmo <i>Tree</i>	62
Tabela 8 - Dados em curto-circuito.	65

LISTA DE ABREVIATURAS E SIGLAS

AVA	Ambientes Virtuais de Aprendizagem
AM	Aprendizagem de máquina
BD	Base de Dados
CRISP-DM	<i>Cross-Industry Process for Data Mining</i>
IoT	<i>Internet of Things</i>
ISEP	Instituto Superior de Engenharia do Porto
KDD	<i>Knowledge Discovery in Database</i>
k-NN	k – Nearest Neighbors
LA	<i>Learning Analytics</i>
MD	Mineração de Dados
PPGTIC	Programa de Pós-Graduação em Tecnologias da Informação e Comunicação
RExLab	<i>Remote Experimentation Lab</i>
ROC	<i>Receiver-operating curve</i>
STEAM	<i>Science, Technology, Engineering, Arts and Mathematics</i>
STEM	<i>Science, Technology, Engineering and Mathematics</i>
STI	Sistemas Tutores Inteligentes
TI	Tecnologia da informação
TIC	Tecnologias de Informação e Comunicação
UFSC	Universidade Federal de Santa Catarina
VISIR	<i>Virtual Instrument Systems in Reality</i>
RELLE	<i>Remote Labs Learning Environment</i>

SUMÁRIO

1	INTRODUÇÃO	15
1.1	CONTEXTUALIZAÇÃO E PROBLEMATIZAÇÃO.....	16
1.1.1	Pergunta de Pesquisa.....	17
1.2	JUSTIFICATIVA	18
1.3	OBJETIVOS.....	18
1.3.1	Objetivo Geral.....	19
1.3.2	Objetivos Específicos	19
1.4	INTERDISCIPLINARIDADE E ADERÊNCIA AO PROGRAMA DE PÓS-GRADUAÇÃO.....	19
1.5	ORGANIZAÇÃO DA DISSERTAÇÃO.....	20
2	REFERENCIAL TEÓRICO	21
2.1	LABORATÓRIOS REMOTOS	21
2.2	LABORATÓRIO REMOTO VISIR	24
2.2.1	Classificação de erros experimentais no VISIR.....	25
2.2.2	Classificação de erros no cenário de 1:1	26
2.3	REXLAB COM PROJETO VISIR.....	28
2.4	LEARNING ANALYTICS	29
2.4.1	Mineração de Dados	30
<i>2.4.1.1</i>	<i>Mineração de Dados Educacionais.....</i>	<i>31</i>
<i>2.4.1.2</i>	<i>Descoberta de conhecimento em bases de dados (KDD).....</i>	<i>31</i>
<i>2.4.1.3</i>	<i>Etapa de pré-processamento</i>	<i>32</i>
<i>2.4.1.3.1</i>	<i>Limpeza dos dados</i>	<i>32</i>
<i>2.4.1.3.2</i>	<i>Transformação dos Dados</i>	<i>33</i>
<i>2.4.1.3.3</i>	<i>Conversão Simbólico-Numérico</i>	<i>33</i>
<i>2.4.1.3.4</i>	<i>Conversão Numérico-Simbólica.....</i>	<i>33</i>

2.4.1.3.5	Transformação de Atributos Numéricos.....	33
2.4.1.4	<i>Mineração</i>	34
2.4.1.4.1	Algoritmo de Classificação.....	34
2.4.1.4.2	k – nearest neighbors	36
2.4.1.4.3	Tree	38
3	PROCEDIMENTOS METODOLÓGICOS	41
3.1	CLASSIFICAÇÃO DA PESQUISA	41
3.2	OBTENÇÃO DOS DADOS.....	42
3.3	DELIMITAÇÃO DA PESQUISA.....	42
3.4	ETAPAS DA PESQUISA	42
4	MODELO PROPOSTO	44
4.1	PRÉ-PROCESSAMENTO DOS DADOS	45
4.1.1	Regra de tratamento	45
4.2	Análise dos Algoritmos	50
4.2.1	Algoritmo k-NN	50
4.2.2	Análise dos dados com <i>Tree</i>	51
5	Análise dos resultados	54
5.1	ANÁLISE PRELIMINAR DA BASE.....	54
5.2	ANÁLISE DOS RESULTADOS COM ALGORITMO	57
5.2.1	Análise dos Dados com k-NN	58
5.2.2	Análise dos dados com <i>Tree</i>	62
6	DISCUSSÃO DOS RESULTADOS	65
7	CONCLUSÃO	67
	REFERÊNCIAS	69

1 INTRODUÇÃO

As tecnologias digitais evoluem constantemente e com elas diversas possibilidades se constroem ao longo do tempo. Para agilizar a vida humana, diversos equipamentos eletrônicos fazem parte do cotidiano. E com eles nos conectamos ao mundo digital onde um grande volume de dados e informações são trocadas. O comércio eletrônico cresceu 83,68% em 2020 na comparação com o ano anterior segundo o índice desenvolvido pelo Comitê de Métricas da Câmara Brasileira da Economia Digital (câmara-e.net) (FORBES, 2021).

Com o avanço de conectividade entre dispositivos, inicia-se o conceito de *Internet of Things* (IoT). O nascimento da IoT iniciou-se entre 2008 e 2009, quando o número de dispositivos conectados ultrapassou o número de pessoas. No ano de 2010 a população mundial era de 6,8 bilhões para 12,5 bilhões de dispositivos conectados (EVANS, 2011).

Dentro de tantas possibilidades as tecnologias de informação e comunicação (TIC) vem crescendo expressivamente nos últimos anos. Com o surgimento da pandemia de COVID-19 diversas escolas com modelos tradicionais de ensino, tiveram que se adaptar a uma nova realidade e entrar de vez para era digital.

A tecnologia da informação (TI) de acordo com uma publicação na FORBES em 2021, é um dos caminhos para resolver uma das maiores necessidades sociais do mundo, a aprendizagem, pois as experiências que constroem conhecimentos e habilidades devem estar acessíveis para todas as pessoas e em todas as fases de suas vidas (RODRIGUES, 2021).

Em um contexto geral de ensino, principalmente no ensino de engenharia, torna-se fundamental o desenvolvimento de competências experimentais (GONÇALVES *et-al.*, 2018).

Como um novo conceito de ensino mediado pelas tecnologias digitais e pela conectividade podem entrar no novo conceito de ensino? Para isso existem alguns modelos de laboratórios, entre os mais conhecidos estão os laboratórios online (remotos e virtuais). O estudo em questão utilizou o modelo de experimento remotos.

A Universidade Federal de Santa Catarina (UFSC) integra um núcleo de pesquisa conhecido como RExLab onde foi fundado em 1997, objetivando difundir ciência e tecnologia. Em seu ambiente existem diversos experimentos remotos para diversas áreas de estudo. Em sua estrutura encontra-se instalado o laboratório remoto *Virtual Instrument Systems in Reality* (VISIR) (SILVA, 2019). O estudo da pesquisa utiliza a base de dados do laboratório VISIR.

O laboratório remoto VISIR foi desenvolvido pelo professor Ingvar em *Blekinge Institute of Technology* (BTH), Suécia. É um laboratório para experimentos com circuitos

elétricos e eletrônicos, considerado o laboratório mais estudado em termos de abordagens didáticas (GARCÍA-ZUBÍA *et-al.*, 2019) (SILVA *et-al.*, 2021) (GARCÍA-ZUBÍA *et-al.*, 2021) COSTA-LOBO *et-al.*, 2011) (SALAH *et-al.*, 2016) (PEREIRA *et-al.*, 2017).

Os estudos acadêmicos, principalmente em circuitos elétricos e eletrônicos consistem inicialmente em compreender conceitos teóricos para posteriormente a prática. Adota-se dois modelos de prática: laboratórios físicos podem ser chamados de laboratório tradicional e o laboratório remoto denominado como não tradicional.

O estudo nas áreas *Science, Technology, Engineering and Mathematics* (STEM) durante o ciclo acadêmico, com o propósito de torná-lo capaz de aplicar seus conhecimentos em resoluções de problemas para facilitar a vida humana (ALEXANDER; SADIKU, 2013).

Os alunos quando estão em contatos com as disciplinas práticas, costumam cometer erros manuais e conceituais (GARCÍA-ZUBÍA *et-al.*, 2019). Quando se trata de circuitos simples, se torna fácil a identificação dos erros, agora quando se tem circuitos complexos, se torna mais trabalhoso de identificá-los.

O trabalho desenvolvido buscará contribuir para o desenvolvimento futuro de uma plataforma com técnica de *Learning Analytics* aplicada no laboratório remoto VISIR, facilitando análise do desempenho do aluno.

1.1 CONTEXTUALIZAÇÃO E PROBLEMATIZAÇÃO

As organizações vêm gerando e armazenando uma grande quantidade de dados. Em 2020 a previsão era de 44 trilhões de gigabytes (44 zettabytes). As possibilidades trazidas pelas análises dessas informações nos últimos anos não imensas. (MINITAB, 2019).

As TIC estão em crescente desenvolvimento principalmente adotando medidas de estudo fora do sistema tradicional. De acordo com Carvalho (2016) após a explicação teórica em sala de aula, os modelos matemáticos que abrangem circuitos elétricos, normalmente são testados em laboratórios, podendo ser validada a relação entre teoria e prática.

O ensino direcionado a laboratórios pode seguir dois modelos, laboratório tradicional e laboratório não tradicional. O laboratório não tradicional conhecidos como laboratórios online e laboratório de instrumentação remota, são utilizados como possibilidade de aperfeiçoamento de ensino nas aulas (TAVARES *et-al.*, 2020).

A pesquisa em questão utilizou o laboratório remoto VISIR. Seu grande objetivo está na propagação do uso de laboratórios online, e cooperando com cursos *Science, Technology, Engineering, Arts and Mathematics* (STEAM).

O projeto VISIR abrange a área da Engenharia Elétrica e Eletrônica, com objetivo de aprendizagem práticas de circuitos. A disciplina de circuitos eletrônicos está construída em pilares: prático, virtual e remotos. O pilar denominado de remoto, está ligado diretamente ao laboratório remoto Virtual Instruments Systems in Reality (VISIR+, 2022).

Como em todo laboratório erros podem ser cometidos, para isso deve-se entender melhor a ocorrência dos erros. Segundo Tavares *et-al.* (2020) o laboratório VISIR possui alguns travamentos que protegem seu sistema de erros. Os erros podem ser manuais e conceituais, mas o que seria um erro?

Segundo Carvalho (2016), um erro é considerado quando seu valor obtido se diverge do resultado esperado. Existem erros manuais e conceituais, os erros manuais o VISIR consegue informar (esse procedimento deve ser configurado no ambiente remoto) ao usuário o procedimento incorreto, o erro conceitual vem ao encontro da falha entre teoria e prática.

O contexto da verificação dos erros pode facilitar muito na análise do professor sobre sua classe. Mas as análises de erros conceituais ainda não integram em um ambiente *Learning Analytics* (LA) para a observação do professor.

Pensando na dificuldade entre alunos e professores no uso do laboratório remoto VISIR, abre-se o estudo de análise dos experimentos realizados, por possuir informação da execução gravada em arquivo de log, no decorrer do trabalho será descrita como base de dados. Diversos estudos acadêmicos podem utilizar as informações, e o principal a este trabalho seria como identificar a classificação dos experimentos realizados no laboratório remoto VISIR utilizando técnicas de *Learning Analytics*.

1.1.1 Pergunta de Pesquisa

A partir do exposto nas seções que trataram da contextualização e problematização, foi elaborada a seguinte pergunta de pesquisa.

Como utilizar técnica de *Learning Analytics* para classificação de experimentos em circuitos elétricos e eletrônicos com utilização do laboratório remoto VISIR?

1.2 JUSTIFICATIVA

As instituições de ensino começam a observar o potencial dos dados armazenados. Com o surgimento da pandemia de COVID-19 muitas rotinas foram alteradas, e aperfeiçoadas. Laboratórios tradicionais ocorreram algumas restrições em seus acessos, e outros tipos de laboratórios começam a ganhar espaço. Os processos de integração das tecnologias administrativas e acadêmicas devem causar uma revolução educacional com demonstrativos de métricas eficientes de ensino-aprendizagem (DAHLSTROM *et-al.*, 2015).

Técnicas de LA conseguem apresentar dados de desempenho educacional com a utilização da base de dados (BD) de aprendizagem. A BD permite os sistemas realizarem a implementação de LA, seus resultados são importantes para analisar e informar dados sobre alunos (SIMÃO, 2018).

O processo de LA segue um ciclo de quatro etapas que são representados por: Alunos, Dados, Análise e Intervenções (CLOW, 2012).

Segundo Rienties, Cross e Grajek (2017), a técnica LA pode ser usada para avaliar a implementação de novas tecnologias de ensino e currículos, além de informações como frequência de acesso, grau de comprometimento entre outros. Pode-se obter o grau de dedicação nos estudos e como o estudante realiza suas atividades. Toda ação realizada em um ambiente de estudo pode-se gerar a reação do professor com a interpretação dos dados coletados.

De acordo com Avella *et-al.* (2016) os experimentos remotos conseguem fornecer uma imensa quantidade de dados, provenientes das interações dos estudantes durante as atividades. A etapa da execução dos experimentos remotamente chama-se de experimentação remota (BOHUS *et-al.*, 1996). Os dados gerados nos experimentos são de suma importância pelo objetivo de identificação de padrões de comportamentos, auxiliando na melhoria do processo de aprendizado.

1.3 OBJETIVOS

A partir do contexto inicial apresentado surgiu o tema da pesquisa. Onde encontra-se delimitado para apresentação dos objetivos.

1.3.1 Objetivo Geral

Utilizar técnica de *Learning Analytics* para análises de experimentos executados no laboratório remoto VISIR, com objetivo de explorar os experimentos e realizar sua classificação de acordo com análise de circuitos elétricos.

1.3.2 Objetivos Específicos

Com o objetivo geral definido derivam-se os objetivos específicos, que consistem para realização da conclusão do objetivo.

- Analisar padrões de classificação de erros experimentais no VISIR.
- Estudar a base de dados do VISIR para entender a melhor forma de tratar as informações no pré-processamento dos dados.
- Aplicar técnica de LA utilizando método de mineração de dados.
- Apresentar resultados quantitativos das classificações realizadas para validação do algoritmo testado.

1.4 INTERDISCIPLINARIDADE E ADERÊNCIA AO PROGRAMA DE PÓS-GRADUAÇÃO

O Programa de Pós-Graduação em Tecnologias da Informação e Comunicação (PPGTIC) da UFSC, encontra-se estruturado na área de concentração Tecnologia e inovação, com linhas de pesquisa em Tecnologia Educacional, Tecnologia Computacional e Gestão e Inovação (PPGTIC, 2022).

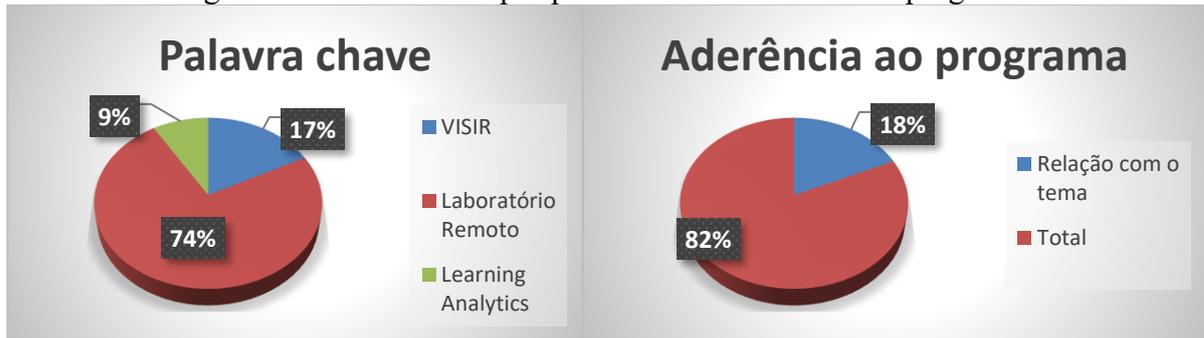
Segundo Yu e Zhou (2015 apud SILVA, 2019, p. 39), “a gestão do conhecimento e suas fases são sustentadas pelas tecnologias da informação e comunicação (TICs), pois as TICs são capazes de melhorar a eficiência dos recursos do conhecimento”.

O presente trabalho encontra-se classificado na linha de Tecnologia Computacional.

“O objetivo da linha é desenvolver modelos, técnicas e ferramentas computacionais auxiliando na resolução de problemas de natureza interdisciplinar. Especificamente, esta linha de pesquisa procura desenvolver novas tecnologias computacionais para aplicação nas áreas de educação e gestão.” (PPGTIC, 2022, on-line).

Consultando o repositório de dissertação do programa e pesquisando por três palavras-chave “VISIR, Laboratório Remoto e *Learning Analytics*”, foi obtido um resultado de vinte e duas referências, tendo uma aderência no programa de 18% de todos os trabalhos publicados até o momento. A Figura 1 apresenta o resultado obtido no repositório de dados do programa.

Figura 1 - Resultados da pesquisa no banco de dados do programa.



Fonte: Elaborada pelo autor.

1.5 ORGANIZAÇÃO DA DISSERTAÇÃO

A dissertação está organizada entre os seguintes tópicos:

- Referencial teórico;
- Procedimento metodológico;
- Modelo proposto;
- Análise dos resultados;
- Discussão dos resultados; e
- Conclusão.

Em referencial teórico, apresenta-se os principais conceitos da análise do estudo proposto. No procedimento metodológico se discute o tipo de metodologia de trabalho científico empregado na dissertação. A próxima etapa inicia o tema proposto, mas subjetivo pois apresenta-se um cenário desejável. Dando sequência aos estudos é o momento de falar da análise dos resultados obtidos na execução do algoritmo.

Após as análises é o momento de finalizar com discussão dos resultados e sua conclusão.

2 REFERENCIAL TEÓRICO

O mundo está mudando e com ele novas possibilidades surgem a cada momento. Este capítulo apresenta um contexto para apoiar o estudo baseado na literatura. Seu principal objetivo no apoio a pesquisa é fornecer um quadro de referências para interpretação dos resultados obtidos.

Sua estrutura está dividida em seções e subseções, iniciando com laboratório remoto (seção 2.1) seguindo com laboratório remoto VISIR (seção 2.2), apresentando RExLab com projeto VISIR (seção 2.3), e finalizando com LA (seção 2.4).

2.1 LABORATÓRIOS REMOTOS

Algumas disciplinas, principalmente as que integram o ensino de engenharia desenvolvem sua grade curricular composta de teoria e prática (FEISEL; PETERSON, 2002). O ensino a distância já adotou diversas tecnologias à medida que surgiam. Mas o problema do ensino em laboratório à distância não foi resolvido. Surge a Internet, onde conseguiu-se conectar as tecnologias quase instantaneamente, e o ensino em laboratórios remotos começou a ganhar espaço entre alunos e professores (FEISEL; ROSA, 2005).

Com o avanço tecnológico e sua conectividade ultrapassando barreiras, os laboratórios remotos começam a ganhar espaços dentro da sala de aula. Considerando que o assunto de laboratórios on-line ganhou espaço nas pesquisas educacionais a pouco tempo, sua aplicação começou a surgir por volta do ano de 1995, segundo o *Second Best to Being There (SBBT)*, na universidade de Oregon State recebeu o reconhecimento como o primeiro laboratório remoto no mundo a entrar em operação em 1995 (BOHUS *et-al.*, 1996).

A utilização da Internet foi fundamental para resolver o problema do ensino a distância. Com ela, alunos conseguiram acesso remoto a aparelhos físicos de laboratório. Geralmente os sistemas são síncronos, possibilitando a sensação real do experimento (FEISEL; ROSA, 2005).

O laboratório remoto é composto de sensores e atuadores, sua responsabilidade é executar ação do experimento e informar o resultado obtido (SIMÃO, 2018).

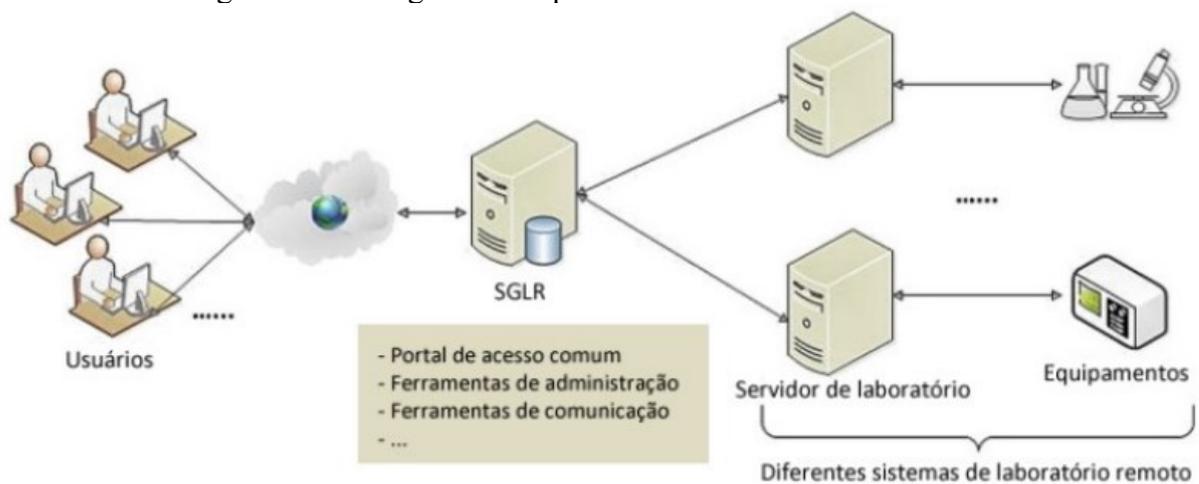
Os laboratórios de instrumentação remota, são conhecidos na literatura como laboratórios online ou laboratórios não tradicionais. Sua utilização torna-se uma grande alternativa para aulas práticas, aplicando os conceitos teóricos aprendidos em sala de aula.

Consiste em conduzir remotamente experimentos reais onde o usuário encontra-se fora do ambiente físico (TAVARES *et-al.*, 2020).

A interação do usuário com o laboratório remoto geralmente ocorre por interface web, mas outros meios vêm surgindo como desenvolvimentos de jogos, mundos virtuais e aplicativos para dispositivos móveis (SIMÃO, 2018) (ROCHADEL *et-al.*, 2013).

A Figura 2 apresenta a arquitetura de um laboratório remoto, sua composição pode ser dividida em quatro etapas: Usuários, portal de acesso, servidor de laboratório e equipamentos.

Figura 2 - Visão geral da arquitetura de um laboratório remoto.



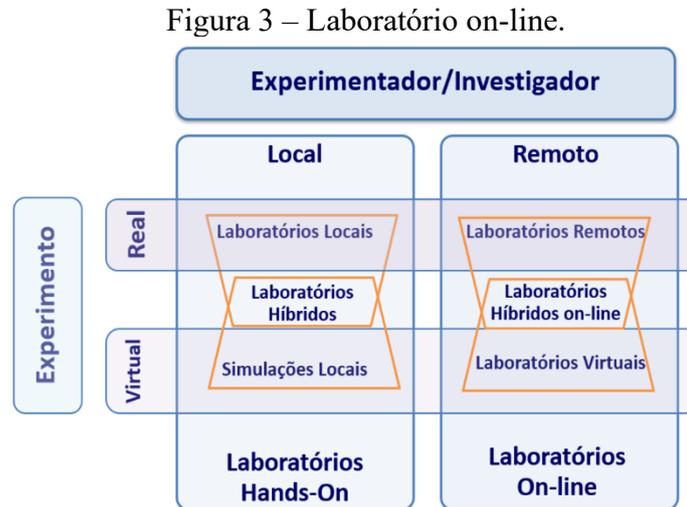
Fonte: Adaptado de Lowe *et-al.* (2016), traduzido por Lima (2016), citado por Simão (2018).

O termo Laboratório abrange um imenso caminho de estudos acadêmicos. O assunto da pesquisa se concentra em laboratórios didáticos para o ensino na Educação Básica e Ensino Superior com experimentos elétricos e eletrônicos. O conceito de “laboratório” compreende-se em duas divisões, são elas: instalações físicas e equipamentos utilizados para a experimentação. O conceito de experimentação está ligado ao acesso do ambiente para realização dos experimentos. A ocorrência da realização de atividade em laboratório remoto é conhecida como experimentação remota. Seu grande objetivo está na interação do usuário (online) com mundo real (ambiente de experimentos), com utilização de sistemas de controle e monitoramento eletrônico, com acessos a sistemas com interface homem-máquina (computador, celular, tablet entre outros) (BOHUS *et-al.*, 1996).

As atividades experimentais, podem ser classificadas com sustentação em dois critérios: acesso ao recurso experimental e a natureza do recurso experimental. O acesso ao recurso experimental pode ser dividido em duas formas, acesso remoto (utiliza Internet) e

acesso local (utiliza Intranet), independentemente do meio de acesso, ambos conseguem realizar manipulação dos experimentos remotos (BENCOMO, 2004).

A Figura 3 apresenta a classificação entre laboratórios hands-on e laboratórios online (ZUTIN *et-al.*, 2010).



Fonte: Zutin *et-al.* (2010).

Existem quatro tipos de ambientes experimentais abordando os critérios e a Figura 3 (SILVA *et-al.*, 2020), são eles: Laboratório Locais (onde na literatura são conhecidos como laboratórios tradicionais); Simulações Locais (laboratórios em rede local, onde executam simulações de experimentos); Laboratórios Remotos (consistem em estrutura de hardware, para execução do experimento com interface Web entre usuário e máquina); e Laboratórios Virtuais (integra o conceito de laboratório remoto e simulações locais, com dados simulados e acesso remoto).

Os quatro modelos de laboratórios descritos apresentam vantagens e desvantagens seguindo o ponto de vista de utilizadores, instituições e objetivos de aprendizagem. Para alguns autores a comparação direta torna-se difícil, pois existe falta de critério de avaliação da eficácia dos laboratórios (LANG *et-al.*, 2007), (NAEF, 2006), (MA; NICKERSON, 2006). Analisando do ponto de vista do ensino, cada laboratório proporciona diferentes benefícios de aprendizagem. Ambos se complementam para realização do ensino ao objetivo de aprendizagem (NAEF, 2006), (LINDSAY; GOOD, 2005).

2.2 LABORATÓRIO REMOTO VISIR

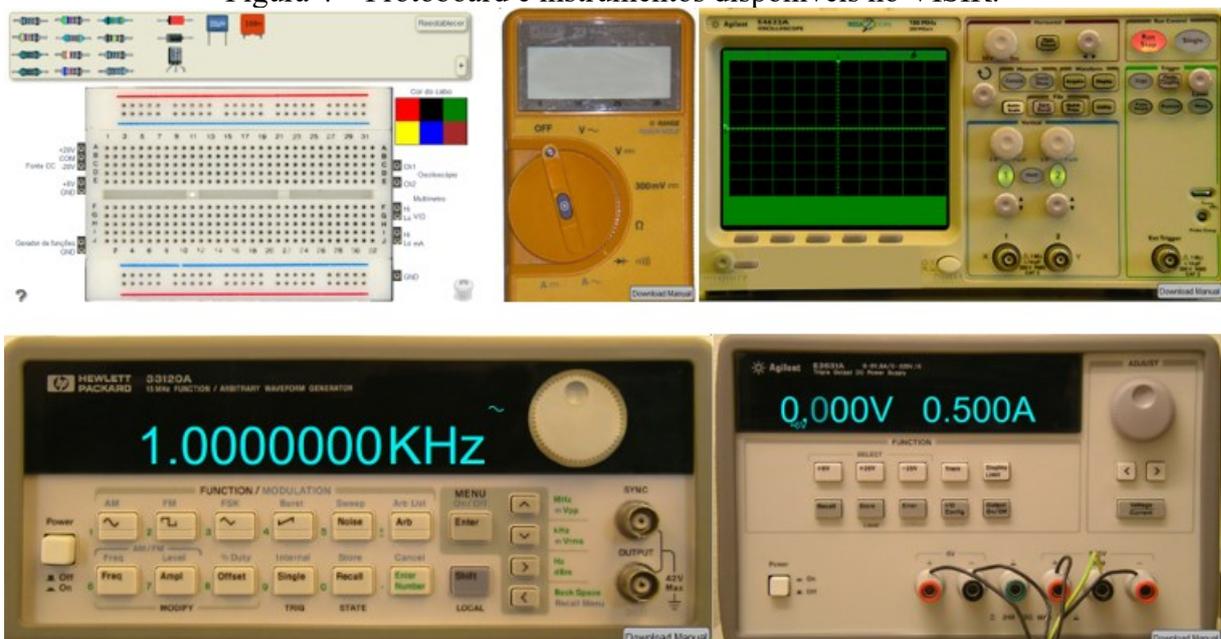
O laboratório remoto, conhecido como Plataforma VISIR OpenLab é um laboratório remoto que possibilita o aprendizado em circuitos elétricos e eletrônicos com medições com utilização de uma bancada de trabalho operada remotamente (TAWFIK *et-al.*, 2013). A plataforma VISIR pode ser utilizada em outras experiências fora da elétrica e eletrônica, como para as áreas de acústica e vibrações mecânicas no BTH (GUSTAVSON; ZACKRISSON; LUNDBERG, 2014).

A estrutura do VISIR permite a realização de experimentos remotos para construção de circuitos elétricos e eletrônicos utilizando uma matriz tridimensional, onde componentes são pré-instalados e relés de comutação fazem seus chaveamentos para construção dos circuitos.

O VISIR permite ao utilizador o acesso remoto através da Internet e que este possa, criar, conectar e medir circuitos eletrônicos numa protoboard, tendo à disposição muitos componentes eletrônicos (CLAESSON; HAKANSSON, 2012). O sistema VISIR pode ser visto como uma bancada de trabalho remota, equipada com os mesmos instrumentos que existem numa bancada de um laboratório hands-on para montar e testar circuitos elétricos e eletrônicos.

Na Figura 04 são mostrados a *protoboard* e instrumentos disponíveis no VISIR.

Figura 4 – Protoboard e instrumentos disponíveis no VISIR.



Fonte: Elaborada pelo autor.

A análise dos modelos de laboratórios remotos acaba sendo interessante por dois motivos principais:

- Análise de desempenho da classe; e
- Liberdade de estudo a qualquer momento.

Os laboratórios remotos, são uma ferramenta alternativa para aulas práticas, como forma de aplicar os conceitos teóricos aprendidos anteriormente. Consiste em conduzir remotamente experimentos reais (TAVARES *et-al.*, 2020).

O VISIR ainda se denomina um projeto em constante desenvolvimento, onde diversos estudos acadêmicos abordam sua estrutura e sua federação (ALVES *et-al.*, 2022) (KULESZA *et-al.*, 2017).

A composição de sua estrutura física é idêntica ao laboratório tradicional ou *off-line*. Ambos possuem os mesmos equipamentos.

Tratando de um laboratório remoto, alguns experimentos estão previamente preparados, outros precisam ser instalados conforme necessidade, assim as instruções para o instrutor virtual são criadas em conjunto pelos professores e pelos técnicos de laboratório (TAVARES *et-al.*, 2020).

2.2.1 Classificação de erros experimentais no VISIR

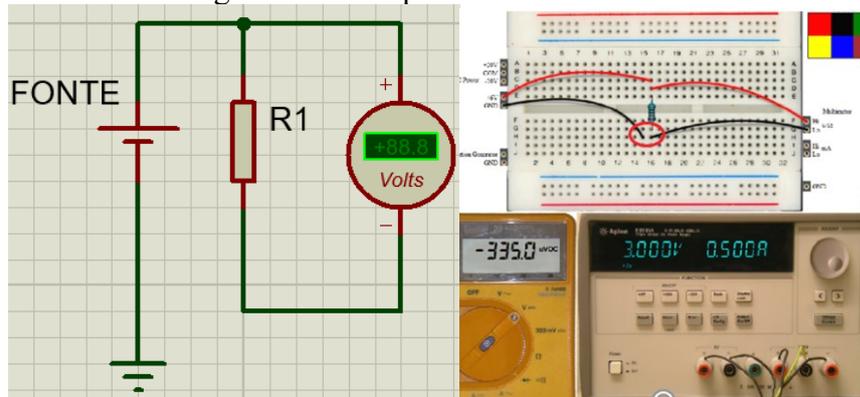
As classificações de erros são muito importantes ao professor da disciplina, com a informação consegue instruir melhor sua classe. Em análise macro da estrutura, são divididos em duas classes de erros: manuais e conceituais.

Erros manuais são definidos em três modelos, circuito aberto, curto-circuito e equipamento não configurado corretamente (GARCÍA-ZUBÍA *et-al.*, 2019).

Os erros conceituais são ocorridos quando a execução das atividades contradiz a lei de análise de circuito e suas medições (GARCÍA-ZUBÍA *et-al.*, 2019).

A figura 5 apresenta um circuito com erro manual (circuito aberto), onde a conexão do GND da fonte não foi conectada ao circuito, assim a medição da tensão não aconteceu corretamente.

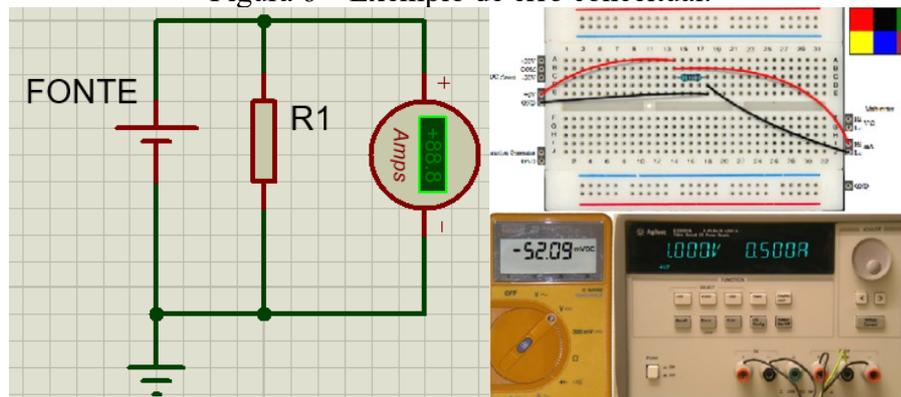
Figura 5 – Exemplo de erro manual.



Fonte: Elaborada pelo autor.

Já o próximo modelo representado pela Figura 6, demonstra um erro conceitual. O exemplo mostra um circuito medindo tensão utilizando as pontes de um amperímetro, neste caso coloca-se o sistema em um curto-circuito.

Figura 6 – Exemplo de erro conceitual.



Fonte: Elaborada pelo autor.

2.2.2 Classificação de erros no cenário de 1:1

A pesquisa atual, estuda o cenário considerado o mais simples conhecido como 1:1. O modelo representa uma fonte de energia para um componente elétrico.

Sua sequência usual é 1:1; 1:n; n:1 e n:n, considerando o primeiro termo a quantidade de fonte de energia e o segundo termo o componente consumidor de energia, denominado carga (GARCÍA-ZUBÍA *et-al.*, 2019).

Quando ocorre uma adição de um analisador de dados elétricos (exemplo: voltímetro e amperímetro) os dados do circuito 1:1 são alterados. Considerando equipamento ideais, os instrumentos de medição podem ser considerados nulos ao efeito do circuito. Na prática, os

elementos possuem uma resistência não ideal (infinita no caso do voltímetro, nula no caso do amperímetro), considerando então que o medidor dispersa um pedaço da potência do circuito nele. Nesse novo cenário o circuito de 1:1 se transforma em um circuito de 1:n (GARCÍA-ZUBÍA *et-al.*, 2019).

A Tabela 1 apresenta um conjunto de erros com objetivo de descrever os possíveis erros ao tentar medir tensão, corrente elétrica e resistência em um cenário de 1:1 (GARCÍA-ZUBÍA *et-al.*, 2019).

Tabela 1 - Possíveis erros ao medir a tensão, corrente elétrica e resistência em cenário de circuito 1:1

	Componente	DMM Connection	CABLE	CONTROL KNOB					
				OFF	VAC/AAV	DIODE	VCC(V/mV)	Ω	A
POWER ON	Resistor ON	Serial	Ω/V				E1	E9	E17
			A				E2	E10	E18
		Parallel	Ω/V				E3	E11	E19
			A				E4	E12	E20
POWER OFF	Resistor ON	Serial	Ω/V				E5	E13	E21
			A				E6	E14	E22
		Parallel	Ω/V				E7	E15	E23
			A				E8	E16	E24

Fonte: García-Zubía *et-al.* (2019).

Dentro da classificação dos erros E1 até E24 possui duas classificações corretas, são elas: E3 e E18.

A classificação do E3 apresenta um circuito medidor de tensão, já a classificação E18 é um circuito medidor de corrente.

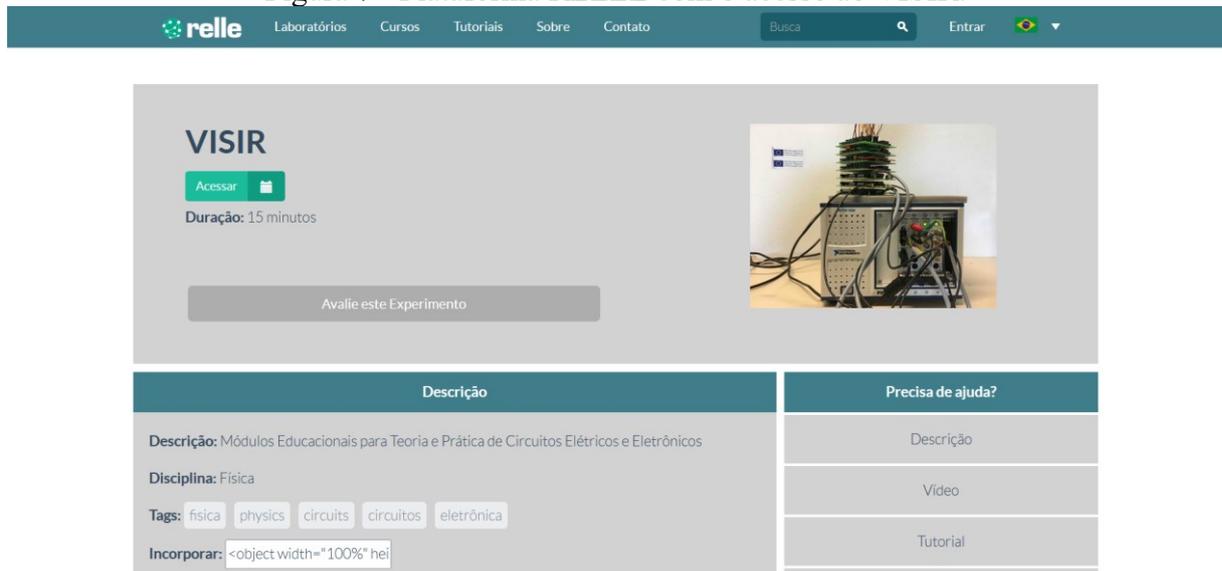
2.3 REXLAB COM PROJETO VISIR

O projeto iniciou em 2015, com o propósito de disseminar o laboratório remoto pela América Latina, inicialmente os países envolvidos eram Portugal, Espanha, Suécia, Áustria, Brasil e Argentina (LIMA, 2016) (SILVA *et-al.*, 2018a).

A Universidade Federal de Santa Catarina, foi uma das instituições brasileira selecionadas para o projeto, representado pelo núcleo de pesquisa do RExLab (ROQUE *et-al.*, 2018).

O RExLab disponibiliza o acesso ao laboratório remoto VISIR pela sua plataforma *Remote Labs Learning Environment* (RELLE). De acordo com Simão (2018), a plataforma é responsável pelo gerenciamento no controle de acesso aos experimentos do núcleo e emissão de relatórios em sua plataforma. O RELLE atualmente gerencia o acesso de 23 laboratórios remotos. A Figura 7 apresenta a plataforma RELLE com o acesso ao VISIR.

Figura 7 - Plataforma RELLE com o acesso ao VISIR.



Fonte: Elaborada pelo autor.

O RExLab possui uma vasta experiência na disseminação da tecnologia. Nos últimos quatro anos o núcleo seguiu com cinco pesquisas, são elas:

- Plataforma aberta modular para construção, gestão e disponibilização de laboratórios remotos na Educação Básica;
- Proposta de Comunidade de Práticas para professores usuários do Laboratório Remoto VISIR;

- Proposta de framework para integração tecnologias digitais na Educação Básica, inspirado na cultura maker;
- Curso de integração de tecnologias digitais na educação; e
- VISIR+.

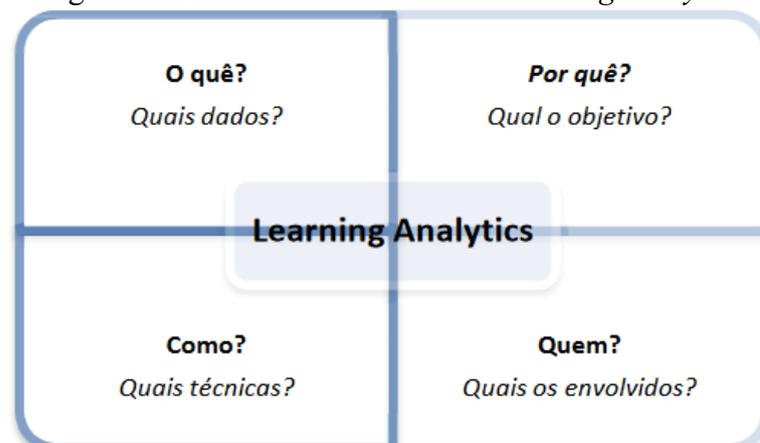
2.4 LEARNING ANALYTICS

De acordo com Diniz (2020), LA apresenta dados de desempenho educacional utilizando base de dados de aprendizagem. Segundo Clemente *et-al.* (2018 apud SIMÃO, 2018, p. 24) a BD permite os sistemas realizarem a implementação de LA, seus resultados são importantes para analisar e informar dados sobre alunos.

Segundo Moissa (2014), o modelo de *Learning Analytics* deve responder quatro perguntas ilustrados na Figura 8:

- O quê?: Refere-se ao tipo de dado coletado.
- Quem?: Está relacionado para quem os resultados serão direcionados.
- Por quê?: Deve-se responder qual objetivo/resultado da análise.
- Como?: Está ligada a técnica que será aplicada no processo. Uma delas está relacionada a essa pesquisa com a utilização de mineração de dados educacionais.

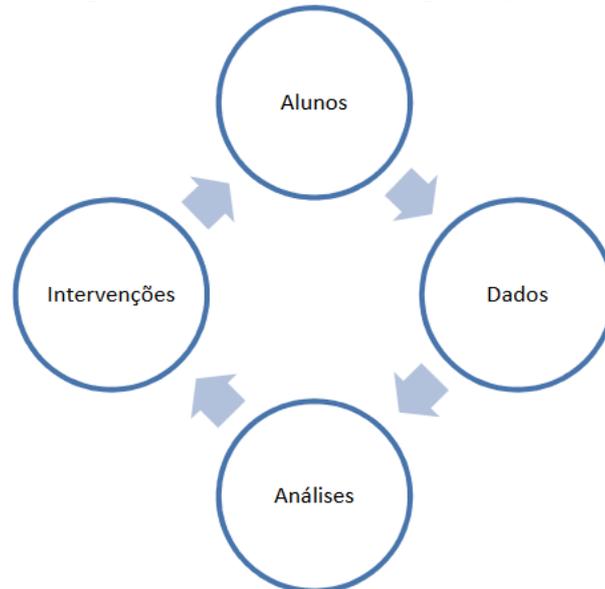
Figura 8 - Modelo de referência de *Learning Analytics*



Fonte: Chatti et-al. (2012), citado por Moissa (2014).

Segundo Moissa (2014), os modelos de referência citados, estão relacionada a uma etapa de ciclo proposto, onde são representados por: Alunos, Dados, Análise e Intervenções, apresentados na Figura 9.

Figura 9 - Ciclos de *Learning Analytics*.



Fonte: Clow (2012), citado por Moissa (2014).

2.4.1 Mineração de Dados

Nos últimos anos o mundo mudou, empresas que estavam no topo da pirâmide da estrutura financeira possuíam sua matéria prima o petróleo. Atualmente empresas tecnológicas conquistaram essa posição, não que o petróleo perdeu valor de mercado, mas as empresas de tecnologia foram ganhando espaço. A matéria prima que possui maior valor são os dados armazenados, informações de usuários, dados de rede sociais, dados clínicos entre diversas informações.

O nome mineração de dados (MD), faz criar uma relação a mineração de minérios, como também a relação de mineração de criptomoedas, um tema atual. Segundo Castro e Ferrari (2016) o processo de mineração quer indicar a extração de um material precioso.

“O termo mineração de dados (MD) foi cunhado como alusão ao processo de mineração [...], uma vez que se explora uma base de dados (mina) usando algoritmos (ferramentas) adequados para obter conhecimentos (minérios preciosos).” (CASTRO; FERRARI, 2016, p 4).

Pensando em análise estatística, modelos preditivos e explanatórios conseguem obter-se tomadas de decisões baseados nos dados armazenados. Segundo Ferguson (2012 apud

CARLOS, 2020, p 42), “O uso de *Learning Analytics* tem crescido nos últimos anos. Tal ação ocorre devido a fatores como o crescimento de áreas de *big-data*, *E-Learning* e preocupações políticas em mensurar a aprendizagem.”.

2.4.1.1 Mineração de Dados Educacionais

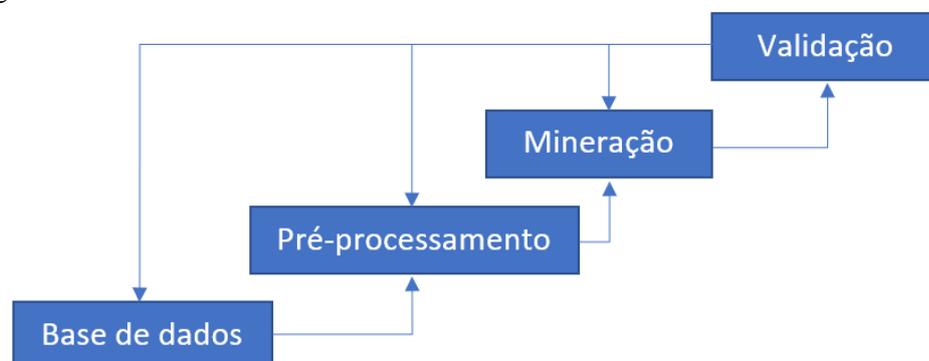
Segundo Costa *et-al.* (2012), a área de MD educacionais tem como objetivo adaptar métodos e algoritmos de mineração existente, resultando em dados de análise educacional, com captura de informação de interação dos alunos e os professores em ambientes AVAs, Sistemas Tutores Inteligentes (STIs), entre outros. O resultado da análise deverá auxiliar no entendimento do aluno e seu processo de aprendizagem.

2.4.1.2 Descoberta de conhecimento em bases de dados (KDD)

KDD conhecido como *Knowledge Discovery in Database* se refere ao processo de extração de conhecimentos a partir de dados, representados na Figura 10. Segundo Castro e Ferrari (2016) esse processo possui quatro etapas onde denominam:

- Base de dados;
- Preparação ou pré-processamento de dados;
- Mineração de dados; e
- Avaliação ou validação do conhecimento.

Figura 10 – Processamento de descoberta de conhecimento em bases de dados.



Fonte: Castro e Ferrari (2016).

2.4.1.3 Etapa de pré-processamento

A seção de pré-processamento apresentará etapas de como deve-se estruturar melhor os dados. Onde dados bem estruturados, deixa o processo de mineração mais eficiente e eficaz (CASTRO; FERRARI, 2016). De acordo com Lidwell, Holden e Butler (2010 apud CASTRO e FERRARI, 2016, p. 27), “Neste mesmo sentido, existe um princípio conhecido em tecnologia da informação e comunicação que diz: “Lixo colocado para dentro, lixo colocado para fora””. Representa muito bem o significado do pré-processamento, dados ruidosos podem gerar saída ruidosa.

2.4.1.3.1 Limpeza dos dados

A etapa da análise e limpeza dos dados é de grande importância. Deve-se entender o conjunto da base para compreender os dados a serem tratados. A informação contida na BD deve ser analisada e optar por sua utilização ou não da informação. O grau de satisfação está no entendimento entre dados e resultado a ser obtido.

“Conjunto de dados podem também apresentar dificuldades relacionadas à qualidade dos dados. Exemplos mais frequentes dessas dificuldades são dados ruidosos (que possuem erros ou valores que são diferentes do esperado), inconsistentes (que não combinam ou contradizem valores de outros atributos do mesmo objeto), redundantes (quando dois ou mais objetos têm os mesmos valores para todos os atributos ou dois ou mais atributos têm os mesmos valores para dois ou mais objetos) ou incompletos (com ausência de valores para alguns dos atributos em parte dos dados).” (FACELI *et-al.*, 2015, p. 34).

Segundo Castro e Ferrari (2016) existe uma grande probabilidade do conjunto de dados possuírem uma baixa qualidade, assim seu principal método para correção é a aplicação das seguintes técnicas:

- Ignorar o objeto;
- Imputar manualmente os valores ausentes;
- Usar uma constante global para imputar o valor ausente;
- Imputação do tipo *hot-deck*;
- Imputar de acordo com a última observação (*last observation carried forward*);
- Usar a média ou moda de um atributo para imputar o valor ausente;
- Usar a média ou moda de todos os objetos da mesma classe para imputar o valor ausente; e
- Usar modelos preditivos para imputar o valor ausente.

2.4.1.3.2 *Transformação dos Dados*

Transformação de dados muitas vezes são para facilitar a utilização de algoritmos onde sua base possui unidade de medidas diferentes. “Uma das normalizações mais utilizadas é a Max-Min no intervalo [0,1]. Para aplicar essa normalização, os valores categóricos que estão em texto serão substituídos pelos seus números [...]” (CASTRO; FERRARI, 2016, p. 57).

“Várias técnicas de aprendizagem de máquina (AM) estão limitadas à manipulação de valores de determinados tipos, por exemplo, apenas valores numéricos ou apenas valores simbólicos. Adicionalmente, algumas técnicas têm seu desempenho influenciados pelo intervalo de variação dos valores numéricos.” (FACELI *et-al.*, 2015, p. 41).

A técnica de transformação de dados se torna muito importante conforme o algoritmo a ser aplicado. Para obter um ganho no resultado deve-se entender os tipos de dados que a base possui, qual o modelo a ser aplicado e tratar os dados conforme a escolha do modelo.

2.4.1.3.3 *Conversão Simbólico-Numérico*

A forma do tratamento dos dados interfere no algoritmo de execução. “[...] vários algoritmos de agrupamento lidam apenas com dados numéricos. [...] quando o conjunto de dados utilizado por essas técnicas apresenta atributos simbólicos, os valores desses atributos devem ser convertidos para valores numéricos.” (FACELI *et-al.*, 2015, p. 41).

2.4.1.3.4 *Conversão Numérico-Simbólica*

Segundo FACELI *et-al.* (2015), alguns algoritmos foram desenvolvidos para operar com valores qualitativos como algoritmos de classificação e de associação. Quando esses mesmos algoritmos necessitam operar em dados quantitativos seu desempenho fica reduzido.

2.4.1.3.5 *Transformação de Atributos Numéricos*

Valores numéricos nem sempre estão preparados para serem analisados.

“Isso geralmente ocorre quando os limites inferior e superior de valores dos atributos são muito diferentes, o que leva a uma grande variação de valores, ou ainda quando vários atributos estão em escalas diferentes.” (FACELI *et-al.*, 2015, p. 44).

2.4.1.4 Mineração

De acordo com Castro e Ferrari (2016), mineração de dados serve para especificar os tipos de informações que devem ser obtidas na etapa de mineração. As principais tarefas da mineração de dados são descritas nesta seção, será apresentado três modelos principais para esse trabalho.

Análise descritiva: “Permite uma sumarização e compreensão dos objetos da base e seus atributos” (CASTRO; FERRARI, 2016, p. 8). De acordo com FACELI *et-al.* (2015), a tarefa pode ser dividida em sumarização, associação e agrupamento.

Predição: É uma técnica que classifica objetos não rotulados a partir de dados rotulados onde se constrói um estimador (FACELI *et-al.*, 2015) (CASTRO; FERRARI, 2016).

Análise de grupos: Agrupamento ou *clustering* como é conhecido o processo de separar objetos em grupos. Os dados de entrada não são rotulados como o de classificação, assim sua entrada não possui os grupos conhecidos.

Como o objetivo da pesquisa é classificar, criar grupos de classificação dos experimentos, o modelo utilizado será o preditivo.

2.4.1.4.1 Algoritmo de Classificação

Algoritmo de classificação, como o próprio nome já diz, são classificadores de grupos de objetos, sua função é rotular o objeto de entrada conforme seu treinamento. Isso resume que um excelente algoritmo consegue realizar sua classificação e atualizar seu modelo de classificação, isso significa que com o passar do tempo o próprio modelo pode ir aprendendo com as informações recebidas.

Classificar um dado é atribuir a ele um grupo, chamado classe, de acordo com a categoria à qual ele pertence. Um algoritmo de classificação é usado na construção de um modelo, também chamado de classificador, o qual é construído por um conjunto de treinamento de classes, ou seja, um conjunto de entrada-saída (CASTRO; FERRARI, 2016).

Existem diversos classificadores na literatura, podendo ser dividido em seis grupos conforme sua estrutura. São eles: baseados em conhecimento, baseados em árvores, conexionistas, baseados em distância, baseados em função e probabilísticos (CASTRO; FERRARI, 2016).

O modelo baseado em conhecimento deve-se seguir um conjunto de regras predefinidas para obter sua classificação (CASTRO; FERRARI, 2016).

Modelo baseado em árvore ou método baseado em procura, constitui de uma técnica utilizada muito em mineração de dados, utiliza a estratégia de dividir para conquistar. Sua estrutura é construída na obtenção do nó raiz e os nós intermediários, onde representam validações de um atributo, já os resultados são representados pelos ramos até a classificação da classe representada por folha. Existem diversos sistemas que incorporam o método são eles: Microsoft SQL Server, R, Statistica, entre outros (CASTRO; FERRARI, 2016) (FACELI *et-al.*, 2015) (RUSSEL; NORVIG, 2013).

A palavra conexionista, ou método baseado em otimização, lembra muito sobre conexão, assim, são modelos baseados em rede de unidade, pertence ao um tipo de grafo. A técnica mais conhecida são as Redes Neurais Artificiais (CASTRO; FERRARI, 2016) (FACELI *et-al.*, 2015) (RUSSEL; NORVIG, 2013).

Modelo baseado em distância deseja classificar um objeto comparando sua distância com objetos já rotulados. A classificação será atribuída pelo objeto de menor distância. O algoritmo K – *nearest neighbors*, segundo a literatura é a técnica mais simples entre todos os algoritmos de Aprendizado de Máquina. (CASTRO; FERRARI, 2016) (FACELI *et-al.*, 2015).

Modelo baseado em função possuem parâmetros predefinido, onde podem ser ajustados em treinamento (CASTRO; FERRARI, 2016).

Probabilístico permite um objeto pertencer a mais de uma classe com sua probabilidade. Algoritmo dedicado, mas quando os dados são incompletos ou imprecisos (CASTRO; FERRARI, 2016) (FACELI *et-al.*, 2015).

O algoritmo de classificação pode ser executado em diversos programas (softwares, frameworks e biblioteca), com modalidade paga ou não (CASTRO; FERRARI, 2016). O software Weka foi desenvolvido em linguagem Java com licença gratuita, com interface gráfica, possui bibliotecas para integrar códigos desenvolvidos em Java, dando, mas controle ao usuário. O Matlab é considerado uma linguagem de programação de alto nível com modalidade paga. Sua estrutura integra diversas bibliotecas para processos de mineração de dados. A linguagem R possui o mesmo seguimento do Matlab, considerado uma das ferramentas importante para estatística e visualização de dados, seu sistema é gratuito e possui código aberto. O software Orange é um sistema gratuito, onde seu desenvolvimento está estruturado para construção visual com utilização de blocos e fluxograma para processos de análise e mineração de dados.

O desenvolvimento da pesquisa optou-se por trabalhar com software Orange pelo fácil desenvolvimento e entendimento das etapas construído por diagrama de blocos.

A seção 2.3.1.8.1 e 2.3.1.8.2 apresenta os dois modelos de classificação utilizadas na pesquisa, algoritmo k-NN e *tree*. O algoritmo k-NN possui sua técnica do modelo baseado em distância, já o algoritmo *tree* é baseado em árvores. A escolha desses dois modelos ocorreu pelo fato de a literatura tratá-los como um dos algoritmos mais fáceis, observando que o trabalho busca classificar experimentos, se faz todo sentido da utilização de classificação por modelo de árvore e do modelo de vizinhos mais próximos.

Os modelos de classificação utilizados foram utilizados técnicas de validação cruzada e amostragem aleatória para o treinamento. A validação cruzada, serve para:

“particionar a base em conjuntos de treinamento e testes de modo que os dados de treinamento sejam usados para ajustar os parâmetros livres do modelo e os dados de testes sejam usados para fornecer uma estimativa de como o modelo vai generalizar para dados não usados no treinamento.” (CASTRO; FERRARI, 2016, p.157).

A amostragem aleatória conhecida como *random sampling*, realiza a divisão dos dados em dois grupos, geralmente 70% e 30%, onde o grupo que possui 70% dos dados realiza o treinamento, e o grupo de 30% realiza o teste de validação do algoritmo (PRETNAR, 2019).

2.4.1.4.2 *k – nearest neighbors*

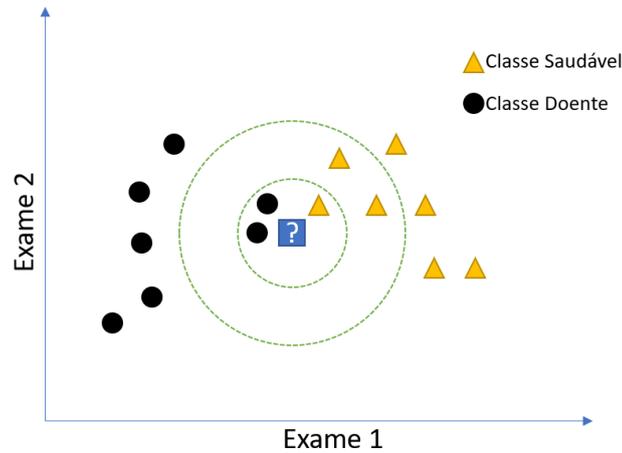
K – nearest neighbors conhecido pela sigla k-NN ou pela sua tradução de k – vizinhos mais próximos “[...] é um dos classificadores não paramétricos em distância mais simples e conhecido na literatura” (CASTRO; FERRARI, 2016, p.167).

“O k-NN opera da seguinte maneira: dado um objeto x_0 cuja classe deseja inferir, encontram-se os k objetos x_i $i = 1, \dots, k$ da base que estejam mais próximos a x_0 e, depois, se classifica o objeto x_0 como pertencente à classe da maioria dos k vizinhos. Empates são decididos aleatoriamente.” (CASTRO; FERRARI, 2016, p. 167).

Segundo FACELI *et-al.* (2015), descreve que o algoritmo que responde ao 1-NN possui uma extensão imediata. Quando atribui valor de k maior que 1, se obtém k vizinhos para cada ponto de teste.

A Figura 11 demonstra classes de saudável e de doente, se $k = 3$, o objeto seria classificado como doente, agora se colocar $k = 5$ o elemento seria classificado como saudável. (FACELI *et-al.*, 2015).

Figura 11 – Impacto do valor de k no algoritmo k-NN.



Fonte: FACELI *et-al.* (2015).

“A escolha do valor de k mais apropriado para um problema de decisão específico pode não ser trivial. O valor de k é definido pelo usuário. Frequentemente, o valor de k é pequeno e ímpar: $k = 3, 5, \dots$ Em problemas de classificação, não é usual utilizar $k = 2$ ou valores pares, para evitar empates.” (FACELI *et-al.*, 2015, p. 62).

Segundo Castro e Ferrari (2016), mesmo o algoritmo sendo simples, seus bons rendimentos acontecem em diversos cenários, principalmente em classes que possuem diversos elementos. A seguir o Quadro 1 representa o pseudocódigo do algoritmo.

A forma que os dados são constituídos torna-se crucial para um excelente resultado, “algoritmos que utilizam medidas de similaridade ou distância, como o k-NN, os dados nominais precisam ser transformados em números e é recomendável que eles sejam normalizados no intervalo $[0,1]$ ” (CASTRO; FERRARI, 2016, p. 191).

Quadro 1 - Pseudocódigo do algoritmo.

<p>Entrada</p> <p>k : número de vizinhos</p> <p>data : base de dados com n objetos e m atributos (n x m)</p> <p>classe : vetor contendo a classe de cada objeto da base (n x 1)</p> <p>obj : objeto que deve ser classificado (1 x m)</p> <p>Saída</p> <p>C : rótulo indicativo da classe do objeto</p> <p>Passos</p> <pre>// Calcular a distância entre base de dados e o objeto D(n x 1) D = dist(data,obj); Obj = Ø; // Determinar os k objetos mais próximos Para i=1:k Faça { aux = D[1]; pos = 1; Para j=2:n Faça { Se (aux > D[j]) e (j ∉ obj == Ø) Então { aux = D[j]; pos = j; } } Objs.Add(pos); } // Pegar a classe dos k objetos mais próximos Ck = classe[objs]; // Determinar a classe mais frequente C = moda(Ck);</pre>

Fonte: Castro e Ferrari (2016).

O método de aprendizado do algoritmo ocorre pelo armazenamento dos dados de treinamento e aguardam até que seja necessária a classificação do registro, assim analisando similaridade entre o objeto de entrada com o treinamento (Luz, 2019).

2.4.1.4.3 Tree

O algoritmo *Tree* conhecido como árvore de decisão, é denominado uma das formas mais simples, possui sua estrutura em formato de ramificação de uma árvore.

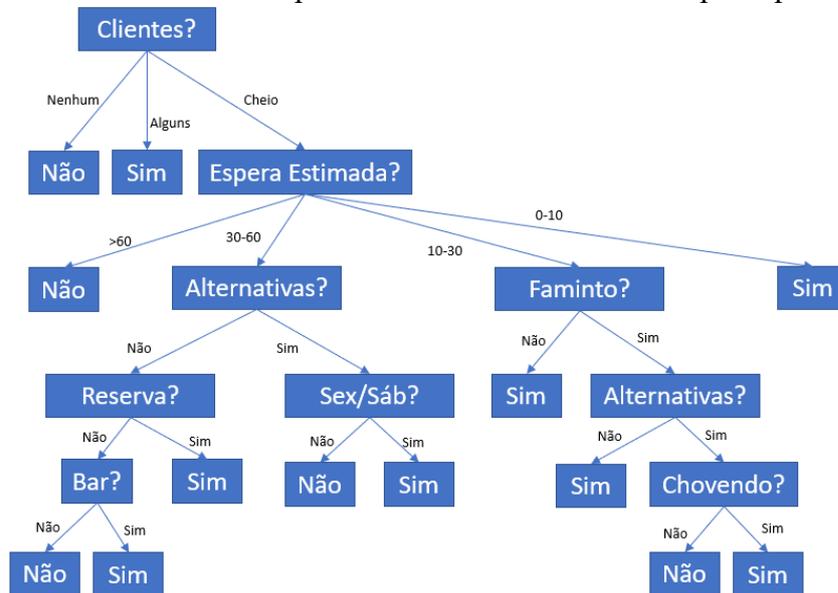
“[...] é uma estrutura em formato de árvore na qual cada nó interno corresponde a um teste de um atributo, cada ramo representa um resultado do teste e os nós folhas representam classes ou distribuição de classes. O nó mais elevado da árvore é conhecido como nó raiz, e cada caminho da raiz até um nó folha corresponde a uma regra de classificação”. (CASTRO; FERRARI, 2016, p. 170).

“Uma árvore de decisão representa uma função que toma como entrada um vetor de valores de atributos e retorna uma “decisão” – um valor de saída único. Os valores de entrada e saída podem ser discretos ou contínuos.” (RUSSEL; NORVIG, 2013, p. 609).

“Uma árvore de decisão alcança sua decisão executando uma sequência de testes. Cada nó interno na árvore corresponde a um teste do valor de um dos atributos de entrada, A_i , e as ramificações dos nós são classificadas com os valores possíveis do atributo, $A_i=V_{ik}$. Cada nó de folha na árvore especifica o valor a ser retornado pela função. A representação de árvores de decisão parece ser muito natural para os seres humanos.” (RUSSEL; NORVIG, 2013, p. 609).

Segundo Faceli *et-al.* (2015) a árvore de decisão apresentada na Figura 12, usa a estratégia dividir para conquistar com objetivo de solucionar problemas de decisão. Se divide os problemas em complexidade menores, aos quais recursivamente é aplicada a mesma estratégia. O Quadro 2 apresenta o pseudocódigo para árvore de decisão.

Figura 12 – Árvore de decisão para definir se vamos ou não esperar por uma mesa.



Fonte: Russell e Norvig (2013).

“Para uma grande variedade de problemas, o formato da árvore de decisão gera um resultado agradável e conciso. Mas algumas funções não podem ser representadas de forma concisa. Por exemplo, a função da maioria, que retorna verdadeiro se e somente se mais da metade das entradas for verdadeira, exige uma árvore de decisão exponencialmente grande. Em outras palavras, as árvores de decisão são boas para alguns tipos de funções e ruins para outros.” (RUSSELL; NORVIG, 2013, p. 610).

Quadro 2 - Algoritmo para construção de uma árvore de decisão.

<p>Entrada: Um conjunto de treinamento $\mathbf{D} = \{(x_i, y_i), i = 1, \dots, n\}$</p> <p>Saída: Árvore de Decisão</p> <p>/*Função GeraÁrvore(D) */;</p> <p>se critério de parada(D) = Verdadeiro então</p> <p style="padding-left: 20px;">Retorna: um nó folha rotulado com a constante que minimiza a função perda;</p> <p>fim</p> <p>Escolha o atributo que maximiza o critério de divisão \mathbf{D};</p> <p>para cada partição dos exemplos D_i baseado nos valores do atributo escolhido faça</p> <p style="padding-left: 20px;">Induz uma subárvore $\hat{Árvore}_i = \text{GeraÁrvore}(D_i)$;</p> <p>fim</p> <p>Retorna: Árvore contendo um nó de decisão baseado no atributo escolhido, e descendentes $\hat{Árvore}_i$;</p>
--

Fonte: FACELI *et-al* (2015).

3 PROCEDIMENTOS METODOLÓGICOS

Este capítulo tem como propósito apresentar os procedimentos metodológicos da pesquisa, e apresentar as etapas percorridas para alcançar os objetivos propostos. O capítulo será dividido em tópicos referentes a classificação da pesquisa (3.1), obtenção de dados (3.2), e por fim, delimitação da pesquisa (3.3).

3.1 CLASSIFICAÇÃO DA PESQUISA

A definição de pesquisa em âmbito geral de acordo com Gil (2002, p. 17), “procedimento racional e sistemático que tem como objetivo proporcionar respostas aos problemas que são propostos”. Segundo Marconi e Lakatos (2003 apud SIMÃO, 2018, p. 65), “não há ciência sem o emprego de métodos científicos”.

A natureza da pesquisa e seu contexto permitem classificá-la como aplicada, pois, tem como objetivo, a utilização das informações disponíveis para a criação de novas abordagens e métodos de avaliação e monitoramento.

Segundo Silva e Menezes (2005, p. 20), “Pesquisa Aplicada: objetiva gerar conhecimento para aplicação prática e dirigidos à solução de problemas específicos. Envolve verdades e interesses locais”.

Sobre a abordagem do problema, classificou-se como pesquisa quantitativa, onde seu foco principal está na demonstração de métricas de ensino no laboratório remoto VISIR.

“Pesquisa Quantitativa: considera que tudo pode ser quantificável, o que significa traduzir em números opiniões e informações para classificá-las e analisá-las. Requer o uso de recursos e de técnicas estatísticas (percentagem, média, moda, mediana, desvio-padrão, coeficientes de correlação, análise de regressão, etc.)” (SILVA; MENEZES, 2005, p. 20).

Com isso a Quadro 3 representa resumidamente a natureza de classificação da pesquisa proposta.

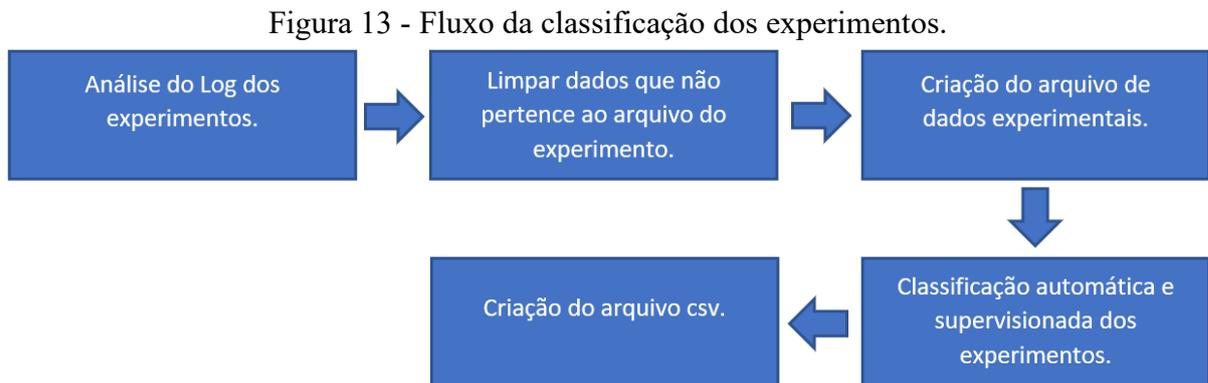
Quadro 3 - Classificação da pesquisa.

ASPECTOS	CLASSIFICAÇÃO
Natureza	Pesquisa Aplicada
Abordagem do Problema	Pesquisa Quantitativa

Fonte: Elaborada pelo autor.

3.2 OBTENÇÃO DOS DADOS

Os dados analisados nesta pesquisa, foram obtidos pelo ISEP. Optou-se por essa escolha pelo fato da expertise do instituto em relação ao laboratório VISIR. Os dados armazenados dos experimentos ficam registrado em arquivos de Log. A Figura 13 apresenta o fluxo do processo para realização da classificação.



Fonte: Elaborada pelo autor.

3.3 DELIMITAÇÃO DA PESQUISA

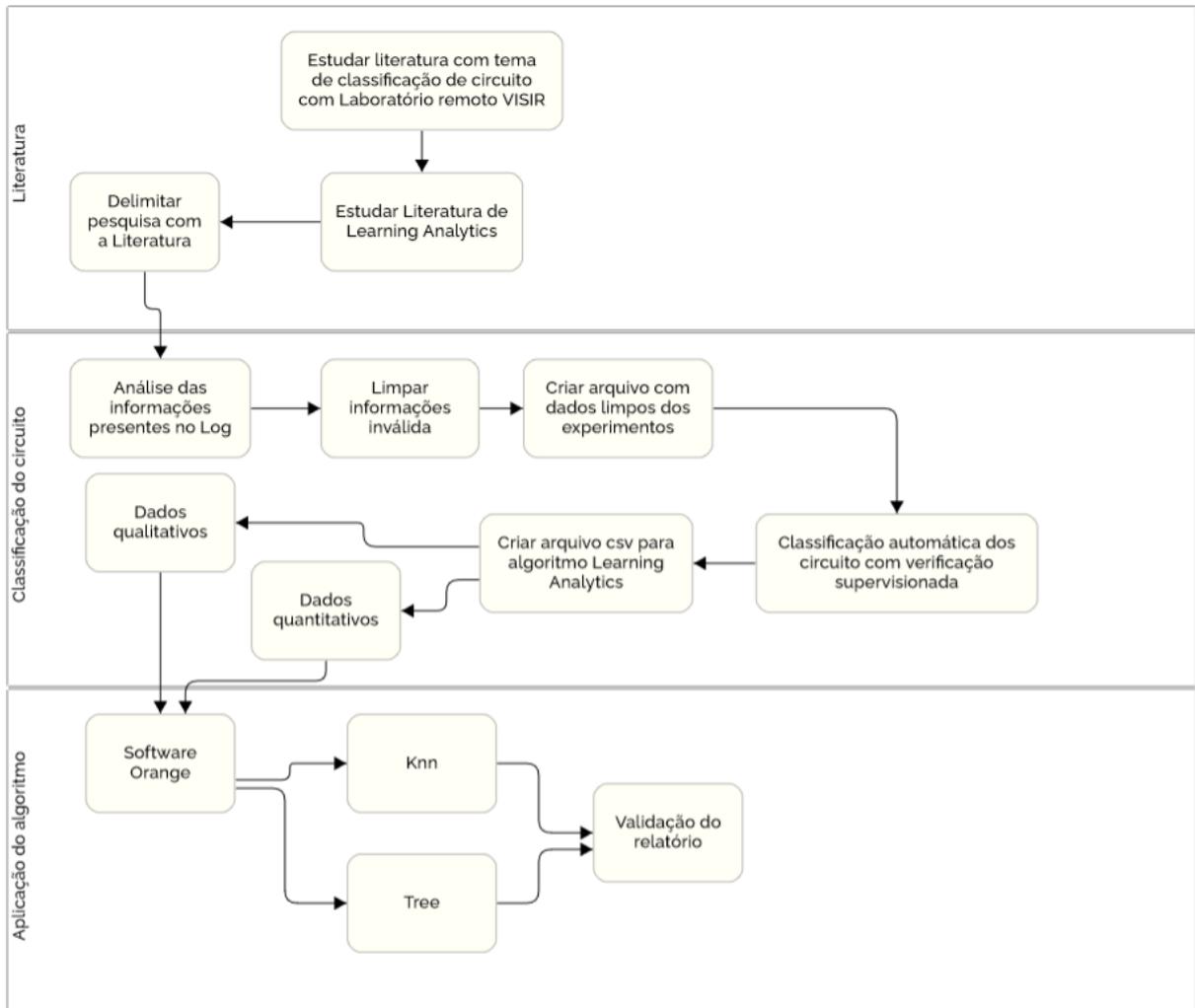
A delimitação da pesquisa é de suma importância, pois com ela o pesquisador consegue visualizar até onde seu trabalho pode ir, assim, o pesquisador tem o dever de acordo com Marconi e Lakatos (2003, p. 162) “estabelecer limites para investigação”.

Esta pesquisa foi realizada no contexto do projeto VISIR, sendo que, o tema visa em modo global na utilização do laboratório remoto VISIR, com criação de classificação utilizando técnicas de LA. A delimitação da pesquisa está relacionada, com a utilização dos modelos de circuitos elétricos 1:0, 0:1 e 1:1.

3.4 ETAPAS DA PESQUISA

Com o objetivo de alcançar resultados satisfatório da pesquisa, foi realizado um fluxo de atividades. A Figura 14 apresenta fluxograma que mostra as etapas a serem realizados.

Figura 14 - Fluxograma das etapas da pesquisa.



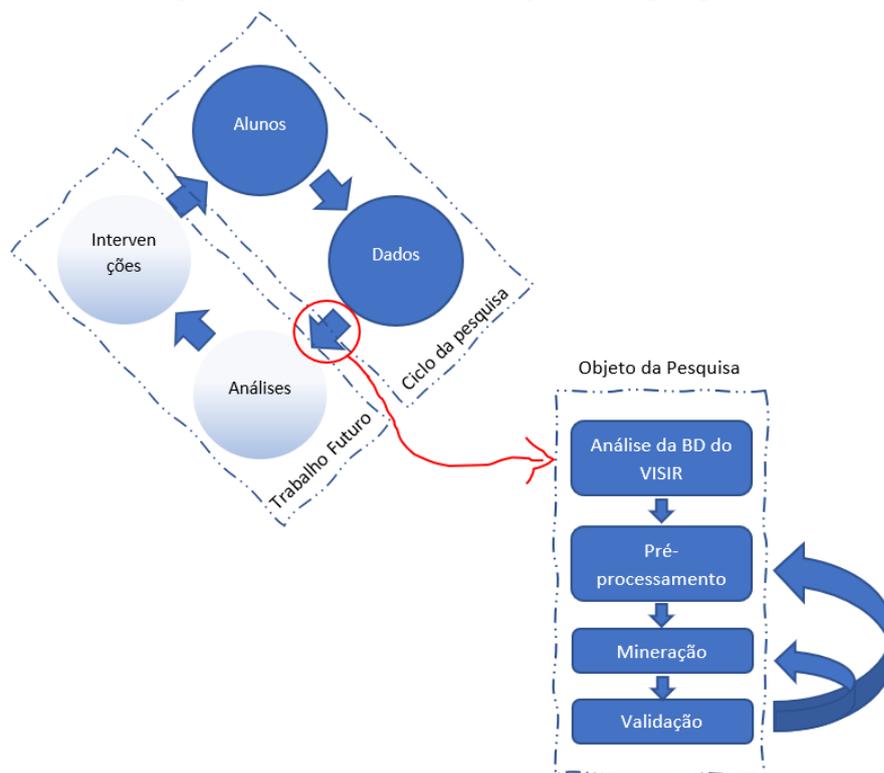
Fonte: Elaborada pelo autor.

4 MODELO PROPOSTO

A seção 2.4 apresenta o ciclo de LA considerando a execução do aluno até possível intervenção do professor. O modelo proposto trabalha entre os dados gerados pela plataforma de operação até antes da execução de análise do professor. Sua construção possui o mesmo ciclo da Figura 15. O ensino da elétrica e eletrônica vai além de modelos 1:1, 0:1 e 1:0, por esse motivo a pesquisa está limitada no momento para realização do ciclo completo.

Quando todo ciclo de LA estiver concluído, os dados podem ser analisados em tempo real de execução do experimento. Assim informando ao aluno e ao professor a classificação do circuito.

Figura 15 – Ciclo LA com objetivo da pesquisa.



Fonte: Elaborada pelo autor.

Segundo Chapman *et-al* (2000) o processo de mineração de dados abrange um ciclo completo de um projeto (início, tarefas e relações entre tarefas).

A metodologia empregada possui seis etapas, de acordo com Chapman *et-al.* (2000) são elas: Entendimento do negócio (Análise aos requisitos do projeto); Entendimento dos dados (entendimento global do problema e validação da qualidade dos dados); Pré-processamentos

dos dados (construção da base que será utilizado no algoritmo), Avaliação (avaliação dos dados gerados pelo algoritmo); e Desenvolvimento (apresentação dos dados gerados).

A seguir as subseções do presente capítulo apresentam os dados de pré-processamento e análise do algoritmo para aplicação da BD.

4.1 PRÉ-PROCESSAMENTO DOS DADOS

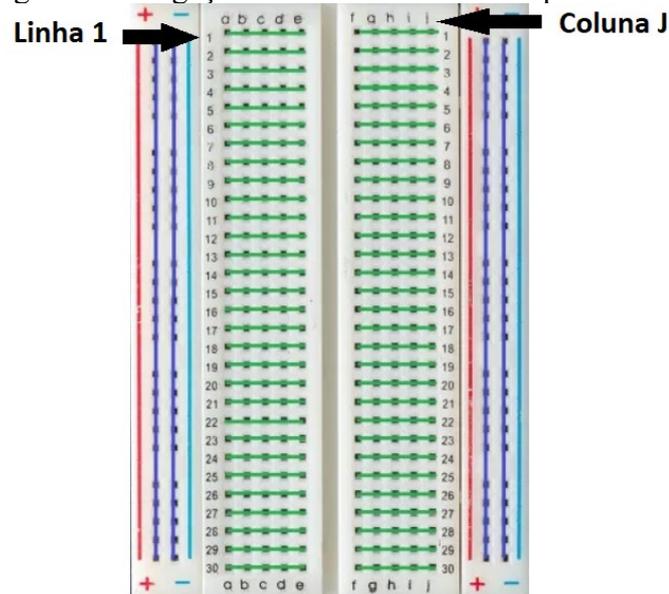
Este capítulo apresenta o modelo de pré-processamento dos dados obtidos pelo arquivo de log.

4.1.1 Regra de tratamento

Agora inicia-se uma das partes mais importante do tratamento de dados para esse estudo. É de grande importância para a continuação da pesquisa explicar como funciona a construção da *protoboard* denominada no trabalho de base de ligação.

A base de ligação representada na Figura 16, é constituída de forma matricial linhas por colunas no exemplo a seguir possui 30 linhas e 10 colunas (a até j). As colunas são divididas para esse modelo em dois grupos, primeiro grupo (a até e), e o segundo grupo (f até j). Suas linhas são independentes não possuem ligação com outras linhas. Já as colunas possuem a mesma ligação do grupo por linha, então a linha 1 tem o mesmo ponto da coluna (a até e).

Figura 16 – Ligação da matriz utilizando a protoboard.



Fonte: <https://portal.vidadesilicio.com.br/protoboard>.

Conhecendo a base de ligação, inicia-se o estudo dos dados, nela possui diversas *tags* e o que irá limitar o experimento são duas *tags*:

- `<protocol version="1.3">`; e
- `<\protocol>`

A representação da construção do circuito está contida dentro da *tag circuitlist*, toda construção realizada na base de ligação se concentra nesse grupo. O Quadro 4 apresenta o circuito com um multímetro ligado em paralelo com uma resistência.

Pontos de conexão da base de ligação são considerados dados ruidosos assim devem ser tratados com a utilização da técnica de circuito equivalente.

Quadro 4 - Dados do circuito com *tag circuitlist*.

```
<circuitlist>
W_X A13 DMM_VLO
W_X A17 DMM_VHI
R_X A13 A17 10k
</circuitlist>
```

Fonte: Elaborada pelo autor.

A Tabela 2 apresenta a nomenclatura de cada objeto utilizado para a construção do circuito, são elas:

Tabela 2 - Nomenclatura dos componentes do circuito.

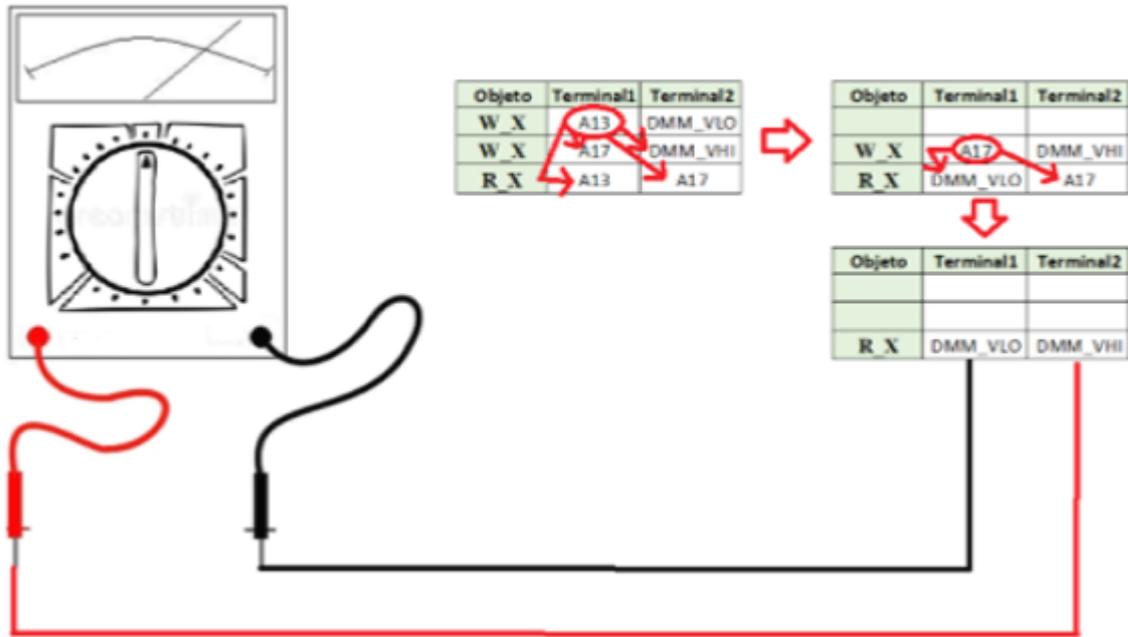
Objeto	Especificação
W_X	fio de ligação
R_X	Resistor
DMM_VLO	Multímetro V/ Ω ponteira negativa
DMM_VHI	Multímetro V/ Ω ponteira positiva
A13	Ponto da matriz da base de ligação
A17	Ponto da matriz da base de ligação
10k	Valor do resistor

Fonte: Elaborada pelo autor.

O tratamento das ligações realizadas se torna fundamental pelo motivo que um aluno pode realizar a construção do circuito em uma parte da matriz e o outro aluno em outra parte da matriz, e o foco não é saber onde ligou e sim analisar o circuito em questão. Com isso os circuitos possuem a mesma base de nomenclaturas das ligações.

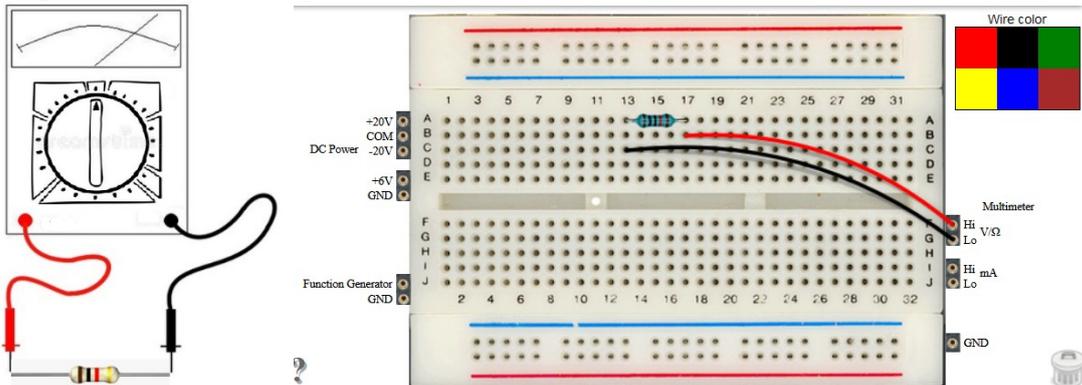
As Figuras 17 e 19 apresenta as etapas de como limpar os dados desnecessários sem perder a análise do circuito. Onde as Figuras 18 e 20 representa a ligação do circuito de um laboratório tradicional e remoto.

Figura 17 – Explicação das etapas de otimização de circuito (circuito medindo resistência).



Fonte: Elaborada pelo autor.

Figura 18 – Circuito equivalente entre laboratório remoto VISIR e laboratório convencional.



Fonte: Elaborada pelo autor.

Os dados apresentados a seguir representam a otimização do circuito realizando medição de corrente. O Quadro 5 apresenta a ligação do circuito com a *tag circuitlist*. A Tabela 3 apresenta as nomenclaturas dos componentes utilizados no Quadro 5. O Quadro 6 apresenta o pseudocódigo para minimizar o circuito.

Quadro 5 - Dados do circuito com *tag circuitlist*.

```

<circuitlist>
W_X A3 DC_+25V
W_X A7 DMM_AHI
W_X DMM_ALO DC_COM
R_X A3 A7 1k
</circuitlist>

```

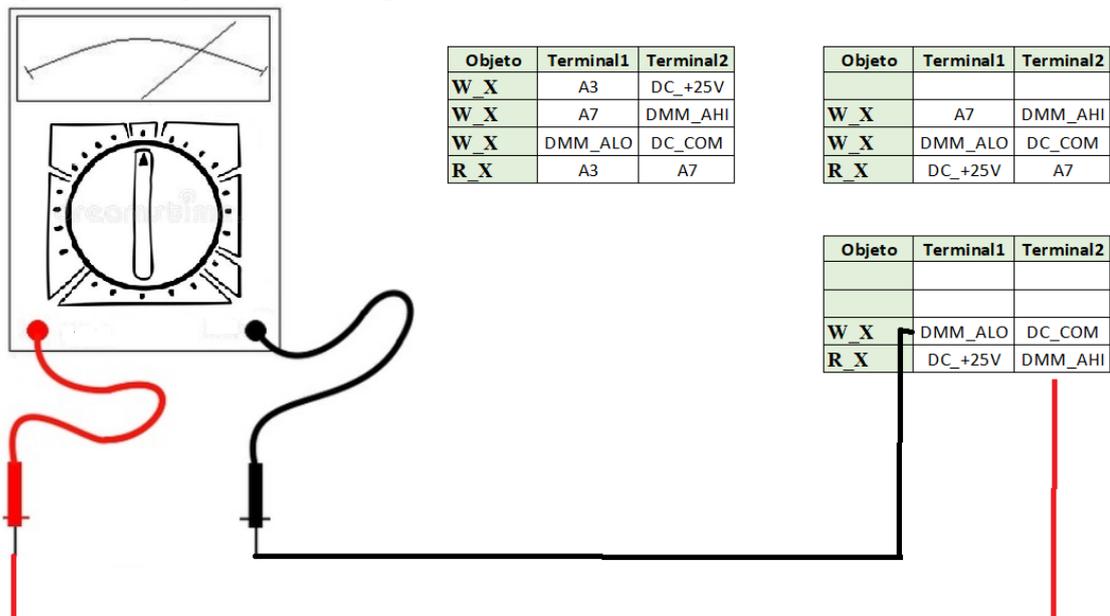
Fonte: Elaborada pelo autor.

Tabela 3 - Nomenclatura dos componentes do circuito.

Objeto	Especificação
W_X	fio de ligação
R_X	Resistor
DMM_ALO	Multímetro A ponteira negativa
DMM_AHI	Multímetro A ponteira positiva
A3	Ponto da matriz da base de ligação
A7	Ponto da matriz da base de ligação
1k	Valor do resistor
DC_COM	GND

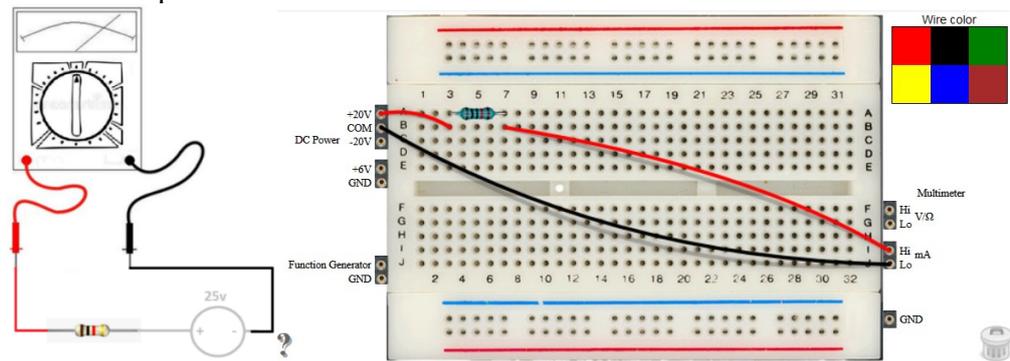
Fonte: Elaborada pelo autor.

Figura 19 – Explicação das etapas de otimização de circuito (circuito medindo corrente).



Fonte: Elaborada pelo autor.

Figura 20 – Circuito equivalente entre laboratório remoto VISIR e laboratório convencional.



Fonte: Elaborada pelo autor.

Quadro 6 - Pseudocódigo para análises.

```

Entrada
classe : lista de objetos circuitos
Passos
  obj = null;
  //obter tamanho da lista
  tamanhoClasse = classe.size();

  Para i=0 : i < tamanhoClasse Faça
  {
    obj = classe.get(i).getCircuito();
    Para j=0 : j < obj.qtdLinhas() Faça
    {
      terminal1 = obj.getTerminal1(j);
      terminal2 = obj.getTerminal2(j);

      // Verifica se o Terminal pertence a conexão da protoboard
      Se (pontoProtoboard(terminal1))
      {
        // Faz a substituição do terminal1 pelo terminal2 em todo circuito e retorna se executou
        status = SubstituirPontos(terminal1, terminal2);
      }
      Se Não (pontoProtoboard(terminal2))
      {
        // Faz a substituição do terminal2 pelo terminal1 em todo circuito e retorna se executou
        status = SubstituirPontos(terminal2, terminal1);
      }

      // Se for verdadeiro remove a linha;
      Se (status)
      {
        obj.get(i).removerLinha(j);
      }
    }
  }
}

```

Fonte: Elaborada pelo autor.

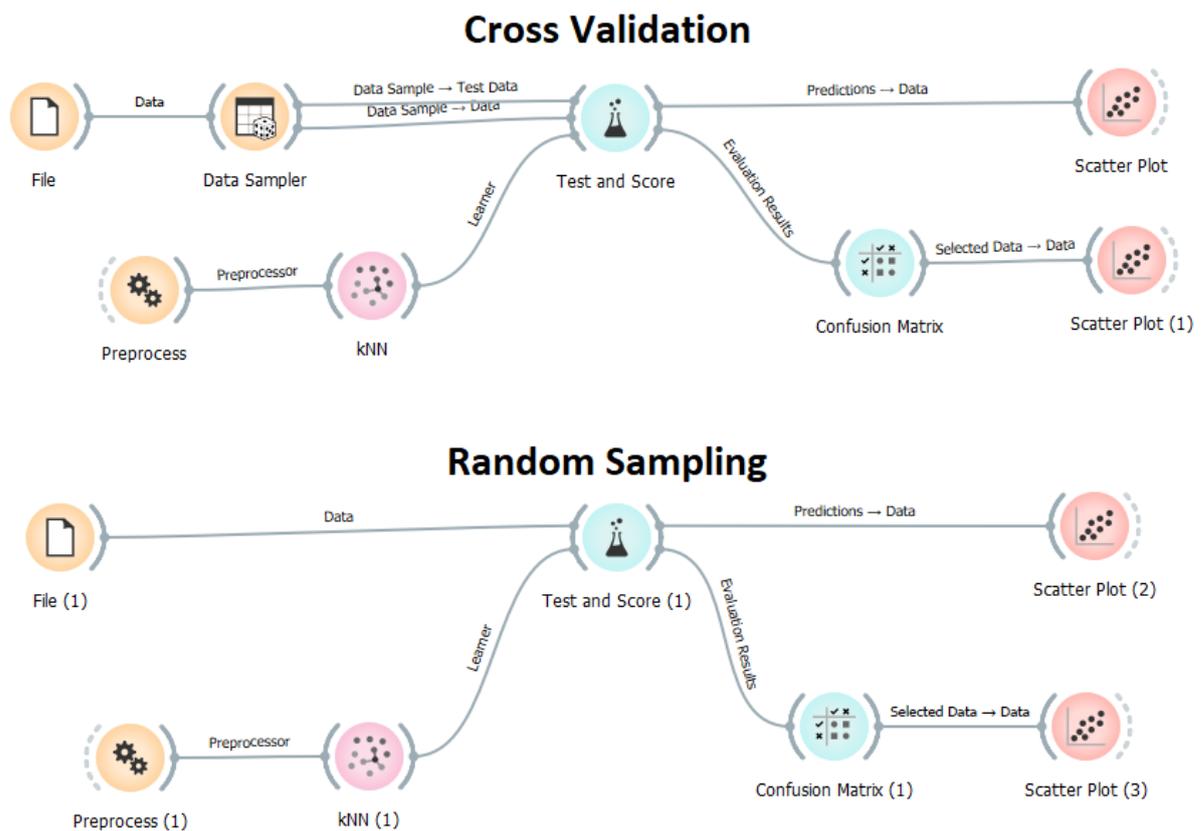
4.2 ANÁLISE DOS ALGORITMOS

Após o pré-processamento da BD é o momento da aplicação do algoritmo para posteriormente a validação dos dados. Os testes realizados ocorreram em dois modelos de algoritmo, k-NN e *tree*.

4.2.1 Algoritmo k-NN

A seção apresenta o modelo de classificação do algoritmo k-NN. Com a técnica foi criado o modelo de leitura da BD apresentado na Figura 21.

Figura 21 – Modelo de execução do k-NN.

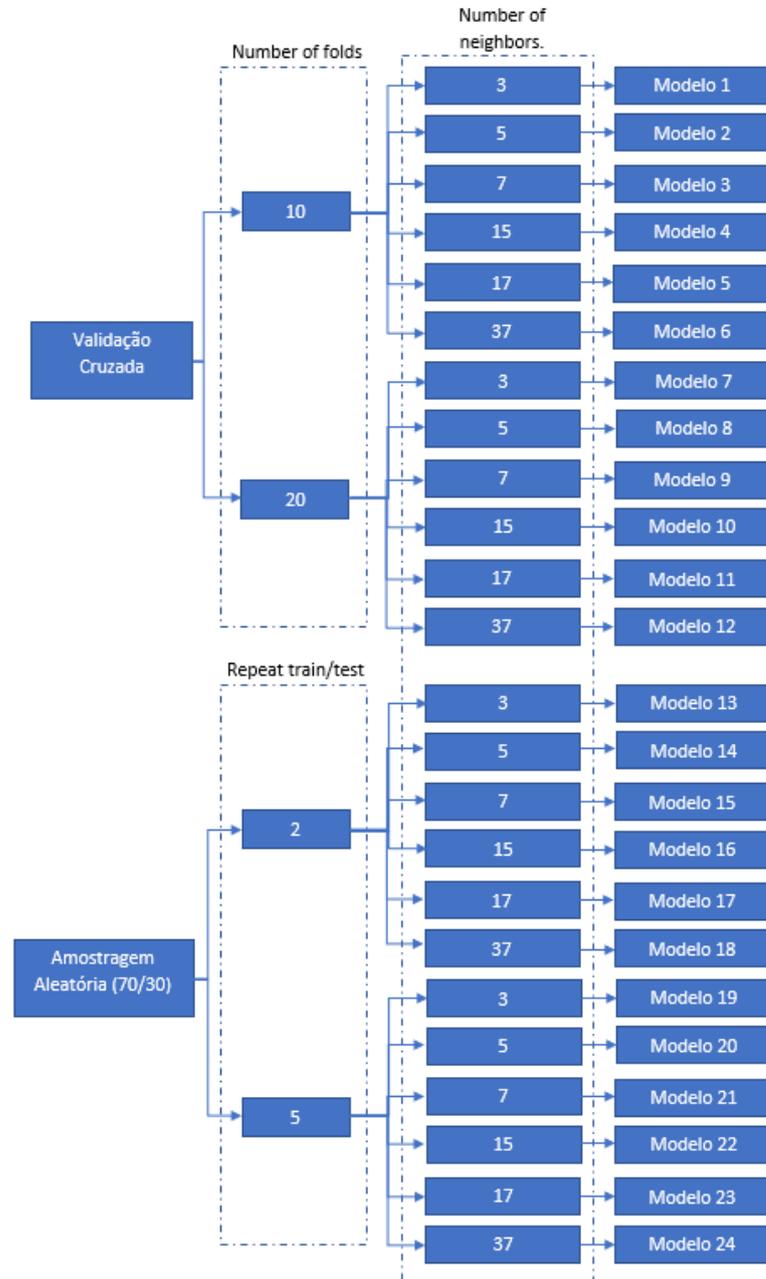


Fonte: Elaborada pelo autor.

Os experimentos realizados foram executados de duas maneiras, com dados categóricos qualitativos e quantitativos, onde os quantitativos foram normalizados em um intervalo $[0, 1]$. Os tratamentos dos testes foram realizados por dois aspectos de parâmetros, com validação cruzada e amostragem aleatória. Obteve uma configuração do algoritmo de

métrica euclidiana e seu peso uniforme, todos os pontos ponderados igualmente. A Figura 22 apresenta a divisão dos modelos para realização dos testes.

Figura 22 - Classificação do experimento do k-NN.

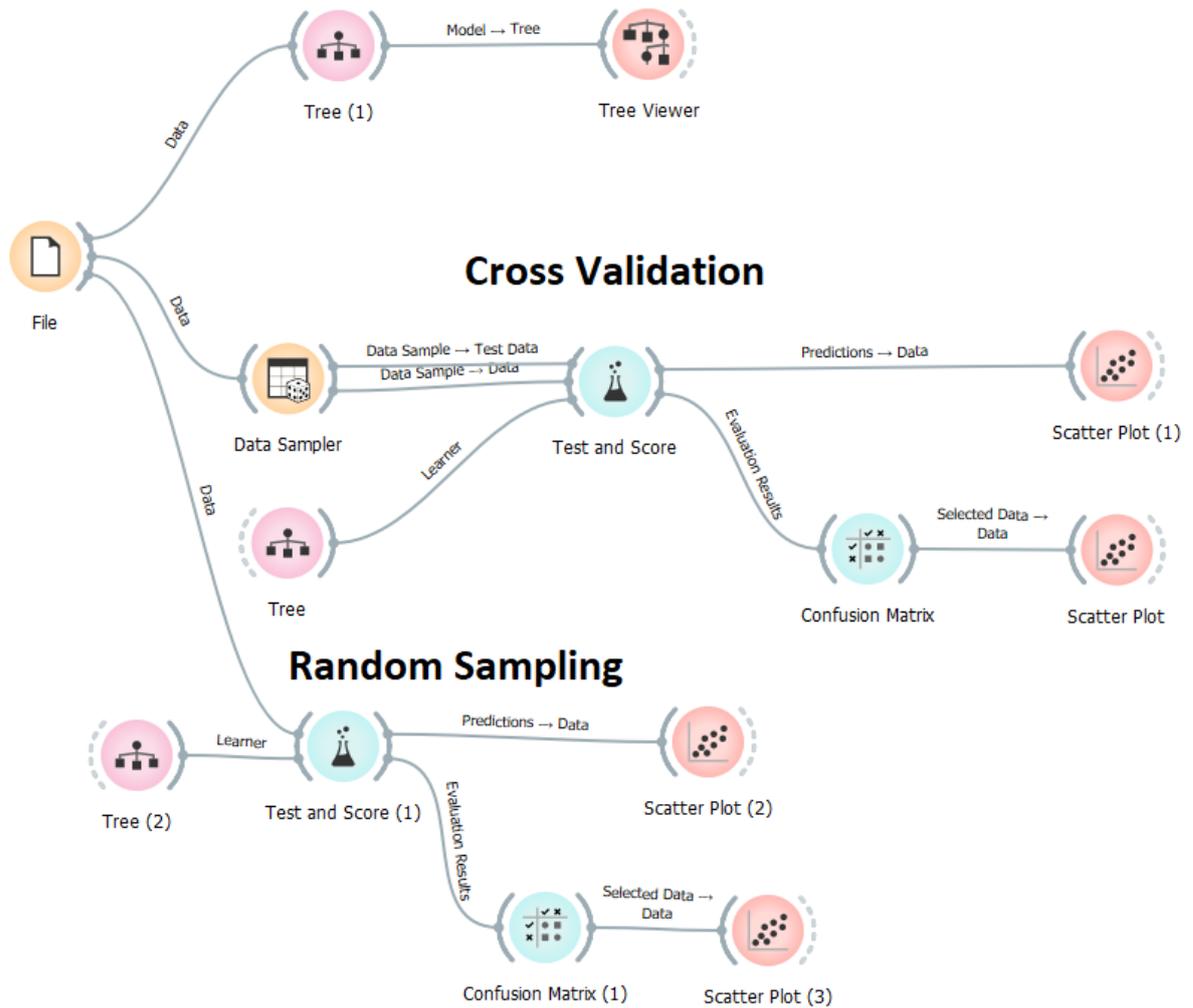


Fonte: Elaborada pelo autor.

4.2.2 Análise dos dados com *Tree*

A execução do algoritmo da árvore de decisão será apresentada nesse tópico. A Figura 23 apresenta o diagrama de execução do modelo.

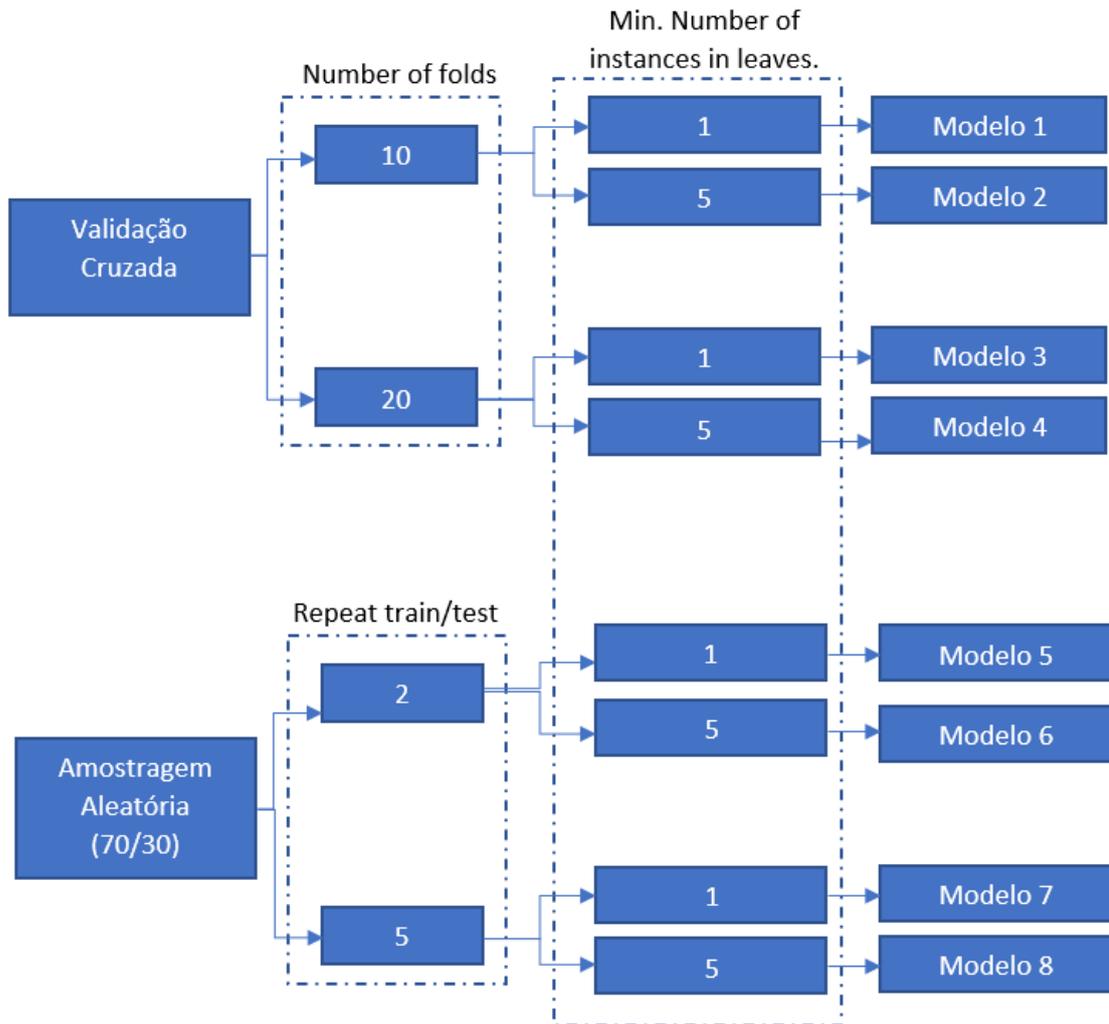
Figura 23 - Modelo de execução da *tree*.



Fonte: Elaborada pelo autor.

A execução do modelo será dividida em seis modelos de conjuntos. Esse formato se torna interessante para validação dos parâmetros de entrada. Sua estrutura segue descrita na Figura 24.

Figura 24 - Classificação do experimento da *tree*.



Fonte: Elaborada pelo autor.

Esse modelo não diferencia seu desempenho em dados quantitativos ou qualitativos, com isso a análise utilizará dados qualitativos. Em sua execução, foi testado com o parâmetro de parada quando a maioria atingir 100%.

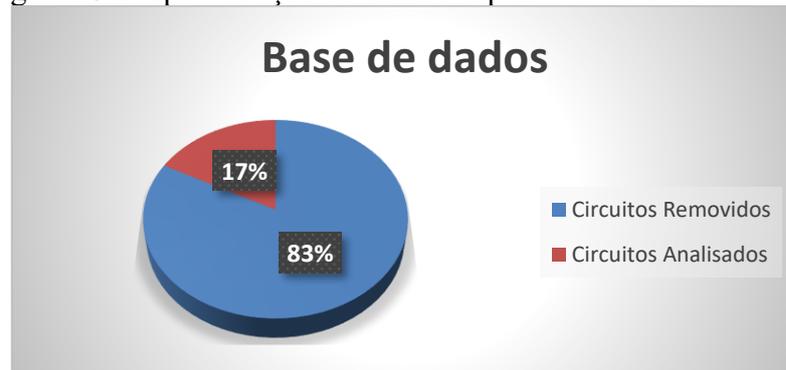
5 ANÁLISE DOS RESULTADOS

Os dados serão analisados nessa seção. Nela será mostrado os dados quantitativos da base e os dados de classificação dos algoritmos.

5.1 ANÁLISE PRELIMINAR DA BASE

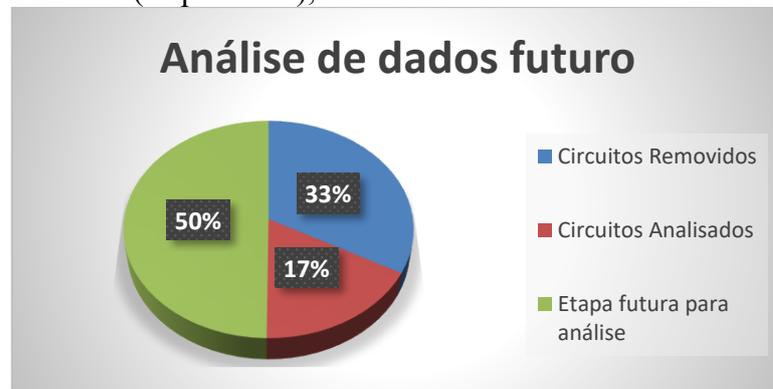
Analisando as informações obtidas na BD, foram adquiridos duzentos e cinquenta e cinco mil, duzentos e trinta experimentos realizados (255.230), mas o foco do estudo da pesquisa é analisar circuitos 1:1, 1:0 e 0:1. Desta forma, será trabalhado somente com 17,24% das informações, ou seja, 43.992 experimentos. As Figuras 25 e 26 apresentam o levantamento em forma de gráfico da base.

Figura 25 – Apresentação da análise superficial na base de dados.



Fonte: Elaborada pelo autor.

Figura 26 – Análise da amostra de dados entre informações trabalhadas, informações a serem trabalhadas (etapa futura), e dados desconsiderados no momento.



Fonte: Elaborada pelo autor.

Com a realização da classificação supervisionada foram analisados inicialmente os experimentos da Tabela 1 do E1 ao E24 a mesma tabela foi explicada no capítulo da revisão da literatura. Posteriormente obteve o incremento de mais análises assim anexando até o E56 conforme a Figura 27, e a Figura 28 apresenta graficamente a quantidade de classificação dos experimentos.

Figura 27 – Tabela de classificação de experimentos 1:1, 1:0 e 0:1.

Power	Componente Resistor	DMM Connection	CABLE	CONTROL KNOB					
				OFF	VAC/AAV	DIODE	VCC(V/mV)	Ω	A
O N	O N	Serial	Ω/V	E31 0			E1 172	E9 66	E17 370
			A	E32 23			E2 318	E10 66	E18 1808
		Parallel	Ω/V	E33 3			E3 301	E11 55	E19 13
			A	E34 0			E4 9	E12 4	E20 13
O F F	O N	Serial	Ω/V	E35 8			E5 33	E13 363	E21 103
			A	E36 19			E6 73	E14 45	E22 603
		Parallel	Ω/V	E37 7			E7 159	E15 84	E23 22
			A	E38 0			E8 3	E16 4	E24 11
N O	O N	Serial / Parallel	Ω/V	E39 134			E25 2055	E27 18300	E29 729
			A	E40 4			E26 73	E28 217	E30 165
O N	O F F	Serial / Parallel	Ω/V	E41 11			E42 984	E43 0	E44 29
			A	E45 5			E46 24	E47 6	E48 158
O F F	O F F	Serial / Parallel	Ω/V	E49 27			E50 613	E51 23	E52 30
			A	E53 1			E54 9	E55 0	E56 35

Fonte: Elaborada pelo autor.

Figura 28 – Apresentação da classificação em forma de gráfico dos experimentos.



Fonte: Elaborada pelo autor.

Os experimentos corretos correspondem aos códigos E3 (medindo tensão), E18 (medindo corrente), E27 (medindo resistência) e E42 (medindo tensão da fonte). O gráfico da Figura 29 apresenta que 86% dos circuitos corretos representam a medição de resistência, como a análise é de 1:1, 1:0 e 0:1 o E3 e o E42 ambos estão medindo tensão da fonte então 6% dos circuitos corretos estão medindo tensão da fonte e 8% medindo corrente.

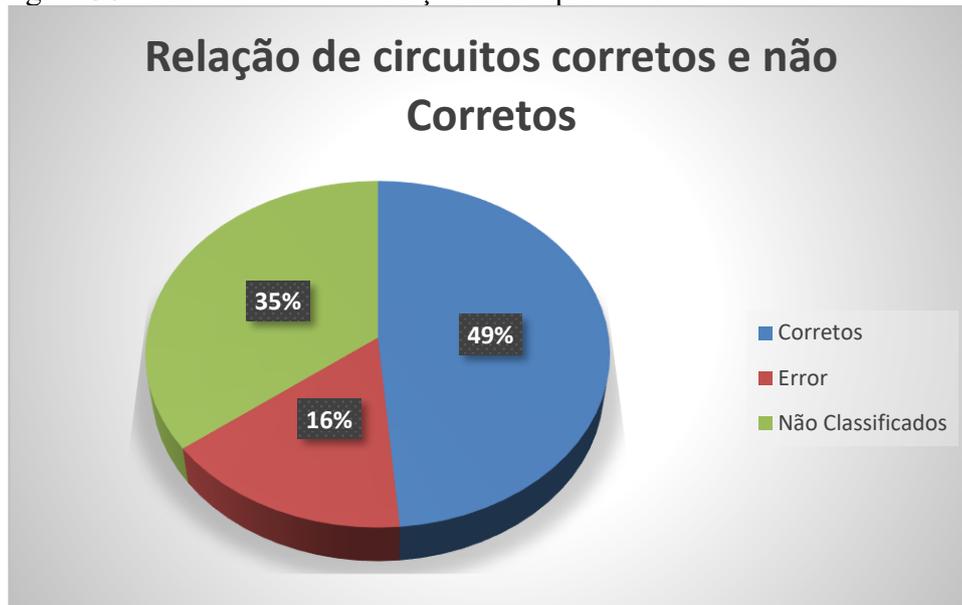
Figura 29 – Análise da proporção dos experimentos corretos.



Fonte: Elaborada pelo autor.

A tabela de classificação representado na Figura 27, não conseguiu classificar todos os 43.992 experimentos. Assim, 35% dos experimentos não conseguiram ser classificados. A classificação dos experimentos como “não classificados”, não foram analisados. A Figura 30 apresenta graficamente a divisão.

Figura 30 – Análise da classificação dos experimentos corretos e não corretos.



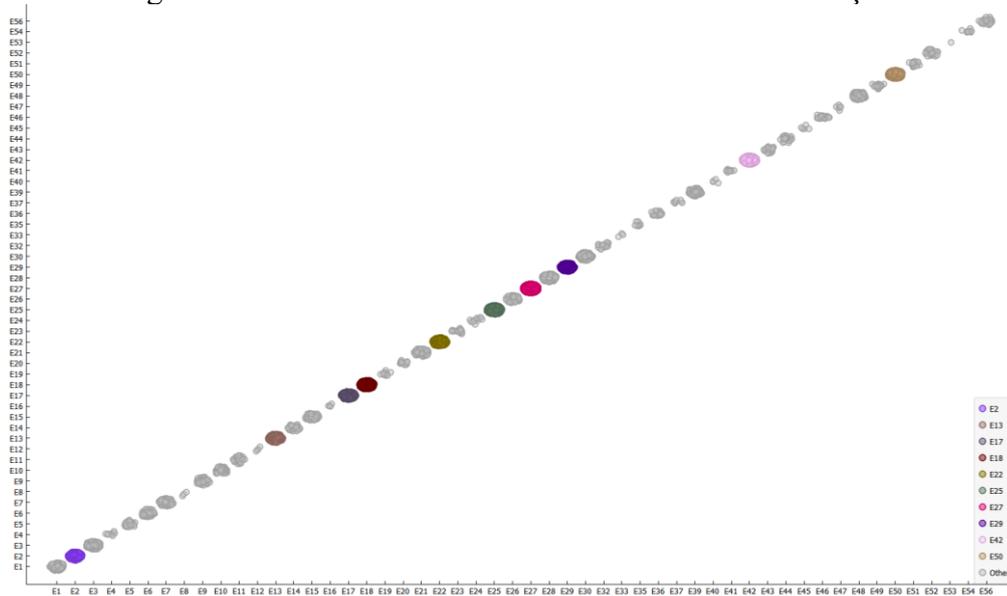
5.2 ANÁLISE DOS RESULTADOS COM ALGORITMO

Este capítulo apresenta as informações coletadas nas execuções dos algoritmos de classificação. A discussão dos resultados será apresentada no próximo capítulo da pesquisa.

A Figura 31 representa como seria a classificação com 100% de acerto, mas, é difícil um modelo ter esse grau de assertividade, então os dados serão divididos em três classes:

- Maior e igual que 0.9 (Excelente).
- Maior e igual a 0.7 e menor que 0.9 (Bom).
- Menor que 0.7 (Ruim).

Figura 31 - Gráfico demonstrativo de 100% de classificação.



Fonte: Elaborada pelo autor.

5.2.1 Análise dos Dados com k-NN

O estudo do algoritmo k-NN será apresentado em uma tabela constituída de cinco tópicos, são eles:

- Área sob ROC: representa a área sob a curva de operação.
- Acurácia: informa a porporção dos exemplos classificados corretamente.
- F-1: apresenta uma média harmônica ponderada.
- Precisão: é a proporção de verdadeiros entre os casos classificados como verdadeiros.
- *Recall*: é a proporção de verdadeiros positivos entre todas as instâncias positivas nos dados.

Os dados dos primeiros 24 modelos, apresentados na Tabela 4, estão analisados em informações qualitativa, esse modelo deverá fazer o algoritmo reduzir seu desempenho. A relação dos 24 modelos de teste serve para analisar os resultados a cada treinamento.

Tabela 4 - Análise dos testes do algoritmo k-NN com dados qualitativos.

Modelo	AUC	CA	F1	<i>Precision</i>	<i>Recall</i>
Modelo 1	0.995	0.987	0.986	0.986	0.987
Modelo 2	0.996	0.984	0.983	0.982	0.984
Modelo 3	0.996	0.982	0.979	0.979	0.982
Modelo 4	0.997	0.971	0.966	0.965	0.971
Modelo 5	0.997	0.969	0.964	0.961	0.969
Modelo 6	0.998	0.957	0.949	0.949	0.957
Modelo 7	0.995	0.988	0.986	0.986	0.988
Modelo 8	0.996	0.985	0.983	0.983	0.985
Modelo 9	0.996	0.982	0.980	0.980	0.982
Modelo 10	0.997	0.972	0.968	0.967	0.972
Modelo 11	0.997	0.970	0.965	0.964	0.970
Modelo 12	0.998	0.957	0.950	0.949	0.957
Modelo 13	0.994	0.987	0.986	0.986	0.987
Modelo 14	0.995	0.983	0.981	0.981	0.983
Modelo 15	0.996	0.980	0.978	0.977	0.980
Modelo 16	0.997	0.971	0.966	0.966	0.971
Modelo 17	0.997	0.968	0.963	0.962	0.968
Modelo 18	0.997	0.955	0.946	0.946	0.955
Modelo 19	0.995	0.987	0.986	0.985	0.987
Modelo 20	0.995	0.984	0.982	0.982	0.984
Modelo 21	0.996	0.981	0.979	0.979	0.981
Modelo 22	0.997	0.971	0.966	0.966	0.971
Modelo 23	0.997	0.970	0.965	0.965	0.970
Modelo 24	0.998	0.956	0.948	0.948	0.956

Fonte: Elaborada pelo autor.

Os outros 24 modelos apresentados na Tabela 5, estão em formato quantitativo normalizado.

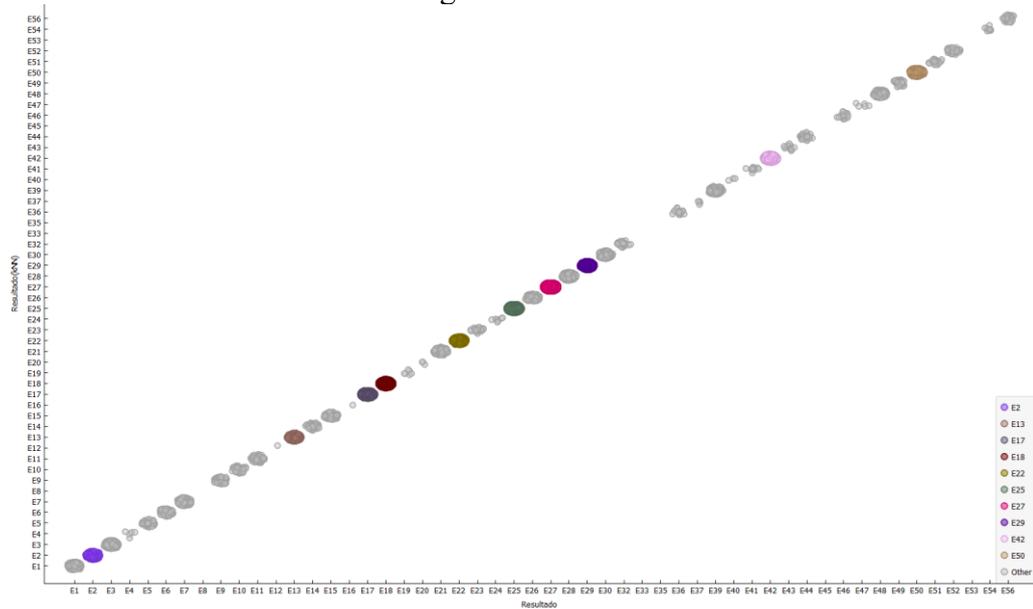
Tabela 5 - Análise dos testes do algoritmo k-NN com dados quantitativos.

Model	AUC	CA	F1	Precision	Recall
Modelo 1	0.996	0.991	0.990	0.990	0.991
Modelo 2	0.996	0.988	0.987	0.986	0.988
Modelo 3	0.997	0.987	0.985	0.984	0.987
Modelo 4	0.997	0.980	0.976	0.975	0.980
Modelo 5	0.997	0.977	0.974	0.972	0.977
Modelo 6	0.998	0.964	0.957	0.954	0.964
Modelo 7	0.996	0.991	0.991	0.990	0.991
Modelo 8	0.996	0.988	0.987	0.987	0.988
Modelo 9	0.997	0.987	0.985	0.984	0.987
Modelo 10	0.997	0.980	0.977	0.975	0.980
Modelo 11	0.997	0.978	0.975	0.973	0.978
Modelo 12	0.998	0.966	0.959	0.956	0.966
Modelo 13	0.996	0.990	0.990	0.990	0.990
Modelo 14	0.997	0.988	0.987	0.988	0.988
Modelo 15	0.997	0.986	0.985	0.985	0.986
Modelo 16	0.998	0.979	0.976	0.977	0.979
Modelo 17	0.998	0.977	0.974	0.973	0.977
Modelo 18	0.998	0.965	0.959	0.956	0.965
Modelo 19	0.996	0.991	0.991	0.991	0.991
Modelo 20	0.997	0.989	0.988	0.988	0.989
Modelo 21	0.997	0.986	0.985	0.985	0.986
Modelo 22	0.998	0.979	0.976	0.976	0.979
Modelo 23	0.998	0.977	0.974	0.973	0.977
Modelo 24	0.998	0.966	0.960	0.956	0.966

Fonte: Elaborada pelo autor.

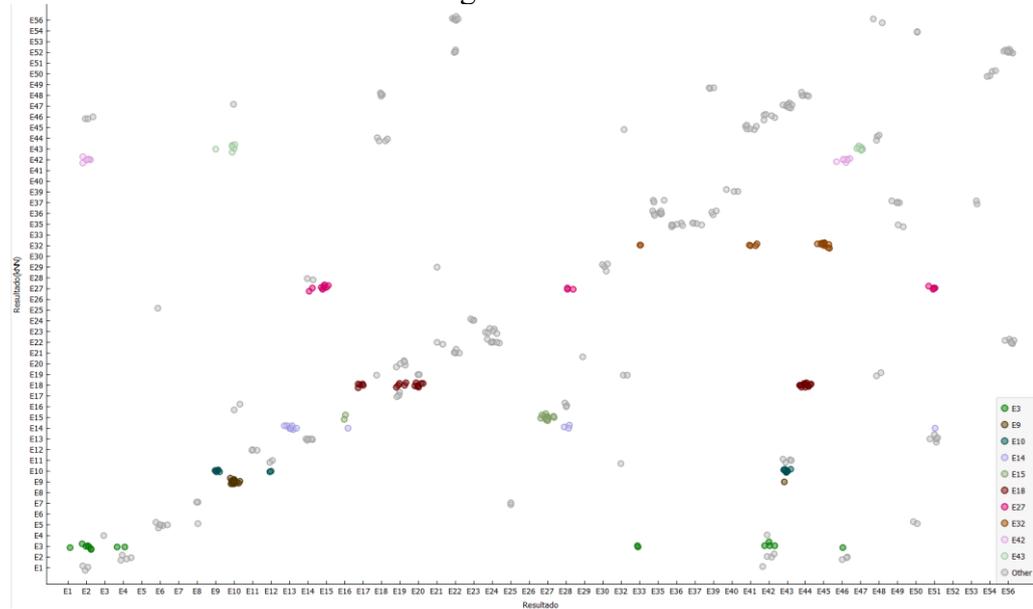
Analisando os dois modelos, tanto o qualitativo e o quantitativo, os seus resultados foram semelhantes. Em toda análise sua acurácia resultou maior que 0.9 podendo ser classificada como um excelente resultado. Consegue-se observar que conforme aumente k o algoritmo fica menos preciso, fazendo que aumente seu erro. As Figuras 32, 33 e 34 representam o resultado do modelo 19 quantitativo, ficando com melhor desempenho.

Figura 32 - Dados classificados corretamente pelo modelo 19 do experimento quantitativo do algoritmo k-NN.



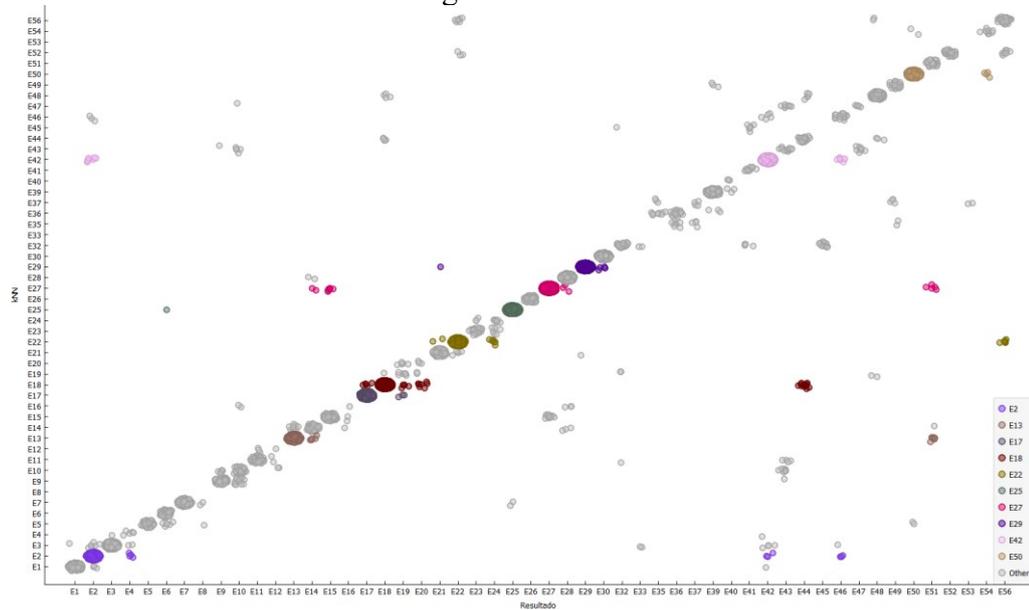
Fonte: Elaborada pelo autor.

Figura 33 - Dados classificados incorretamente pelo modelo 19 do experimento quantitativo do algoritmo k-NN.



Fonte: Elaborada pelo autor.

Figura 34 - Resultado da classificação geral pelo modelo 19 do experimento quantitativo do algoritmo k-NN



Fonte: Elaborada pelo autor.

5.2.2 Análise dos dados com *Tree*

A seguir apresenta-se duas tabelas (Tabelas 6 e 7), a primeira apresenta valores do treinamento de cada modelo, a segunda apresenta o número de nós e folhas para um valor mínimo de instâncias em cada folha.

Tabela 6 - Análise do desempenho de classificação do algoritmo *Tree*.

Model	AUC	CA	F1	Precision	Recall
1	0.992	0.990	0.990	0.990	0.990
2	0.994	0.977	0.975	0.973	0.977
3	0.992	0.990	0.990	0.990	0.990
4	0.995	0.978	0.976	0.974	0.978
5	0.991	0.990	0.989	0.989	0.990
6	0.992	0.973	0.970	0.970	0.973
7	0.992	0.990	0.990	0.990	0.990
8	0.994	0.977	0.975	0.974	0.977

Fonte: Elaborada pelo autor.

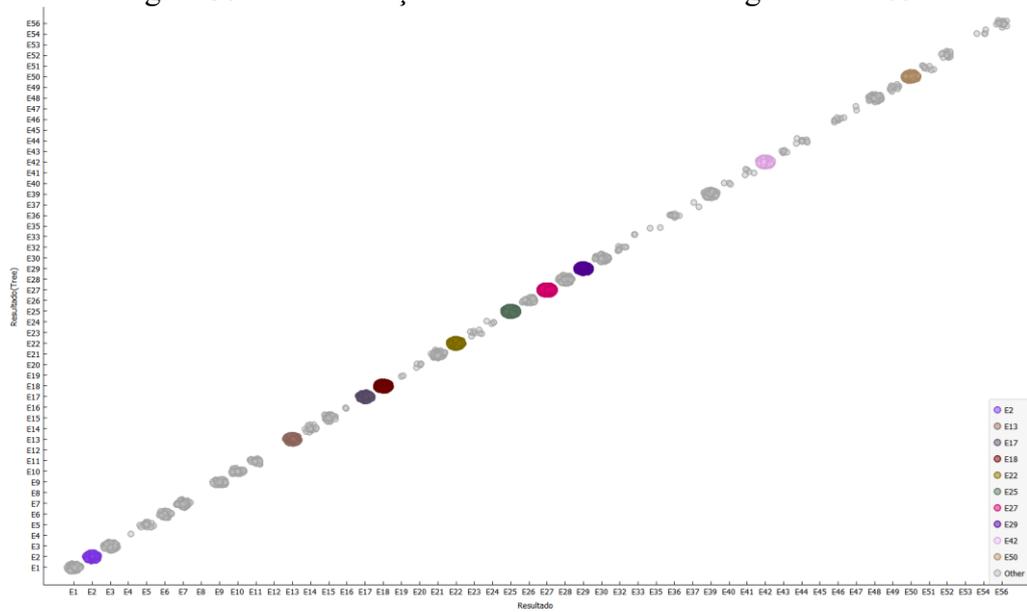
Tabela 7 - Análise de nós e folhas pelo algoritmo *Tree*.

Min. Number of Instances in Leaves	Nós	Folhas
1	890	549
5	437	262

Fonte: Elaborada pelo autor.

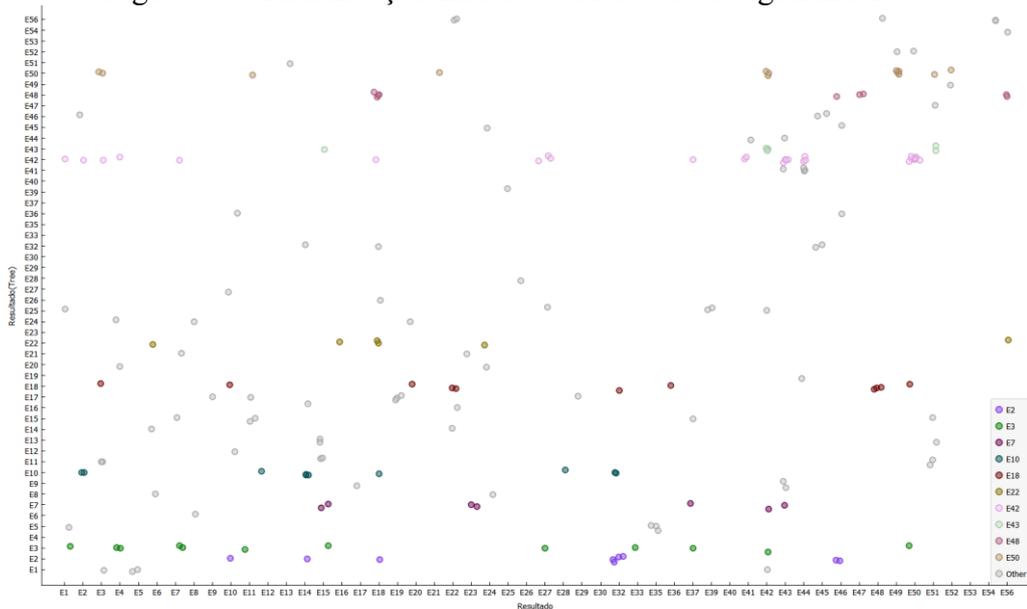
O algoritmo *Tree* nos 8 modelos de treinamento conquistou dados semelhantes entre eles. A acurácia ficou acima de 0.9, classificando o algoritmo como excelente para esse modelo. Os modelos 1, 3 e 7 conquistaram o melhor resultado, e o modelo 6 conquistou o pior. As Figuras 35, 36 e 37 apresentam o resultado em formato gráfico do modelo 1, sua apresentação está orientada pela seguinte sequência: somente classificação dos dados corretos; somente classificação dos dados incorretos; e classificação mista.

Figura 35 - Classificação correta dos dados do algoritmo *Tree*.



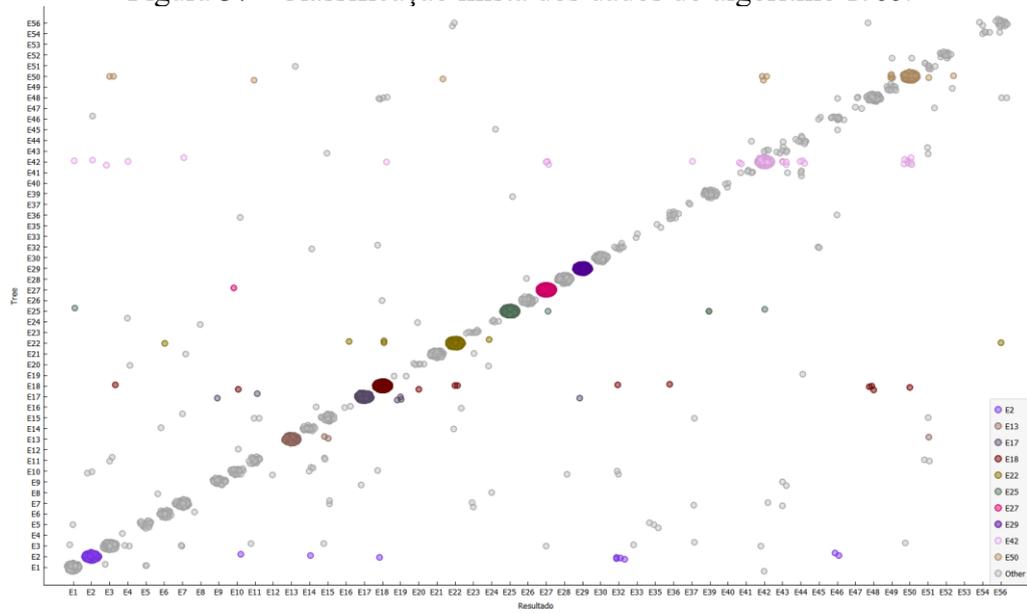
Fonte: Elaborada pelo autor.

Figura 36 - Classificação incorreta dos dados do algoritmo *Tree*.



Fonte: Elaborada pelo autor.

Figura 37 - Classificação mista dos dados do algoritmo *Tree*.



Fonte: Elaborada pelo autor.

6 DISCUSSÃO DOS RESULTADOS

O propósito da pesquisa está em adquirir informações da BD que podem ajudar em tomadas de decisões. A Tabela 8 apresenta os erros em que colocaria o sistema em curto-circuito. O erro em curto-circuito representa 1.68% dos dados. O experimento crítico ocorre quando a fonte está conectada e ligada. Considerando essas informações em um laboratório tradicional, ocorreria risco de danificar os equipamentos ao longo do tempo. O VISIR é facilmente configurável para não executar experimentos que venha a gerar um curto-circuito em seu *hardware* (geralmente a corrente de saída da fonte fica limitada em 100 mA), assim informando o problema ao usuário.

Tabela 8 - Dados em curto-circuito.

DMM Connection	CABLE	CONTROL KNOB					
		OFF	VAC/AAV	DIODE	VCC(V/mV)	Ω	A
Serial	Ω/V						
	A						
Parallel	Ω/V						
	A	E34			E4	E12	E20
Serial	Ω/V						
	A						
Parallel	Ω/V						
	A	E38			E8	E16	E24
Serial / Parallel	Ω/V						
	A	E40			E26	E28	E30
Serial / Parallel	Ω/V						
	A	E45			E46	E47	E48
Serial / Parallel	Ω/V						
	A	E53			E54	E55	E56

Fonte: Elaborada pelo autor.

Apresentou-se dois modelos de classificação: *k*-NN e *Tree*. Pode-se denominar o *k*-NN como uma caixa preta, onde sua classificação é executada por cálculos matemáticos, onde não se tem uma saída por regra de classificação. A *tree* consegue entregar o modelo de classificação, no qual o modelo fica visível.

Independente do modelo, os dados são classificados como excelentes, todas as acurácias dos algoritmos ficaram acima de 0.9. A escolha do melhor modelo se torna

complicada no momento, por duas questões: existem diversos circuitos a serem classificados em trabalhos futuros, e ambos tiveram ótimos desempenhos.

A escolha desses dois modelos ocorreu pelo fato de serem os mais simples e práticos, outros algoritmos poderiam ser testados, como regressão logística ou rede neural entre outros. Mas como os valores foram excelentes não foi necessário testar mais modelos no momento.

A comparação do algoritmo k-NN executando modelo qualitativo para o modelo quantitativo apresentados nas Figuras 38 e 39, realmente perde muito desempenho quando se trata de modelos qualitativos. O desempenho ficou 84.62% menor no tempo de testes e 60.07% menor no tempo de treino com o modelo quantitativo.

Figura 38 - Resultados de análise do modelo qualitativo com algoritmo k-NN.

Model	Train time [s]	Test time [s]	AUC	CA	F1	Precision	Recall
kNN	40.018	132.415	0.998	0.957	0.950	0.949	0.957

Fonte: Elaborada pelo autor.

Figura 39 - Resultados de análise do modelo quantitativo com algoritmo k-NN.

Model	Train time [s]	Test time [s]	AUC	CA	F1	Precision	Recall
kNN	15.979	20.359	0.998	0.966	0.959	0.956	0.966

Fonte: Elaborada pelo autor.

7 CONCLUSÃO

Com a necessidade dos avanços das aplicações TIC no mundo global, principalmente no ensino-aprendizagem que anos atrás não se preocupavam com a integração tecnológica, hoje é possível obter ferramentas de análise de desempenhos superiores do que simplesmente notas de avaliações. Com elas podem ser analisadas tempo de estudo em ambiente virtual, desempenho nas atividades, suas maiores dificuldades de aprendizado, rotina de estudo, entre tantas outras informações.

Os laboratórios remotos facilitam sua instalação com o fato de ocupar um pequeno espaço físico e suas atividades podem ser compartilhadas com a comunidade. O núcleo do RExLab (UFSC) possui diversos experimentos remotos compartilhados com a comunidade com a utilização da plataforma RELLE, no qual o laboratório remoto VISIR está disponível na plataforma.

Este trabalho teve o objetivo proposto, de analisar a base de dados do laboratório remoto VISIR e classificar os experimentos iniciais de 1:1, 1:0 e 0:1, pode-se dizer que os seus resultados foram excelentes. Em todos os modelos de classificação sua acurácia ficou acima de 0.9 mantendo uma grande porcentagem de acerto.

Antes de qualquer aplicação de algoritmo de aprendizagem de máquina ser aplicado, é de extrema importância que os dados sejam analisados. Existe dois modelos para analisar, a primeira é por técnicas estatísticas e a segunda no qual utilizamos no trabalho foi a de visualização. Conseguindo uma melhor compreensão dos dados e como poderia ser trabalhado.

O formato de tratamento de dados ignorando os pontos da matriz fazendo a otimização dos circuitos resultou na obtenção de um resultado excelente. Isso ocorreu pelo fato de não importar a região da ligação dos componentes, caso considerasse os algoritmos tentaria mapear um padrão de região de ligação e não o modelo de análise de circuitos que é o foco da pesquisa.

O processo de aprendizado do algoritmo torna-se fácil, se os dados forem pré-processados corretamente. Foi comprovado que o algoritmo k-NN perde drasticamente seu desempenho trabalhando com dados qualitativos, assim, deve-se entender qual tipo de algoritmo será utilizado e realizar seu pré-processamento das informações corretamente.

O objetivo desse estudo era compreender os experimentos que estavam sendo realizados no laboratório remoto VISIR, pois, atualmente não se possui indicadores. Com isso foi possível realizar 56 classificações dentro da delimitação da pesquisa. Observando que conseguiu-se obter os experimentos que poderiam criar um curto-circuito na bancada remota.

A próxima etapa da pesquisa futura é analisar os dados com mais de um resistor e realizar sua classificação. Fazendo uma analogia com os dados atuais e os dados futuros, a classificação de medição de resistência deve prevalecer com maior volume, pelo fato de realizar uma medição de resistência para cada resistor e comparar sua tolerância com comprovação dos dados teóricos sendo justificados em relatórios de laboratórios.

O laboratório remoto não veio para disputar lugar com um laboratório tradicional, ele veio para agregar ferramentas para capacitação de alunos. Com sua utilização, facilita a coletas de dados, utilizando técnicas de mineração, faz com que as informações coletadas se transformem em dados analíticos que informa ao professor o aprendizado do aluno.

REFERÊNCIAS

- ALEXANDER, Charles K.; SADIKU, Matthew N. O.. **Fundamentos de Circuitos elétricos**. 5. ed. Porto Alegre: Amgh Editora Ltda, 2013. 874 p.
- ALVES, G. R., GARCÍA-ZUBÍA, J., CASTRO, M., FIDALGO, A., HERNÁNDEZ-JAYO, U., GARCÍA-LORO, F. MARQUES, M. A., KREITER, C. (2022) **A roadmap for the VISIR remote lab**. European Journal of Engineering Education. In press.
<https://doi.org/10.1080/03043797.2022.2054312>
- AVELLA, J., KEBRITCHI, M., NUNN, S. & KANAI, T. (2016). **Learning analytics methods, benefits and challenges in higher education**. A systematic literature review. Online Learning Journal, 20(2), p-17.
- BENCOMO, S.D. (2004). **Control learning: present and future**. *Annual Reviews in Control*. Vol. 28(1), pp. 115-136.
- BOHUS, C., CROWL, L.A., AKTAN, B. e SHOR, M.H. (1996). **Running Control Engineering Experiments Over the Internet**. In IFAC Proceedings Volumes. Volume 29 (1), pp. 2919-2927.
- CARLOS, Lucas Mellos. **Arquitetura para análise de aprendizagem no uso de laboratórios remotos..** 2020. 112 f. Dissertação (Mestrado) - Curso de Programa de Pós-Graduação em Tecnologias da Informação e Comunicação, Universidade Federal de Santa Catarina, Araranguá, 2020. Disponível em:
<https://repositorio.ufsc.br/handle/123456789/216386?msckid=d1b86062b34f11eca04938f3682c34e6>. Acesso em: 1 mar. 2022.
- CARVALHO, Rafael. (2016). **Design instrucional: tudo sobre o design educacional**. Edools. Disponível em: <https://www.edools.com/design-instrucional/>. Acesso em: 26/03/2022.
- CASTRO, Leandro Nunes de; FERRARI, Daniel Gomes. **Introdução à Mineração de Dados: conceitos básicos, algoritmos e aplicações**. São Paulo: Saraiva, 2016. 351 p.
- CHAPMAN, Pete et-al. **CRISP-DM 1.0**. Disponível em: <https://www.the-modeling-agency.com/crisp-dm.pdf>. Acesso em: 1 mar. 2022.
- CHATTI, M. A.; DYCKHOFF, A. L.; SCHROEDER, U.; THÜS, H. A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, v. 4, n. 5, p. 318-331, 2012.
- CLAESSON, L. e HAKANSSON, L. (2012). **Using an On-line Remote Laboratory for Electrical Experiments in Upper Secondary Education**. *International Journal of On-line Engineering (iJOE)*, 8 (2), pp. 24-30.

CLEMENTE, F. J. G. et al. Collecting experience data from remotely hosted learning applications. In: AUER, M. E.; ZUTIN, D. G. (Ed.). *Online Engineering & Internet of Things: Proceedings of the 14th International Conference on Remote Engineering and Virtual Instrumentation REV 2017*. Cham: Springer International Publishing, 2018. p. 170–181. ISBN: 978-3-319-64352-6.

CLOW, Doug. The learning analytics cycle. **Proceedings Of The 2Nd International Conference On Learning Analytics And Knowledge**, [S.L.], 29 abr. 2012. ACM. <http://dx.doi.org/10.1145/2330601.2330636>. Disponível em: <http://oro.open.ac.uk/34330/1/LAK12-DougClow-personalcopy.pdf>. Acesso em: 26 fev. 2022.

COSTA, Evandro; BAKER, Ryan S.J.D.; AMORIM, Lucas; MAGALHÃES, Jonathas; MARINHO, Tarsis. **Mineração de Dados Educacionais: Conceitos, Técnicas, Ferramentas e Aplicações**. 2012. Disponível em: <https://www.br-ie.org/pub/index.php/pie/article/view/2341>. Acesso em: 27 mar. 2022.

COSTA-LOBO, M.C., Alves, Gustavo R., Marques, M. A., Viegas, C., Barral, R. G., Couto, R. J., Jacob, F. L., Ramos, C. A., Vilão, G. M., Covita, D. S., Alves, Joaquim, Guimarães. P. S., Gustavsson, Ingvar (2011). Using remote experimentation in a large undergraduate course: initial findings. *Proceedings 41st ASEE/IEEE Frontiers in Education Conference (FIE 2011)*, USB version, IEEE Catalog Number: CFP11FIE-USB. ISBN: 978-1-61284-467-1.

DAHLSTROM, E., BROOKS, C., GRAJEK, S., REEVES, J. (2015). **Study of students and information technology**. Research Report. Lousville, CO: ECAR.

DINIZ, Yasmine. **Entenda o que é Learning Analytics e para que serve**. 2020. Disponível em: <https://educacao.imagine.com.br/learning-analytics/>. Acesso em: 27 mar. 2022.

EVANS, Dave. **A Internet das Coisas: como a próxima evolução da internet está mudando tudo**. Como a próxima evolução da Internet está mudando tudo. 2011. Disponível em: https://www.cisco.com/c/dam/global/pt_br/assets/executives/pdf/internet_of_things_iot_ibsg_0411final.pdf. Acesso em: 20 mar. 2022.

FACELI, Katti *et al.* **Inteligência Artificial: uma abordagem de aprendizado de máquina**. Rio de Janeiro: Ltc, 2015. 394 p.

FEISEL, L.D., e PETERSON, G.D. (2002). A Colloquy on Learning Objectives for Engineering Education Laboratories. In *Proceedings of the 2002 American Society for Engineering Education Annual Conference & Exposition (Montreal, Canada, 2002)*.

FEISEL, L.D., e ROSA, A. (2005). The role of the laboratory in undergraduate engineering education. *Journal of Engineering Education*, Vol. 94, Issue 1, pp. 121-130.

FERGUSON, Rebecca. Learning analytics: Drivers, developments and challenges. **International Journal of Technology Enhanced Learning**, Inderscience Publishers, v. 4, n. 5-6, p. 304–317, set. 2012. ISSN 17535255. DOI: 10.1504/IJTEL.2012.051816. Disponível em: <http://dx.doi.org/10.1504/IJTEL.2012.051816>.

FORBES. **Um ano de pandemia: como a tecnologia impactou a vida de seis executivos da indústria de inovação:** do trabalho à resolução de problemas pessoais, eles recorreram a ferramentas de todos os tipos para enfrentar a crise sanitária. 2021. Disponível em: <https://forbes.com.br/forbes-tech/2021/03/um-ano-de-pandemia-como-a-tecnologia-impactou-a-vida-de-seis-executivos-da-industria-de-inovacao/>. Acesso em: 14 fev. 2022.

GARCÍA-ZUBÍA, J. et al. (2021). Using VISIR Remote Lab in the Classroom: Case of Study of the University of Deusto 2009–2019. In: Auer, M., May, D. (eds) Cross Reality and Data Science in Engineering. REV 2020. Advances in Intelligent Systems and Computing, vol 1231. Springer, Cham. https://doi.org/10.1007/978-3-030-52575-0_7

GARCÍA-ZUBÍA, Javier; ALVES, Gustavo R.; HERNANDEZ-JAYO, Unai; CUADROS, Jordi; SERRANO, Vanessa; FIDALGO, Andre. A framework for interpreting experimental errors in VISIR. **2019 5th Experiment International Conference (Exp.At'19)**, [S.L.], jun. 2019. IEEE. <http://dx.doi.org/10.1109/expat.2019.8876568>. Disponível em: <https://ieeexplore.ieee.org/document/8876568/authors#authors>. Acesso em: 19 fev. 2022.

GIL, Antônio Carlos. Como elaborar projetos de pesquisa. 4. ed. São Paulo: Atlas, 2002.

GONÇALVES, Alexandre L. *et al.* **Learning Analytics and Recommender Systems toward Remote Experimentation.** 2018. Disponível em: https://www.researchgate.net/publication/327347249_Learning_Analytics_and_Recommender_Systems_toward_Remote_Experimentation. Acesso em: 14 fev. 2022.

GUSTAVSON, I., ZACKRISSON, J. e LUNDBERG, J. (2014). **VISIR work in progress.** In Proceedings of 2014 IEEE Global Engineering Education Conference (EDUCON 2014). (Istanbul, Turkey, 2014).

KULESZA, W. et al., "A federation of VISIR remote laboratories through the PILAR Project," 2017 4th Experiment@International Conference (exp.at'17), 2017, pp. 28-32, doi: 10.1109/EXPAT.2017.7984407.

LANG, D., MENGELKAMP, C., JÄGER, R., GEOFFROY, D., BILLAUD, M. e ZIMMER, T. (2007). **Pedagogical evaluation of remote laboratories in eMerge project.** *European Journal of Engineering Education*, Vol. 32(1), pp. 57-72.

LIMA, Natércia et al. VISIR's usage as an educational resource. Proceedings Of The Fourth International Conference On Technological Ecosystems For Enhancing Multiculturality - Teem '16, [s.l.], p.1-2, 2016. ACM Press. <http://dx.doi.org/10.1145/3012430.3012623>.

LINDSAY, E. e GOOD, M. (2005). **Effects of Laboratory Access Modes Upon Learning Outcomes.** *IEEE Transactions on Education*, 48(4), pp. 619 – 631.

LUZ, Filipe. **Algoritmo KNN para classificação.** 2019. Disponível em: <https://inferir.com.br/artigos/algoritmo-knn-para-classificacao/>. Acesso em: 24 maio 2022.

MA, J. e NICKERSON, J. (2006). **Hands-on, simulated, and remote laboratories.** *ACM Computing Surveys*, 38(3), n. pag.

MARCONI, M. de A.; LAKATOS, E. M. *Fundamentos de Metodologia Científica*. 5. ed. São Paulo: Editora Atlas, 2003. 311 p.

MINITAB. **Por que a análise de dados é fundamental para o sucesso do seu negócio:** software analítico permite extrair inteligência de grande volume de informações e, consequentemente, tomar melhores decisões. 2019. Disponível em: <https://forbes.com.br/brand-voice/2019/06/por-que-a-analise-de-dados-e-fundamental-para-o-sucesso-do-seu-negocio/>. Acesso em: 14 fev. 2022.

MOISSA, Barbara; GASPARINI, Isabela; KEMCZINSKI, Avanilde. **Learning Analytics: um mapeamento sistemático**. 2014. Disponível em: http://www.tise.cl/volumen10/TISE2014/tise2014_submission_262.pdf. Acesso em: 27 mar. 2022.

NAEF, O. (2006). Real Laboratory, virtual laboratory or remote laboratory: what is the most effective way? *Journal of On-line Engineering*, Vol 2 (3), n. pag.

PEREIRA, Josiel, Nardi da Silva, Isabela, José Pedro Schardosim Simão, Lucas Mellos Carlos, Juarez Bento Silva, Simone Meister Sommer Bilessimo, João Bosco da Mota Alves, Gustavo R. Alves. (2017) Modelo de Repositórios de Práticas Didáticas de Circuitos Elétricos e Eletrônicos Utilizando o Laboratório Remoto VISIR. XLV Congresso da Associação Brasileira de Ensino de Engenharia (COBENGE2017), Joinville, SC, Brasil, 26-29 Setembro 2017

PPGTIC. **Linhas de Pesquisa**. [S.I]. Disponível em: <https://ppgtic.ufsc.br/linhas-de-pesquisa/>. Acesso em: 14 fev. 2022.

PPGTIC. **Programa de Pós-graduação em Tecnologias da Informação e Comunicação**. [S.I]. Disponível em: <https://ppgtic.ufsc.br/sobre-o-ppgtic/>. Acesso em: 14 fev. 2022.

PRETNAR, Ajda. **The Mystery of Test & Score**. 2019. Disponível em: <https://orangedatamining.com/blog/2019/1/28/the-mystery-of-test-and-score/>. Acesso em: 25 maio 2022

RIENTIES, B., CROSS, S. & ZDRAHL, Z. (2017). **Implementing a learning analytics intervention and evaluation framework: what works?**. In: Kei Daniel B. (eds) *Big Data and Learning Analytics in Higher Education*. Springer. DOI https://doi.org/10.1007/978-3-319-06520-5_10.

ROCHADEL, W., da Silva, J. B., Simão, J. P. S. and Alves, G. R. (2013). Educational application of remote experimentation for mobile devices. 2013 10th International Conference on Remote Engineering and Virtual Instrumentation (REV), 2013, pp. 1-6, doi: 10.1109/REV.2013.6502905.

RODRIGUES, Haroldo. **Como a tecnologia pode impactar a educação no século 21**. 2021. Disponível em: <https://forbes.com.br/forbes-collab/2021/08/haroldo-rodriques-como-a-tecnologia-pode-impactar-a-educacao-do-seculo-21/>. Acesso em: 14 fev. 2022.

ROQUE, Gabriela Rocha et al. **CONSTRUÇÃO DO MODELO SECI NO PROJETO VISIR+: UM ESTUDO DE CASO DAS PRÁTICAS E INICIATIVAS DE COMPARTILHAMENTO DE CONHECIMENTO INTERORGANIZACIONAL**. In: CONGRESSO BRASILEIRO DE EDUCAÇÃO EM ENGENHARIA, 46., 2018, Salvador/ba. ANAIS DO COBENGE 2018. Salvador/ba: Abenge, 2018. p. 1 - 10.

RUSSELL, Stuart; NORVIG, Peter. **Inteligência Artificial**. 3. ed. Rio de Janeiro: Elsevier, 2013. 1016 p.

SALAH, Razwan Mohmed, Alves, Gustavo R., Guerreiro, Pedro, Gustavsson, Ingvar. Using UML Models to Describe the VISIR System. 2016. International Journal of Online Engineering (iJOE), Vol. 12, No 06, pp. 34-42. <https://doi.org/10.3991/ijoe.v12i06.5707>

SILVA, Edna Lúcia da; MENEZES, Estera Muszkat. **Metodologia da Pesquisa e Elaboração de Dissertação**. 2005. Disponível em: https://tccbiblio.paginas.ufsc.br/files/2010/09/024_Metodologia_de_pesquisa_e_elaboracao_d_e_teses_e_dissertacoes1.pdf. Acesso em: 28 fev. 2022.

SILVA, Isabela Nardi da. **COMUNIDADE INTERNACIONAL DE PRÁTICAS PARA COMPARTILHAMENTO DE EXPERIÊNCIAS ENTRE DOCENTES USUÁRIOS DO LABORATÓRIO REMOTO VISIR**. 2019. 153 f. Dissertação (Mestrado) - Curso de Programa de Pós-Graduação em Tecnologias da Informação e Comunicação, Universidade Federal de Santa Catarina, Araranguá, 2019. Disponível em: <https://bases.bu.ufsc.br/repositorio-institucional-ufsc/>. Acesso em: 14 fev. 2022.

SILVA, J. B.; BILESSIMO, S.M.S. ; SCHEFFER, G. R. ; SILVA, I. N. . **Laboratórios Remotos como Alternativa para Atividades Práticas em Cursos na Modalidade EAD**. EAD EM FOCO, v. 10, p. 1-26, 2020. Homepage: <http://https://eademfoco.cecierj.edu.br/index.php/Revista/article/view/942>; Série: 2; ISSN/ISBN: 21778310.

SILVA, Mariana M.; FIDALGO, André V.; MARQUES, Arcelina; ALVES, Gustavo R.; SALAH, Razwan M.; JACOB, Frederico L., "A comprehensive VISIR bibliographical reference", in WEEF-GEDC 2021, World Engineering Education Forum; International Federation of Engineering Education Societies; The Global Engineering Deans Council, 15-18 November 2021, Madrid, Spain. <https://doi.org/10.1109/WEEF/GEDC53299.2021.9657332>

SILVA, I. N., PEREIRA, J., BILESSIMO, S., ALVES, G., & SILVA, J. B. (2018a). O LABORATÓRIO VISIR: UMA REVISÃO SISTEMÁTICA. ANAIS DO II SIMPÓSIO ÍBERO AMERICANO DE TECNOLOGIAS EDUCACIONAIS (pp. 1-10). ARARANGUÁ/SC: HARD TECH INFORMÁTICA.

SIMÃO, José Pedro Schardosim. **MODELO PARA REGISTRO DE DADOS DE EXPERIÊNCIA DE APRENDIZAGEM EM LABORATÓRIOS REMOTOS**. 2018. 115 f. Monografia (Especialização) - Curso de Programa de Pós-Graduação em Tecnologias da Informação e Comunicação, Universidade Federal de Santa Catarina, Araranguá, 2018. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/191118/PTIC0031-D.pdf?sequence=-1&isAllowed=y>. Acesso em: 19 fev. 2022.

TAVARES, Isabelli Sasdelli; ALVES, Gustavo R.; VALENTE JUNIOR, Wilson; SCHLICHTING, Luis Carlos Martinhago. CARACTERIZAÇÃO DE ERROS EXPERIMENTAIS EM CIRCUITOS ELETRÔNICOS NO LABORATÓRIO REMOTO VISIR+. In: Educação Contemporânea – Vol. 15 – Ensino Superior/ Organização: Gonçalves, Maria Célia da Silva; Jesus, Bruna Guzman de – Belo Horizonte– MG: Poisson, ISBN: 978-65-5866-057-6, Capítulo 23, <https://doi.org/10.36229/978-65-5866-057-6.CAP.23>, pp. 283-297. Fevereiro 2021.

TAWFIK, M., SANCRISTOBAL, E., MARTIN, S., GIL, R., DIAZ, G., COLMENAR, A., PEIRE, J., CASTRO, M. e NILSSON, K. (2013). **Virtual Instrument Systems in Reality (VISIR) for Remote Wiring and Measurement of Electronic Circuits on Breadboard**. IEEE Transactions on Learning Technologies, vol. 6 (1), pp. 60-72.

VISIR+. **Proceedings Of The XLVIII Brazilian Congress Of Engineering Education**, [S.L.], dez. 2020. Associação Brasileira de Educação em Engenharia. <http://dx.doi.org/10.37702/cobenge.2020.2964>. Disponível em: http://abenge.org.br/sis_artigo_doi.php?e=COBENGE&a=20&c=2964. Acesso em: 19 fev. 2022.

VISIR+ PROJECT. Disponível em:

<https://www2.isep.ipp.pt/visir/#:~:text=VISIR%2B%20brings%20together%20the%20power%20of%20the%20best,from%20Argentina%2C%20Austria%2C%20Brazil%2C%20Portugal%20Spain%2C%20and%20Sweden.?msslkid=d6667e23abdc11ec88c358f4d7cffe2>. Acesso em: 26 fev. 2022.

YU, Dengke; ZHOU, Rong. Intellectual Management: An Integrative Theory. Journal Of The Knowledge Economy, [s.l.], v. 8, n. 3, p.929-956, 10 out. 2015. Springer Nature. <http://dx.doi.org/10.1007/s13132-015-0305-0>.

ZUTIN, D.G., AUER, M.E., MAIER, C. e NIEDERSTÄTTER, M. (2010). **Lab2go: A repository to locate educational on-line laboratories**. In *Proceedings of IEEE Educon 2010 Conference*. (Madrid, Spain, 2010).