



UNIVERSIDADE FEDERAL DO AMAZONAS - UFAM
INSTITUTO DE COMPUTAÇÃO- ICOMP
PROGRAMA PÓS-GRADUAÇÃO EM INFORMÁTICA - PPGI

Proposta de um Dashboard para Análise e Visualização de Dados Educacionais de Sistemas Tutores Inteligentes

Matheus Freitas de Menezes

Manaus - AM

Junho / 2024

Matheus Freitas de Menezes

Proposta de um Dashboard para Análise e Visualização de Dados Educacionais de Sistemas Tutores Inteligentes

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Universidade Federal do Amazonas, como requisito parcial para a obtenção do título de Mestre em Informática.

Orientador

Dr. José Francisco de Magalhães Netto

Universidade Federal do Amazonas - UFAM

Instituto de Computação- IComp

Manaus - AM

Junho / 2024

Ficha Catalográfica

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

M543p

Menezes, Matheus Freitas de

Proposta de um Dashboard para Análise e Visualização de Dados Educacionais de Sistemas Tutores Inteligentes / Matheus Freitas de Menezes . 2024

125 f.: il. color; 31 cm.

Orientador: José Francisco de Magalhães Netto
Dissertação (Mestrado em Informática) - Universidade Federal do Amazonas.

1. Mineração de dados educacionais. 2. Sistema tutor inteligente.
3. Dashboard. 4. Visualização de dados. I. Netto, José Francisco de Magalhães. II. Universidade Federal do Amazonas III. Título



Ministério da Educação
Universidade Federal do Amazonas
Coordenação do Programa de Pós-Graduação em Informática

FOLHA DE APROVAÇÃO

"PROPOSTA DE UM DASHBOARD PARA ANÁLISE E VISUALIZAÇÃO DE DADOS EDUCACIONAIS DE SISTEMAS TUTORES INTELIGENTES"

MATHEUS FREITAS DE MENEZES

DISSERTAÇÃO DE MESTRADO DEFENDIDA E APROVADA PELA BANCA EXAMINADORA CONSTITUÍDA PELOS PROFESSORES:

Prof. Dr. José Francisco de Magalhães Netto - PRESIDENTE

Profa. Dra. Fabíola Guerra Nakamura - MEMBRO INTERNO

Prof. Dr. Márcio Aurélio dos Santos Alencar - MEMBRO EXTERNO

MANAUS, 17 de junho de 2024.



Documento assinado eletronicamente por **José Francisco de Magalhaes Netto, Professor do Magistério Superior**, em 26/06/2024, às 05:58, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Márcio Aurélio dos Santos Alencar, Usuário Externo**, em 26/06/2024, às 18:04, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Fabíola Guerra Nakamura, Professor do Magistério Superior**, em 03/07/2024, às 07:51, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Maria do Perpétuo Socorro Vasconcelos Palheta, Secretária em exercício**, em 03/07/2024, às 17:31, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufam.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2094042** e o código CRC **8DF21143**.

Avenida General Rodrigo Octávio, 6200 - Bairro Coroadão I Campus Universitário
Senador Arthur Virgílio Filho, Setor Norte - Telefone: (92) 3305-1181 / Ramal 1193
CEP 69080-900, Manaus/AM, coordenadorppgi@icomp.ufam.edu.br

Referência: Processo nº 23105.025839/2024-63

SEI nº 2094042

Dedico este trabalho à memória da minha tia, Cláudia Mendonça de Freitas, cuja presença iluminou nossas vidas. Sua bondade, amor e incentivo foram fundamentais para a minha vida. Que este trabalho sirva como uma humilde homenagem a tudo que você representou e continua representando para mim. Seu legado de carinho e dedicação jamais será esquecido.

AGRADECIMENTOS

Em primeiro lugar, expresso minha profunda gratidão aos meus pais, Jocicleide e Virlebaldo, cujo empenho incansável em me proporcionar as melhores condições tornou possível a conclusão desta etapa. Sem eles, nada disso faria sentido.

Agradeço também às minhas tias Karla Mendonça e Claudia Mendonça, e à minha avó Cleide Mendonça, pelo amor, apoio e incentivo constantes ao longo desta jornada.

Aos meus irmãos, Milton Menezes, Yuri Assayag e Maria Eduarda, pelo apoio e incentivo inestimáveis.

À minha amada Beatriz Bruna, por todo o amor, apoio e compreensão incondicionais. Sua presença constante, palavras encorajadoras e paciência infinita foram essenciais para a realização deste trabalho. Você acreditou em mim nos momentos mais difíceis e me motivou a seguir em frente. Sem você, essa jornada não teria sido possível. Te amo eternamente e sou grato por ter você ao meu lado.

Ao Prof. Dr. José Francisco Magalhães Netto, meu orientador, por suas valiosas contribuições, orientação sábia e apoio ao longo desta jornada acadêmica. Suas orientações e ensinamentos foram fundamentais para a realização deste trabalho.

Sou grato à Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e à Universidade Federal do Amazonas (UFAM) pelo suporte financeiro e institucional.

Finalmente, meus sinceros agradecimentos aos meus companheiros de jornada, Arcanjo Miguel, Ronilson Cavalcante, Fabiann Matthaus, Joethe Carvalho e Angela Vieira, por compartilharem seus conhecimentos e por todas as contribuições.

"Educação não transforma o mundo. Educação muda as pessoas. Pessoas transformam o mundo."

Paulo Freire

Proposta de um Dashboard para Análise e Visualização de Dados Educacionais de Sistemas Tutores Inteligentes

Autor: Matheus Freitas de Menezes

Orientador: Dr. José Francisco de Magalhães Netto

Resumo

Esta dissertação propõe a criação do dashboard Edudata Visualizer, destinado à análise e visualização de dados educacionais provenientes de Sistemas Tutores Inteligentes (STIs). Para alcançar este objetivo, empregamos técnicas de visualização de dados, estatística e Mineração de Dados Educacionais (MDE) em um STI desenvolvido especificamente para o ensino de álgebra. O propósito foi extrair informações úteis desses dados, proporcionando suporte aos professores no processo de tomada de decisões educacionais. Utilizando a metodologia de Descoberta de Conhecimento em Bancos de Dados e por meio do desenvolvimento de um dashboard web, buscamos criar uma ferramenta útil e intuitiva para análise e interpretação dos dados educacionais. Os resultados obtidos foram positivos, como evidenciado por duas avaliações conduzidas utilizando a técnica *System Usability Scale* (SUS) com um total de 43 professores. Estes resultados demonstraram um desempenho promissor da ferramenta no contexto educacional, destacando seu potencial para fornecer dados importantes que podem informar e enriquecer as práticas de ensino.

Palavras-chave: Mineração de Dados Educacionais, Sistema Tutor Inteligente, Dashboard.

Proposta de um Dashboard para Análise e Visualização de Dados Educacionais de Sistemas Tutores Inteligentes

Autor: Matheus Freitas de Menezes

Orientador: Dr. José Francisco de Magalhães Netto

Abstract

This dissertation proposes the creation of the Edudata Visualizer dashboard, aimed at analyzing and visualizing educational data from Intelligent Tutoring Systems (ITSs). To achieve this goal, we employed data visualization techniques, statistics, and Educational Data Mining (EDM) in an ITS specifically developed for teaching algebra. The purpose was to extract useful information from this data, providing support to teachers in the educational decision-making process. Using the Knowledge Discovery in Databases methodology and through the development of a web dashboard, we aimed to create a useful and intuitive tool for analyzing and interpreting educational data. The results obtained were positive, as evidenced by two evaluations conducted using the System Usability Scale (SUS) technique with a total of 43 teachers. These results demonstrated a promising performance of the tool in the educational context, highlighting its potential to provide important data that can inform and enrich teaching practices.

Keywords: Educational Data Mining, Intelligent Tutoring System, Dashboard.

LISTA DE ILUSTRAÇÕES

Figura 1 – Visão geral da Metodologia	26
Figura 2 – Representação do histórico da pesquisa.	29
Figura 3 – Relação entre as áreas	32
Figura 4 – Representação dos Módulos (HURTATIZ et al., 2015).	34
Figura 5 – Interface do Estudante (LOPES, 2019).	36
Figura 6 – Representação das atividades que envolvem o agente inteligente.	40
Figura 7 – Etapas do MS	43
Figura 8 – Incidência de STIs por área de estudo.	47
Figura 9 – Distribuição de STIs entre os níveis educacionais.	48
Figura 10 – Técnicas de inteligência artificial identificadas.	50
Figura 11 – Prevalência de <i>dashboards</i> em STIs	51
Figura 12 – Arquitetura - CloudTrail-Tracker	57
Figura 13 – Tela de VLA <i>dashboard</i> - Fonte: (SILVA; NETTO; SOUZA, 2018)	60
Figura 14 – Arquitetura de desenvolvimento <i>Edudata Visualizer</i>	65
Figura 15 – Diagrama de caso de uso: Professor	68
Figura 16 – Diagrama de caso de uso: Agentes	69
Figura 17 – Tela inicial do <i>dashboard</i> no formato Smartphone	71
Figura 18 – Primeira versão da tela inicial	71
Figura 19 – Passos metodológicos	73
Figura 20 – Árvore gerada pelo algoritmo J48	80
Figura 21 – Árvore gerada pelo algoritmo decision tree	81
Figura 22 – Algoritmo gerado pela união das duas árvores	82

Figura 23 – Exibição do classificação dos estudantes no <i>dashboard</i>	83
Figura 24 – Tela inicial do <i>dashboard</i>	84
Figura 25 – <i>dashboard</i> de desempenho dos estudantes	85
Figura 26 – <i>dashboard</i> com as estatísticas das questões	86
Figura 27 – Lista com estatísticas dos estudantes	86
Figura 28 – Tela de login do <i>dashboard Edudata Visualizer</i>	87
Figura 29 – Primeira <i>dashboard</i> do <i>Edudata Visualizer</i>	88
Figura 30 – Tela com os dados de desempenho dos estudantes	89
Figura 31 – Tela de login do <i>dashboard Edudata Visualizer</i>	90
Figura 32 – Tela de previsão de desempenho dos estudantes	91
Figura 33 – Janela do relatório inteligente	92
Figura 34 – Tela de busca individual	93
Figura 35 – Pontuação do teste SUS	97
Figura 36 – Respostas às outras afirmações feitas no questionário.	98
Figura 37 – Divisão de gênero entre os participantes	100
Figura 38 – Respostas sobre a primeira questão do teste SUS	101
Figura 39 – Respostas sobre a segunda questão do teste SUS	102
Figura 40 – Respostas sobre a terceira questão do teste SUS	102
Figura 41 – Respostas sobre a quarta questão do teste SUS	103
Figura 42 – Respostas sobre a quinta questão do teste SUS	104
Figura 43 – Respostas sobre a sexta questão do teste SUS	104
Figura 44 – Respostas sobre a sétima questão do teste SUS	105
Figura 45 – Respostas sobre a oitava questão do teste SUS	106
Figura 46 – Respostas sobre a nona questão do teste SUS	106
Figura 47 – Respostas sobre a décima questão do teste SUS	107
Figura 48 – Pontuação do teste SUS	107

LISTA DE TABELAS

Tabela 1 – Publicações Selecionadas	46
Tabela 2 – Taxa de passos corretos para cada questão	75
Tabela 3 – Taxa de passos certos para cada estudante	76
Tabela 4 – Modelo de classificação do INEP	77

LISTA DE QUADROS

4.1 Quadro comparativo entre as classificações	77
--	----

LISTA DE ABREVIATURAS E SIGLAS

ARFF *Atribute Relation File Format*

AVA Ambiente Virtual de Aprendizagem

AWS *Amazon Web Services*

CE Critérios de Exclusão

CI Critérios de Inclusão

FIE *Frontiers in Education*

GQM *Goal Question-Metric*

IA Inteligência Artificial

ITS *Intelligent Tutoring Systems*

KNN *K-Nearest Neighbor*

LAD *Learning Analytics Dashboard*

LEIA *Learning Algebra*

MDE Mineração de Dados Educacionais

MS Mapeamento Sistemático

PLN Processamento de Linguagem Natural

QP Questão de Pesquisa

STI Sistema Tutor Inteligente

SUS *System Usability Scale*

WEKA *Waikato Environment for Knowledge Analysis*

SUMÁRIO

1	INTRODUÇÃO	19
1.1	Contextualização	19
1.2	Definição do Problema	21
1.3	Justificativa	22
1.4	Objetivos	24
1.4.1	Objetivo Geral	24
1.4.2	Objetivos Específicos	24
1.5	Metodologia	24
1.6	Histórico da Pesquisa	28
1.7	Organização do Trabalho	29
1.8	Conclusão do Capítulo	30
2	FUNDAMENTAÇÃO TEÓRICA	31
2.1	Conceitos Relacionados	31
2.1.1	Mineração de Dados Educacionais	31
2.1.2	Sistema Tutor Inteligente (STI)	33
2.1.3	STI LEIA	35
2.1.4	Visualização de dados	37
2.1.5	<i>Dashboard</i>	37
2.1.6	Agentes Inteligentes	39
2.2	Conclusão do Capítulo	40
3	TRABALHOS CORRELATOS	42
3.1	Mapeamento Sistemático da Literatura	42
3.1.1	Planejamento do Mapeamento	44

3.1.1.1	Objetivos e Questões de Pesquisa	44
3.1.1.2	Fontes, Idioma e Expressões de Busca	44
3.1.1.3	Critérios de Seleção	45
3.1.1.4	Procedimento de Extração de Dados	45
3.1.2	Condução do Mapeamento Sistemático	46
3.2	Resultados e Discussões	46
3.2.1	QP1: Quais as áreas de estudo do STI?	46
3.2.2	QP2: quais camadas educacionais são cobertas pelo STI?	48
3.2.3	QP3: quais as técnicas de Inteligência Artificial utilizadas nos STIs?	49
3.2.4	QP4: quais STIs utilizam <i>dashboards</i> ?	50
3.3	Snowballing	53
3.4	Trabalhos Relacionados	53
3.5	Conclusão do Capítulo	63
4	PROPOSTA	64
4.1	Visão Geral do Sistema	64
4.2	Descrição dos Agentes	66
4.3	Diagramas de Caso de Uso	67
4.4	Ferramentas Utilizadas no Desenvolvimento	70
4.5	Módulo de Análise de Dados	72
4.5.1	Primeira Análise de Dados	74
4.5.1.1	Taxa de passos corretos para cada questão	75
4.5.1.2	Definindo o nível das questões	76
4.5.2	Segunda Análise de Dados: Geração de Árvores de Classificação	78
4.6	Prototipação Inicial	83
4.6.1	<i>Dashboards</i> Projetados	84
4.7	Edudata Visualizer	87
4.8	Conclusão do Capítulo	93
5	RESULTADOS	94

5.1	Avaliação Utilizando o <i>System Usability Scale</i>	95
5.1.1	Resultado da Primeira Avaliação SUS	97
5.1.2	Resultado da Segunda Avaliação SUS	100
5.2	Conclusão do Capítulo	108
6	CONSIDERAÇÕES FINAIS	109
6.1	Contribuições	110
6.2	Limitações	111
6.3	Trabalhos Futuros	111
Referências		113
APÊNDICE A	APÊNDICE	119
B	ANEXOS	125

1

INTRODUÇÃO

Neste capítulo são exemplificados o contexto dos dados educacionais, visualização de dados e os problemas identificados e abordados, tal como os objetivos que guiaram esta pesquisa e a justificativa que motivou o desenvolvimento desta dissertação.

1.1 Contextualização

Com o crescente uso das tecnologias para suporte ao processo de ensino e aprendizagem no campo educacional ([GAŠEVIĆ; TSAI; DRACHSLER, 2022](#)), um grande volume de dados têm sido gerado com base nos diferentes tipos de interação com o sistema, abrangendo principalmente estudantes, tutores e professores. À medida que o ensino à distância se expande nas instituições educacionais, surge um vasto conjunto de dados à disposição. Este conjunto representa um poderoso recurso analítico que pode permitir monitorar, examinar e compreender os processos de aprendizagem dos alunos ([ALENCAR et al., 2020](#)).

No entanto, boa parte desses dados não têm sido analisados, estabelecendo uma lacuna significativa, dada a quantidade de informações que podem ser potencialmente extraídas de tais dados ([RABELO et al., 2017](#)). Paralelamente, professores enfrentam desafios ao monitorar o desempenho dos alunos de maneira eficaz e atualizada. O uso de tecnologias é crucial para aprimorar a apresentação dinâmica e simplificada dessas informações nas atividades docentes ([FALGUERAS, 2022](#)).

Nesse contexto, o futuro apoio das ferramentas de inteligência artificial surge

como uma solução promissora. Com a implementação dessas tecnologias, os professores poderiam potencialmente redirecionar o seu foco principal para a elaboração de planos de aula inovadores, envolvendo-se no desenvolvimento profissional e oferecendo treino e orientação personalizados a cada aluno. Todas essas atividades são fundamentais para melhorar o desempenho de aprendizagem dos alunos em relação às habilidades e desafios do futuro (GRASSINI, 2023).

Propondo-se a trabalhar neste problema, os *dashboards* educacionais são um meio de apresentar informações sobre os alunos e suas interações nas atividades em ambientes de aprendizagem online (SHOAIB et al., 2022). Normalmente, um painel educacional para professores inclui algum tipo de auxílio visual que incentive a reflexão sobre os padrões de comportamento do aluno e a agir de acordo com eles. Na prática, esta ferramenta pode auxiliar os professores a tomar decisões baseadas em dados, apoiando assim o seu crescimento profissional, no entanto, até agora, o uso de painéis de educação por professores tem sido pouco estudado (MICHAELI; KROPARO; HERSHKOVITZ, 2020).

Podendo trabalhar em conjunto com os *dashboards* e visando resolver os problemas relacionados a análise destes dados, as técnicas de aprendizagem de máquina, mineração de dados e estatísticas estão ganhando cada vez mais importância no setor educacional (SHOAIB et al., 2022) (NAFEA, 2018). A mineração de dados educacionais (MDE) atraiu um interesse significativo nos últimos anos na tentativa de personalizar e melhorar o processo de aprendizagem dos alunos (INJADAT et al., 2020)(YAĞCI, 2022). Como em muitas outras áreas, o ensino está descobrindo o potencial impacto dessas técnicas no processo de aprendizagem. Certamente, essas técnicas podem fornecer informações para apoiar o desenvolvimento de modelos educacionais que aumentam a eficiência e a qualidade do ensino e aprendizagem (ALDOWAH; AL-SAMARRAIE; FAUZY, 2019)(SUSNJAK; RAMASWAMI; MATHRANI, 2022).

No contexto do ambiente de uma instituição educacional, as interações envolvem três tipos de atores: professor, aluno e o ambiente. A interação entre esses três atores gera dados volumosos que podem ser sistematicamente agrupados para extrair informações valiosas (DUTT; ISMAIL; HERAWAN, 2017). Com esse propósito, muitas instituições

utilizam o conceito de *dashboards*, de modo a utilizar uma ferramenta poderosa para o suporte à tomada de decisão educacional, desempenhando um papel mais ativo na governança educacional ([FARAHMAND; DEWAN; LIN, 2020](#)).

1.2 Definição do Problema

Diante do cenário atual nas escolas, a identificação das dificuldades individuais ou coletivas pelos professores tem se tornado cada vez mais desafiadora, dada a abundância de variáveis em jogo. O reconhecimento dessas questões muitas vezes só ocorre por meio dos resultados das notas finais, que integram o método avaliativo composto no processo educacional ([VANNESTE et al., 2021](#)).

Os professores enfrentam vários desafios na monitorização do desempenho dos estudantes. Alguns desses desafios incluem:

1. Quantidade de alunos: acompanhar o progresso de uma grande quantidade de estudantes pode ser bastante demorado e desafiador.
2. Diversidade de habilidades: com tantos alunos com diferentes níveis de habilidade e compreensão, pode ser difícil identificar o progresso individual.
3. Falta de dados unificados: muitas vezes, os dados do desempenho dos estudantes estão espalhados em várias fontes, o que dificulta ter uma visão geral.
4. Atualização de dados: manter os dados atualizados é crucial, mas pode ser um desafio para os professores, especialmente quando eles precisam fazer isso manualmente.
5. Identificação de padrões: analisar dados e identificar padrões pode ser difícil sem ajuda tecnológica.

Assim sendo, os professores enfrentam vários obstáculos para acompanhar o desempenho dos estudantes de forma eficiente, precisa e atualizada.

Para enfrentar esse problema, é essencial utilizar tecnologias que permitam a apresentação dinâmica e simplificada das informações em suas atividades ([FALGUE-](#)

RAS, 2022)(MENEZES et al., 2020). Com esse propósito, muitas instituições utilizam o conceito de *dashboards*, criados visando compilar métricas-chave em uma interface simples e fácil de interpretar, para que alunos, professores e administradores possam rapidamente visualizar as informações, da forma mais prática possível. Os sistemas são atualizados automaticamente baseados em dados armazenados nos bancos de dados educacionais (WEST, 2015).

Tendo em vista este quadro, sem acesso à informação precisa sobre o desempenho dos alunos, é difícil identificar pontos fracos e elaborar planos de melhoria. Muitos sistemas educacionais não armazenam os dados de forma correta ou não investem na análise dos dados disponibilizados. Fazendo parte dessa conjuntura, os STIs geram vastos dados sobre o progresso e aprendizado dos alunos, como estimativas de conhecimento por habilidades, erros comuns, e registro de tempo e avanço em atividades. No entanto, muitos STIs carecem de *dashboards* para professores, privando-os de informações cruciais sobre as necessidades dos estudantes em diferentes áreas. Embora essas análises sejam importantes para pesquisas, são raramente adaptadas para atender às demandas práticas dos educadores. (HOLSTEIN et al., 2010).

1.3 Justificativa

Os *dashboards* educacionais oferecem uma ampla gama de vantagens para professores, administradores e alunos, fornecendo uma visão geral e em tempo real do desempenho dos estudantes. Dessa forma, os professores podem identificar rapidamente pontos fracos e desenvolver planos de ação para melhorar o ensino. Além disso, os *dashboards* permitem uma análise mais profunda dos dados, ajudando os professores a avaliarem o impacto de suas estratégias de ensino e a fazer os ajustes necessários (AYALA et al., 2023).

Adicionalmente, os painéis educacionais auxiliam na comunicação entre professores, pais e alunos, permitindo que todos tenham acesso à informação atualizada sobre o desempenho escolar. Consequentemente, isso promove a colaboração e a transparência, melhorando a qualidade do ensino. Paralelamente, eles ajudam na eficiência

administrativa, economizando tempo e recursos ao automatizar processos manuais, como o monitoramento de desempenho, o acompanhamento de progresso e a geração de relatórios. Em resumo, a utilização de *dashboards* educacionais é fundamental para garantir que os sistemas forneçam informações de qualidade e ajudem os estudantes a alcançarem seu máximo potencial ([NG; WANG; HU, 2022](#)).

Levando em conta os *dashboards* como ferramentas importantes no processo de aprendizagem, Permitindo aos estudantes a realização de modificações durante o seu processo de aprendizagem, permitindo que alunos e professores possam fazer mudanças estratégicas nas abordagens de ensino, visando reduzir dificuldades. Portanto, as evidências derivadas da visualização de dados fornecerão uma ferramenta poderosa para o suporte à tomada de decisão educacional, desempenhando um papel mais ativo na governança de seu próprio aprendizado ([FARAHMAND; DEWAN; LIN, 2020](#)).

Diante desse contexto e considerando a vasta quantidade de dados gerados pelos sistemas de suporte ao processo de ensino e aprendizagem dos estudantes, bem como a relevância dos *dashboards* no apoio ao ensino, a presente dissertação detalha o desenvolvimento de uma ferramenta dedicada à análise e visualização de dados provenientes do STI LEIA (*Learning Algebra*) ([LOPES et al., 2019](#)). Adicionalmente, conduzimos experimentos que empregaram técnicas de MDE, aprendizado de máquina e estatística, aplicadas aos dados do STI, visando o embasamento e implementação destes modelos de análise no *dashboard*.

1.4 Objetivos

Para alcançar o propósito desta pesquisa, representamos a seguir o objetivo geral e os objetivos específicos.

1.4.1 Objetivo Geral

O objetivo central deste projeto consiste em **Propor e validar um *dashboard* para suporte à análise e visualização de dados educacionais de STIs**, visando fornecer aos professores uma ferramenta intuitiva e eficaz que facilitará o processo de tomada de decisões pedagógicas.

1.4.2 Objetivos Específicos

- Elucidar e Identificar aspectos técnicos e pedagógicos relacionados aos STIs e *dashboards* educacionais.
- Conduzir uma análise abrangente dos dados educacionais gerados pelo STI LEIA, visando identificar padrões e tendências relevantes para o aprimoramento contínuo da plataforma e otimização da experiência de aprendizado dos usuários.
- Desenvolver e avaliar um *dashboard* dedicado à análise e visualização de dados educacionais provenientes de STIs, com o objetivo de gerar uma interface intuitiva e informativa para professores, facilitando a identificação de informações, padrões de desempenho e áreas de melhoria no processo de ensino-aprendizagem.

1.5 Metodologia

A metodologia de pesquisa é o guia pelo qual os pesquisadores devem conduzir suas investigações, delineando o processo pelo qual as questões e objetivos são formulados, além de apresentar os resultados com base nos dados coletados durante o período de estudo ([SILEYEW, 2019](#)).

Toda pesquisa começa com uma pergunta que surge a partir de uma observação inicial. A partir daí, são adotados procedimentos para preencher essa lacuna de conhecimento. Dentro do campo da metodologia científica, as pesquisas podem ser classificadas de várias maneiras com base em características como abordagem (qualitativa ou quantitativa), natureza (básica ou aplicada), objetivos (exploratórios, descritivos ou explicativos) e procedimentos (experimental, bibliográfico, estudo de caso, transversal, etc.) ([SAMPAIO, 2022](#)).

A natureza do projeto se define por ser uma pesquisa aplicada, com a finalidade de gerar conhecimentos para aplicação prática. A pesquisa aplicada busca resolver um problema específico, esses problemas ou questões podem estar em um nível individual, grupal ou social. Caracteriza-se por ser denominado como “não sistemático”, visto que vai direto para a busca de soluções ([BROWN, 2015](#)).

Para alcançar os objetivos desta pesquisa, optamos por adotar uma abordagem descritiva. Esta técnica visa apresentar características específicas de uma determinada população ou fenômeno. Além disso, permite estabelecer correlações entre variáveis e entender melhor sua natureza. Vale ressaltar que a abordagem descritiva não necessariamente busca explicar os fenômenos que descreve, mas pode fornecer uma base sólida para futuras explicações. Um exemplo comum de pesquisa que se enquadra nessa categoria é a pesquisa de opinião.

No desenvolvimento do projeto, seguimos os passos da pesquisa de intervenção ([OLIVEIRA, 2016](#)). Inicialmente, identificamos o problema a ser solucionado e procedemos com a apresentação e implementação da solução, que neste caso foi o *dashboard Edudata Visualizer*. Em seguida, coletamos dados dos experimentos realizados e conduzimos uma análise detalhada dos resultados, com o objetivo de extrair conclusões sobre a eficácia do *dashboard*.

Para a condução do plano da dissertação, utilizamos o desenho de pesquisa, visando criar uma estrutura geral, servindo de guia para a condução do estudo. Sendo uma etapa crucial na elaboração de qualquer investigação, pois define como os dados serão coletados, analisados e interpretados para responder às questões de pesquisa ou testar hipóteses específicas ([GUILHERME; CHERON, 2021](#)). Segue a Figura 1, que

ilustra as principais etapas necessárias para alcançar o resultado desejado.

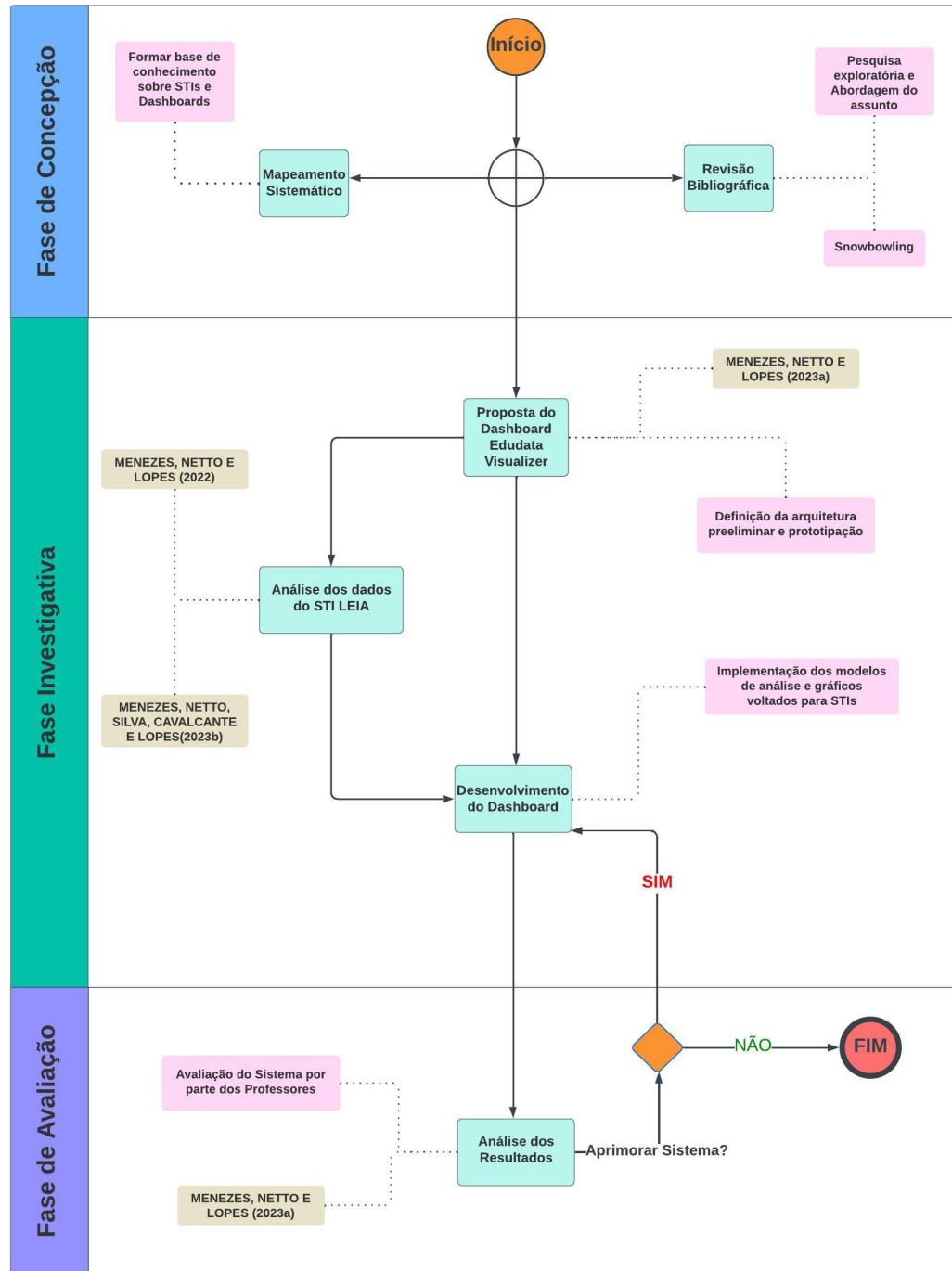


Figura 1 – Visão geral da Metodologia

A seguir, detalham-se as etapas do desenho de pesquisa conforme apresentado na Figura 1.

Fase de Concepção: Nesta etapa do processo de pesquisa, a metodologia adotada foi fundamentada em uma abordagem exploratória abrangente. O objetivo foi realizar um estudo profundo da base teórica relacionada aos principais temas que cons-

tituem o arcabouço conceitual da Dissertação, abrangendo áreas como STIs, *Dashboards*, Visualização de Dados e MDE.

A pesquisa bibliográfica foi essencial, proporcionando acesso a uma ampla variedade de artigos científicos relevantes. Essa abordagem permitiu uma exploração dos conceitos essenciais, teorias e práticas já estabelecidas, estabelecendo assim uma base sólida para o avanço da pesquisa.

Além disso, uma etapa crucial nesta fase foi a condução de um Mapeamento Sistemático da Literatura (Capítulo 2). Este processo envolveu uma abordagem estruturada para identificar, selecionar e analisar criticamente as principais obras e pesquisas relevantes ao tema em questão. Através deste mapeamento, foi possível mapear de forma abrangente o estado atual do conhecimento na área de estudo, identificando lacunas, tendências e áreas de interesse específicas que serviram como base para a formulação dos objetivos de pesquisa.

A Fase de Concepção representou um estágio fundamental no processo de pesquisa, onde a combinação de pesquisa bibliográfica extensiva e Mapeamento Sistemático da Literatura proporcionou uma base sólida para o desenvolvimento do trabalho.

Fase Investigativa: Durante esta fase, foi realizada a elaboração da primeira proposta do *dashboard*, por meio da concepção de uma arquitetura preliminar. Utilizando a técnica de prototipação, foi possível visualizar as ideias de forma prática.

Além disso, em paralelo, foi iniciada a análise dos dados provenientes do STI LEIA. Esta etapa foi importante para obter uma compreensão do panorama de dados disponíveis e para identificar padrões que poderiam orientar o desenvolvimento da ferramenta de forma mais precisa. Para isso, durante a análise foram seguidos processos como a identificação de variáveis relevantes, a avaliação da qualidade dos dados e a aplicação de modelos analíticos para revelar correlações e tendências significativas. Após a condução da análise e concepção da visão geral do sistema, deu-se início ao desenvolvimento da ferramenta (Capítulo 4).

Fase de Avaliação: Nesta etapa, foram realizados os testes de usabilidade com professores e a análise dos dados gerados neste experimento de avaliação. Após os testes, melhorias e contribuições podem ser realizadas no *Edudata Visualizer*, caso não

haja mais ajustes, o *dashboard* é finalizado, concluindo assim a fase avaliativa.

Esta etapa finaliza o desenho de pesquisa, onde a pesquisa se concretiza mediante as avaliações realizadas, nos resultados alcançados, no *feedback*, na experiência e prática vivenciada pelos professores envolvidos no processo.

1.6 Histórico da Pesquisa

O panorama histórico desta pesquisa abrange eventos passados que fundamentam o presente e, por conseguinte, delineiam possíveis cenários futuros, englobando ações e suas causas. A abordagem histórica compreende a análise, compreensão e exposição de eventos passados. Nesse contexto, a finalidade da pesquisa científica reside na compreensão das ocorrências passadas, seus desdobramentos, causas e efeitos, proporcionando uma melhor visão sobre processos, modelos e métodos históricos empregados em contextos específicos. Essa compreensão é essencial para gerar resultados de pesquisa embasados em fatos.

A Figura 2 ilustra as etapas desenvolvidas ao longo dos cinco semestres do programa de Mestrado. A investigação para a dissertação iniciou-se no segundo semestre de 2021, com previsão de conclusão no primeiro semestre de 2024. Para garantir o cumprimento das metas estabelecidas, foi elaborado um planejamento detalhado (cronograma), visando a execução das etapas nos prazos determinados. Vale ressaltar que o avanço e os resultados obtidos durante a pesquisa foram fundamentais para o êxito no cumprimento do cronograma proposto.



Figura 2 – Representação do histórico da pesquisa.

1.7 Organização do Trabalho

O conteúdo desta dissertação está estruturado em capítulos, sendo o presente capítulo o primeiro, apresentando o contexto do estudo para o desenvolvimento da aplicação, bem como as motivações, objetivos e metodologia empregada para alcançar tais objetivos. Os demais capítulos são definidos nos seus seguintes propósitos:

Capítulo 2 – Fundamentação Teórica: este capítulo apresenta os conceitos relacionados a condução da pesquisa, exibindo um panorama sobre os fundamentos da MDE, assim como uma explanação sobre os *dashboards* e STIs.

Capítulo 3 – Trabalhos Correlatos: neste capítulo são descritos os trabalhos identificados na literatura como o estado da arte em *dashboards* e análise de dados educacionais.

Capítulo 4 – Proposta: este capítulo destina-se a apresentar os pontos essenciais para o entendimento do desenvolvimento da proposta. Apresentando uma visão geral do sistema e a exibição do protótipo.

Capítulo 5 – Resultados: neste capítulo, são apresentados os resultados derivados da realização dos testes, abordando a avaliação dos usuários após a interação com o sistema, e também quantificando a qualidade do sistema por meio de uma pontuação.

Capítulo 6 – Considerações Finais: neste capítulo, são destacadas as contribuições da pesquisa, as dificuldades enfrentadas durante o processo e os trabalhos futuros a serem desenvolvidos.

1.8 Conclusão do Capítulo

Nesta introdução, delineamos o propósito principal deste trabalho: o desenvolvimento de um *dashboard* educacional voltado para STIs, visando auxiliar os professores em sua prática pedagógica. Reconhecemos a importância crescente da tecnologia na educação e a necessidade de ferramentas eficazes que permitam aos educadores compreenderem melhor o progresso e as necessidades individuais de seus alunos.

Ao longo deste capítulo, exploramos os desafios enfrentados pelos educadores e a complexidade de lidar com dados educacionais em um ambiente cada vez mais digitalizado. Destacamos a importância de um *dashboard* educacional como uma solução para simplificar a análise e interpretação desses dados, fornecendo visões valiosas para os professores e facilitando a tomada de decisões informadas.

No próximo capítulo, adentraremos na fundamentação teórica necessária para compreender os conceitos subjacentes ao desenvolvimento do *dashboard* educacional. Exploraremos definições e princípios relacionados a termos-chave, como *dashboard*, MDE e visualização de dados, estabelecendo as bases teóricas para a implementação eficaz do sistema proposto.

2

FUNDAMENTAÇÃO TEÓRICA

Este capítulo tem por objetivo contextualizar a proposta, apresentando teorias a respeito dos temas abordados na pesquisa que sustenta o desenvolvimento do projeto. Neste contexto, será apresentada uma explanação a respeito dos conceitos sobre MDE, STI, visualização de dados, *dashboards* e agentes inteligentes.

2.1 Conceitos Relacionados

2.1.1 Mineração de Dados Educacionais

A MDE surge como uma área promissora de pesquisa interdisciplinar, reunindo os campos da educação e da informática. Sua ascensão dentro da área de análise de dados educacionais se justifica pelas inúmeras vantagens que proporciona aos estudantes, gestores e professores. Nesse contexto, diversas técnicas avançadas de mineração de dados têm sido empregadas com o objetivo de aprimorar os resultados do processo de aprendizagem. Essas abordagens exploram extensivamente conjuntos de dados provenientes de ambientes educacionais, permitindo uma análise robusta e detalhada para informar práticas pedagógicas mais eficazes ([BATOOL et al., 2023](#)).

O método de MDE utiliza campos relacionados, como aprendizagem de máquina, mineração de texto (abordagens para encontrar padrões em texto em linguagem natural) e estatística. Outras influências importantes são a psicometria (o estudo de instrumentos psicológicos para medir as habilidades e características humanas) e a análise de log da

web (abordagens para identificar perfis de usuário e padrões de navegação dos usuários do site) ([BAKER; MARTIN; ROSSI, 2018](#)). A MDE pode contribuir para a avaliação de sistemas de aprendizagem, identificar padrões regulares ou incomuns, estratégias de resolução de problemas de alunos e padrões de colaboração bem ou malsucedida, ajudando assim a formular novas hipóteses científicas ([SCHEUER; MCLAREN, 2012](#)) ([PEÑA-AYALA, 2014](#)). A Figura 3 exemplifica a correlação entre as áreas.

Além disso, a MDE auxilia na identificação de fatores de aprendizagem e também ajudando na análise desses fatores, com o propósito de compreender o método de aprendizado apoiado em tecnologias de comunicação e informação ([LAMPROPOULOS, 2023](#)). Deste modo, é possível entender de maneira mais eficaz e apropriada, os alunos, como eles aprendem, o papel do contexto no qual a aprendizagem ocorre, além de outros fatores que influenciam a aprendizagem ([RAMOS et al., 2020](#)).

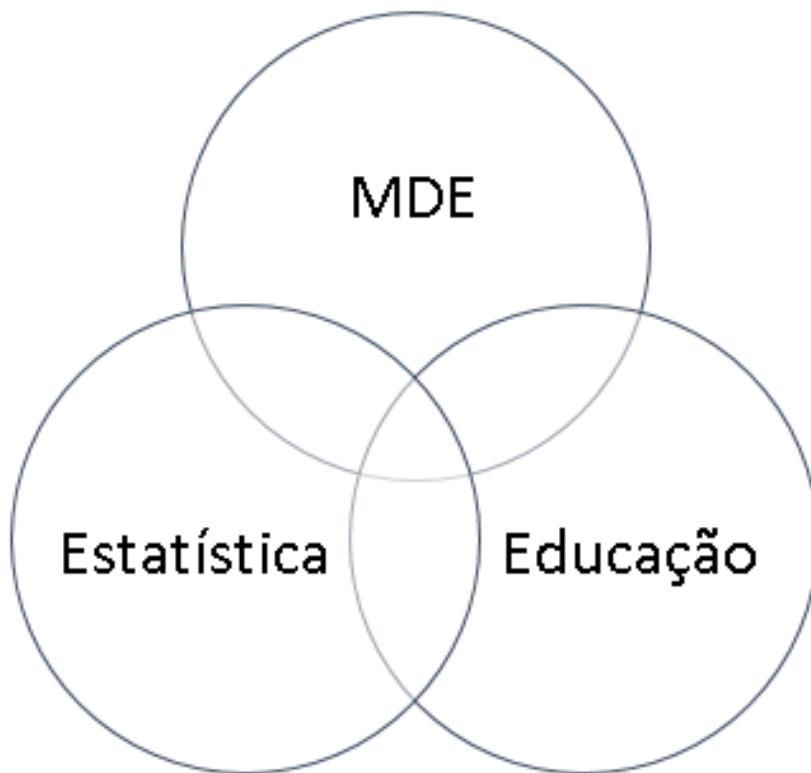


Figura 3 – Relação entre as áreas

Uma ampla gama de métodos de MDE surgiram nos últimos anos. Alguns são mais ou menos semelhantes aos observados no uso de mineração de dados em outros domínios, enquanto outros são exclusivos do MDE. Neste trabalho, é utilizado o modelo

de classificação para desenvolver um modelo que possa compreender um aspecto único dos dados a partir de alguma combinação específica. Na MDE, classificação e regressão são os tipos mais comuns de modelos de previsão (BAKER; INVENTADO, 2018).

A estrutura da maioria dos projetos de MDE pode ser dividida em três partes: coleção, arquivamento e análise. Coleção refere-se às ferramentas e sistemas de tutoria usados para registrar as informações relevantes, sejam pontuações dos alunos, respostas a questionários online ou eventos de um STI. O arquivamento é o processo de armazenamento e integração dos dados coletados. Para uma quantidade menor de dados, este é um problema relativamente menor, mas para a abundância de dados gerados por alguns STIs, esta pode ser uma tarefa mais complexa (KAMATH; KAMAT, 2022).

A análise realiza a MDE nos dados coletados na tentativa de aprofundar a compreensão da aprendizagem do aluno, descobrir as relações entre as questões e, possivelmente, desenvolver uma compreensão quantitativa mais profunda dos processos cognitivos, em geral. Dependendo do projeto de MDE, essas três tarefas mudarão de complexidade e importância relativa, mas todos eles devem ser abordados em qualquer projeto de EDM, mesmo que de forma diferente, dividida entre mais ou menos processos (KAMATH; KAMAT, 2022).

2.1.2 Sistema Tutor Inteligente (STI)

Os STIs fornecem instruções personalizadas e *feedback* aos alunos, se adaptando ao comportamento do aluno, mudando seu modo de interação com cada aluno dependendo das características individuais. Normalmente, consiste em um modelo que envolve o lado do aluno e o lado pedagógico. Os STIs registram todas as interações aluno-tutor, como, por exemplo, cliques do mouse, digitação e fala (ROMERO; VENTURA, 2020). Utilizando técnicas avançadas de Inteligência Artificial, esses sistemas são capazes de inferir informações cruciais sobre o estudante, como conhecimentos prévios, estados afetivos, entre outros, para fornecer instrução e assistência personalizadas (JAQUES, 2023).

O sistema tutor SCHOLAR foi o primeiro STI, introduzido por Jaime R. Carbonell

em 1970, este sistema foi desenvolvido para acompanhar o desempenho do aluno. Nas últimas décadas, o uso dos métodos de inteligência artificial, especialmente o aprendizado de máquina e a mineração de dados, caracterizando os STIs como softwares educacionais que incorporam técnicas de inteligência artificial ([MOUSAVINASAB et al., 2021](#)).

Os STIs são compostos por diversos módulos, que trabalham juntos para fornecer uma experiência de aprendizagem adaptativa e personalizada ([HURTATIZ et al., 2015](#)). Na Figura 4 pode-se observar a representação dos módulos.

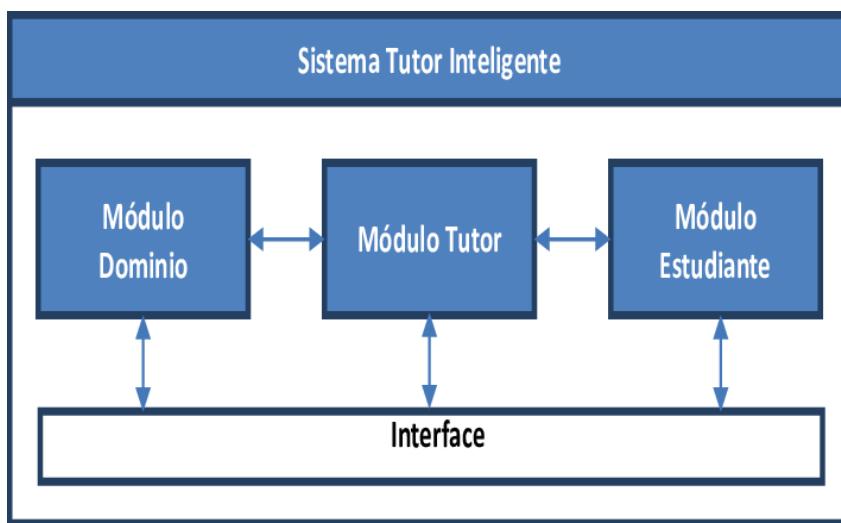


Figura 4 – Representação dos Módulos ([HURTATIZ et al., 2015](#)).

Módulo domínio: responsável por representar e armazenar o conhecimento do domínio que está sendo ensinado. Fundamental para que o sistema possa fornecer instruções precisas e personalizadas ao aluno. Pode incluir modelos conceituais, exemplos de problemas e soluções, entre outras informações relevantes para o aprendizado do domínio em questão.

Módulo tutor: responsável por controlar a interação do sistema com o aluno. Utiliza o conhecimento armazenado no módulo de domínio para fornecer instruções e *feedback* personalizados ao aluno. O tutor também pode utilizar técnicas de adaptação, como a seleção de tarefas apropriadas com base no nível de conhecimento do aluno e a seleção de estratégias de ensino adequadas para o estilo de aprendizagem do aluno.

Módulo interface: responsável por fornecer uma interface de usuário amigável e fácil de usar. Permite que o aluno interaja com o sistema de forma intuitiva e eficiente.

A interface do usuário pode incluir elementos como menus, botões, gráficos e outros elementos visuais para tornar a interação do aluno com o sistema mais agradável e eficaz.

Módulo estudante: responsável por modelar o conhecimento e o desempenho do aluno. Ele coleta informações sobre o desempenho do aluno em atividades de aprendizagem e utiliza essas informações para adaptar a instrução e o *feedback* do sistema às necessidades específicas do aluno. O módulo do estudante pode incluir modelos de conhecimento do aluno, histórico de desempenho e outras informações relevantes para personalizar a experiência de aprendizagem.

Dessa forma, os módulos domínio, tutor, interface e estudante trabalham juntos para fornecer uma experiência de aprendizagem adaptativa e personalizada, utilizando técnicas de inteligência artificial para representar o conhecimento do domínio, dando instruções e *feedback* personalizados ao aluno, exibindo uma interface de usuário amigável e coletando informações sobre o desempenho do aluno.

2.1.3 STI LEIA

O STI LEIA é uma plataforma baseada em tecnologias web que proporciona aos alunos a oportunidade de praticar atividades de ensino e aprendizagem dos conceitos fundamentais de Álgebra, independentemente de sua localização geográfica. O sistema foi desenvolvido utilizando PHP com o framework Laravel e utiliza o MySQL para a persistência dos dados ([LOPES, 2019](#)). Na figura 5 pode-se observar uma das telas do sistema.

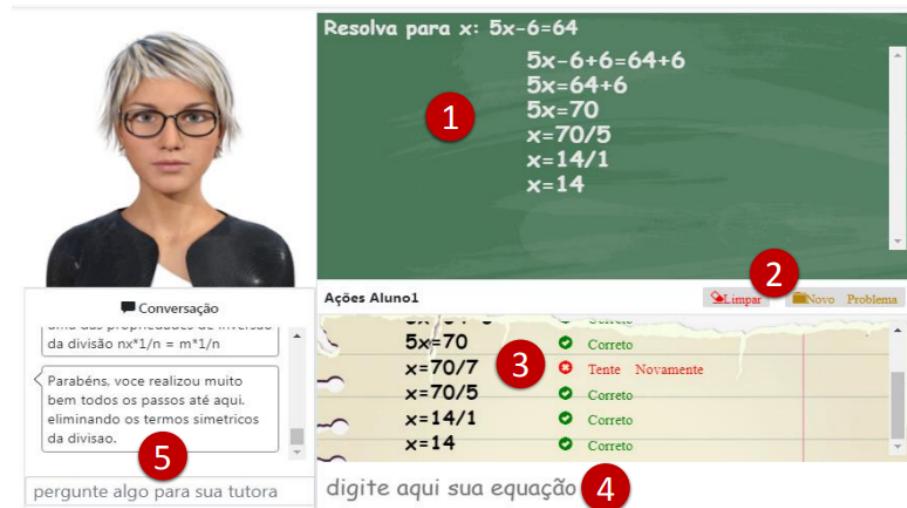


Figura 5 – Interface do Estudante ([LOPES, 2019](#)).

A proposta pedagógica do STI LEIA visa oferecer uma aprendizagem colaborativa, com foco na resolução de equações polinomiais de primeiro grau. Isso é facilitado por um Assistente Virtual Pedagógico e mensagens de feedback que têm o objetivo de ajudar e motivar os alunos ao longo do processo de aprendizagem. Além disso, a plataforma disponibiliza aos professores um ambiente onde podem monitorar o progresso individual de cada estudante. Para alcançar esses objetivos, foi projetada uma arquitetura específica que aborda essas necessidades ([LOPES, 2019](#)).

O STI LEIA é fundamental para a geração de dados essenciais para esta pesquisa, uma vez que permite a coleta de informações detalhadas sobre o desempenho e o progresso dos alunos em atividades de Álgebra. Esses dados são essenciais para a análise quantitativa e qualitativa, fornecendo dados sobre os padrões de aprendizagem, as dificuldades enfrentadas pelos estudantes e a eficácia das estratégias pedagógicas implementadas. Utilizando essas informações, é possível criar um dashboard informativo e interativo, que facilita a visualização e interpretação dos resultados, auxiliando tanto pesquisadores quanto educadores na tomada de decisões informadas e na melhoria contínua dos processos de ensino e aprendizagem.

2.1.4 Visualização de dados

O uso de STIs gera um crescimento exponencial da quantidade de dados a serem analisados, fornecendo uma oportunidade para novas formas de representação destes dados. Para isso, a visualização de dados trabalha na representação das informações em forma de gráficos, tabelas, imagens ou outras representações visuais para tornar a informação mais fácil de compreender e interpretar. Sendo uma parte importante da análise de dados ([SURSE; KOKATE; SANGHAVI, 2022](#)).

Segundo ([GRINSTEIN; WARD, 2002](#)), os benefícios da utilização da visualização de dados incluem a capacidade de destacar tendências, relações e padrões em grandes quantidades de dados, ajudando os usuários a identificar informações valiosas e tomar decisões informadas.

Existem vários tipos de visualizações de dados, incluindo gráficos de barras, gráficos de linha, gráficos de dispersão, mapas e gráficos de setores. Cada tipo de visualização é adequado para diferentes tipos de dados e objetivos, e é importante escolher a visualização adequada para representar a informação de maneira clara e precisa ([SURSE; KOKATE; SANGHAVI, 2022](#)). A visualização de dados também pode ser combinada com tecnologias avançadas, como inteligência artificial e aprendizado de máquina, para criar soluções ainda mais poderosas e eficazes. Tal como, sistemas de visualização de dados podem ser integrados a bancos de dados e fontes de dados em tempo real para fornecer informações atualizadas em tempo real ([UNWIN, 2020](#)).

2.1.5 Dashboard

Para atingir um dos nossos objetivos utilizaremos um *dashboard* nomeado como *Edudata Visualizer*, uma ferramenta operacional utilizada para exibição visual de informações, com foco na análise dos principais dados, de modo a fornecer uma visão geral dos objetivos e o monitoramento das atividades ([FEW, 2006](#)) ([FRIENDLY; WAINER, 2021](#)). Com a velocidade, variabilidade e grande volume de dados provenientes das inúmeras atividades praticadas em um sistema, naturalmente uma abundância de dados é gerada. A *Big Data* utiliza os *Dashboards* para dar suporte à tomada de decisão ([APARICIO;](#)

[COSTA, 2015\).](#)

Os *dashboards* podem exibir dados de desempenho concretos com visualizações de ponta em uma única plataforma, permitindo que usuários acessem e sintetizem informações de forma eficiente, visando melhorar o desempenho, auxiliando a interpretação dos dados, integrando dados de várias fontes de avaliação em torno de competências específicas para demonstrar uma trajetória de aprendizagem ([BOSCARDIN et al., 2018](#)), ([ELLAWAY et al., 2014](#)).

Os *dashboards* educacionais desempenham um papel crucial na simplificação da análise de dados complexos no contexto educacional. Entre os diversos tipos de gráficos empregados para representar informações, os gráficos de barras assumem uma posição de destaque. Utilizados para visualizar comparativamente o desempenho acadêmico em diferentes disciplinas ou categorias, esses gráficos oferecem uma representação visual clara das discrepâncias entre conjuntos de dados discretos.

Os gráficos de linha, por sua vez, são particularmente eficazes na representação de tendências ao longo do tempo. Empregados para ilustrar o progresso acadêmico ao longo de semestres ou anos, esses gráficos proporcionam informações dinâmicas sobre padrões e variações temporais, permitindo uma compreensão mais aprofundada das trajetórias educacionais.

No que diz respeito aos gráficos de setores, embora seu uso seja debatido devido à possível distorção visual, eles são frequentemente adotados para representar a composição percentual de um todo. Nesse contexto educacional, podem ser empregados para visualizar a distribuição percentual de estudantes em diferentes faixas etárias ou programas acadêmicos.

Os gráficos de dispersão, outra ferramenta valiosa, são utilizados para explorar relações entre variáveis. Ao representar graficamente a interdependência entre horas de estudo e desempenho acadêmico, esses gráficos permitem a identificação de padrões de correlação, fornecendo noções sobre fatores que influenciam o sucesso acadêmico.

Além disso, os *box plots*, embora menos comuns, desempenham um papel crucial na representação da distribuição estatística de dados. Ao oferecer uma visão detalhada da variabilidade do desempenho dos alunos, esses gráficos proporcionam informa-

ções sobre a dispersão e a concentração dos dados, auxiliando na compreensão da consistência ou variabilidade dos resultados.

A integração desses diversos tipos de gráficos em *dashboards* educacionais proporciona uma abordagem abrangente na apresentação visual de dados educacionais. Essa abordagem facilita a análise aprofundada e a tomada de decisões informadas no ambiente acadêmico, contribuindo para o aprimoramento contínuo da qualidade do ensino e do aprendizado.

2.1.6 Agentes Inteligentes

Os agentes inteligentes representam uma entidade computacional concebida para efetuar decisões autônomas e realizar ações com base em percepções ambientais. Estes agentes, configurados como aplicações, apresentam a capacidade de interagir de maneira inteligente com o ambiente circundante ([RUSSELL; NORVIG, 2021](#)).

O funcionamento de um agente inteligente implica a realização de diversas etapas sequenciais. Inicialmente, o agente adquire informações do ambiente por meio de sensores, tais como câmeras, microfones, sensores térmicos ou através do monitoramento de dados das variáveis em um determinado banco. Estas percepções são, subsequentemente, processadas mediante algoritmos e modelos internos, promovendo a compreensão da situação em que o agente se encontra ([ROCHA, 2017](#)).

Em etapa subsequente, o agente emprega raciocínio para tomar decisões. Este processo pode demandar a avaliação de diversas opções a fim de selecionar a ação mais apropriada para alcançar os objetivos estipulados. As ações resultantes são então implementadas no ambiente, podendo envolver movimentos físicos, interações com outros sistemas ou a produção de saídas específicas em um software ([ROCHA, 2017](#)). Na Figura 6 pode-se observar a execução das etapas do agente.

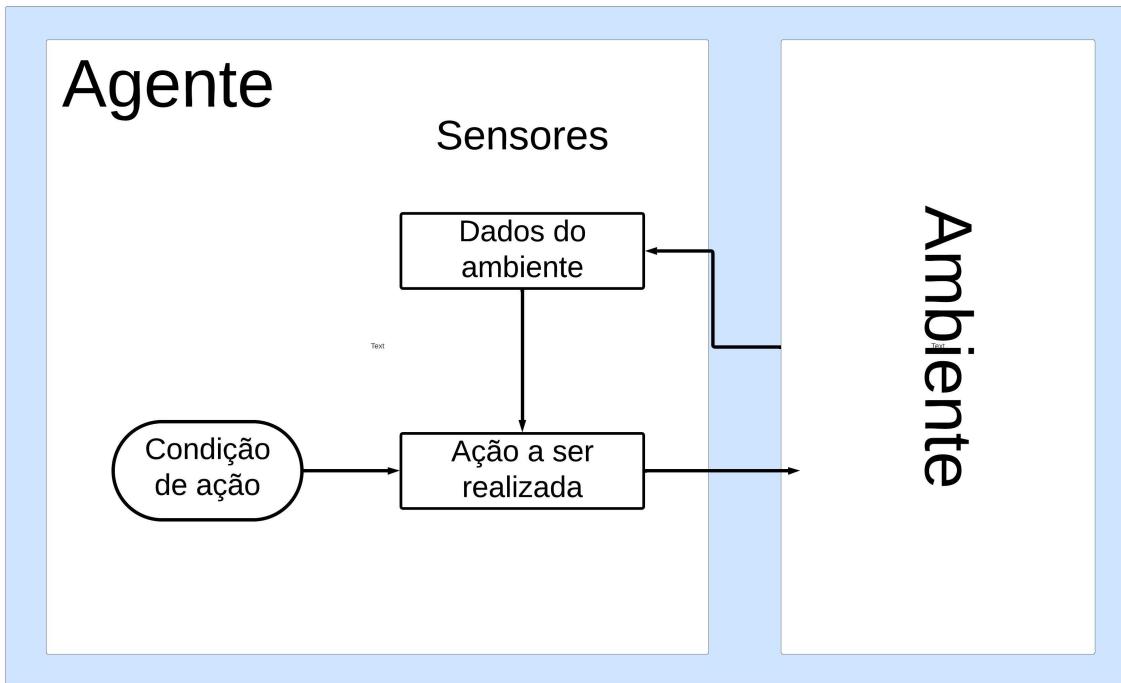


Figura 6 – Representação das atividades que envolvem o agente inteligente.

Os agentes inteligentes encontram aplicação em diversos domínios, desde sistemas de automação residencial e assistentes virtuais até veículos autônomos e robôs industriais. Desempenhando um papel central na evolução da inteligência artificial, os agentes capacitam os sistemas a abordar problemas complexos e a adaptar-se eficazmente a ambientes dinâmicos.

Considerando a complexidade desta tarefa, sua natureza distribuída e a semelhança relativa com problemas previamente resolvidos, optaremos por adotar uma abordagem multiagente como parte da solução.

2.2 Conclusão do Capítulo

Nesta fundamentação teórica, foram explorados conceitos essenciais relacionados ao desenvolvimento da dissertação, abordando temas como mineração de dados educacionais, *dashboard* educacional, visualização de dados e agentes inteligentes. Estabelecendo um entendimento sólido desses conceitos, destacando sua relevância no contexto educacional contemporâneo e sua aplicabilidade para o desenvolvimento da proposta em

questão.

Após a compreensão destes conceitos, o próximo capítulo exibirá os trabalhos relacionados, explorando pesquisas anteriores que serviram como base para o desenvolvimento da proposta. Esta análise comparativa permitirá situar o trabalho no contexto mais amplo da pesquisa acadêmica, identificando lacunas de conhecimento e oportunidades para a contribuição original.

3

TRABALHOS CORRELATOS

Neste capítulo são descritos os trabalhos identificados na literatura, sendo o estado da arte sobre os *dashboards* educacionais, além de investigar as características técnicas e pedagógicas que envolvem a área e como se dá a presença destas ferramentas.

3.1 Mapeamento Sistemático da Literatura

Para compreender as características dos STIs, visando auxiliar a etapa de análise de dados e construção dos *dashboards* para visualização de dados, esta parte do trabalho tem o objetivo de mapear características pedagógicas e técnicas relacionadas aos STIs, identificando aspectos como: nível educacional, plataforma de desenvolvimento, tecnologias envolvidas e a presença de *dashboard*.

A metodologia de pesquisa adotada neste trabalho, para coletar as informações para cumprir o objetivo, está fundamentada nos princípios da Engenharia de Software Experimental que se baseia na condução de um estudo secundário: Mapeamento Sistemático (MS).

O MS provê um espectro geral de um campo de pesquisa, detectando o número, os tipos de pesquisas concretizadas, os resultados disponíveis, além das frequências de publicações ao decorrer do tempo para identificar tendências ([PETERSEN et al., 2008](#)). Kitchenham e Charters (2007) afirmam que estudos em engenharia de software têm sido recomendados, sobretudo para áreas de pesquisa onde é difícil visualizar a gama de materiais relevantes e de alta qualidade, que possam estar disponíveis. Assim sendo,

a escolha do MS como suporte à condução desta pesquisa.

O MS foi fundamentado nas diretrizes desenvolvidas por Kitchenham e Charters (2007) e determinado em três fases: (a) Planejamento do Mapeamento: nessa etapa, os fins da pesquisa foram listados e o protocolo do mapeamento foi decidido; (b) Condução do Mapeamento: durante essa fase, as fontes para o mapeamento foram selecionadas, os estudos foram identificados e avaliados conforme os critérios estabelecidos no protocolo do mapeamento e (c) Resultado do Mapeamento: nessa fase, os dados dos estudos foram extraídos e sintetizados para serem publicados. Na Figura 7 pode-se observar as etapas do MS.

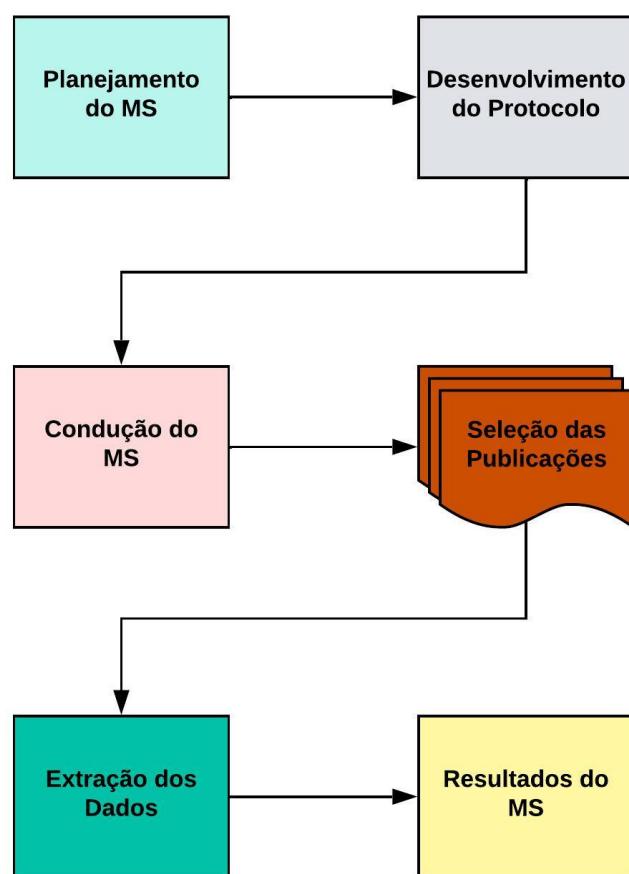


Figura 7 – Etapas do MS

A metodologia bem definida torna menos provável que os resultados encontrados na literatura sejam tendenciosos. Desta forma, este tipo de revisão é eficiente para mapeamento de áreas de estudo, onde é difícil visualizar o conjunto de materiais que possam estar disponíveis (KEELE et al., 2007).

3.1.1 Planejamento do Mapeamento

Nesta etapa foi definido o protocolo de pesquisa, que consiste em definir o objetivo do estudo, especificar as Questões da Pesquisa (QP), formular a expressão de busca, além de mencionar os procedimentos de extração dos dados e os critérios de seleção de cada publicação. A estrutura do protocolo foi baseada nos trabalhos de Kitchenham e Charters (2007).

3.1.1.1 Objetivos e Questões de Pesquisa

O objetivo da pesquisa foi detectar artigos publicados em conferências e periódicos que relatassem sobre as características técnicas e pedagógicas relacionadas aos STIs, para isso algumas QP foram elaboradas para auxiliar na investigação deste processo, sendo estas:

- QP1: quais as áreas de estudo do STI?
- QP2: quais camadas educacionais são cobertas pelo STI?
- QP3: quais as técnicas de Inteligência Artificial utilizadas nos STIs?
- QP4: quais os STIs utilizam *dashboards*?

3.1.1.2 Fontes, Idioma e Expressões de Busca

Os locais de busca definidos para esta pesquisa foram feitos, a partir de buscas automáticas em três bibliotecas digitais internacionais, que contam com grande acervo relacionado ao assunto de interesse, sendo elas: ACM, IEEE *Explore* e Springer *Link*.

Para a realização das buscas, o inglês foi escolhido como idioma devido à sua adoção pela maioria das conferências, periódicos e editoras da área de pesquisa. As buscas foram restringidas usando-se palavras-chave específicas para encontrar as publicações de interesse, a expressão de busca utilizada foi "*Intelligent Tutoring System*".

3.1.1.3 Critérios de Seleção

A seleção dos estudos foi realizada em duas etapas nos locais de busca definidos, se dividindo em duas partes:

1^a Seleção: Foram selecionados trabalhos publicados a partir do ano de 2012, lidos o resumo e as palavras-chave, os quais tiveram relação com o objetivo da pesquisa, foram agrupados em um conjunto preliminar de seleção. Para a escolha destes artigos, foram definidos critérios de inclusão (CI) e critérios de exclusão (CE), sendo eles:

CI1: A publicação deve propor um modelo de STI.

CI2: A publicação deve relatar as características pedagógicas do STI.

CI3: A publicação deve relatar as características técnicas do STI.

CE1: A publicação não aborda conceitos técnicos ou pedagógicos de um STI.

CE2: A publicação menciona um STI já descrito em outro artigo selecionado, baseado nos CI.

2^a Seleção: As publicações do conjunto preliminar foram lidas na íntegra e foram somente selecionadas as que atendiam os CI e CE. Estas publicações passaram para a fase de extração de dados.

3.1.1.4 Procedimento de Extração de Dados

A partir das publicações que passaram da segunda seleção, vários dados foram extraídos e registrados em tabela. A fim de organizar e facilitar o processo de obtenção de informações, estes dados foram categorizados em derivados da publicação, como: título da publicação e o nome dos autores. Além dos derivados do objetivo, como: Nome do STI, área de conhecimento, nível educacional, plataforma de desenvolvimento, tecnologias extras utilizadas no sistema e a presença de *dashboards*.

3.1.2 Condução do Mapeamento Sistemático

O MS foi realizado entre fevereiro de 2022 e abril de 2022 e as publicações foram selecionadas de acordo com critérios de seleção estabelecidos durante o protocolo do MS. Publicações duplicadas e aquelas não disponíveis na Internet foram excluídas. Além disso, foram excluídas as publicações que não se enquadram nos CI e CE, definidos na fase de critérios de seleção.

Inicialmente, foram retornadas 1327 publicações, como mostra a Tabela 1. Após a primeira análise, apenas 108 foram selecionadas para o segundo filtro, pois correspondiam com os critérios de seleção estabelecidos durante o protocolo do MS.

Tabela 1 – Publicações Selecionadas

Fonte	Busca Inicial	Após o 1º Filtro	Após o 2º Filtro
ACM	428	43	17
IEEE Explore	281	49	21
Springer Link	618	16	10

Depois da aplicação de todos os filtros, 48 publicações foram designadas para a próxima etapa, onde as informações serão coletadas, para isso os dados derivados das publicações serão inseridos em tabelas, conforme o processo de extração de dados descrito no protocolo do MS.

3.2 Resultados e Discussões

Com as informações obtidas nas 48 publicações selecionadas, foi possível responder às QP definidas no protocolo. Todas as publicações foram analisadas de forma manual, buscando extrair os dados importantes para o MS. Para mais detalhes, a relação dos STIs com as respectivas contribuições pode ser conferida no link: <https://cutt.ly/GLleJtK>.

3.2.1 QP1: Quais as áreas de estudo do STI?

Dentre os 48 STIs identificados, 45 relataram a disciplina a ser ensinada pelo sistema, ficando evidente o predomínio dos STIs voltados à área da ciência da computação,

compondo 34,1% dos *softwares* detectados. Nota-se que grande parte dos STIs pertencem à área das ciências exatas, 28 dos 39 STIs tem o foco nos campos da matemática e ciência da computação, entre os temas mencionados nos STIs voltados para área da computação, 6 fazem referência ao ensino da programação e apenas 1 menciona a linguagem a ser ensinada, sendo o PHP, através do PHP ITS. Na Figura pode-se observar a quantidade de STIs por campos de estudo.

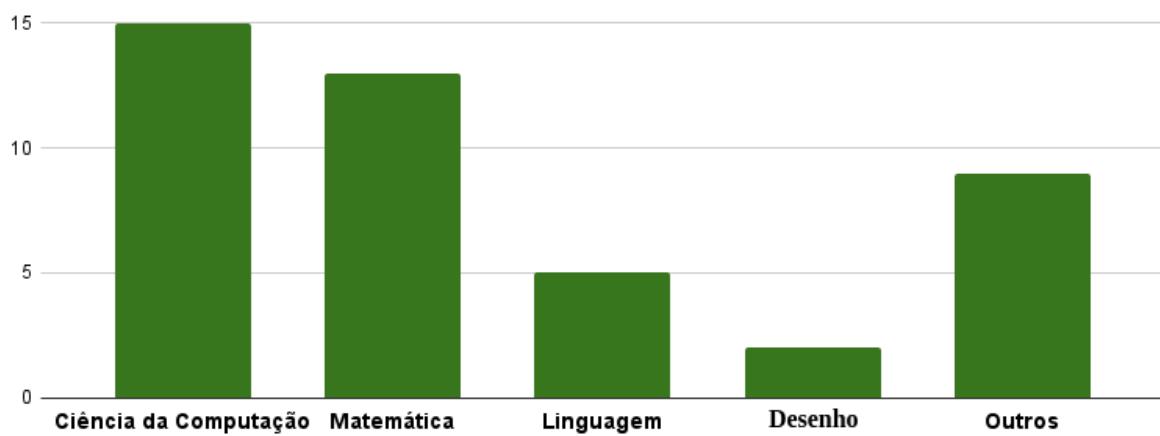


Figura 8 – Incidência de STIs por área de estudo.

No campo da matemática, responsável por 28,9% dos STIs identificados na pesquisa, também pode-se verificar a presença de 3 STIs voltados ao ensino da Algebra e outros 2 com o foco em exponenciação. Além de contar com 4 sistemas orientados ao ensino da matemática básica. 29 dos 42 sistemas encontrados fazem parte da área das ciências exatas, divididos entre tópicos envolvendo matemática, computação, física, tecnologia, eletrônica e eletricidade.

Com relação ao ramo das linguagens, apenas 5 STIs foram enquadrados nesta área, dentre estes, um sistema para o ensino de pontuação e outro para conceitos de pré-leitura, o restante abordou o ensino das línguas inglesa e portuguesa. Além disso, 5 sistemas foram categorizados na classe "outros", um destes abrange tópicos científicos, não relatando em qual subárea o STI foca.

Além destes, foi identificado um STI desenvolvido para o ensino geral. Por fim, apenas dois STI tiveram como temática a produção de desenhos técnicos da área de engenharia, ambos os sistemas fornecem um *feedback* sobre desenhos 2D feitos através

da tela, ajudando a melhorar o desempenho prático dos alunos.

3.2.2 QP2: quais camadas educacionais são cobertas pelo STI?

Os STIs identificados na pesquisa foram aplicados em cinco níveis educacionais: Ensino Fundamental, Ensino Médio, Ensino Superior, Educação Especial e Ensino Técnico.

Dentre os 48 STIs encontrados no MS, apenas 5 não fizeram referência ao nível educacional abordado pelo sistema, sendo eles: *IT2SN*, *STUART*, *MAELITS*, *SMIT* e *AI-Tutor*. Além disso, os STIs *Nsmav-Bot* e *SIMPLIFY* se classificam como voltados para todos os públicos. Outros 5 sistemas (*MathSpring*, *Dragoon*, *SketchTivity*, *Wayang Outpost* e outro STI para o ensino de matemática) consideram que alcançam dois níveis diferentes. Na Figura 9 pode-se constatar como os STIs estão distribuídos entre os níveis educacionais.

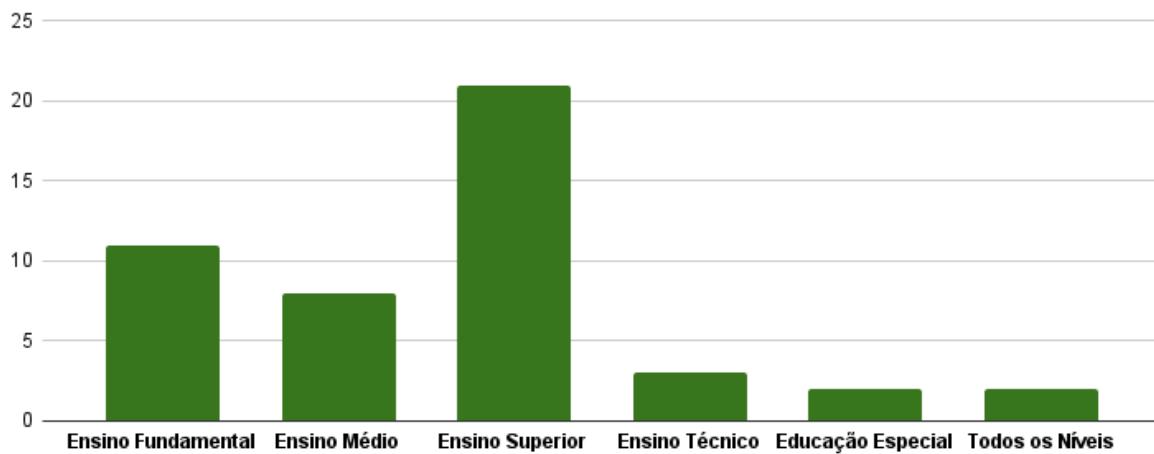


Figura 9 – Distribuição de STIs entre os níveis educacionais.

A partir dos dados extraídos, fica evidente a predominância dos STIs orientados as camadas do ensino superior, correspondendo a 21 sistemas, grande parte destes STIs foram desenvolvidos para o ensino de disciplinas relacionadas a área da ciência da computação, representando 15 dentre os 21 achados, evidenciando a falta de disciplinas envolvendo computação no ensino básico. Os outros STIs voltados ao ensino universitário têm o foco no ensino da matemática e desenho.

Os sistemas desenvolvidos para o Ensino Fundamental representam 10 dentre os 48 identificados no MS, atuando como a segunda camada mais alcançada pelos STIs. Grande parte destes softwares auxilia principalmente no ensino da matemática, cerca de 60%, representando um grande campo de interesse para pesquisadores. O restante dos sistemas abordam a tutoria de linguagem (30%) e Ciência e Tecnologia (10%).

Com relação à categoria Ensino Médio, 6 STIs foram encontrados, dentre estes, metade são focados no ensino da matemática, dando ênfase em álgebra e o ensino básico. O restante aborda disciplinas relacionadas à linguagem e desenho. Além disso, o ensino técnico também foi englobado por 3 STIs, mas apenas dois relataram a área de interesse, sendo eles: *SIMFOR* para formação em gestão de crises e o *Beetle-Grow*, focado em conceitos de eletricidade e eletrônica. Por fim, somente 2 sistemas fizeram referência ao ensino especial.

3.2.3 QP3: quais as técnicas de Inteligência Artificial utilizadas nos STIs?

Em resposta a questão, foram identificadas 35 técnicas utilizadas nos STIs, dentre estas, o Processamento de Linguagem Natural - PLN foi a mais aplicada, presente em 14,2% dos sistemas encontrados durante o MS, aplicado em 2 *chatterbots* nos STIs *Nsmav-Bot* e *STUART*, além de auxiliar no suporte ao ensino de programação e linguagens. Na Figura 10 pode-se observar algumas das técnicas encontradas durante o MS.

O MS também constatou um significativo número de menções ao uso de redes neurais, representando 17,6% dos STIs, um desses sistemas voltado para área da ciência da computação, relatou a importância da técnica relatando uma variedade de tarefas que podem ser adaptadas de diferentes estilos de aprendizagem, sendo possível implementar um modelo que avalia automaticamente o nível de conhecimento do aluno observando seu comportamento (IVE, 2018).

Alguns sistemas utilizaram algoritmos de classificação, como: *Support Vector Machine*, *Adaboost*, *Cluster-Based Classification*, *K-Nearest Neighbor - KNN*. Outros exemplos identificados são os STIs que utilizam Lógica Fuzzy, mencionados em 17,6% das

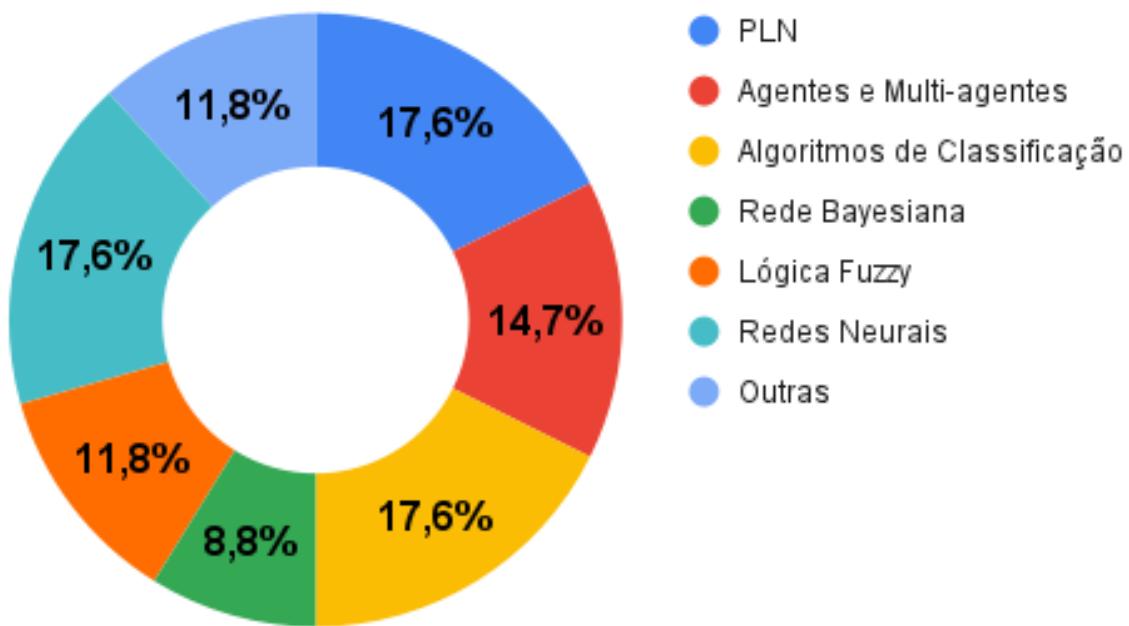


Figura 10 – Técnicas de inteligência artificial identificadas.

publicações.

Além destas técnicas, 14,7% STIs foram baseados na aplicação de agentes e multiagentes, os sistemas *MAELITS*, *FITS*, *SIMFOR*, *AICA* and *STUART* utilizam agentes inteligentes, a fim de analisar as interações no ambiente de aprendizagem e atuar a partir dos dados gerados durante a observação.

Por fim, 3 sistemas mencionaram o uso de Redes Bayesiana, todos voltados à área das ciências exatas. Além disso, outras 3 técnicas foram classificadas na categoria outros, sendo elas: *k-means clustering*, *hidden Markov model* and *Fuzzy Neural Network*.

3.2.4 QP4: quais STIs utilizam *dashboards*?

Dentre os 48 STIs identificados na pesquisa, 8 fizeram referências à exibição de *dashboards*, todos voltados para plataforma web. Na Figura 11 pode-se observar a prevalência de *dashboards* nos STIs.

OS STIs com *dashboard* estão presentes em todos os níveis educacionais escolares, indo do ensino fundamental ao superior. Com relação às áreas educacionais,

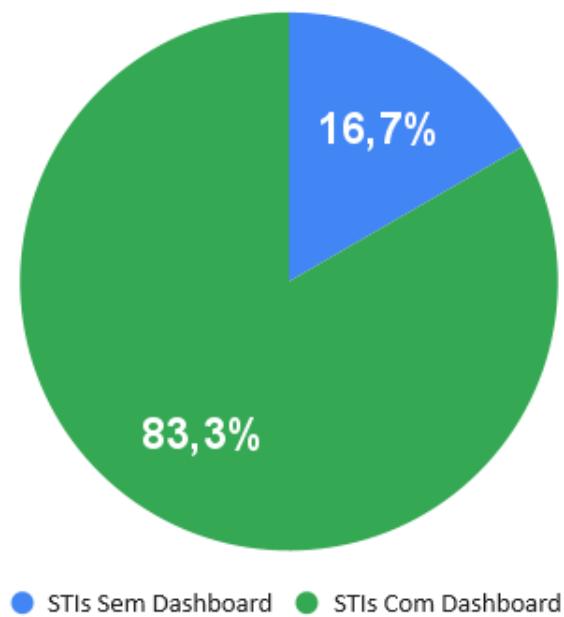


Figura 11 – Prevalência de *dashboards* em STIs

não apresentam diversidade, visto que todos os sistemas são voltados para as ciências exatas. A respeito do uso de ferramentas, o STI *Lynnette* se destaca pela utilização do *dashboard* Luna ([HOLSTEIN et al., 2010](#)), criado para ser utilizado com qualquer STI desenvolvido com a ferramenta *Cognitive Tutor Authoring Tools* (CTAT). Abaixo foram listadas as características dos *dashboards* implementados nos STIs.

MathSpring: O *dashboard* do STI *MathSpring* apresenta duas abas, “Meu Jardim” e “Meu Progresso”, onde os alunos podem acompanhar o progresso de aprendizagem atual, exibindo estatísticas sobre a participação, tópicos visitados, desempenho em cada assunto, número de problemas resolvidos, nível de domínio e detalhes sobre a resolução do problema.

Lynnette: Utilizou o *dashboard* Luna, sendo uma ferramenta de visualização de dados específica para professores que usam sistemas tutores inteligentes, fornecendo aos professores uma visão geral dos estudantes, incluindo o desempenho em atividades específicas, o tempo gasto em atividades e outras métricas relacionadas à aprendizagem. Além disso, permite aos professores personalizar a exibição de informações para atender às suas necessidades específicas, tal como, filtrar informações por grupos de estudantes ou acessar informações detalhadas sobre o progresso de cada estudante.

SMIT: O *dashboard* foi projetado utilizando a análise em três painéis separados: aluno, professor e administrador, empregando indicadores analíticos em três categorias: vídeo, atividade e uso do sistema, calculados com base nas interações dos alunos com o sistema. Os resultados dos indicadores individuais são exibidos no painel, bem como os indicadores relativos à média da turma. Assim, o aluno recebe a oportunidade de comparar sua própria situação segundo a média da turma. Os alunos podem organizar sua própria aprendizagem e estruturar seu processo de ensino e estratégias de estudo conforme o conselho e orientação do STI.

VerilUOC: Utiliza um painel inicial desenvolvido *em web-based*, resumindo todas as ameaças detectadas relacionadas à detecção de plágio, qualidade das submissões e informações estatísticas relacionadas ao número de envios, as corretas/incorrectas. Além disso, os erros não tratados também são exibidos no painel, podendo levar diretamente para os relatórios que ajudam a identificar os problemas encontrados.

AI-Tutor: Exibe o *dashboard* após a conclusão do curso, para que o aluno possa realizar a avaliação cognitiva e diagnóstica, conduzida para obter seu nível de conhecimento e a lista de deficiências. Os resultados do diagnóstico podem ser visualizados no painel de apresentação aos alunos por meio da interface do usuário.

TopoMath: O sistema inclui um sistema próprio de *dashboard* em tempo real e um sistema de gestão de conteúdos de aprendizagem. Os instrutores têm várias opções para controlar, como quando o *TopoMath* dá *feedback* sobre a condição de aprendizagem do aluno.

LEIA: Possui um ambiente de *dashboard* que permite aos usuários analisar dados de diferentes perspectivas, reunindo informações visualmente. No campo educacional, isso pode facilitar a visualização e compreensão dos dados gerados pelos estudantes. Neste contexto, o professor pode facilmente ver o progresso de um aluno em uma área específica, como álgebra, além de verificar sua participação nas atividades e frequência no sistema.

3.3 *Snowballing*

Após a condução da MS, empregamos a técnica de *snowballing* para frente, buscando identificar estudos adicionais relevantes. Essa técnica envolve a análise das citações dos artigos incluídos na revisão sistemática ([SILVA, 2017](#)). Para isso, utilizamos as bases de dados do Google Acadêmico, que permite rastrear quais novos artigos citaram os trabalhos originalmente selecionados.

O processo de *snowballing* para frente começou com a coleta dos artigos citantes, seguido por uma avaliação criteriosa de sua relevância com base no título e resumo. Quando necessário, o texto completo foi analisado para garantir que os artigos fossem pertinentes ao escopo da nossa pesquisa. Esse método permitiu identificar pesquisas mais recentes e relevantes que expandem e complementam os achados da nossa revisão inicial. Após a busca, os artigos encontrados foram adicionados nos trabalhos relacionados.

A aplicação do *snowballing* para frente foi crucial para assegurar a abrangência da nossa revisão da literatura, garantindo que incorporássemos tanto estudos históricos quanto desenvolvimentos recentes na área. Além disso, essa técnica contribuiu para o desenvolvimento do trabalho ao minimizar o risco de omissão de estudos significativos e proporcionar uma visão mais completa do panorama atual da pesquisa.

3.4 Trabalhos Relacionados

Esta subseção oferece uma análise de estudos e pesquisas anteriormente realizados sobre o tema. O propósito é inserir este estudo dentro do contexto mais abrangente da pesquisa existente, destacando lacunas, identificando tendências e reconhecendo contribuições significativas.

Uma das contribuições significativas sobre a relevância dos *dashboards* como instrumento educacional foi elaborada por ([SUSNJAK; RAMASWAMI; MATHRANI, 2022](#)), este estudo aborda as atuais práticas de criação de *dashboards*, destacando desafios enfrentados por profissionais da educação. A análise revela que a maioria dos painéis se limita a análises descritivas superficiais, com incorporação de poucas análises preditivas.

Em resposta, foi proposto um painel de última geração que integra análises descritivas com aprendizado de máquina, permitindo tanto análises preditivas quanto prescritivas. Essa abordagem inovadora busca preencher lacunas identificadas em painéis recentes, visando proporcionar ideias e aumentar a eficácia na adaptação dos estudantes.

O *dashboard* proposto não apenas fortalece a interpretação do comportamento preditivo do modelo, utilizando ferramentas analíticas emergentes, mas também atende aos requisitos regulatórios e aumenta a confiança dos alunos. Além disso, destaca-se a capacidade de fornecer conselhos concretos aos estudantes por meio de análises prescritivas baseadas em dados. Atualmente em fase de testes em uma instituição de ensino superior, este painel é pioneiro em sua amplitude de análises integradas, buscando impulsionar mudanças comportamentais e melhorar a eficácia do processo educacional.

Outro trabalho importante discute a implementação de uma solução de *Learning Analytics* para identificar padrões de comportamento de estudantes em Ambientes Virtuais de Aprendizagem (AVAs) e criar modelos preditivos para evasão e reprovação ([CECHINEL et al., 2023](#)). A estratégia envolve a instalação de um *plugin* no servidor *Modular Object-Oriented Dynamic Learning Environment* (Moodle) da instituição para coletar dados de interações dos estudantes no AVA. O Moodle, sendo um dos AVAs mais utilizados globalmente, facilita a coleta e análise de dados devido à sua estrutura modular e personalizável. Esses dados são processados em uma plataforma na nuvem e apresentados em um *dashboard* interativo, com o objetivo de fornecer informações importantes a educadores e administradores para detectar precocemente tendências de evasão ou reprovação e tomar medidas proativas para melhorar os desempenhos acadêmicos.

A solução foi desenvolvida considerando a necessidade de uma arquitetura em nuvem que exigisse pouco trabalho das equipes técnicas das instituições clientes, garantindo facilidade de uso e escalabilidade. A proteção de dados sensíveis, a gestão de grandes volumes de dados, a falta de dados históricos em algumas instituições e a expansão das funcionalidades do *dashboard* foram desafios enfrentados e considerados para futuras melhorias e aprimoramentos da ferramenta.

De forma semelhante, o artigo de ([KHELIFI et al., 2024](#)) apresenta o EX-LAD, um painel de *Learning Analytics* explicável desenvolvido para o ensino superior. O principal objetivo do EX-LAD é fornecer informações claras e comprehensíveis sobre o desempenho dos alunos, seu engajamento e perseverança no contexto educacional. O painel foi projetado para ser acessível tanto para professores quanto para alunos, permitindo que ambos acompanhem o progresso dos alunos, identifiquem aqueles em situação de risco e tomem decisões informadas para melhorar o desempenho acadêmico.

O estudo incluiu a realização de um caso prático com dados reais coletados de uma escola de engenharia na França, durante o ano letivo 2021-2022. Essa análise de caso serviu como evidência concreta da eficácia e do valor do EX-LAD no contexto educacional. Os resultados obtidos demonstraram como o painel pode auxiliar os professores, ajudando na identificação dos que estão em situação de risco e oferecer suporte direcionado. Além disso, os alunos também podem utilizar o painel para acompanhar sua própria jornada de aprendizado, identificar pontos fortes e fracos e tomar decisões embasadas para aprimorar seu desempenho acadêmico.

Já o trabalho de ([TAMADA; GIUSTI; NETTO, 2021](#)) propôs uma metodologia para prever estudantes em situação de risco no contexto de um curso auxiliado por um Sistema de Gerenciamento de Aprendizado (LMS), utilizando técnicas de Aprendizado de Máquina com base em registros do Ambiente de Aprendizado do Moodle. Esses dados foram utilizados para analisar o comportamento do aluno e criar um modelo que ajude a detectar estudantes em risco.

Os autores realizaram uma análise comparativa entre sete algoritmos com modelos treinados, com a intenção de explorar o equilíbrio entre a detecção precoce e tardia de alunos em risco. Os resultados mostram sete modelos criados para previsão. As descobertas indicaram que o *Random Forest* teve o melhor desempenho na previsão do desempenho do aluno. O estudo forneceu um modelo de previsão totalmente capaz de orientar estudos e o desenvolvimento de ferramentas para reduzir as elevadas taxas de desistência.

Abordando a carência de estudos relacionados à previsão de desempenho nos cursos introdutórios de programação, o trabalho de ([PEREIRA et al., 2021](#)) realizou

uma análise de dados educacionais baseada em logs muito detalhados de 2056 alunos, coletados do sistema *CodeBench*. Os autores extraíram *features* que medem como os alunos lidam com prazos, como corrigem erros, quanto tempo gastam programando, entre outros aspectos. Em seguida, construíram um modelo preditivo que obteve ótimos resultados.

Os autores explicaram de forma detalhada os efeitos das features, visualizando e analisando as previsões individuais, a importância global das features e a identificação de caminhos típicos de previsão. O método utilizado pode ser aplicado a outros modelos competitivos emergentes, à medida que o campo de previsão de disciplinas introdutórias de programação avança, garantindo transparência no processo para as partes interessadas-chave: administradores, professores e alunos.

Em outra pesquisa, ([RUIPÉREZ-VALIENTE et al., 2021](#)) desenvolveram o *Shadowspect*, um jogo de quebra-cabeça geométrico 3D utilizado por professores em um grupo de escolas nos EUA. O sistema foi construído utilizando técnicas de análise de aprendizado para gerar um conjunto de métricas implementadas em um painel ao vivo que visa facilitar aos professores entender a interação dos alunos com *Shadowspect*.

Além disso, os autores forneceram casos de uso que exemplificam como os professores podem usar o painel para compreender o progresso global de sua turma e de cada aluno individualmente, a fim de intervir, adaptar suas aulas e dar *feedback* personalizado quando apropriado.

Um dos principais trabalhos sobre a importância de *dashboards* como ferramenta educacional foi desenvolvido por ([NARANJO et al., 2019](#)), apresentando o *CloudTrail-Tracker*, uma plataforma de código aberto que fornece ao instrutor um painel visual que retrata o uso agregado e individual dos recursos utilizados por todos os alunos durante um determinado período. Para facilitar a autorregulação dos alunos, o painel também mostra a porcentagem de progresso de cada sessão do laboratório e as ações pendentes do aluno. O painel foi integrado em quatro disciplinas que usam diferentes metodologias de aprendizado, seja presencial ou online, destacando positivamente a utilidade da ferramenta para instruções de Nuvem na *Amazon Web Services - AWS*. A ferramenta adquire de forma automática os dados, gerando análises de aprendizado quase em

tempo real, úteis tanto para avaliação semiautomática quanto para autoconsciência do aluno sobre o seu próprio progresso na formação. A Figura 12 exibe a arquitetura do *CloudTrail-Tracker*.

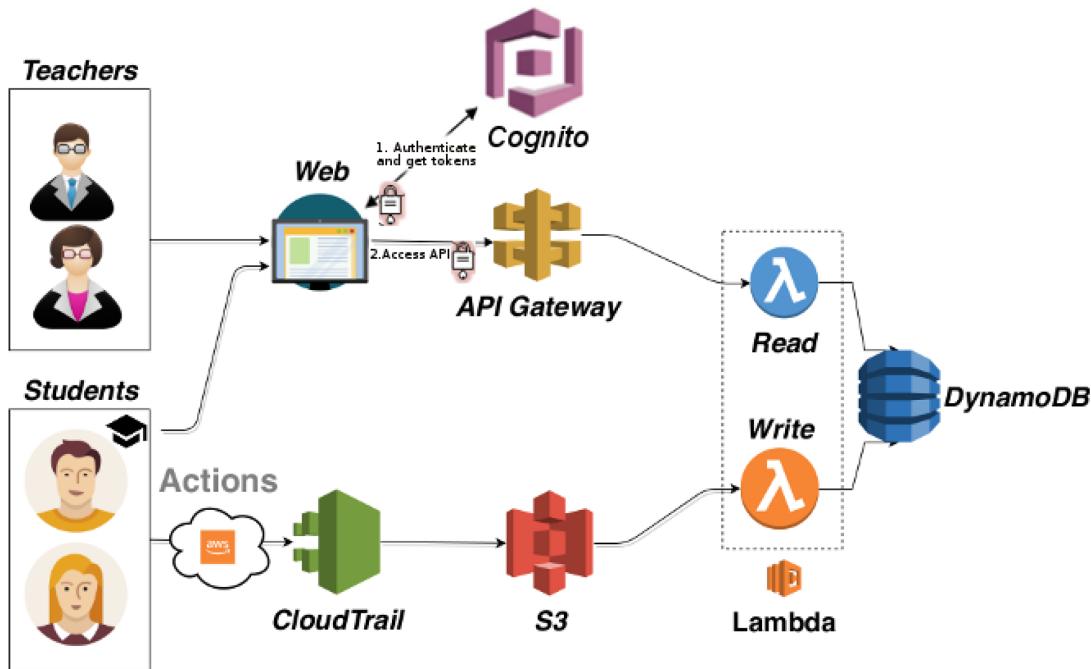


Figura 12 – Arquitetura - CloudTrail-Tracker

Em outro artigo importante ([MAJUMDAR et al., 2019](#)) aborda o desenvolvimento do *Learning Analytics Visualizations and Evidence Widgets - LAViEW*, um *dashboard* para análise de aprendizagem visando auxiliar alunos e professores, fornecendo evidências sobre o comportamento de aprendizagem e o desempenho acadêmico, a ferramenta também disponibiliza métricas como *feedbacks* personalizados e pontuações por engajamento.

O uso do *dashboard* também permite ao professor projetar a intervenção buscando mitigar as dificuldades dos estudantes, em seguida, monitorar a eficácia. Os autores relatam que a tecnologia pode ajudar a captar os detalhes do processo educacional, causando uma reflexão sobre a eficácia das práticas no âmbito acadêmico.

O trabalho de ([PARK; JO, 2015](#)) relatou o desenvolvimento do aplicativo *Learning Analytics dashboard* (LAD), visando dar suporte e exibir o desempenho acadêmico dos alunos, identificando padrões de comportamento em um ambiente virtual de aprendizagem. A aplicação funciona através do rastreio de arquivos de registro dos

alunos, extraindo uma abundância de dados, visando auxiliar a visualização de dados. Os autores também aplicaram um teste de usabilidade, a avaliação da ferramenta foi realizada em um ambiente de pesquisa experimental com um grupo de controle, pesquisas adicionais indicaram que as lições aprendidas com a usabilidade e os testes-piloto demonstram que as informações visualizadas impactam no nível de compreensão dos alunos.

Focado no desenvolvimento de um *dashboard* para análise e visualização de dados educacionais, Paranhos (2021) abordou toda a construção da ferramenta, demonstrando aspectos relacionados aos tipos de dados analisados e como as informações poderiam ser extraídas. A autora utiliza a ferramenta *PowerBi* para dar suporte à análise de dados, criando os recursos visuais necessários para compreensão de todo o cenário educacional.

Após a extração dos dados originados do banco de dados do instituto de Ciências Exatas e Aplicadas da Universidade Federal de Ouro Preto, as informações foram organizadas através da criação de 7 *dashboards* interativos, sendo eles o demográfico, de admissão, de evasão, colação, desempenho aluno, desempenho da matéria e desempenho do departamento ([PARANHOS et al., 2021](#)).

A conclusão do artigo afirma que o painel de análise de aprendizagem é uma ferramenta valiosa para ajudar os educadores a avaliar o desempenho dos estudantes e tomar decisões informadas sobre como melhorar a educação. Além disso, o artigo destaca a importância de fornecer informações sobre o desempenho dos estudantes de uma forma acessível e fácil de entender, o que pode ajudar a melhorar a educação e o desempenho escolar. O trabalho de Paranhos difere deste em relação às fontes de dados e às ferramentas de desenvolvimento utilizadas, visto que este projeto foca nos dados voltados para STIs e também voltados para a plataforma web.

O trabalho de Schmitt et al. 2019 apresentou o desenvolvimento de um aplicativo móvel (*iPermanency*) que dispõe das informações geradas por meio de MDE, para que gestores interessados possam refletir sobre seus resultados. A arquitetura proposta utiliza as bases de dados que dispõe das informações utilizadas no processo de MDE, sendo elas: dados de comunicação do MOODLE e dados do sistema institucional.

Na outra parte o trabalho utiliza componentes que realizam a transformação

e MDE, sendo eles: uma base de dados única, fruto da integração entre os dados do MOODLE e o sistema institucional; uma aplicação Java que utiliza a API do WEKA para efetuar a técnica de MDE e armazenar os dados resultantes do processo; e um *WebService* que disponibiliza os dados resultantes do processamento para serem acessados pelo aplicativo ([SCHMITT; BERNARDI; KANTORSKI, 2019](#)).

Na análise do sistema, ficou esclarecido que a disponibilização dessas informações pode ser grande aliada da gestão acadêmica, em busca da atenuação dos números elevados de evasão escolar ([SCHMITT; BERNARDI; KANTORSKI, 2019](#)). O trabalho de Schmitt et al. 2019 difere deste projeto, pois busca analisar dados provenientes de um ambiente virtual de aprendizagem, enquanto o projeto tem o foco em um STI, logo existe uma mudança nos tipos de dados, considerando os aspectos pedagógicos.

O trabalho de ([SILVA; NETTO; SOUZA, 2018](#)) propõe a criação de um *dashboard* para monitorar e avaliar o desempenho dos estudantes de matemática no ensino médio. A ferramenta trabalha no apoio aos professores, permitindo-lhes ter uma visão geral dos resultados dos estudantes em diferentes tarefas e avaliações. Além disso, o painel permite aos estudantes e pais terem acesso a informações sobre o progresso e desempenho escolar de forma clara e fácil de entender.

O VLA *dashboard* inclui gráficos e tabelas interativas para exibir informações como notas, frequência, participação em tarefas e avaliações, e outras métricas relevantes. Além disso, o painel conta com recursos de filtragem e classificação para permitir que os usuários vejam os dados de uma forma mais específica e personalizada. A Figura 13 exibe uma das telas do VLA *dashboard*.



Figura 13 – Tela de VLA *dashboard* - Fonte: ([SILVA; NETTO; SOUZA, 2018](#))

O artigo de ([SILVA; CARVALHO; MACIEL, 2021](#)) descreve a criação de um *dashboard* de análise de aprendizagem para modelos de mineração de dados educacionais, com foco em fornecer uma ferramenta visual para apresentar os resultados de modelos de mineração de dados educacionais de uma forma clara e fácil de entender.

O *dashboard* apresentado no artigo é projetado para ser interativo e permitir aos usuários explorar e analisar os dados de aprendizagem de várias perspectivas diferentes. Ele inclui gráficos e tabelas para exibir informações como desempenho do estudante, padrões de uso do sistema e outras métricas relevantes. Além disso, o *dashboard* permite aos usuários filtrar e classificar os dados de acordo com suas necessidades, ajudando a fornecer uma visão mais clara e personalizada dos resultados.

A conclusão do artigo afirma que o *dashboard* de análise de aprendizagem desenvolvido é uma ferramenta útil para a apresentação dos resultados dos modelos de MDE. Além disso, os autores destacam a importância de fornecer informações sobre o desempenho dos estudantes de uma forma acessível e fácil de entender, o que pode ajudar a melhorar a educação e o desempenho escolar.

Os seguintes artigos publicados durante o mestrado, serviram de base para a conclusão da pesquisa, as publicações ([MENEZES; NETTO; LOPES, 2022](#)) ([MENEZES](#)

et al., 2023a) (MENEZES et al., 2023b) detalham o processo de análise dos dados gerados por um STI, sendo possível a descoberta de informações importantes, como a definição do nível das questões propostas pelo sistema e a classificação do desempenho dos estudantes. Em uma das análises, foi constatado que 64% das questões matemáticas aplicadas para os estudantes estavam niveladas incorretamente, fazendo com que o processo evolutivo do aluno fosse prejudicado.

Esta proposta se diferencia das demais ao direcionar sua atenção para os STIs e suas características distintivas como foco principal e tipos de dados a serem analisados. Os STIs apresentam complexidade e heterogeneidade, compartilhando características únicas que os tornam desafiadores para análise. Além disso, o projeto incorpora agentes inteligentes com o propósito de desenvolver uma ferramenta avançada de visualização de dados, possibilitando aos usuários explorar e compreender os dados de STIs de maneira mais eficaz. Essa abordagem visa proporcionar conhecimentos mais profundos sobre esses sistemas complexos. No Quadro 1 pode-se observar as características dos *dashboards* relatados na sessão e o *dashboard* proposto.

Quadro 1: Comparativo dos trabalhos relacionados com a proposta

#	Ferramenta	Ano	Origem dos Dados	Alvo	Técnicas	Tipos de Gráficos	Tipo de Sistema
1	LAD	2015	Arquivo de registro dos alunos	Professores e Estudantes	Visualização de dados e MDE	Barra	Moodle
2	VLA Dashboard	2018	Arquivo de registro dos alunos e logs de sistema	Professores	Visualização de dados	Barra e setores	Moodle
3	CloudTrail-Tracker	2019	Arquivo de registro dos alunos e logs de sistema	Professores e Estudantes	Visualização de dados	Barra e setores	LMS
4	LAView	2019	Arquivo de registro dos alunos	Professores e Estudantes	Visualização de dados	Barra	LMS
5	iPermanency	2019	Arquivo de registro dos alunos e logs de sistema	Gestores	Visualização de dados e MDE	Barra	Moodle
6	Dashboard - ICEA	2021	Arquivo de registro dos alunos	Professores e Gestores	Visualização de dados	Barra e setores	Não Declarado
7	LAD MMDE	2021	Arquivo de registro dos alunos	Professores	Visualização de dados e MDE	Tabelas e cards	LMS
8	Susnjak LAD	2022	Arquivo de registro dos alunos e logs de sistema	Estudantes	Visualização de dados e MDE	Barra, diagrama de caixa e linha.	Moodle
9	LANSE	2023	Arquivo de registro dos alunos	Professores e Gestores	Visualização de dados e MDE	Barra e linha.	Moodle
10	EX-LAD	2024	Arquivo de registro dos alunos	Professores e Estudantes	Visualização de dados e MDE	Barra, linha e radar.	LMS
11	Edudata Visualizer	2024	Arquivo de registro dos alunos e logs de sistema	Professores	Visualização de dados, MDE e Classificação.	Barra, barra de progresso, cards, setores, tabelas e linhas.	STI

3.5 Conclusão do Capítulo

Na conclusão deste capítulo de trabalhos relacionados, foi conduzido um mapeamento sistemático para compreender a presença e a eficácia dos STIs e a incidência de *dashboards* nesses sistemas. A análise revelou uma lacuna significativa em relação à implementação de *dashboards* educacionais em STIs, indicando uma oportunidade para contribuição neste campo de pesquisa.

Além disso, identificaram-se diversos trabalhos que exerceram influência sobre a pesquisa, fornecendo base e abordagens que ajudaram a moldar a proposta do *dashboard* educacional aqui apresentada. Essa revisão da literatura permitiu a contextualização da pesquisa dentro do corpo existente de conhecimento.

No próximo capítulo, é apresentado o desenvolvimento da proposta do *dashboard* educacional. Este capítulo detalhará os aspectos técnicos e metodológicos da nossa solução, demonstrando como ela se alinha com as necessidades identificadas na literatura e oferece uma contribuição significativa para o avanço do campo de sistemas tutores inteligentes na educação.

4

PROPOSTA

Este capítulo visa fornecer uma análise abrangente dos elementos fundamentais para a compreensão do processo de desenvolvimento do *Dashboard Edudata Visualizer*. Serão apresentados detalhes sobre a arquitetura do sistema, as ferramentas adotadas durante o processo de desenvolvimento, bem como a fase crucial de prototipação e análise dos dados. Estas etapas foram conduzidas, visando fornecer uma base sólida e informada para a construção e implementação eficazes da ferramenta. O entendimento completo desses aspectos é crucial para uma visão holística do projeto e para a apreciação de sua contribuição no contexto educacional.

4.1 Visão Geral do Sistema

A proposta desta pesquisa é conceber e implementar um *dashboard* interativo e intuitivo, destinado à visualização dinâmica do desempenho dos estudantes em STIs. Para a primeira versão deste sistema, deliberou-se pela utilização do STI LEIA, baseando-se na capacidade de registrar informações detalhadas dos usuários em seu banco de dados. Além disso, o STI LEIA proporciona a flexibilidade necessária para personalizações específicas, essenciais na avaliação contínua da aprendizagem dos estudantes. É importante ressaltar que este projeto está sendo conduzido no âmbito do Laboratório de Sistemas Inteligentes da Universidade Federal do Amazonas. A Figura 16 exibi a arquitetura proposta para o desenvolvimento do *dashboard*.

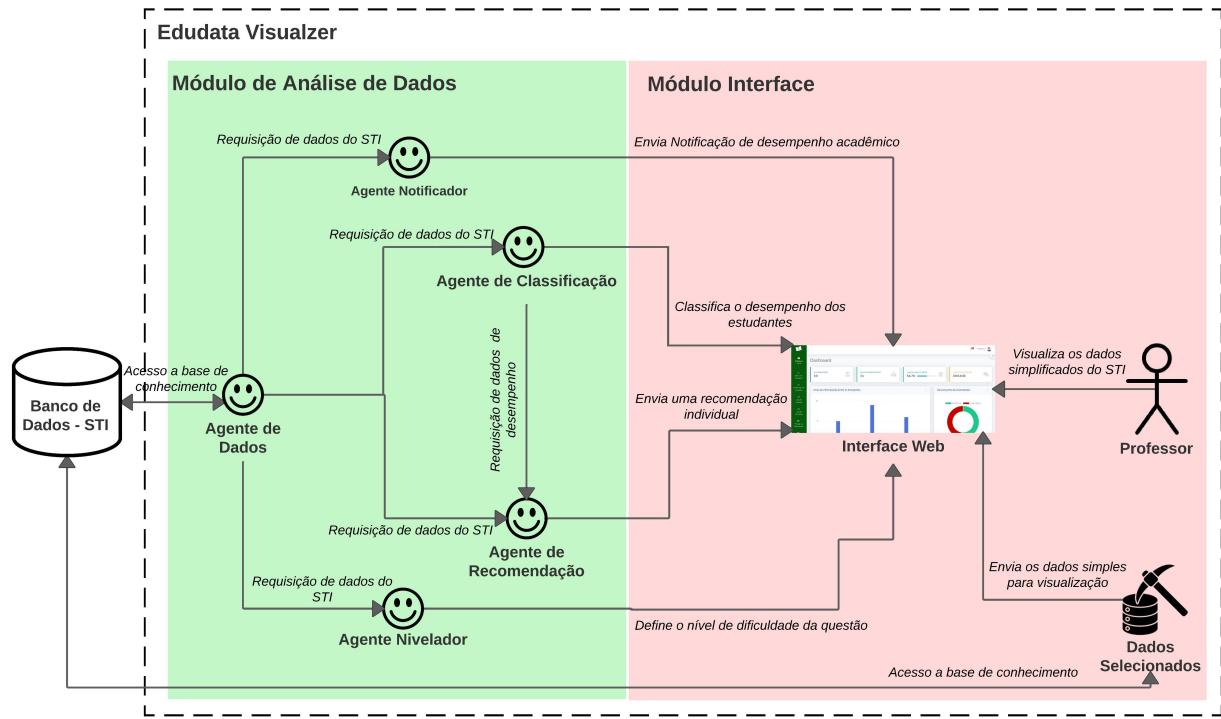


Figura 14 – Arquitetura de desenvolvimento *Edudata Visualizer*

A arquitetura proposta consiste em duas partes distintas, cada uma desempenhando um papel fundamental no processo. Inicialmente, os dados são gerados utilizando o STI LEIA. Esses dados são então registrados em um banco de dados e encaminhados para dois módulos. No Módulo de Análise de Dados, ocorre o processo de Mineração de Dados Educacionais (MDE), que visa extrair as informações mais relevantes e complexas. Este processo envolve a conversão de abordagens analíticas em código. Neste contexto, cinco agentes inteligentes operam para realizar a análise de dados, por exemplo, o agente de classificação atua através da conversão da análise realizada pelo algoritmo de Árvore de Decisão em uma função em PHP. Essa função atua como um agente inteligente e é posteriormente integrada ao código principal do *Dashboard*, agregando valor ao sistema como um todo.

Após a conclusão da análise de dados, o Módulo de Interface atua recebendo os dados que não requerem tratamento adicional diretamente do banco de dados do STI LEIA. Esses dados incluem variáveis essenciais, como o nome do estudante, informações das questões, a precisão da resposta, o tempo dedicado à questão e o nível de dificuldade da questão. Além disso, a Interface também é responsável por

receber os resultados gerados pelo Módulo de Análise, que incluem a classificação do desempenho dos estudantes, a avaliação do nível de dificuldade das questões e um relatório detalhado de recomendações educacionais.

Dentro do Módulo de Interface, esses dados são processados e apresentados de forma clara e acessível aos professores. Isso envolve a criação de gráficos, tabelas e outras representações visuais para facilitar a compreensão e a tomada de decisões informadas.

Além disso, a Interface também permite a interação direta com os dados, possibilitando aos usuários explorar informações específicas, gerar relatórios personalizados. Portanto, o Módulo de Interface desempenha um papel crucial na transformação dos dados brutos em informações valiosos e acionáveis, capacitando os envolvidos na educação a tomar medidas eficazes para melhorar o desempenho dos estudantes e otimizar o processo educacional na totalidade.

4.2 Descrição dos Agentes

A otimização das operações do sistema é alcançada por meio da integração de cinco agentes inteligentes, os quais desempenham funções cruciais na análise de dados e na emissão de notificações, oferecendo suporte essencial ao processo de tomada de decisão. Esses agentes inteligentes assumem responsabilidades específicas para capturar, processar e interpretar informações relevantes, visando fornecer conhecimentos significativos que fundamentam decisões informadas.

Adicionalmente, ao fornecerem notificações proativas e personalizadas, esses agentes facilitam a comunicação eficiente de informações críticas, permitindo uma resposta rápida e assertiva às condições em constante evolução. Dessa forma, a presença e contribuição desses agentes inteligentes revelam-se fundamentais para a eficiência e eficácia do sistema, assegurando sua capacidade de adaptação e desempenho otimizado. O texto abaixo descreve os 5 agentes desenvolvidos para o *Edudata Visualizer* e os seus respectivos papéis.

Agente de Dados (A): atua como a conexão primária com o banco de dados,

permitindo que outros agentes interpretem informações cruciais. Estes incluem: taxa de acerto, tempo online, acerto das questões, nível de dificuldade das questões e tempo de resolução das atividades.

Agente Notificador (B): desempenha uma função essencial ao receber e formatar os dados do Agente (A), monitorando a taxa de acerto e o tempo de atividade da turma. Notifica imediatamente o professor por meio de alertas visíveis na área superior direita de todas as telas do sistema.

Agente de Classificação (C): utiliza o método de árvore de decisão para avaliar o desempenho individual de cada aluno. Analisa os dados fornecidos pelo Agente (A), os quais são inseridos na árvore de decisão, gerando assim uma classificação específica para cada conjunto de dados.

Agente de Recomendação (D): gera relatórios detalhados sobre o desempenho dos estudantes e oferece recomendações específicas para melhorar sua atuação acadêmica. Este agente interage com o Agente (A) para obter os dados necessários e, em seguida, interpreta esses dados com base nas variáveis e classificações do Agente (C).

Agente Nivelador (E): desempenha um papel crucial na avaliação da dificuldade das questões propostas pelo sistema. Utiliza a escala do INEP para classificar as questões como: muito fáceis, fáceis, médias, difíceis ou muito difíceis. As variáveis como questão, precisão e nível de dificuldade são obtidas por meio da interação com o Agente (A), e então categorizadas adequadamente. Este processo permite uma melhor adaptação do nível de dificuldade das atividades aos diferentes níveis de habilidade dos alunos.

4.3 Diagramas de Caso de Uso

O diagrama de caso de uso é uma das principais ferramentas utilizadas na Linguagem de Modelagem Unificada (UML) para capturar os requisitos de um sistema de *software* do ponto de vista do usuário. Descreve as interações entre os usuários (ou atores) e o sistema, destacando os diferentes cenários de uso ([LETAW, 2024](#)).

A UML utilizada neste projeto é uma linguagem de modelagem padronizada utilizada na engenharia de *software* para visualizar, projetar, especificar e documentar

sistemas complexos. Ela oferece um conjunto de notações gráficas que permitem aos desenvolvedores representar de forma visual os diversos aspectos de um sistema de software, como sua estrutura, comportamento, interações e processos (LETAW, 2024).

Para exemplificar as interações entre os atores do sistema, foi empregado dois diagramas de caso de uso, visando a demonstração detalhada das interações entre o ator "professor" e a interface do sistema. Além disso, os diagramas foram utilizados para representar os 5 agentes inteligentes e as suas respectivas interações com o sistema. Essa abordagem proporciona uma visão abrangente e clara das funcionalidades oferecidas aos usuários, bem como dos fluxos de interação entre os diferentes elementos do sistema, auxiliando no entendimento do projeto. Na Figura 15 é exibido o diagrama de caso de uso do ator professor.

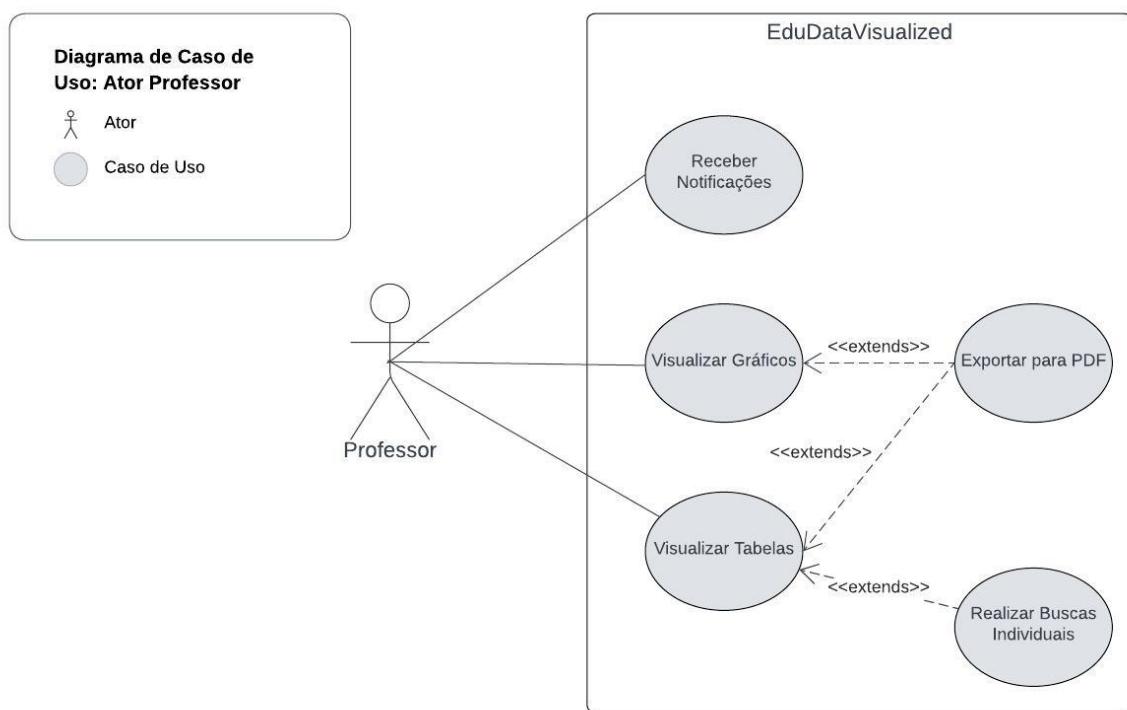


Figura 15 – Diagrama de caso de uso: Professor

Descrição de cada caso de uso do ator professor:

- **Receber Notificações:** o ator professor é regularmente atualizado pelo sistema com uma variedade de notificações que detalham o desempenho atual dos estudantes.

- **Visualizar Gráficos:** a interpretação do ambiente educacional pelo Ator professor é facilitada pela presença de uma variedade de gráficos, que permitem uma visualização clara e eficaz dos dados.
- **Visualizar Tabelas:** O ator professor pode visualizar as informações por meio de tabelas que mostram dados relacionados às questões e aos alunos, incluindo a classificação do desempenho educacional.
- **Exportar para PDF:** Cada tela de visualização de dados pode ser exportada para o formato PDF, permitindo que o ator professor realize o download dos arquivos.
- **Realizar Buscas Individuais:** Na tela de dados individuais, o ator professor pode buscar dados individuais de cada estudante, tendo uma visão mais ampla da trajetória de aprendizagem do aluno.

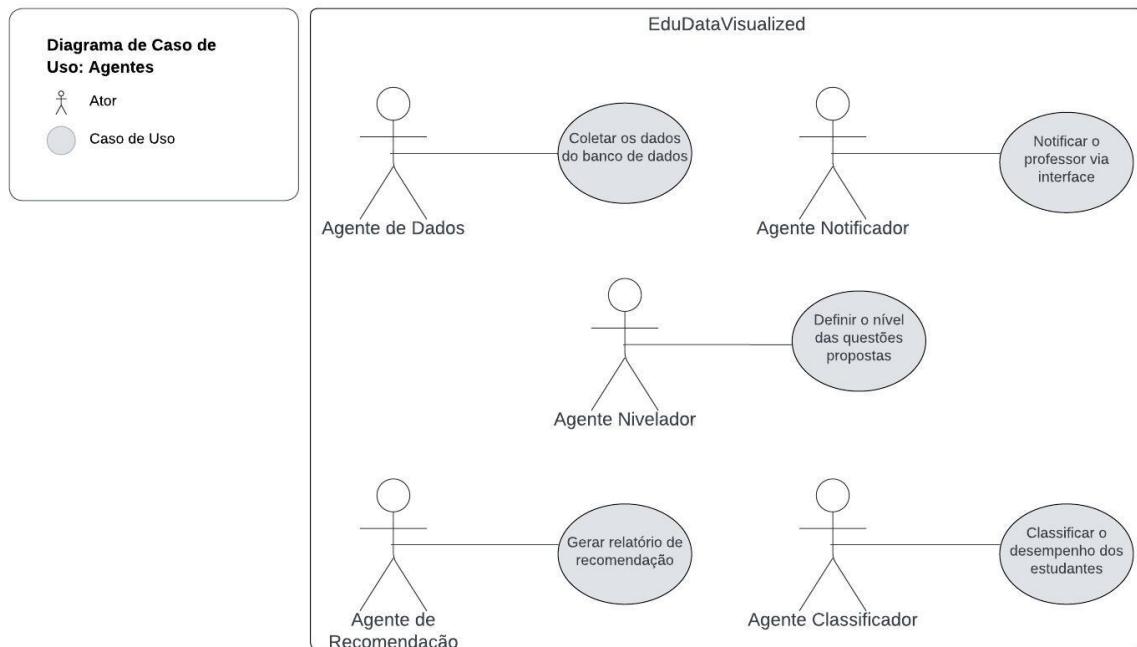


Figura 16 – Diagrama de caso de uso: Agentes

Descrição dos casos de uso dos agentes:

- **Agente de dados - Coletar os dados do banco de dados:** realiza a coleta das variáveis de interesse diretamente do banco de dados do STI LEIA, tornando-as

disponíveis para os demais agentes e assegurando o pleno funcionamento dos recursos gráficos.

- **Agente Notificador - Notificar professor via interface:** envia notificação via barra superior direita na interface, informando a situação atual do desempenho dos usuários.
- **Agente Nivelador - Definir o nível das questões propostas:** atua na tabela de questões, alterando a classificação de dificuldade das atividades propostas pelo sistema.
- **Agente de Recomendação - Gerar relatório de recomendação:** atua na tabela de questões, alterando a classificação de dificuldade das atividades propostas pelo sistema.
- **Agente Classificador - Classificar o desempenho dos estudantes:** realiza a classificação de desempenho dos estudantes, baseando-se na análise feita com árvore de decisão.

4.4 Ferramentas Utilizadas no Desenvolvimento

Para o desenvolvimento do front-end, foi utilizado o framework Bootstrap, visando dar suporte a criação de interfaces de usuário com recursos intuitivos, modernos e fáceis de usar, além de contar com um grande acervo de informações sobre a utilização da ferramenta, se destacando pela facilidade de integração com outros aplicativos. Ademais, o uso desta arquitetura responsiva permite que o sistema seja escalável para qualquer tipo de aparelho, como mostra a Figura 17.

Para apoiar a parte de visualização de dados, foram implementados modelos gráficos disponibilizados pelo serviço Web Google Charts, fornecendo gráficos interativos, por meio de uma simples ligação expressa em JavaScript. Além disso, também foi utilizada a biblioteca JavaScript Chart.js, que gera gráficos baseados em HTML, sendo uma das bibliotecas de visualização mais simples para JavaScript. Na Figura 18 pode-se observar a primeira versão do *dashboard* inicial.



Figura 17 – Tela inicial do *dashboard* no formato Smartphone

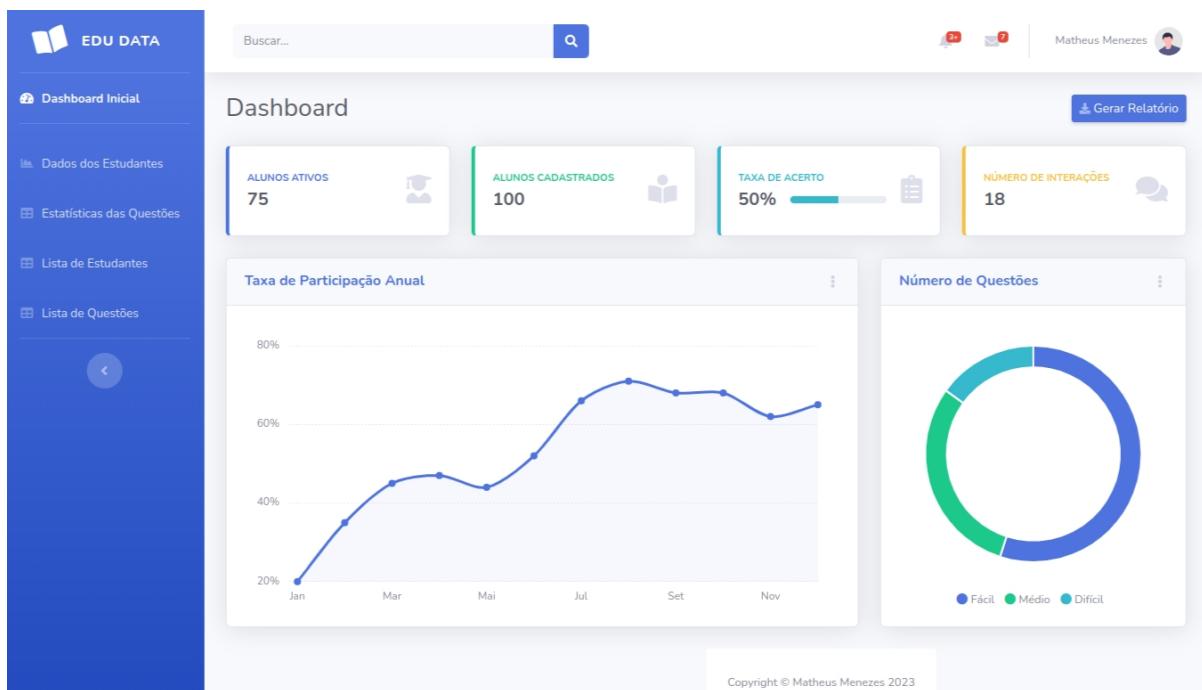


Figura 18 – Primeira versão da tela inicial

No que diz respeito ao back-end, foi empregado o phpMyAdmin, uma aplicação web de código aberto desenvolvida em PHP para facilitar a administração do MySQL pela Internet. Através deste sistema, foi obtida a capacidade de criar e eliminar bases de dados, gerenciar tabelas, inserir, excluir e modificar campos, e executar consultas SQL. Para estabelecer conexões com o banco de dados e implementar as funcionalidades

nos agentes inteligentes, foi utilizado o PHP, uma linguagem de script open source amplamente utilizada, especialmente reconhecida por sua adequação ao desenvolvimento web. Neste caso, o PHP foi integrado ao HTML para aprimorar a interatividade e dinamismo da interface.

Para conduzir as operações de aprendizado de máquina e mineração de dados, foi empregada a ferramenta Weka, cujo propósito foi realizar experimentos de análise de dados, com o intuito de identificar o modelo mais adequado a ser implementado no sistema.

4.5 Módulo de Análise de Dados

O projeto para o desenvolvimento do Módulo de Análise de Dados foi inspirado na metodologia de descoberta do conhecimento. O processo de descoberta de conhecimento é mostrado na Figura 19 como uma sequência iterativa das seguintes etapas ([HAN; KAMBER; PEI, 2022](#)):

- Limpeza de dados para remover ruídos e dados inconsistentes.
- Integração de dados onde várias fontes de dados podem ser combinadas.
- Seleção de dados onde os dados relevantes para a tarefa de análise são recuperados do banco de dados.
- Transformação de dados onde os dados são transformados e consolidados em formas apropriadas para mineração, executando operações de resumo ou agregação.
- Mineração de dados, sendo um processo essencial onde métodos inteligentes são aplicados para extrair padrões de dados.
- Avaliação de padrões para identificar os padrões que representam o conhecimento com base em medidas de interesse.
- Apresentação do conhecimento, onde técnicas de visualização e representação de conhecimento são usadas para apresentar as informações.

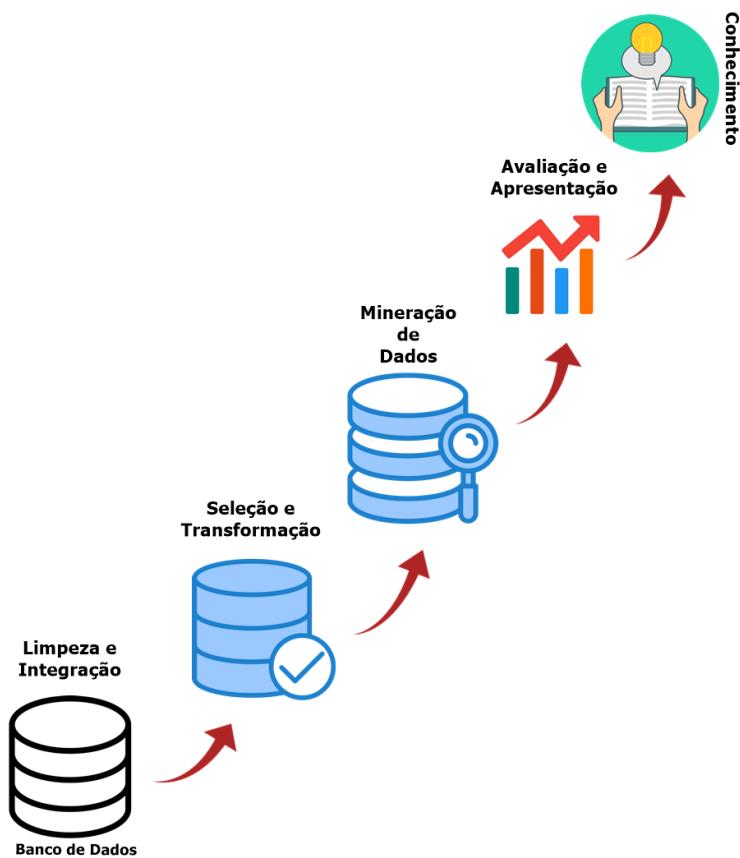


Figura 19 – Passos metodológicos

Os dados utilizados no processo de descoberta do conhecimento foram originados do STI LEIA, a partir disso as etapas da metodologia escolhida foram aplicadas, iniciando pela limpeza de dados, onde as variáveis irrelevantes foram selecionadas e excluídas com base nas informações a serem extraídas.

Posteriormente, os dados relevantes foram selecionados e transformados para serem apropriadamente aplicados no processo de mineração de dados. Após a realização da etapa de mineração de dados, as informações extraídas foram analisadas e organizadas para identificar os padrões, buscando evidências relacionadas ao processo de aprendizagem dos estudantes com a finalidade de aprimorar os STIs. Em seguida, foi desenvolvido um *dashboard* para melhor visualização dos dados e apresentar o conhecimento adquirido.

Foram realizadas duas etapas de análise de dados em ([MENEZES; NETTO; LOPES, 2022](#)) e ([MENEZES et al., 2023b](#)), onde foi explicado parte do processo de MDE desenvolvido no projeto. Em ambas etapas, foram utilizados os dados de experimentos

iniciais do STI LEIA. Os dados aplicados no processo de descoberta do conhecimento, foram obtidos do banco de dados do STI LEIA. Estes dados foram obtidos por meio de um experimento, onde 20 alunos de uma escola pública utilizaram o sistema, tentando resolver um conjunto de questões propostas pelo STI. A partir disto as etapas do processo foram aplicadas, iniciando pela limpeza de dados onde as variáveis irrelevantes foram selecionadas e excluídas com base nas informações a serem extraídas, por exemplo, neste momento o STI contabilizava a utilização do recurso ajuda de forma incorreta, então a variável que representa esta função foi excluída da mineração, os demais dados relevantes foram integrados ao processo.

Após a conclusão do processo de limpeza e integração, os dados relevantes foram selecionados e transformados para serem apropriadamente aplicados no processo de análise dos dados. As informações extraídas foram analisadas e organizadas para identificar os padrões, buscando evidências relacionadas ao processo de aprendizagem dos estudantes.

4.5.1 Primeira Análise de Dados

Tomando-se como base o conjunto de dados com os atributos descritos, alguns experimentos foram realizados, com a finalidade de avaliar o desempenho dos alunos, mensurar o nível de dificuldade das questões e comparar com a classificação adotada pelo sistema, além de compreender o caminho que o aluno percorre até acertar ou errar uma atividade. Abaixo, pode-se entender as variáveis utilizadas no processo.

- **Questões:** Representa as questões algébricas disponibilizadas no STI LEIA.
- **Estudantes:** Representa os estudantes que participaram do experimento, utilizando o STI LEIA.
- **Acurácia:** Representa os erros e acertos em cada atividade.
- **Nível:** Nível de dificuldade das questões propostas.

4.5.1.1 Taxa de passos corretos para cada questão

No experimento 1 de (MENEZES; NETTO; LOPES, 2022), a busca foi centrada em identificar a taxa de passos corretos para acertar cada questão, servindo como exemplo, para resolver a questão 1, o aluno 1 precisou de dois passos certos para chegar à conclusão, ou seja, 100% de passos corretos na questão 1. Na Tabela 2 pode-se identificar a taxa de acerto para cada questão.

Tabela 2 – Taxa de passos corretos para cada questão

Questões	Taxa de Acerto
q1	81,15%
q2	77,14%
q3	92,95%
q4	31,03%
q5	46,53%
q6	82,35%
q7	58,57%
q8	88,53%
q9	88,37%
q10	84,93%
q11	18,18%
q12	47,36%
q13	50,00%
q15	70,37%
q16	66,66%
q17	53,84%
q18	68,57%

As questões q14, q19 e q20 foram desconsideradas no processo de mineração, visto que poucos alunos conseguiram responder às questões, devido ao nível de dificuldade elevada.

Levando em consideração os dados da tabela, foi percebido que apenas as questões q4 e q11 receberam uma pequena taxa de acerto, abaixo de 35%, demonstrando uma clara dificuldade dos alunos no desenvolvimento dos passos matemáticos.

Outras 6 questões obtiveram uma taxa de passos certos superior a 80%, sendo elas: q1, q3, q6, q8, q9 e q10. Indicando uma maior proficiência dos alunos sobre as questões. As demais questões ficaram entre os 35% e 80%, sinalizando uma taxa média.

A Tabela 3 exibe a taxa de acerto de passes matemáticos para cada aluno participante do experimento, demonstrando que a metade dos estudantes conseguiu atingir

uma taxa de acerto superior a 70%, além disso, apenas 2 estudantes atingiram menos de 50% de acertos.

Tabela 3 – Taxa de passos certos para cada estudante

estudantes	Taxa de passos certos
estudante1	91,66%
estudante2	89,47%
estudante3	88,23%
estudante4	94,73%
estudante5	39,74%
estudante6	87,14%
estudante7	89,58%
estudante8	34,61%
estudante9	44,30%
estudante10	58,73%
estudante11	62,00%
estudante12	76,08%
estudante13	58,62%
estudante14	85,71%
estudante15	42,85%
estudante16	72,50%
estudante17	64,38%
estudante18	58,06%
estudante19	59,64%
estudante20	78,43%

4.5.1.2 Definindo o nível das questões

No experimento 2, executado em ([MENEZES; NETTO; LOPES, 2022](#)), visando a definição do nível das questões, foi utilizado o índice de facilidade, desenvolvido pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), para ser aplicado durante a realização do Exame Nacional de Desempenho dos Estudantes. Este índice classifica as questões em cinco níveis de dificuldade através da taxa de acerto, sendo: muito fácil, fácil, média, difícil e muito difícil. A tabela 4 exibe a taxa correspondente a cada nível.

Um dos objetivos desta etapa é comparar o nível de dificuldade atribuído a cada questão com o seu nível real, baseado no índice de facilidade, visto que os alunos sentiram dificuldades em algumas questões classificadas como fáceis. O STI LEIA

Tabela 4 – Modelo de classificação do INEP

Taxa de acerto	Classificação para o ENADE
>0,86	Muito fácil
0,61 - 0,85	Fácil
0,41 - 0,60	Normal
0,16 - 0,40	Difícil
<0,15	Muito Difícil

distribui as questões entre três categorias: fácil, normal e difícil, baseando-se no tamanho das equações. Quanto mais elementos a equação apresentava, maior o seu nível de dificuldade. Para realizar a comparação, foi obtida a taxa de acerto de cada questão, depois categorizada dentro dos padrões do índice de facilidade e realizada a comparação entre os métodos.

Após a classificação dos dados, foram obtidas as taxas de acerto para cada questão, com isso foi possível realizar uma comparação entre os classificadores de dificuldade das questões. No Quadro 4.1 pode-se observar o índice de acertos e a categorização das questões.

Questão	Classificação - INEP	Classificação - STI
q1	Muito Fácil	Fácil
q2	Fácil	Fácil
q3	Fácil	Fácil
q4	Normal	Fácil
q5	Difícil	Fácil
q6	Difícil	Fácil
q7	Fácil	Normal
q8	Fácil	Normal
q9	Difícil	Normal
q10	Normal	Normal
q11	Muito Difícil	Normal
q12	Difícil	Normal
q13	Difícil	Difícil
q15	Normal	Fácil
q16	Normal	Difícil
q17	Fácil	Difícil
q18	Difícil	Difícil

Quadro 4.1: Quadro comparativo entre as classificações

Considerando a taxa de acertos dos passos dados para resolver uma questão, foi possível observar algumas diferenças substanciais entre os tipos de acertos, exemplificando através da q6 onde a taxa de acertos dos passos obtidos foi de 0,82 e a taxa

de acertos de questões foi de 0,36, ou seja, mesmo com um número elevado na taxa de acerto de etapas, poucos alunos conseguiram finalizar a questão, essa diferença também se justifica pela desistência no momento da resolução, quando o aluno pula a questão o STI conta como abandono, classificando como erro.

O outro exemplo é a questão q17, que em relação ao índice de acerto das etapas atingiu um índice de 0,53, mas em relação ao índice de acerto da questão, teve mais sucesso, chegando a 0,71, isso se justifica pelo fato de alguns alunos conseguirem resolver o problema mesmo pulando ou errando algumas etapas, provavelmente devido à ajuda do STI, no momento o sistema não armazena os dados relacionados ao uso do recurso de ajuda, sendo impossível detectar se a ajuda foi útil ou não, essa funcionalidade será implementada no futuro.

Quanto à comparação entre os modelos de classificação, algumas considerações devem ser feitas, como a adequação das categorias para serem iguais em número, pois o método do índice de facilidades possui 5 conjuntos e a classificação do STI contém apenas 3, portanto, a categoria muito fácil e fácil, correspondem a fácil do STI. Consequentemente, os tipos muito difícil e difícil correspondem ao difícil do STI, a categoria normal permanece a mesma.

Após o ajuste das nomenclaturas, nota-se que a maioria das questões não possui o mesmo nível de dificuldade. 11 questões não se enquadram na classificação estabelecida pelo STI, sendo elas: q4, q5, q6, q7, q8, q9, q11, q12, q15, q16 e q17. Dentre essas, destacam-se as questões q5 e q6, categorizadas como fácil pelo STI, porém é classificada como difícil pelo índice de facilidade, expondo uma discrepância na rotulagem, assim como na q11, considerada uma questão Normal pelo STI, no entanto, possui apenas 0,05 no coeficiente de acerto, avaliado como muito difícil.

4.5.2 Segunda Análise de Dados: Geração de Árvores de Classificação

Nesta segunda análise, o propósito foi categorizar o desempenho dos estudantes utilizando a técnica da árvore de decisão, visando assim emitir sinais de alerta individuais

para cada cenário identificado. Os sinais foram divididos em 5 categorias: desempenho muito baixo, baixo, médio, alto e muito alto.

- **Questões:** Representa as questões algébricas disponibilizadas no STI LEIA.
- **Estudantes:** Representa os estudantes que participaram do experimento, utilizando o STI LEIA.
- **Acurácia:** Representa os erros e acertos em cada atividade.
- **Tempo:** Conta o tempo de resolução das questões propostas pelo STI.
- **Tempo de estudo:** Conta o tempo que o aluno permanece online no sistema.
- **Taxa de Acerto:** Representa a porcentagem de acerto total do estudante, é calculada baseado no atributo acurácia, transformando os acertos e erros das questões em uma porcentagem de taxa de acerto total.
- **Desempenho:** Sinaliza se o aluno está no caminho da aprovação ou reaprovação baseado na sua porcentagem de acerto.

Após definir e coletar os dados, utilizou-se o algoritmo J48 visando desenvolver um modelo de classificação em árvore. Esse modelo foi construído através da análise dos dados de treinamento, gerando assim um modelo capaz de classificar dados ainda não categorizados. Para avaliar o desempenho do algoritmo J48, foram utilizados os percentuais de Instâncias Corretamente Classificadas (ICC) e de Instâncias Erroneamente Classificadas (IEC).

A aplicação do algoritmo de classificação J48, com base em seu desempenho, resultou em um ICC de 85,7143% e um IEC de 14,2857%. A árvore resultante categorizou os estudantes reprovados como apresentando desempenho baixo. Em seguida, realizou uma segmentação entre os estudantes aprovados, levando em consideração o tempo dedicado ao sistema. Os usuários aprovados com tempo online abaixo e dentro da média foram classificados como desempenho médio. Por outro lado, os estudantes com tempo de dedicação acima da média e uma taxa de acerto menor ou igual a 80% foram identificados como desempenho alto, enquanto aqueles com uma taxa de acerto

superior a 80% foram classificados como desempenho muito alto. A Figura 20 exibe a árvore gerada.

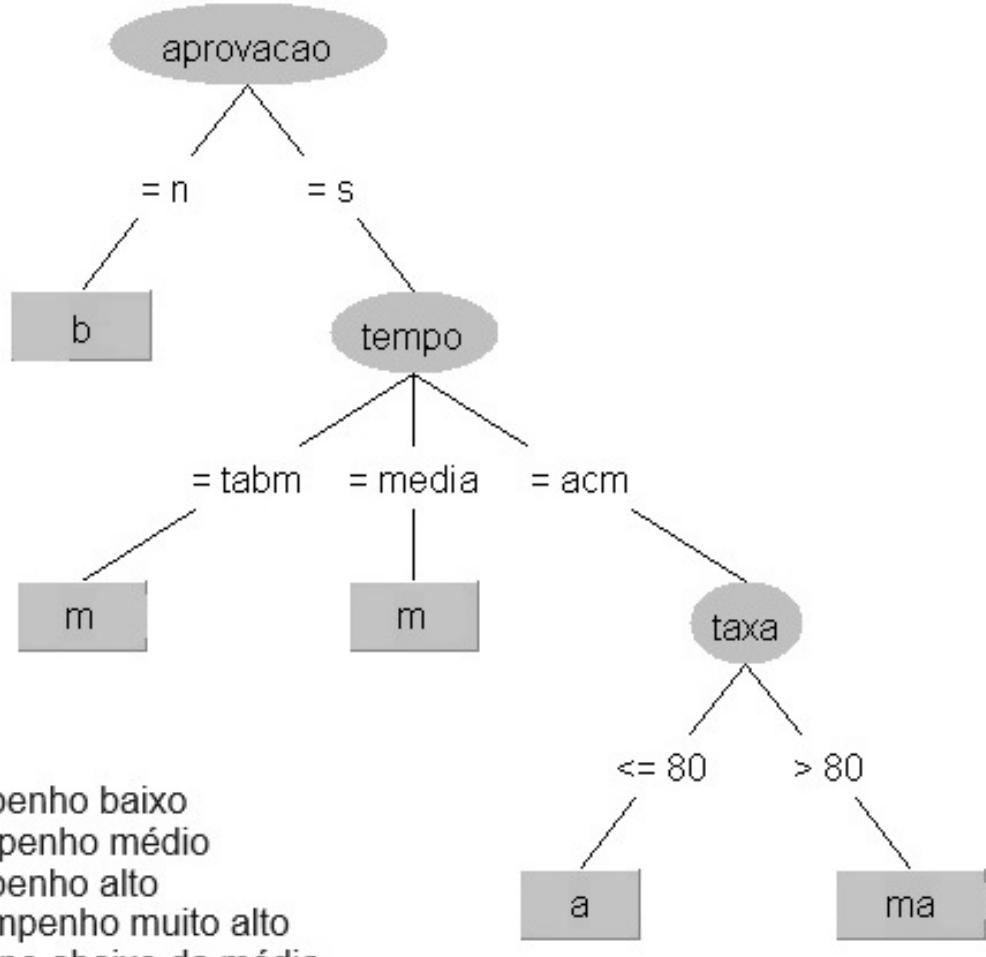


Figura 20 – Árvore gerada pelo algoritmo J48

Em uma segunda avaliação na mesma base de dados, empregou-se o algoritmo de árvore de decisão para analisar a eficácia da técnica. Assim como na primeira análise, utilizaram-se os percentuais de ICC e de IEC como métricas de desempenho. Os resultados indicaram um ICC de 76,1905% e um IEC de 23,8095%.

A árvore gerada durante o experimento classificou os estudantes com base na taxa de acerto nas questões, resultando em uma categorização mais simples. Ao contrário da primeira análise, esta identificou os estudantes com desempenho muito baixo, considerando aqueles que obtiveram pontuações inferiores a 38,09. A Figura 21

exibe a árvore gerada.

A análise de dados de ambas as árvores foi realizada a partir de uma base de dados de 20 estudantes que utilizaram o sistema LEIA em sala de aula. Consequentemente, é importante estar ciente das limitações associadas ao tamanho reduzido do conjunto de dados. Modelos de árvore de decisão podem ser sensíveis a conjuntos de dados pequenos, e a capacidade do modelo em generalizar para novos dados pode ser comprometida. Portanto, novas análises devem ser realizadas em virtude do aumento ou modificação dos dados.

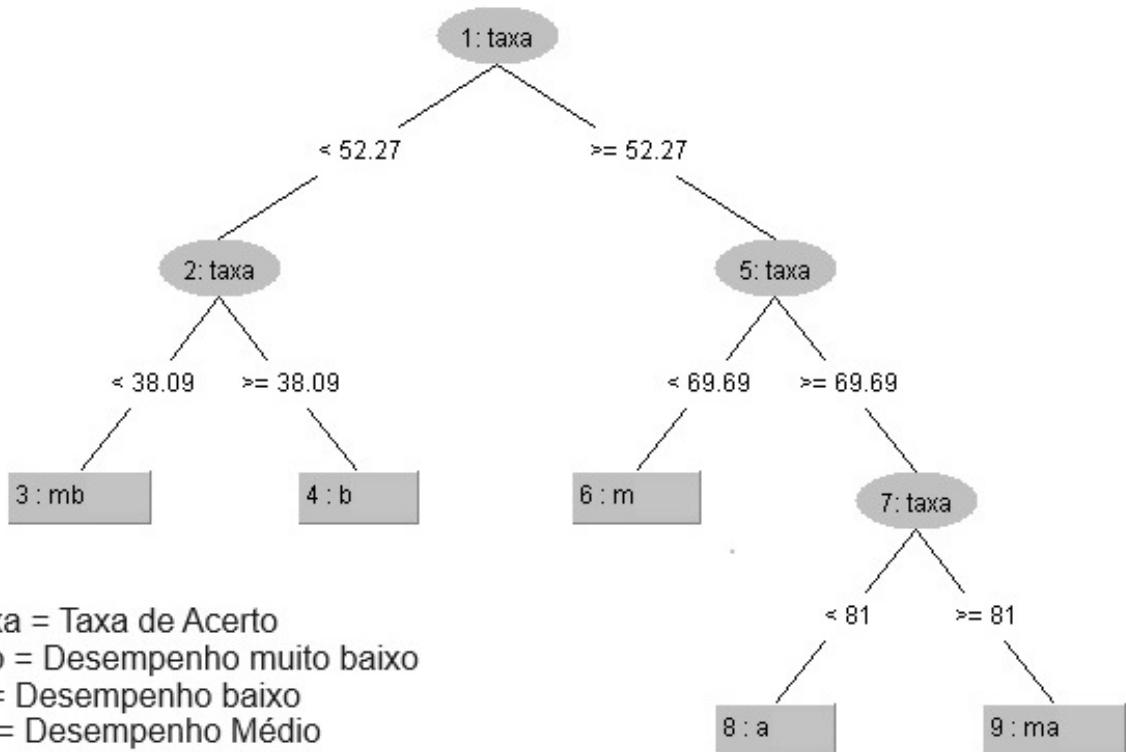


Figura 21 – Árvore gerada pelo algoritmo decision tree

Após conduzir o experimento que resultou nas duas árvores de decisão, uma criada pelo algoritmo J48 e outra pelo decision tree, optou-se por integrá-las em uma única função na linguagem PHP. Essa função foi projetada para receber os dados dos estudantes como entrada e, com base nas análises previamente realizadas, emitir alertas correspondentes a cada cenário identificado. Essa abordagem consolidada proporciona uma ferramenta mais abrangente e eficaz para avaliar e categorizar o desempenho dos estudantes, simplificando a interpretação dos resultados e facilitando a tomada

de decisões com base nas características específicas de cada caso. A Figura 22 exibe o código utilizado descrito em português de forma estruturada.

```

função calcularDesempenho(aprovacao, taxa_acerto, tempo_online) {
    // Verifica se a aprovação é 'n' (não)
    se (aprovacao == 'n') {
        // Se a aprovação for 'n' e a taxa de acerto for menor que
        38.09%, retorna "DESEMPENHO MUITO BAIXO"
        se (taxa_acerto < 38.09) {
            retorno "DESEMPENHO MUITO BAIXO";
        }
        // Se a aprovação for 'n' e a taxa de acerto for maior ou igual
        a 38.09%, retorna "DESEMPENHO BAIXO"
        senão {
            retorno "DESEMPENHO BAIXO";
        }
    }
    // Verifica se a aprovação é 's' (sim)
    senão se (aprovacao == 's') {
        // Se a aprovação for 's' e a taxa de acerto for menor que
        69.69%, e o tempo online for menor que 75% ou 125% da média de tempo
        total, retorna "DESEMPENHO MÉDIO"
        se (taxa_acerto < 69.69 e (tempo_online < mediatotaltempo *
        0.75 ou tempo_online > mediatotaltempo * 1.25)) {
            retorno "DESEMPENHO MÉDIO";
        }
        // Se a aprovação for 's' e o tempo online for maior que 125%
        da média de tempo total
        senão se (tempo_online > mediatotaltempo * 1.25) {
            // Se a taxa de acerto for menor ou igual a 80%, retorna
            "DESEMPENHO ALTO", senão retorna "DESEMPENHO MUITO ALTO"
            se (taxa_acerto <= 80) {
                retorno "DESEMPENHO ALTO";
            }
            senão {
                retorno "DESEMPENHO MUITO ALTO";
            }
        }
    }
    // Se nenhuma das condições anteriores for satisfeita, retorna
    "Classificação Inválida"
    retorno "Classificação Inválida";
}

// Exemplo de chamada da função
aprovacao = 's';
taxa_acerto = 75;
tempo_online = 2500;

resultado = calcularDesempenho(aprovacao, taxa_acerto, tempo_online);
escreva(resultado);

```

Figura 22 – Algoritmo gerado pela união das duas árvores

A função calcularDesempenho serve de base para o funcionamento do agente Classificador, implementado na janela de análise de desempenho do *dashboard*. A classificação individual foi adicionada em uma tabela de exibição, junto de um ícone específico para cada cenário. Na Figura 23 pode-se observar a atuação do agente.

444123	80%	29 Minutos	Alto 
444456	33.33%	38 Minutos	Muito Baixo 
444789	50%	25 Minutos	Baixo 
555123	42.85%	25 Minutos	Baixo 
555456	55.55%	44 Minutos	Medio 
666123	33.33%	29 Minutos	Muito Baixo 
666789	72.72%	41 Minutos	Alto 
777123	57.15%	41 Minutos	Medio 

Figura 23 – Exibição do classificação dos estudantes no *dashboard*

4.6 Prototipação Inicial

Na fase inicial do sistema, o protótipo foi concebido com base na análise da literatura especializada e nas variáveis do STI LEIA. O design foi elaborado por meio da ferramenta Omnipage, buscando otimizar sua eficácia e funcionalidade.

Segundo (SAEED et al., 2019), a prototipação é uma abordagem utilizada na indústria de desenvolvimento de *software* que visa criar versões iniciais ou modelos simplificados do *software* antes da sua implementação completa. Esta abordagem permite que os interessados tenham uma visão do *software* e possam realizar testes antes de investir tempo e recursos na sua produção final. A prototipação pode incluir modelos de papel, wireframes interativos ou protótipos funcionais utilizados para identificar possíveis problemas e soluções, resultando em uma melhora na qualidade e eficiência do produto final. Além disso, a prototipação é uma forma eficaz de comunicação de ideias e obtenção de *feedback* dos usuários, garantindo que o *software* seja desenvolvido de acordo com suas necessidades e expectativas.

O protótipo foi pensado inicialmente com cinco *dashboards* e uma página inicial

referente a tela de login. Dois *dashboards* são referentes ao desempenho, considerando a perspectiva do estudante e as estatísticas das questões propostas pelo STI, além de um *dashboard* inicial, com informações relativas às atividades do sistema. As últimas telas abordam as listas de estudantes e questões.

4.6.1 Dashboards Projetados

Iniciando o desenho do projeto, a tela inicial do *dashboard* é a primeira interface com a qual os usuários interagem quando acessam o sistema. Desenvolvida para ser clara, intuitiva e de fácil navegação, para que os usuários possam rapidamente encontrar as informações e recursos de que precisam. A tela inclui informações gerais sobre como gráficos e dados destacados para visualização das informações, como estudantes, estudantes ativos e desempenho nas atividades.

O *layout* possui um menu com botões bem definidos para facilitar a navegação. A tela é responsiva e ajuste-se facilmente a diferentes tamanhos e dispositivos, para que os usuários possam acessá-la de qualquer lugar. Na Figura 24 pode-se observar o protótipo da tela inicial.

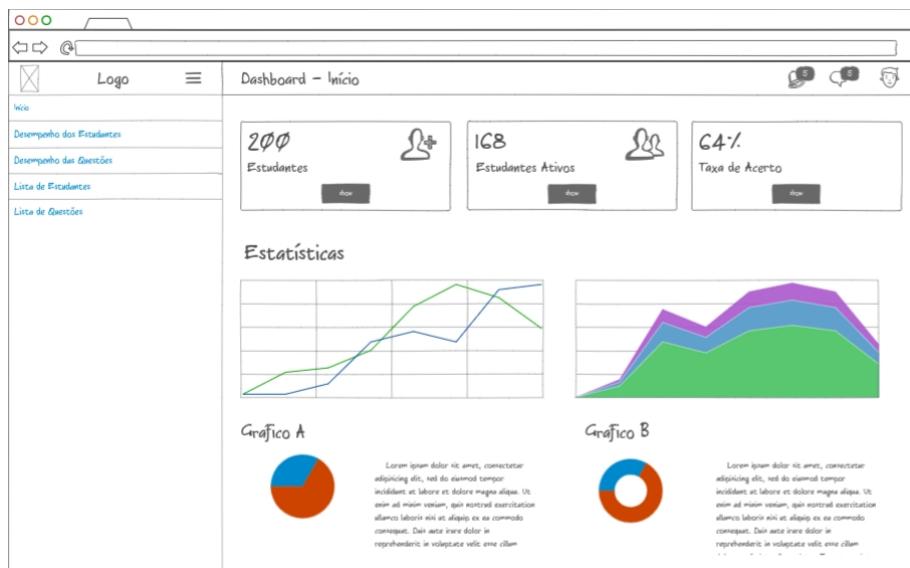


Figura 24 – Tela inicial do *dashboard*

A segunda tela trata dos dados de desempenho de estudantes, sendo uma interface que apresenta informações importantes sobre o progresso acadêmico do estudante,

durante o uso do STI. Sendo uma ferramenta valiosa para monitorar a atuação dos estudantes e tomar decisões educacionais. O *dashboard* inclui informações como taxas de acerto, participação, acompanhamento dos passos matemáticos para acertar determinada questão e o desempenho em avaliação. Os dados devem ser apresentados em forma de gráficos, tabelas ou outros formatos visualmente atraentes. Na Figura 25 pode-se observar o protótipo da tela de desempenho dos estudantes. O *dashboard* também inclui recursos adicionais, como notificações para ajudar os usuários a monitorar o progresso dos estudantes de perto e tomar medidas adequadas.

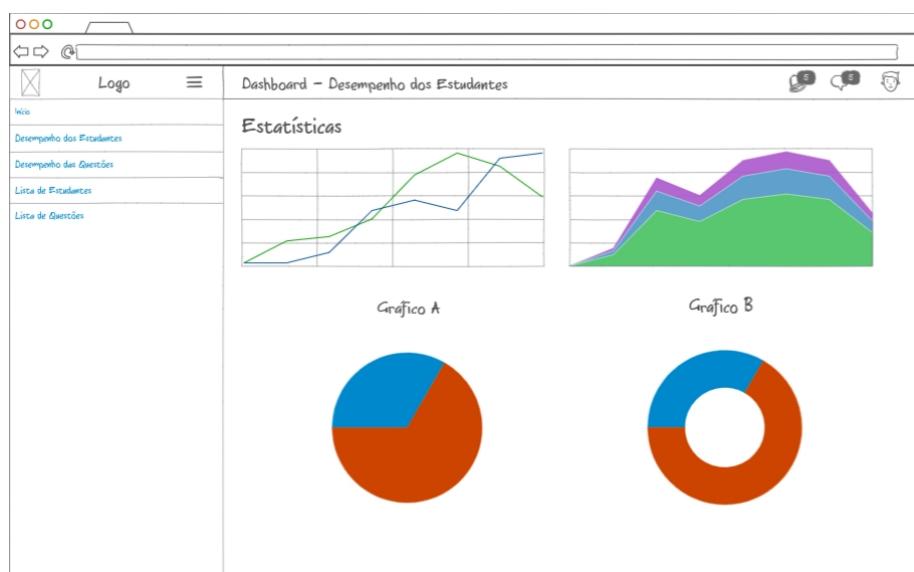


Figura 25 – *dashboard* de desempenho dos estudantes

A terceira tela apresenta informações sobre as questões matemáticas propostas pelo STI, permitindo monitorar o como os estudantes estão desenvolvendo as questões e identificar pontos fortes e fracos na compreensão das atividades. Incluindo informações como o número de questões respondidas corretamente, o tempo médio de resposta e o desempenho por nível de dificuldade. Os dados podem ser apresentados em forma de gráficos, tabelas ou outros formatos visualmente atraentes. Na Figura 26 pode-se observar o protótipo da tela com as estatísticas das questões.

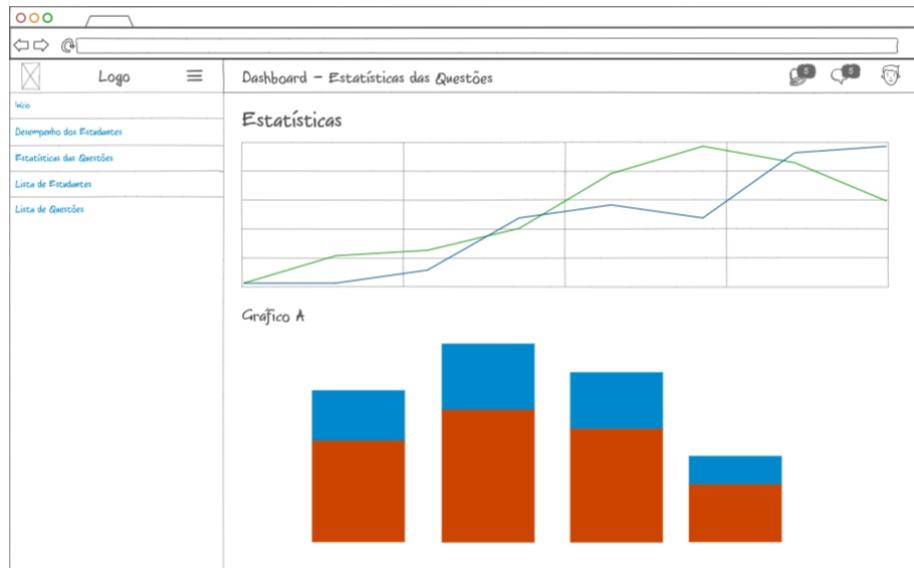


Figura 26 – *dashboard* com as estatísticas das questões

A quarta e quinta telas exibem listas com dados dos estudantes e questões, sendo possível identificar individualmente cada um. A lista de estudantes e questões detalham informações como: taxa de acerto da questão, tempo de resolução das questões. Na Figuras 27 pode-se observar o protótipo da tela com a lista de estudantes.

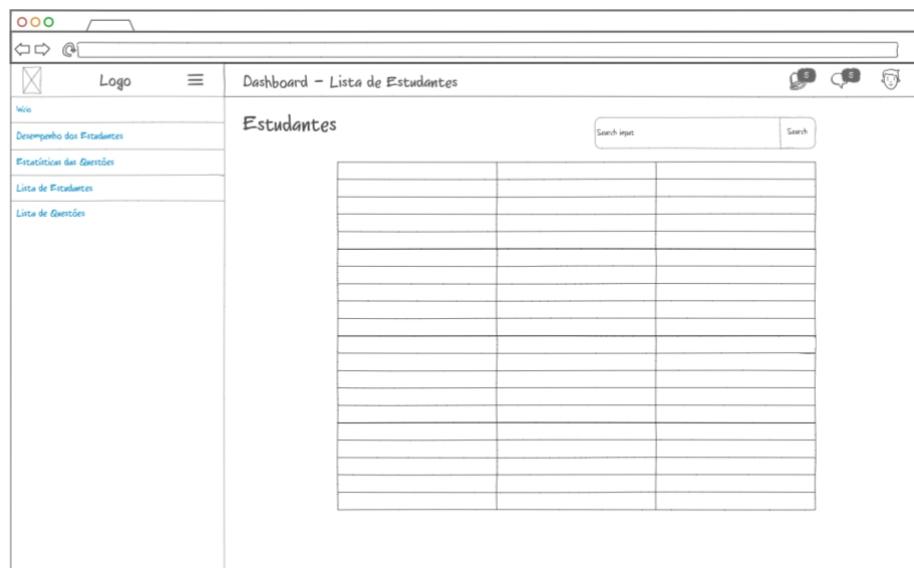


Figura 27 – Lista com estatísticas dos estudantes

4.7 *Edudata Visualizer*

Esta seção apresenta o *dashboard Edudata Visualizer*, ferramenta de visualização de dados educacionais desenvolvida de forma responsiva para plataforma *web*, sendo possível utilizar em computador pessoal, tablet e *smartphones*. O *Edudata Visualizer* consiste em fornecer recursos para exibição e análise de dados de STIs, utilizando MDE e visualização.

A aplicação foi projetada para ser integrada em qualquer STI que compartilhe a estrutura de variáveis comuns encontradas na maioria dos STIs. Essa estrutura abrange uma ampla gama de dados, incluindo informações sobre o desempenho individual dos alunos e as atividades propostas pelo sistema. Dessa forma, a aplicação pode ser adaptada aos diferentes contextos educacionais, proporcionando uma solução flexível e abrangente para a análise e o acompanhamento do progresso dos estudantes. A Figura 28 exibe a tela de *login*, esta etapa é responsável por dar acesso à primeira janela do *dashboard* mediante ao cadastro ou uma conta pré-cadastrada no STI de origem.

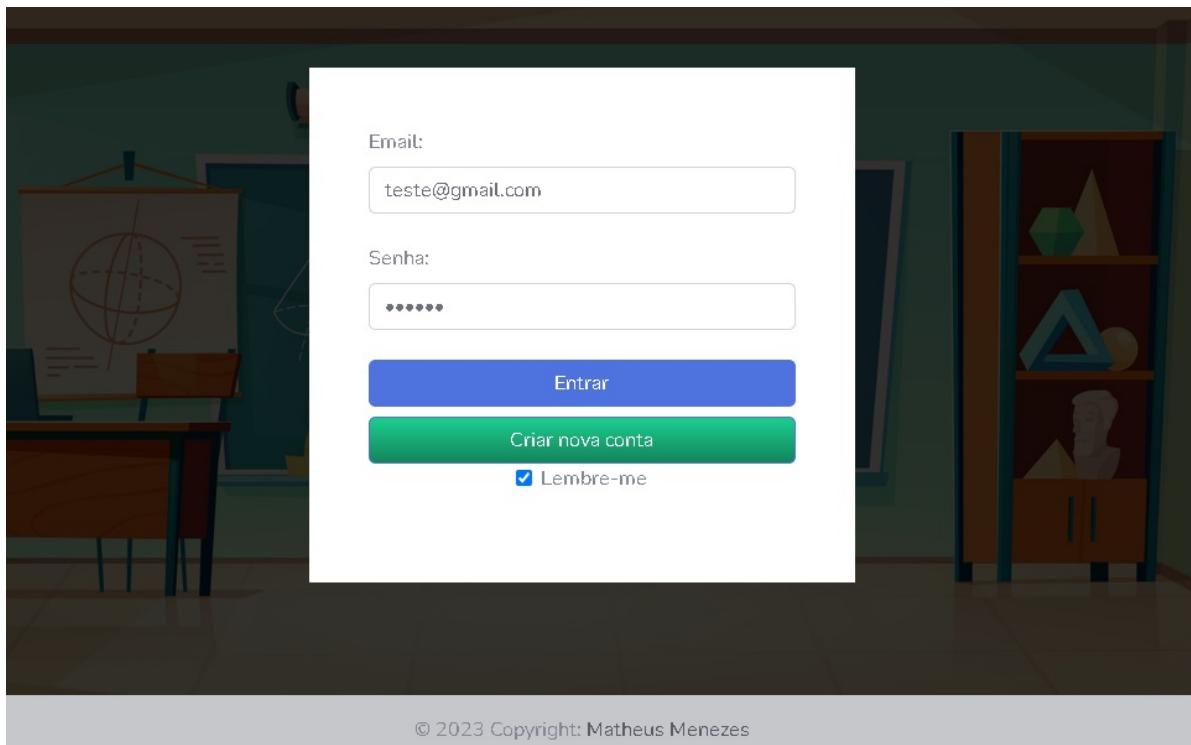


Figura 28 – Tela de login do *dashboard Edudata Visualizer*

Após o login ou acesso direto através do STI, o usuário é direcionado para a primeira interface que apresenta dados gerais sobre os estudantes. Esses dados incluem

o número total de alunos cadastrados, o número de alunos ativos (aqueles que participaram de atividades no STI), a média de acertos nas atividades, o tempo total de utilização do sistema pelos estudantes, o índice de participação dos alunos (calculado com base no tempo de utilização e atividades realizadas) e a classificação do desempenho dos usuários, que pode ser categorizado como satisfatório ou insatisfatório. Esta classificação é determinada com base em um valor predefinido pelo painel, sendo neste caso estabelecido em 50% de acerto. Além disso, o painel também apresenta um alerta que surge no canto superior direito, informando sobre a condição atual da turma.

A interface também exibe uma barra lateral, que pode ser diminuída ou expandida, esta barra contém os *links* para outras páginas. A Figura 29 apresenta o estágio atual de desenvolvimento da tela.

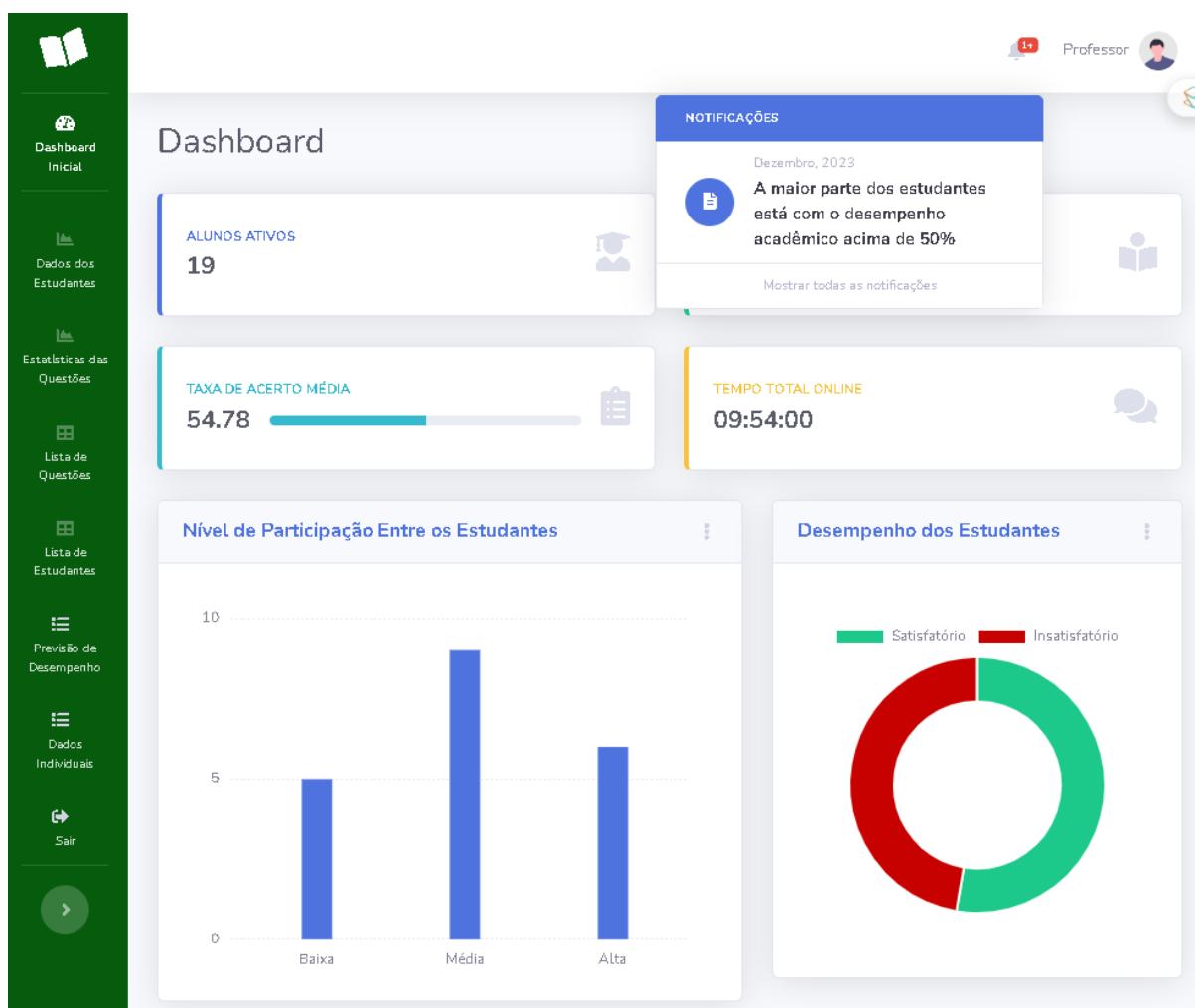


Figura 29 – Primeira dashboard do *Edudata Visualizer*

Na segunda janela de visualização, são apresentadas informações mais especí-

ficas sobre o desempenho geral dos usuários. Isso inclui a evolução da taxa de acerto ao longo do tempo, o gráfico de setores que mostra a satisfação com o desempenho e barras de progresso indicando a porcentagem de acerto nas atividades propostas pelo STI. A Figura 30 exibe o *dashboard* com os dados de rendimento dos estudantes.



Figura 30 – Tela com os dados de desempenho dos estudantes

A terceira opção da barra lateral exibe as estatísticas das questões propostas pelo STI, nesta interface é apresentada uma tabela com dados específicos como: conteúdo

da questão, nível de dificuldade, número de acertos e erros, taxa de acerto e na última coluna, o agente nivelador realiza o cálculo para definir o nível de dificuldade da questão com relação à métrica do INEP, classificando como: muito fácil, fácil, médio, difícil ou muito difícil. A Figura 31 mostra uma parte da tabela de questões do sistema.

The screenshot shows a dark-themed dashboard interface. On the left is a vertical sidebar with icons and labels: 'Painel Inicial' (Dashboard), 'Dados dos Estudantes' (Student Data), 'Estatísticas das Questões' (Question Statistics), 'Lista de Questões' (List of Questions), 'Lista de Estudantes' (List of Students), 'Previsão de Desempenho' (Performance Prediction), 'Dados Individuais' (Individual Data), and 'Sair' (Logout). At the top right, there is a user icon labeled 'Professor'. The main area has a light gray background and displays a table titled 'Tabela de Questões' (Table of Questions). The table has columns: Questão (Question), Dificuldade (Difficulty), Acertos (Correct Answers), Erros (Errors), Taxa de Acerto (Accuracy Rate), and Dificuldade - INEP (INEP Difficulty). Below the table, a note states: 'A tabela de estudantes exibe os dados individuais relativos a cada questão cadastrada no sistema, como: questão, taxa de acerto, tempo dedicado e classificação da dificuldade.' (The student table displays individual data for each question registered in the system, such as: question, accuracy rate, time dedicated and difficulty classification.)

Questão	Dificuldade	Acertos	Erros	Taxa de Acerto	Dificuldade - INEP
$2x=8$	fácil	20	0	100%	Muito Fácil
$3x=99$	fácil	14	9	61%	Fácil
$6x=92$	fácil	9	18	33%	Difícil
$9x=18$	fácil	13	6	68%	Fácil
$3x=9$	fácil	13	6	68%	Fácil
$7x=49$	fácil	7	12	37%	Média
$7x=55$	fácil	0	2	0%	Muito Difícil
$5x=71$	fácil	0	1	0%	Muito Difícil
$5x+6=46$	média	10	8	56%	Fácil

Figura 31 – Tela de login do *dashboard Edudata Visualizer*

A quarta e quinta janelas exibem dados de desempenho dos alunos por meio de tabelas detalhadas. Uma delas apresenta informações mais específicas, como o número de acertos e erros.

A interface de previsão de desempenho dos estudantes oferece um diferencial através da atuação de dois agentes inteligentes: o agente classificador e o agente de recomendação. Na coluna de desempenho, são exibidos alertas atribuídos com base em uma árvore de decisão gerada a partir da análise dos dados do STI, neste caso, o STI LEIA. Os dados dos estudantes são direcionados para uma função que classifica o desempenho do estudante com base nessa árvore de decisão.

A quinta-coluna apresenta a opção "Ver Relatório", na qual o professor pode acessar um relatório específico para cada situação do estudante. Esse relatório descreve

a situação atual e sugere melhorias para que o professor possa atuar na melhora do desempenho individual de cada estudante. A Figura 32 exibe a tela de previsão de desempenho dos estudantes.

Estudante	Taxa de Acerto	Tempo Dedicado	Desempenho	Relatório Inteligente
111123	50%	34 Minutos	Baixo ⚠️	<button>Ver Relatório</button>
111456	61.53%	29 Minutos	Médio ⚠️	<button>Ver Relatório</button>
111789	58.33%	43 Minutos	Médio ⚠️	<button>Ver Relatório</button>
222123	44.4%	31 Minutos	Baixo ⚠️	<button>Ver Relatório</button>
222456	50%	28 Minutos	Baixo ⚠️	<button>Ver Relatório</button>
222789	50%	23 Minutos	Baixo ⚠️	<button>Ver Relatório</button>
333123	54.54%	30 Minutos	Médio ⚠️	<button>Ver Relatório</button>
333456	62.5%	22 Minutos	Médio ⚠️	<button>Ver Relatório</button>

Figura 32 – Tela de previsão de desempenho dos estudantes

Para a geração dos textos de recomendação, foram utilizadas como base as recomendações da Secretaria da Educação (SEDU) (SEDU, 2024) e o chatGPT, um modelo de linguagem desenvolvido pela OpenAI, projetado para compreender e gerar texto em linguagem natural em uma ampla variedade de contextos. Ele pode ser usado para uma variedade de propósitos, incluindo responder a perguntas, gerar texto criativo, traduzir entre idiomas e muito mais. Quanto ao uso do ChatGPT para fornecer dicas para melhorar o desempenho de aprendizado individual, pode ser utilizado para fornecer dicas baseadas em cada contexto (OPARA; THERESA; ADUKE, 2023). No caso do *Edudata Visualizer*, foi pedido ao chatbot para dar um panorama sobre a situação individual de cada estudante, foram utilizadas as saídas de texto do chat na janela de

relatório.

É importante ressaltar que, embora o ChatGPT possa ser uma ferramenta útil para complementar o aprendizado individual, ele não substitui a orientação de um professor qualificado ou de um tutor humano. A Figura 33 exibe um dos relatórios individuais gerados pelo *dashboard*.



Figura 33 – Janela do relatório inteligente

No penúltimo link da barra lateral, encontra-se a área de busca individual Figura 34. Esta janela possibilita a pesquisa de cada estudante, revelando detalhes sobre seu desempenho, bem como as questões resolvidas por ele. Isso proporciona um panorama mais detalhado sobre o processo de aprendizagem de cada aluno.

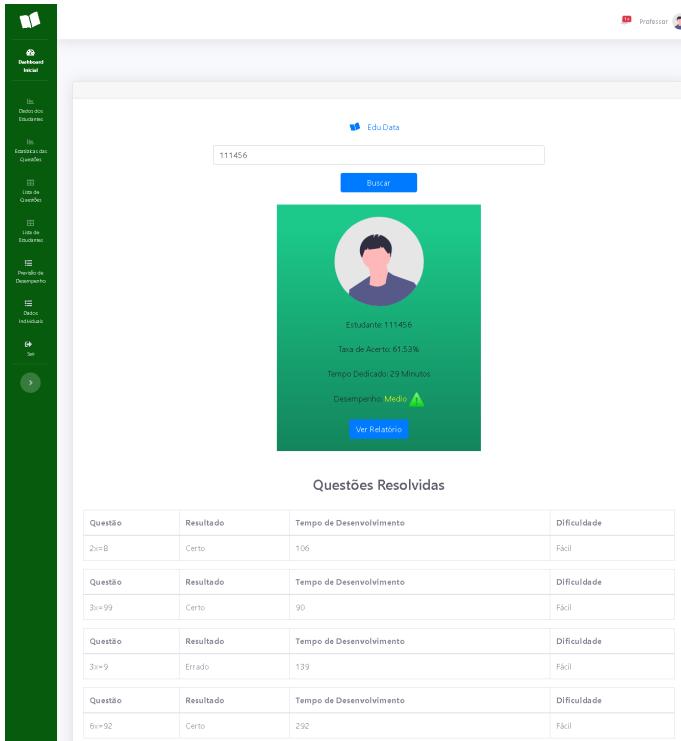


Figura 34 – Tela de busca individual

4.8 Conclusão do Capítulo

Neste capítulo dedicado à proposta do *dashboard* educacional, foi abordado o processo de desenvolvimento da solução, oferecendo uma visão geral do sistema, incluindo seu protótipo e as ferramentas utilizadas durante o desenvolvimento. Destacou-se a importância dessa proposta como uma ferramenta útil para auxiliar educadores no contexto dos STIs, proporcionando uma análise mais abrangente e eficiente dos dados educacionais.

O sistema proposto visa preencher uma lacuna identificada na literatura, fornecendo aos professores uma plataforma intuitiva e acessível para monitorar o progresso e as necessidades dos alunos. A integração de recursos de visualização de dados e funcionalidades de mineração de dados educacionais promete oferecer recursos que podem informar práticas pedagógicas mais eficazes.

No próximo capítulo, serão apresentados os resultados do experimento do teste de usabilidade, proporcionando uma avaliação empírica da eficácia e da usabilidade do *dashboard* educacional desenvolvido. Essa etapa é fundamental para validar a utilidade prática do sistema e identificar possíveis áreas de melhoria.

5

RESULTADOS

Neste capítulo, foi explorada a condução de duas etapas da avaliação. A primeira etapa, conduzida logo após o desenvolvimento da primeira versão do Dashboard, envolveu um grupo reduzido de participantes. Essa fase inicial permitiu uma análise detalhada dos aspectos de usabilidade e funcionalidade, proporcionando ideias para implementações futuras.

Posteriormente, o trabalho avançou para a segunda fase de avaliação, realizada após correções e aprimoramentos, bem como a implementação de novas funcionalidades. Nessa etapa, o sistema foi testado com um número substancialmente maior de professores de matemática, visando não apenas validar as melhorias implementadas, mas também garantir uma maior representatividade da comunidade de usuários.

O aumento no número de participantes, especialmente entre os educadores de matemática, foi planejado para enriquecer a diversidade de perspectivas e experiências, ampliando assim a abrangência e confiabilidade dos resultados obtidos durante o teste. Essa abordagem reforçada não apenas ofereceu uma visão mais abrangente da eficácia do *dashboard*.

Em ambos experimentos, foi utilizado o termo de consentimento livre e esclarecido (TCLE) em um teste de usabilidade SUS, fundamental para garantir a ética e a transparência na condução da pesquisa, disponibilizado nos Anexos. O TCLE é um documento que apresenta de forma clara e compreensível todos os detalhes relevantes sobre a participação do usuário no estudo, incluindo os objetivos da pesquisa, os procedimentos envolvidos, os potenciais riscos e benefícios, bem como os direitos do

participante.

Ao solicitar o consentimento dos usuários por meio do TCLE, é garantido a compreensão dos propósitos e as implicações da pesquisa, permitindo-lhes tomar uma decisão informada sobre sua participação. Isso promove o princípio da autonomia do participante, respeitando seu direito de consentir ou recusar participar do estudo, sem qualquer forma de coerção ou pressão.

5.1 Avaliação Utilizando o *System Usability Scale*

Os sistemas de *dashboard* educacional desempenham um papel crucial na facilitação do monitoramento e na visualização de dados relevantes para a tomada de decisões no contexto educacional. A avaliação da usabilidade desses sistemas é fundamental para garantir sua eficácia e aceitação pelos usuários. Neste contexto, o *System Usability Scale (SUS)* é uma ferramenta amplamente reconhecida e validada para avaliação da usabilidade de sistemas ([SAURO, 2011](#)).

Nesta fase do estudo, o objetivo foi investigar a aplicação do SUS na avaliação da usabilidade de um sistema de *dashboard* educacional, conduzindo um estudo empírico utilizando uma abordagem mista, envolvendo tanto métodos quantitativos quanto qualitativos. Os participantes foram recrutados entre professores de matemática e outros profissionais da área educacional. Os participantes foram solicitados a interagir com o *Edudata Visualizer* e, em seguida, preencheram o questionário SUS para avaliar sua usabilidade percebida. Além disso, foram conduzidas perguntas para coletar opiniões qualitativas sobre a experiência do usuário e possíveis áreas de melhoria do sistema. O questionário SUS é baseado em 10 questões que podem ser respondidas através da escala Likert, sendo um modelo de escala de questionários comumente empregado em pesquisas de opinião e de satisfação. Essa ferramenta possibilita a mensuração dos pontos de vista e das posturas dos usuários de maneira graduada, variando, por exemplo, de "discordo totalmente" a "concordo totalmente".

Está são as 10 questões que compõe o SUS:

1. Eu acho que gostaria de usar esse sistema frequentemente.

2. Achei o sistema desnecessariamente complexo.
3. Achei o sistema fácil de usar.
4. Acredito que precisaria de ajuda de um especialista para usar esse sistema.
5. Achei as várias funções do sistema bem integradas.
6. Achei que o sistema apresenta muita inconsistência.
7. Aprendi rapidamente a usar o sistema.
8. Achei o sistema muito complicado.
9. Me senti confiante ao usar o sistema.
10. Precisei aprender várias coisas antes de começar a usar o sistema.

Para calcular a pontuação do SUS, as respostas são tratadas de maneira específica. Para as perguntas ímpares (1, 3, 5, 7 e 9), a pontuação é calculada subtraindo 1 da classificação escolhida pelo participante. Para as perguntas pares (2, 4, 6, 8 e 10), a pontuação é calculada subtraindo a classificação escolhida pelo participante de 5.

Após calcular as pontuações para todas as perguntas, elas são somadas para cada participante. O resultado pode variar de 0 a 40, sendo 0 a pontuação mais baixa possível e 40 a pontuação mais alta possível. No entanto, para facilitar a interpretação dos resultados, essa pontuação bruta precisa ser convertida para uma escala de 0 a 100. Para converter a pontuação bruta para a escala de 0 a 100, utiliza-se a seguinte fórmula: pontuação bruta * 2,5. Portanto, a pontuação resultante indica a usabilidade percebida do sistema em uma escala de 0 a 100, onde pontuações mais altas indicam uma melhor usabilidade percebida.

Os resultados deste estudo fornecem uma compreensão aprofundada da usabilidade do sistema proposto, identificando pontos fortes e áreas de melhoria. A pontuação média do SUS foi analisada para determinar a usabilidade geral do sistema, enquanto as entrevistas qualitativas foram analisadas para identificar temas emergentes e dicas adicionais.

5.1.1 Resultado da Primeira Avaliação SUS

Na avaliação de usabilidade, foi empregado o SUS para avaliar a facilidade de uso do sistema com um grupo de 11 participantes (7 professores de matemática e 4 graduados em licenciatura em computação).

O resultado da pontuação do teste alcançou uma média de 77,95, indicando uma experiência positiva do usuário. Na escala SUS, que varia de 0 a 100, este número posiciona o sistema acima do ponto médio (50), sugerindo que, em média, os participantes consideraram o sistema utilizável e fácil de usar. Na Figura 35, pode-se observar as respostas dos participantes e o resultado do cálculo do escore, as colunas indicam os questionamentos do teste e as linhas mostram as respostas de cada um dos participantes, sendo o número 1 representando o "discordo totalmente" e o 5 como "concordo totalmente".

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
1	3	2	4	1	4	1	4	1	5	2
2	4	1	3	2	4	1	4	1	4	2
3	3	1	3	2	3	3	4	2	4	2
4	5	3	4	1	4	2	5	2	5	1
5	4	3	4	3	4	2	5	2	4	1
6	3	1	3	3	3	4	4	1	5	2
7	4	2	4	2	4	1	4	1	5	1
8	3	1	4	2	3	2	3	2	3	2
9	3	3	5	1	4	2	4	1	4	3
10	4	2	4	1	5	2	4	1	3	2
11	5	1	3	2	4	1	5	1	4	1

Results

Mean SUS score:

77.95

Accuracy

Score range [i](#)

90%

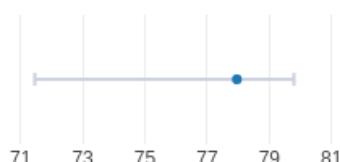


Figura 35 – Pontuação do teste SUS

Durante a condução do questionário SUS, além das perguntas padrão, também foram apresentadas aos participantes outras afirmações específicas, buscando ampliar a gama de opiniões sobre a ferramenta. Essas afirmações foram avaliadas utilizando a escala Likert, permitindo uma análise mais abrangente da percepção dos usuários em relação ao sistema.

A Figura 36 apresenta as respostas registradas pelos participantes para essas afirmativas adicionais. Esta abordagem complementar ajudou na compreensão da experiência do usuário, contribuindo significativamente para futuros aprimoramentos e desenvolvimentos do sistema.

1. Eu utilizaria o *Edudata Visualizer* na minha rotina profissional.
2. O uso do *Edudata Visualizer* tornaria o trabalho mais fácil.
3. O *Edudata Visualizer* oferece recursos úteis para visualização de dados.
4. O *Edudata Visualizer* facilitaria a interpretação e análise de dados.
5. Eu me sentiria motivado a utilizar o *Edudata Visualizer* regularmente

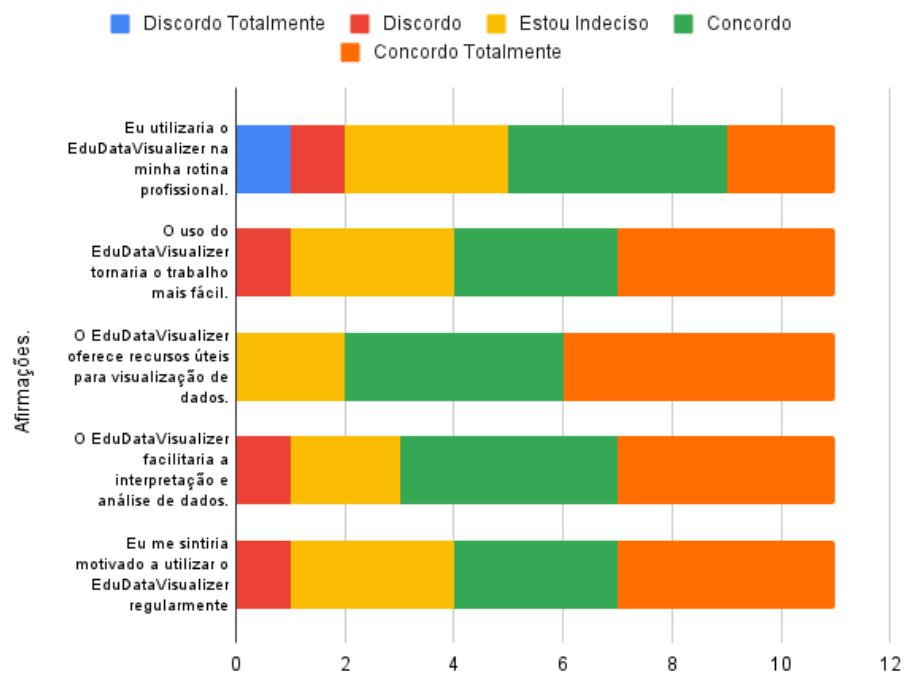


Figura 36 – Respostas às outras afirmações feitas no questionário.

A interpretação dos dados apresentados revela percepções variadas sobre o *Edudata Visualizer* e seu potencial impacto na rotina profissional. O conjunto de afirmações aborda diferentes aspectos relacionados à utilidade, facilidade de uso, recursos oferecidos e motivação para utilização regular da ferramenta.

Em relação à disposição para utilizar o *Edudata Visualizer* na rotina profissional, 54.45% dos participantes expressaram avaliações favoráveis. No entanto, 27.27% demonstraram indecisão. Essa diversidade de opiniões sugere uma ambiguidade quanto à disposição para utilização da ferramenta.

Quanto à percepção sobre o potencial do *Edudata Visualizer* em facilitar o trabalho, há uma tendência mais positiva, com a maioria concordando ou concordando totalmente que a ferramenta poderia tornar as tarefas mais fáceis. Isso sugere que, embora haja incerteza sobre sua utilização na rotina, muitos reconhecem seu potencial para simplificar o processo de trabalho.

Quando se trata dos recursos oferecidos para visualização de dados, a maioria concorda ou concorda totalmente que o *Edudata Visualizer* oferece recursos úteis. Isso sugere que, embora possa haver hesitação quanto à sua utilização, os respondentes reconhecem a utilidade das funcionalidades oferecidas pela ferramenta.

Além disso, em relação à capacidade do *Edudata Visualizer* de facilitar a interpretação e análise de dados, novamente há uma tendência positiva. Isso sugere que muitos percebem a ferramenta como uma potencial aliada na compreensão e análise de conjuntos de dados.

Em termos de motivação para utilizar o *Edudata Visualizer* regularmente, a maioria dos respondentes parece estar propensa a fazê-lo, com uma quantidade significativa concordando ou concordando totalmente que se sentiriam motivados a utilizar a ferramenta de forma consistente.

Portanto, a interpretação geral dos dados sugere que embora haja uma divisão de opiniões sobre a utilidade e aplicabilidade do *Edudata Visualizer* na rotina profissional, há um reconhecimento geral de seu potencial em facilitar o trabalho, oferecer recursos úteis para visualização de dados, facilitar a interpretação e análise de dados, e motivar sua utilização regular.

5.1.2 Resultado da Segunda Avaliação SUS

No segundo experimento SUS, dentre os 32 participantes, a maioria, totalizando 25, possui formação em licenciatura em matemática. Três apresentam também formação em engenharia, complementada por especialização no ensino da matemática. Um participante possui título de bacharel em engenharia mecânica e completou a formação pedagógica em matemática, reconhecida pelo Ministério da Educação (MEC) como equivalente à licenciatura na área. Adicionalmente, um dos participantes tem formação em estatística, sendo especialista em ensino de ciências e matemática. Três participantes possuem bacharelado em matemática, sendo um deles especialista no ensino da disciplina.

No que tange à declaração de gênero, destaca-se que entre os participantes, 21 foram identificados como do sexo masculino, enquanto 11 declararam ser do sexo feminino. Esta distribuição revela uma diversidade de gênero na amostra, contribuindo para uma representação mais abrangente e inclusiva. Esse equilíbrio entre os sexos promove uma dinâmica enriquecedora nos resultados obtidos, considerando as diferentes perspectivas e contribuições que podem surgir a partir das experiências únicas de cada participante, independente do gênero. A Figura 37 exibe o gráfico que demonstra a distribuição entre participantes do sexo masculino e feminino.

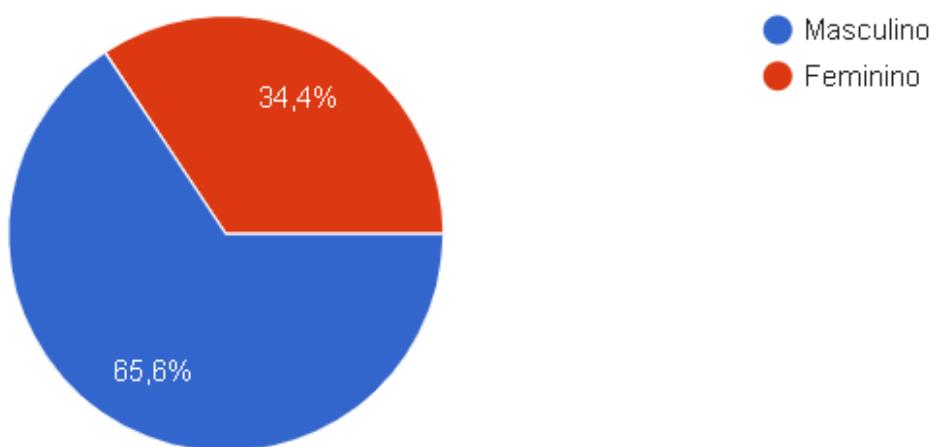


Figura 37 – Divisão de gênero entre os participantes

Com relação às respostas do experimento SUS, a primeira questão proposta, "Eu acho que gostaria de usar esse sistema com frequência", recebeu uma avaliação positiva de 56,3% dos participantes, indicando uma inclinação significativa para a integração do *dashboard* em suas rotinas profissionais. Além disso, 11 participantes manifestaram uma resposta neutra, enquanto apenas 3 apresentaram uma visão negativa em relação ao sistema. Estas informações são visualmente representadas na Figura 38, que demonstra a distribuição das respostas obtidas no teste.

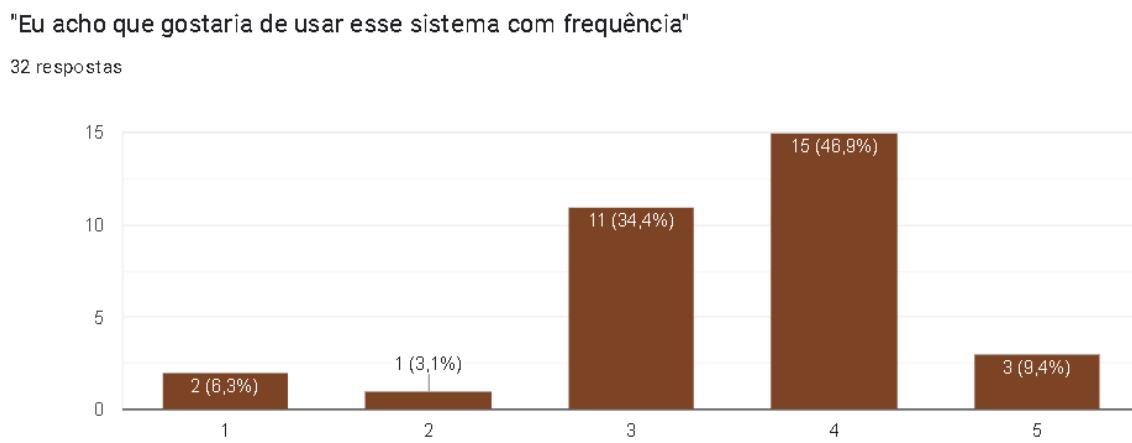


Figura 38 – Respostas sobre a primeira questão do teste SUS

No que concerne à complexidade do sistema, uma parcela mínima, somente 3 participantes, expressou encontrar alguma dificuldade, o que corresponde a menos de 10% do total de respondentes. Essa constatação sugere que a maioria dos participantes percebeu o sistema como acessível e de fácil compreensão, refletindo uma tendência positiva em relação à usabilidade do mesmo. Estas informações são visualmente representadas na Figura 39, que demonstra a distribuição das respostas obtidas no teste.

"Eu acho o sistema desnecessariamente complexo"

32 respostas

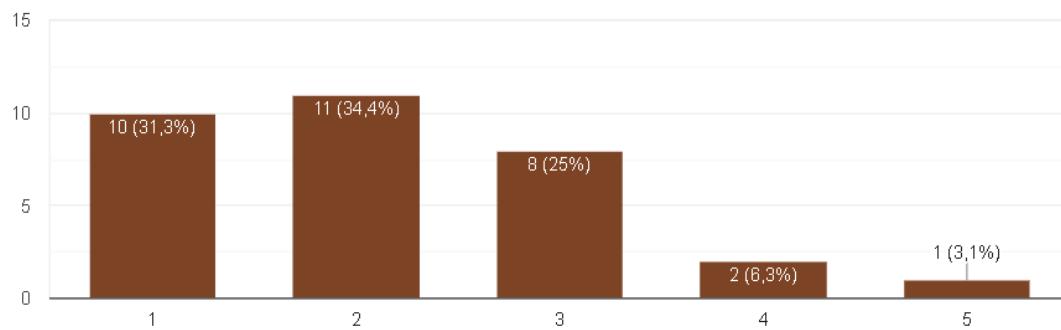


Figura 39 – Respostas sobre a segunda questão do teste SUS

Em relação à facilidade de utilização, uma maioria expressiva de 59,4% dos participantes demonstrou concordância ou concordância total com a afirmação proposta, indicando uma percepção predominantemente positiva sobre a usabilidade da ferramenta. Notavelmente, apenas dois participantes relataram experiências negativas com a ferramenta. Esses resultados são claramente ilustrados na Figura 40, que apresenta a distribuição visual das respostas obtidas no teste.

"Eu achei o sistema fácil de usar"

32 respostas

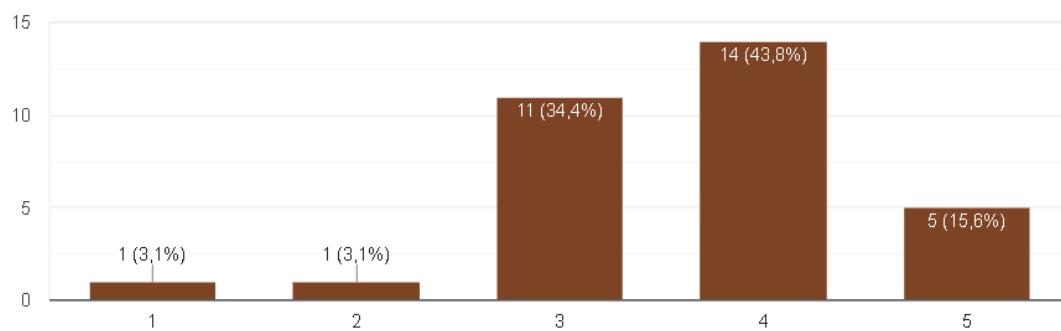


Figura 40 – Respostas sobre a terceira questão do teste SUS

Na questão "Eu acho que precisaria de ajuda de uma pessoa com conhecimentos técnicos para usar o sistema", a maioria, representando 65,6% dos professores participantes, discordou da afirmação, evidenciando uma confiança considerável na capacidade de utilizar o sistema sem assistência técnica. Por outro lado, 28,1% indicaram neutralidade, o que sugere uma certa insegurança ou dúvida em relação à utilização dos recursos do

sistema sem auxílio técnico. Este panorama retrata uma importante percepção sobre a acessibilidade e a autonomia na utilização da ferramenta. Esses resultados são claramente ilustrados na Figura 41, que apresenta a distribuição visual das respostas obtidas no teste.

"Eu acho que precisaria de ajuda de uma pessoa com conhecimentos técnicos para usar o sistema"

32 respostas

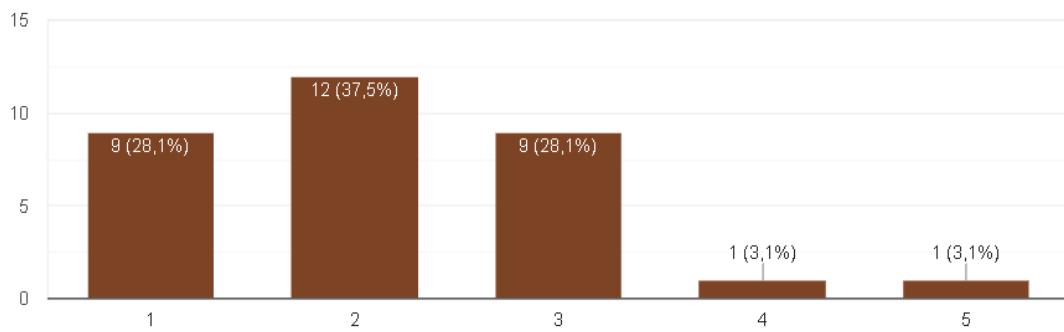


Figura 41 – Respostas sobre a quarta questão do teste SUS

Na quinta questão, que aborda a integração entre as funcionalidades do sistema, observou-se que a maioria dos respondentes expressou respostas positivas, indicando uma percepção favorável em relação à integração das diferentes funcionalidades. No entanto, é importante destacar que também foi identificado um grupo significativo, correspondendo a 34,4% dos participantes, que optou por respostas neutras. Essa proporção sugere a presença de dúvidas ou incertezas entre os usuários quanto à efetividade, ou eficiência da integração do sistema. Essa análise revela a complexidade da percepção dos usuários em relação à interconexão das funcionalidades da ferramenta. Esses resultados são claramente ilustrados na Figura 42, que apresenta a distribuição visual das respostas obtidas no teste.

"Eu acho que as várias funções do sistema estão muito bem integradas"

32 respostas

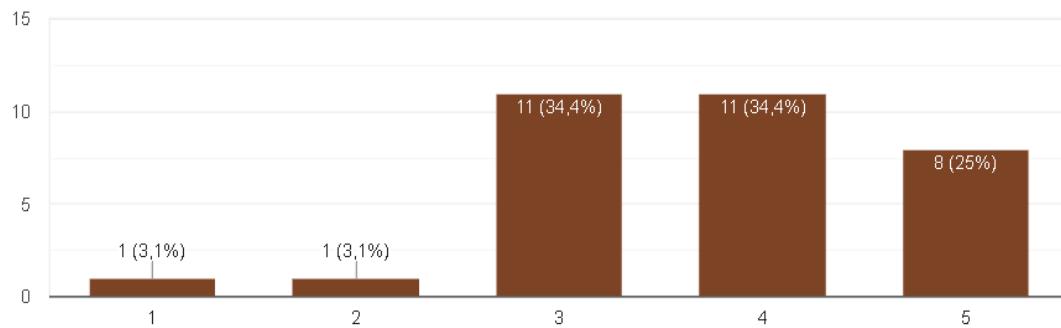


Figura 42 – Respostas sobre a quinta questão do teste SUS

Na sexta questão, notou-se que a grande maioria dos participantes, correspondendo a 78,2%, discordou da afirmação "Eu acho que o sistema apresenta muita inconsistência". Isso sugere que a maioria dos usuários teve uma experiência consistente e confiável durante a utilização do sistema, refletindo uma percepção positiva da sua estabilidade e funcionamento coerente. Por outro lado, apenas 6,2% dos participantes manifestaram algum nível de concordância com essa opinião, indicando uma minoria que percebeu inconsistências no sistema. Esse resultado destaca uma tendência geral de satisfação e confiabilidade dos usuários em relação à consistência do sistema. Esses resultados são claramente ilustrados na Figura 43, que apresenta a distribuição visual das respostas obtidas no teste.

"Eu acho que o sistema apresenta muita inconsistência"

32 respostas

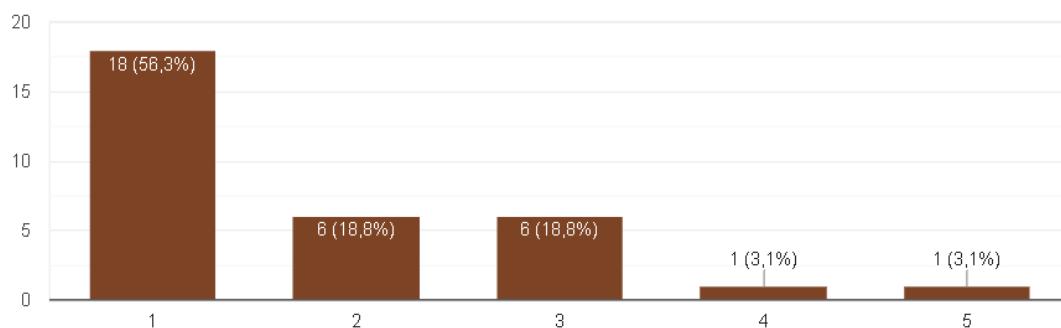


Figura 43 – Respostas sobre a sexta questão do teste SUS

Na questão "Eu imagino que as pessoas aprenderão como usar esse sistema rapi-

damente", uma parcela significativa, equivalente a 62,5% dos professores participantes, optou por responder de forma neutra. Essa escolha sugere uma hesitação ou incerteza generalizada em relação à facilidade de aprendizado da utilização do sistema. Essa constatação indica a necessidade de uma avaliação mais aprofundada sobre os recursos de treinamento e suporte disponíveis para os usuários, visando aprimorar a experiência de aprendizagem e garantir uma transição suave para a utilização do sistema. Esses resultados são claramente ilustrados na Figura 44, que apresenta a distribuição visual das respostas obtidas no teste.

"Eu imagino que as pessoas aprenderão como usar esse sistema rapidamente"

32 respostas

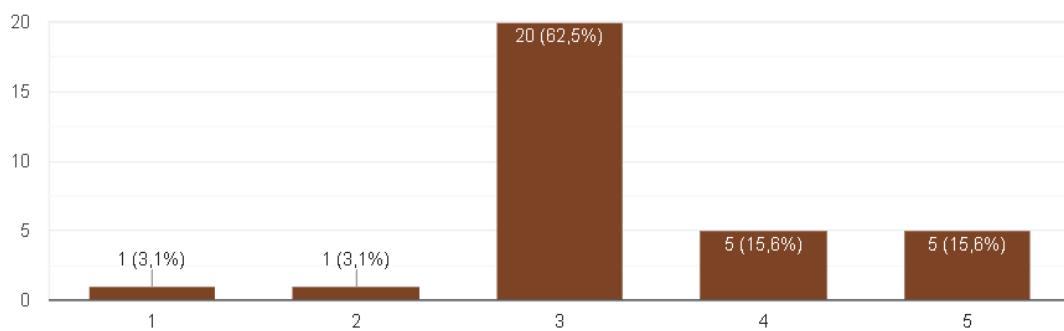


Figura 44 – Respostas sobre a sétima questão do teste SUS

Abordando a questão da usabilidade, na oitava pergunta, "Eu achei o sistema atrapalhado de usar", uma maioria expressiva de 71,9% dos entrevistados discordou da afirmação. Isso indica que uma parcela significativa dos participantes teve uma experiência positiva do ponto de vista da usabilidade, sugerindo que o sistema foi percebido como intuitivo e fácil de usar. Essa resposta reflete uma importante avaliação positiva da eficiência e da acessibilidade do sistema pelos usuários, o que pode contribuir para a sua aceitação e adoção contínua. Esses resultados são claramente ilustrados na Figura 45, que apresenta a distribuição visual das respostas obtidas no teste.

"Eu achei o sistema atrapalhado de usar"

32 respostas

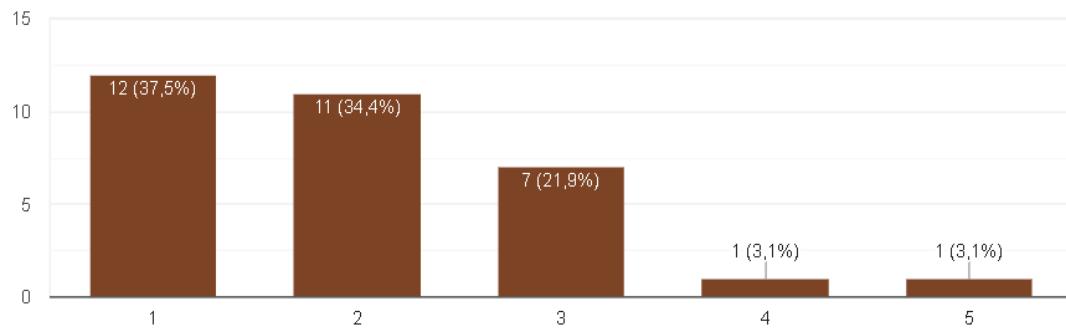


Figura 45 – Respostas sobre a oitava questão do teste SUS

Em relação à afirmação "Eu me senti confiante ao usar o sistema", uma maioria expressiva de 75% dos respondentes concordou em algum nível. Essa resposta reflete críticas majoritariamente favoráveis aos aspectos de utilização do sistema, sugerindo que os usuários se sentiram seguros e confortáveis ao interagir com a ferramenta. Essa confiança percebida é crucial, pois indica que o sistema inspira credibilidade e eficácia em sua funcionalidade, o que pode contribuir para uma adoção mais ampla e uma experiência geral positiva por parte dos usuários. Esses resultados são claramente ilustrados na Figura 46, que apresenta a distribuição visual das respostas obtidas no teste.

"Eu me senti confiante ao usar o sistema"

32 respostas

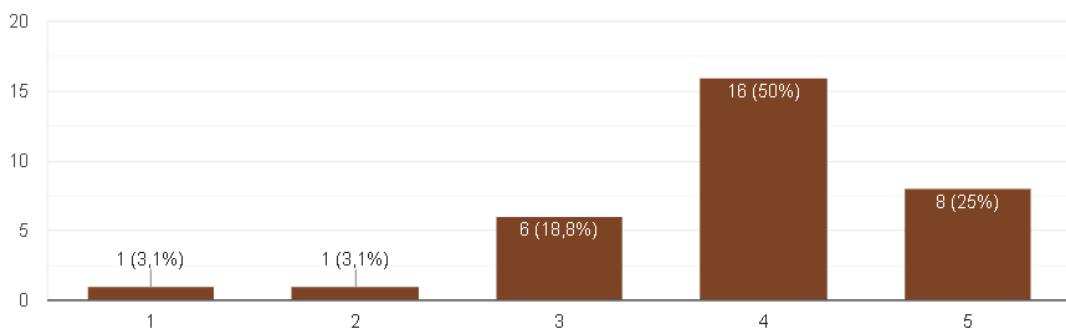


Figura 46 – Respostas sobre a nona questão do teste SUS

Levando em consideração aspectos sobre conhecimento prévio ao utilizar o sistema, constatou-se que apenas 6,2% dos participantes consideraram necessário aprender

coisas novas para utilizar o sistema. Essa parcela pequena sugere que uma minoria enfrentou dificuldades na interpretação da ferramenta devido à sua complexidade ou à falta de familiaridade com determinados conceitos. Esses resultados indicam, em geral, uma boa adaptabilidade do sistema, com a maioria dos usuários conseguindo utilizá-lo sem a necessidade de adquirir novos conhecimentos específicos. Esses resultados são claramente ilustrados na Figura 47, que apresenta a distribuição visual das respostas obtidas no teste.

"Eu precisei aprender várias coisas novas antes de conseguir usar o sistema"

32 respostas

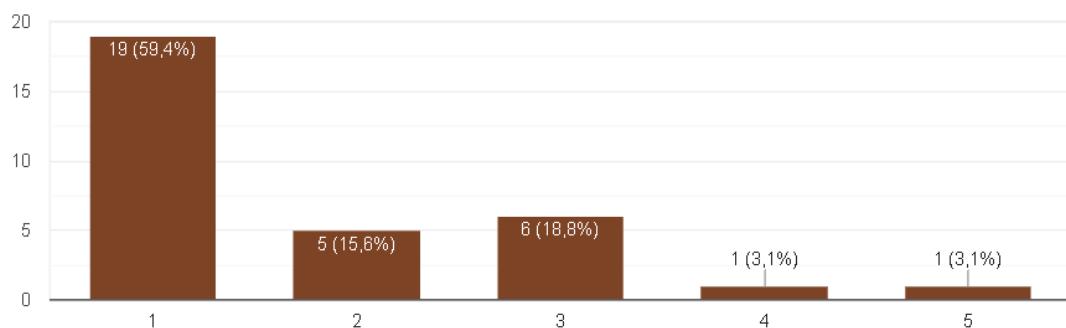


Figura 47 – Respostas sobre a décima questão do teste SUS

O resultado da pontuação do teste revelou uma média de 72,03, apontando para uma experiência positiva dos usuários. A Figura 48, mostra a pontuação do teste e a precisão.

Pontuação média SUS:

72,03

Precisão (i)

Faixa de pontuação (i)

100%

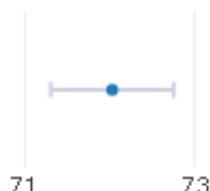


Figura 48 – Pontuação do teste SUS

Na escala SUS, que abrange de 0 a 100, esse valor coloca o sistema acima da média, sugerindo que, em geral, os participantes o consideraram utilizável e de fácil navegação. Em relação à pontuação, o segundo teste registrou uma pontuação inferior ao primeiro. No entanto, vale ressaltar que teve uma participação significativamente maior e demonstrou uma precisão mais elevada, medida pela consistência das respostas. Isso sugere que o segundo teste gerou resultados mais confiáveis e consistentes em comparação com o primeiro.

5.2 Conclusão do Capítulo

Na conclusão deste capítulo de resultados, foram apresentados os resultados obtidos a partir dos dois experimentos do teste de usabilidade SUS, conduzidos com intuito de avaliar a eficácia e a usabilidade do *dashboard* educacional proposto. Esses experimentos forneceram visões valiosas sobre a experiência dos usuários ao interagir com o sistema, permitindo uma análise abrangente e detalhada de diversos aspectos relacionados à sua usabilidade e funcionalidades.

Os resultados revelaram aspectos positivos, como a facilidade de uso do sistema, a clareza das informações apresentadas no *dashboard* e a capacidade de fornecer dados relevantes aos educadores. Além disso, foram identificadas áreas que podem ser aprimoradas para melhor atender às necessidades dos usuários e garantir uma experiência ainda mais satisfatória, tais como a navegabilidade, a organização das informações e a resposta do sistema.

É importante ressaltar que os dados coletados foram analisados levando em consideração não apenas as métricas quantitativas, mas também as percepções qualitativas dos participantes. Essa abordagem proporcionou uma compreensão mais completa do desempenho e da aceitação do *dashboard* educacional.

No próximo capítulo, serão apresentadas as considerações finais do trabalho de proposta do *dashboard* educacional, discutindo os principais achados deste estudo, bem como as implicações práticas e as sugestões para futuras pesquisas na área.

6

CONSIDERAÇÕES FINAIS

O trabalho teve como objetivo a criação de um dashboard, projetado para oferecer suporte na análise e visualização de dados educacionais provenientes de STIs. Este dashboard foi concebido com o propósito de aprimorar o desempenho do sistema educacional, enquanto simultaneamente proporciona assistência aos alunos e professores. Utilizando técnicas de visualização de dados, estatísticas e mineração de dados, visando extrair informações que impulsionassem a evolução do cenário educacional.

Para atingir o objetivo, realizamos uma investigação abrangente, mapeando o estado da arte dos STIs, para compreender as suas nuances técnicas e pedagógicas. Além disso, procedemos à avaliação da prevalência dos painéis educacionais existentes, identificando lacunas e oportunidades para a evolução da proposta.

Após mapear os trabalhos da área, avançamos para o desenvolvimento da primeira versão do sistema e conduzimos um teste de usabilidade para validar a eficácia e a experiência do usuário com a ferramenta proposta.

Atualmente, estamos em pleno desenvolvimento de uma versão aprimorada dessa ferramenta, implementando melhorias e trabalhando na sua integração com outros STIs. Este estágio permitirá uma análise mais robusta e detalhada, viabilizando a descoberta de informações até então desconhecidas, com o potencial de contribuir significativamente para o contexto educacional.

Assim sendo, esta pesquisa representa um marco importante na busca por uma educação mais eficaz e personalizada, proporcionando aos alunos e professores as ferramentas necessárias para otimizar o processo de aprendizagem e, por conseguinte,

fomentar um ambiente educacional mais dinâmico e produtivo.

6.1 Contribuições

A pesquisa conduzida teve como objetivo propor uma ferramenta inovadora: o Edudata Visualizer, voltada para a visualização dinâmica do desempenho dos estudantes em STIs. Esta ferramenta, fundamentada na análise de dados educacionais, visou facilitar a identificação de alunos que enfrentam dificuldades em áreas específicas, capacitando os professores a aplicar estratégias pedagógicas mais eficazes. Embora o STI LEIA tenha sido o principal foco deste estudo, é válido destacar que a mesma abordagem pode ser adaptada para outros STIs com mínimas alterações. O Edudata Visualizer representa um avanço no aprimoramento do processo de ensino-aprendizagem, oferecendo aos educadores uma ferramenta para auxiliar na compreensão e atender às necessidades individuais dos estudantes.

A pesquisa gerou quatro artigos, sendo três publicados na conferência *IEEE - Frontiers in Education*, um deles apresentado em 2022 de forma online para o evento realizado em Uppsala na Suécia e outros dois apresentados em 2023 de forma presencial na cidade de *College Station* nos Estados Unidos. Além destes, um artigo foi aceito para publicação na Revista Novas Tecnologias na Educação. Ademais, durante o mestrado contribuí com um artigo desenvolvido no laboratório de Sistemas Inteligentes. Seguem os títulos dos trabalhos:

- *IEEE FIE 2022: Educational Data Mining: Analysis Based On an Intelligent Tutoring System for Teaching Algebra*
- *IEEE FIE 2023: Educational Data Mining Applied to Performance Prediction in an Intelligent Tutoring System for Teaching Algebra*
- *IEEE FIE 2023: The Proposal of a Dashboard for Analysis and Visualization of Educational Data of Intelligent Tutoring Systems*
- *IEEE FIE 2023: ERPLab: Remote Laboratory for Teaching Robotics and Programming*

- RENOTE 2024: Visualização de Dados Educacionais: Desenvolvendo uma Ferramenta para Análise de Dados de Sistemas Tutores Inteligentes.

6.2 Limitações

Durante o desenvolvimento da pesquisa, foi constatada uma série de desafios que demandaram ajustes na abordagem. Inicialmente, o planejamento estava fundamentado na execução de um experimento de escopo mais amplo, visando utilizar uma base de dados mais extensa. No entanto, devido a restrições logísticas e temporais, adaptamos a estratégia e utilizamos os dados disponíveis na base do STI LEIA.

Ao realizar uma análise de dados educacionais com técnicas de mineração de dados, enfrentamos a limitação de uma base de dados pequena, o que compromete a validade e a generalização dos resultados. A insuficiência de dados dificulta a captura de variáveis importantes e a identificação de padrões complexos. Para mitigar essas limitações, utilizamos a técnica de árvore de decisão, que se adaptou melhor à base de dados menor, oferecendo resultados mais interpretáveis e menos suscetíveis ao *overfitting*. No entanto, a validação dos resultados ainda foi desafiadora devido à divisão dos dados reduzidos.

Embora essa mudança tenha limitado momentaneamente a análise, planejamos realizar futuras investigações para explorar novas perspectivas na visualização de dados. Essas análises adicionais não apenas enriquecerão nossos resultados, mas também contribuirão para a evolução contínua do *Edudata Visualizer*, fortalecendo sua capacidade de fornecer informações aos usuários.

6.3 Trabalhos Futuros

Como parte do processo de avanço contínuo do projeto, planejamos disponibilizar a nossa ferramenta por meio de uma plataforma de hospedagem de código-fonte e arquivos com controle de versão. Este esforço pretende garantir a manutenção do sistema e facilitar a disponibilização de novas versões para os desenvolvedores interessados.

Além disso, como parte dos planos futuros, iremos expandir a implementação do dashboard para outros sistemas educacionais dentro do Laboratório de Sistemas Inteligentes. Esta expansão envolverá a realização de novos experimentos e adaptações para diferentes tipos de plataformas educativas, focando em explorar novas oportunidades de aplicação e em garantir que a ferramenta seja acessível e útil em uma variedade de contextos educacionais. Este processo de expansão e adaptação visa manter o processo de inovação contínua e com o fornecimento de soluções eficazes para os desafios enfrentados no campo da educação.

Outra abordagem estudada é a implementação no dashboard de um modelo computacional desenvolvido para conduzir a Análise de Sentimentos em textos postados por estudantes nas interações com o STI. Permitindo identificar os sentimentos expressos pelos estudantes, destacar o sentimento predominante em uma turma e revisar o histórico de sentimentos de um aluno específico. Esta abordagem já foi desenvolvida em outro contexto no Laboratório de Sistemas Inteligentes, através do trabalho Modelo para Análise de Sentimentos no AVA: Um estudo de Caso ([ALENCAR; NETTO, 2022](#)).

REFERÊNCIAS

- ALDOWAH, H.; AL-SAMARRAIE, H.; FAUZY, W. M. Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, Elsevier, v. 37, p. 13–49, 2019. [20](#)
- ALENCAR, M. A. d. S. et al. Uma abordagem computacional para detectar emoções de alunos em cursos online. Universidade Federal do Amazonas, 2020. [19](#)
- ALENCAR, M. A. dos S.; NETTO, J. F. de M. Modelo para análise de sentimentos no ava: Um estudo de caso. *Revista Novas Tecnologias na Educação*, v. 20, n. 1, p. 143–152, 2022. [112](#)
- APARICIO, M.; COSTA, C. J. Data visualization. *Communication design quarterly review*, ACM New York, NY, USA, v. 3, n. 1, p. 7–11, 2015. [38](#)
- AYALA, M. V. A. M. V. et al. Dashboard for the improvement of school management in educational institutions. *Ciencia Latina Revista Científica Multidisciplinar*, v. 7, n. 1, p. 3018–3037, 2023. [22](#)
- BAKER, R. S.; INVENTADO, P. S. Educational data mining and learning analytics. In: *Learning analytics*. [S.I.]: Springer, 2018. p. 61–75. [33](#)
- BAKER, R. S.; MARTIN, T.; ROSSI, L. M. Educational data mining and learning analytics. *The Wiley handbook of cognition and assessment: Frameworks, methodologies, and applications*, Wiley Online Library, p. 379–396, 2018. [32](#)
- BATOOL, S. et al. Educational data mining to predict students' academic performance: A survey study. *Education and Information Technologies*, Springer, v. 28, n. 1, p. 905–971, 2023. [31](#)
- BOSCARDIN, C. et al. Twelve tips to promote successful development of a learner performance dashboard within a medical education program. *Medical Teacher*, Taylor & Francis, v. 40, n. 8, p. 855–861, 2018. [38](#)
- BROWN, T. A. *Confirmatory factor analysis for applied research*. [S.I.]: Guilford publications, 2015. [25](#)
- CECHINEL, C. et al. Learning analytics para moodle em uma arquitetura na nuvem: uma solução escalável para predição de risco acadêmico. In: SBC. *Anais do II Workshop de Aplicações Práticas de Learning Analytics em Instituições de Ensino no Brasil*. [S.I.], 2023. p. 128–137. [54](#)

- DUTT, A.; ISMAIL, M. A.; HERAWAN, T. A systematic review on educational data mining. *Ieee Access*, IEEE, v. 5, p. 15991–16005, 2017. [20](#)
- ELLAWAY, R. H. et al. Developing the role of big data and analytics in health professional education. *Medical teacher*, Taylor & Francis, v. 36, n. 3, p. 216–222, 2014. [38](#)
- FALGUERAS, P. B. Dashboard for environmental awareness education. 2022. [19, 22](#)
- FARAHMAND, A.; DEWAN, M. A. A.; LIN, F. Student-facing educational dashboard design for online learners. In: IEEE. 2020 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech). [S.I.], 2020. p. 345–349. [21, 23](#)
- FEW, S. *Information dashboard design: The effective visual communication of data*. [S.I.]: O'Reilly Media, Inc., 2006. [37](#)
- FRIENDLY, M.; WAINER, H. *A history of data visualization and graphic communication*. [S.I.]: Harvard University Press, 2021. [37](#)
- GAŠEVIĆ, D.; TSAI, Y.-S.; DRACHSLER, H. Learning analytics in higher education–stakeholders, strategy and scale. *The internet and higher education*, Elsevier Science Inc., v. 52, n. 2022, p. 100833, 2022. [19](#)
- GRASSINI, S. Shaping the future of education: exploring the potential and consequences of ai and chatgpt in educational settings. *Education Sciences*, MDPI, v. 13, n. 7, p. 692, 2023. [20](#)
- GRINSTEIN, G. G.; WARD, M. O. Introduction to data visualization. *Information visualization in data mining and knowledge discovery*, Morgan Kaufmann Publishers San Francisco, v. 1, p. 21–45, 2002. [37](#)
- GUILHERME, A. A.; CHERON, C. *Guia prático de pesquisa em Educação*. [S.I.]: Ediucs, 2021. [25](#)
- HAN, J.; KAMBER, M.; PEI, J. *Data mining concepts and techniques 4th edition*. [S.I.]: Morgan kaufmann, 2022. [72](#)
- HOLSTEIN, K. et al. Luna: a dashboard for teachers using intelligent tutoring systems. *Education*, v. 60, n. 1, p. 159–171, 2010. [22, 51](#)
- HURTATIZ, Y. E. M. et al. Sistemas tutores inteligentes como apoyo en el proceso de aprendizaje. *Redes de Ingeniería*, v. 6, n. 1, p. 25–44, 2015. [10, 34](#)
- INJADAT, M. et al. Systematic ensemble model selection approach for educational data mining. *Knowledge-Based Systems*, Elsevier, p. 105992, 2020. [20](#)
- IVE, I. T. M. T. Intelligent tutor for programming system using multiple intelligences. *IEEE Latin America Transactions*, IEEE, v. 16, n. 2, p. 634–638, 2018. [49](#)
- JAQUES, P. A. Da teoria à prática em sala de aula: experiências e insights de uma pesquisadora em inteligência artificial aplicada à educação. *Tecnologias, Sociedade e Conhecimento*, v. 10, n. 2, p. 124–152, 2023. [33](#)

- KAMATH, R.; KAMAT, R. *Educational Data Mining with R and Rattle*. [S.l.]: CRC Press, 2022. [33](#)
- KEELE, S. et al. *Guidelines for performing systematic literature reviews in software engineering*. [S.l.], 2007. [43](#)
- KHELIFI, T. et al. Ex-lad: an explainable learning analytics dashboard in higher education. In: *Proceedings of 36th International Conference on*. [S.l.: s.n.], 2024. v. 97, p. 38–51. [55](#)
- LAMPROPOULOS, G. Educational data mining and learning analytics in the 21st century. In: *Encyclopedia of Data Science and Machine Learning*. [S.l.]: IGI Global, 2023. p. 1642–1651. [32](#)
- LETAW, L. Unified modeling language class and sequence diagrams. *Handbook of Software Engineering Methods*, Oregon State University, 2024. [67, 68](#)
- LOPES, A. M. M. et al. Improving students skills to solve elementary equations in k-12 programs using an intelligent tutoring system. In: IEEE. *2019 IEEE Frontiers in Education Conference (FIE)*. [S.l.], 2019. p. 1–9. [23](#)
- LOPES, A. M. M. L. Um sistema de tutoria para apoiar o desenvolvimento do conhecimento matemático algébrico na resolução de equação polinomial de 1º grau. Universidade Federal do Amazonas, 2019. [10, 35, 36](#)
- MAJUMDAR, R. et al. Laview: Learning analytics dashboard towards evidence-based education. In: SOCIETY FOR LEARNING ANALYTICS RESEARCH (SOLAR). *Companion Proceedings of the 9th International Conference on Learning Analytics and Knowledge*. [S.l.], 2019. [57](#)
- MENEZES, M. F. D.; NETTO, J. F. D. M.; LOPES, A. M. M. Educational data mining: Analysis based on an intelligent tutoring system for teaching algebra. In: IEEE. *2022 IEEE Frontiers in Education Conference (FIE)*. [S.l.], 2022. p. 1–7. [60, 73, 75, 76](#)
- MENEZES, M. F. d. et al. Um mapeamento sistemático sobre softwares de gestão e planejamento educacional. 2020. [22](#)
- MENEZES, M. F. de et al. Applying educational data mining to classify students in an intelligent tutoring system for algebra instruction. In: IEEE. *2023 IEEE Frontiers in Education Conference (FIE)*. [S.l.], 2023. p. 1–8. [61](#)
- MENEZES, M. F. de et al. The proposal of a dashboard for analysis and visualization of educational data of intelligent tutoring systems. In: IEEE. *2023 IEEE Frontiers in Education Conference (FIE)*. [S.l.], 2023. p. 1–8. [61, 73](#)
- MICHAELI, S.; KROPARO, D.; HERSHKOVITZ, A. Teachers' use of education dashboards and professional growth. *International Review of Research in Open and Distributed Learning*, Athabasca University Press (AU Press), v. 21, n. 4, p. 61–78, 2020. [20](#)
- MOUSAVINASAB, E. et al. Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments*, Taylor & Francis, v. 29, n. 1, p. 142–163, 2021. [34](#)

- NAFEA, I. T. Machine learning in educational technology. *Machine learning-advanced techniques and emerging applications*, IntechOpen, p. 175–183, 2018. [20](#)
- NARANJO, D. M. et al. A visual dashboard to track learning analytics for educational cloud computing. *Sensors*, MDPI, v. 19, n. 13, p. 2952, 2019. [56](#)
- NG, J. T.-D.; WANG, Z.; HU, X. Needs analysis and prototype evaluation of student-facing la dashboard for virtual reality content creation. In: *LAK22: 12th International Learning Analytics and Knowledge Conference*. [S.l.: s.n.], 2022. p. 444–450. [23](#)
- OLIVEIRA, F. *TCF: O Projeto de Pesquisa Intervenção - Pesquisa de Intervenção: Aspectos e considerações sobre sua natureza e formas de Trabalho de Conclusão Final no PROFIAP*. [S.l.]: Universidade Federal Rural do Semi-Árido (UFERSA), 2016. [25](#)
- OPARA, E.; THERESA, A. M.-E.; ADUKE, T. C. Chatgpt for teaching, learning and research: Prospects and challenges. *Opara Emmanuel Chinonso, Adalikwu Mfon-Ette Theresa, Tolorunleke Caroline Aduke* (2023). *ChatGPT for Teaching, Learning and Research: Prospects and Challenges*. *Glob Acad J Humanit Soc Sci*, v. 5, 2023. [91](#)
- PARANHOS, H. P. et al. Desenvolvimento de um dashboard para análise e visualização dos dados educacionais dos discentes do instituto de ciências exatas e aplicadas da ufop. 2021. [58](#)
- PARK, Y.; JO, I.-H. Development of the learning analytics dashboard to support students' learning performance. *Journal of Universal Computer Science*, IICM, v. 21, n. 1, p. 110, 2015. [57](#)
- PEÑA-AYALA, A. Educational data mining: A survey and a data mining-based analysis of recent works. *Expert systems with applications*, Elsevier, v. 41, n. 4, p. 1432–1462, 2014. [32](#)
- PEREIRA, F. D. et al. Explaining individual and collective programming students' behavior by interpreting a black-box predictive model. *IEEE Access*, v. 9, p. 117097–117119, 2021. [55](#)
- PETERSEN, K. et al. Systematic mapping studies in software engineering. In: *12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12*. [S.l.: s.n.], 2008. p. 1–10. [42](#)
- RABELO, H. et al. Utilização de técnicas de mineração de dados educacionais para predição de desempenho de alunos de ead em ambientes virtuais de aprendizagem. In: *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*. [S.l.: s.n.], 2017. v. 28, n. 1, p. 1527. [19](#)
- RAMOS, J. L. C. et al. Crisp-edm: uma proposta de adaptação do modelo crisp-dm para mineração de dados educacionais. In: SBC. *Anais do XXXI Simpósio Brasileiro de Informática na Educação*. [S.l.], 2020. p. 1092–1101. [32](#)
- ROCHA, J. *Multi-agent Systems*. Rijeka: IntechOpen, 2017. ISBN 978-953-51-3536-4. Disponível em: <<https://doi.org/10.5772/66595>>. [39](#)
- ROMERO, C.; VENTURA, S. Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Wiley Online Library, v. 10, n. 3, p. e1355, 2020. [33](#)

- RUIPÉREZ-VALIENTE, J. A. et al. Ideating and developing a visualization dashboard to support teachers using educational games in the classroom. *IEEE access*, IEEE, v. 9, p. 83467–83481, 2021. [56](#)
- RUSSELL, S.; NORVIG, P. Artificial intelligence: a modern approach, 4th us ed. *University of California, Berkeley*, 2021. [39](#)
- SAEED, S. et al. Analysis of software development methodologies. *International Journal of Computing and Digital Systems*, University of Bahrain, v. 8, n. 5, p. 446–460, 2019. [83](#)
- SAMPAIO, T. B. Metodologia da pesquisa. Brasil, 2022. [25](#)
- SAURO, J. *A Practical Guide to the System Usability Scale: Background, Benchmarks Best Practices*. [S.I.]: Createspace Independent Publishing Platform, 2011. [95](#)
- SCHEUER, O.; MCLAREN, B. M. Educational data mining. *Encyclopedia of the Sciences of Learning*, Springer, Springer, 2012. [32](#)
- SCHMITT, J. A.; BERNARDI, G.; KANTORSKI, G. Z. ipermanency: uma ferramenta para gestão da evasão na ead baseada em mineração de dados educacionais. *RENOTE-Revista Novas Tecnologias na Educação*, v. 17, n. 1, p. 306–315, 2019. [59](#)
- SEDU. Dicas de como se organizar para estudar em casa. In: GOVERNO DO ESTADO DO ESPIRITO SANTO. *ESCOLAR: Atividade pedagógica não presencial*. [S.I.], 2024. [91](#)
- SHOAIB, M. et al. Prediction of an educational institute learning environment using machine learning and data mining. *Education and Information Technologies*, Springer, v. 27, n. 7, p. 9099–9123, 2022. [20](#)
- SILEYEW, K. J. *Research design and methodology*. [S.I.]: Cyberspace, 2019. v. 7. [24](#)
- SILVA, C. R. Q. Critérios para priorização de estudos primários identificados por snowballing com conjunto inicial gerado por string de busca. Universidade Federal de São Carlos, 2017. [53](#)
- SILVA, E. V. da; NETTO, J. F. de M.; SOUZA, R. A. L. de. Vla dashboard: Um mecanismo para visualização do desempenho de estudantes de matemática no ensino médio. *RENOTE-Revista Novas Tecnologias na Educação*, v. 16, n. 2, p. 484–494, 2018. [10, 59, 60](#)
- SILVA, G. L. B.; CARVALHO, J. A. de; MACIEL, A. M. A. Desenvolvimento de um learning analytics dashboard a partir de modelos de mineração de dados educacionais. *Revista de Engenharia e Pesquisa Aplicada*, v. 6, n. 3, p. 59–69, 2021. [60](#)
- SURSE, M.; KOKATE, M.; SANGHAVI, M. Dashboard for monitoring the real time data in educational organization. *International Journal of Advances in Engineering and Management (IJAE)*, 2022. [37](#)
- SUSNJAK, T.; RAMASWAMI, G. S.; MATHRANI, A. Learning analytics dashboard: a tool for providing actionable insights to learners. *International Journal of Educational Technology in Higher Education*, Springer, v. 19, n. 1, p. 12, 2022. [20, 53](#)
- TAMADA, M. M.; GIUSTI, R.; NETTO, J. F. de M. Predicting student performance based on logs in moodle lms. In: *2021 IEEE Frontiers in Education Conference (FIE)*. [S.I.: s.n.], 2021. p. 1–8. [55](#)

UNWIN, A. Why is data visualization important? what is important in data visualization? *Harvard Data Science Review*, v. 2, n. 1, p. 1, 2020. [37](#)

VANNESTE, P. et al. Computer vision and human behaviour, emotion and cognition detection: A use case on student engagement. *Mathematics*, MDPI, v. 9, n. 3, p. 287, 2021. [21](#)

WEST, D. M. Big data for education: Data mining, data analytics, and web dashboards. *Governance studies at Brookings*, v. 4, n. 1, p. 1–10, 2015. [22](#)

YAGCI, M. Educational data mining: prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, Springer, v. 9, n. 1, p. 11, 2022. [20](#)

A

APÊNDICE

Nesta seção adicional, apresentamos os artigos publicados na conferência Frontiers in Education, juntamente com o artigo submetido à Revista Novas Tecnologias na Educação.

Apêndice A - Primeiro artigo publicado: Educational Data Mining: Analysis Based On an Intelligent Tutoring System for Teaching Algebra.

Educational Data Mining: Analysis Based On an Intelligent Tutoring System for Teaching Algebra

Matheus Freitas de Menezes
Institute of Computing (IComp)
Federal University of Amazonas
Manaus, AM, Brazil
matheus.menezes@icomp.ufam.edu.br

José Francisco de M. Netto
Institute of Computing (IComp)
Federal University of Amazonas
Manaus, AM, Brazil
jnetto@icomp.ufam.edu.br

Arcanjo Miguel Mota Lopes
Institute of Computing (IComp)
Federal University of Amazonas
Manaus, AM, Brazil
amml@icomp.ufam.edu.br

Abstract—This complete article of the research category presents proposes the application of Educational Data Mining (EDM) techniques, based on data generated by a web-based Intelligent Tutor System (ITS), developed for the teaching of algebra, which helps students to develop fundamental knowledge of mathematics, in addition to carrying out a comparative study between the level of difficulty of the proposed algebraic questions. For this, we use the knowledge discovery methodology to perform the following steps from the data: cleaning, integration, selection, transformation, mining, evaluation and presentation of information. The practical results reveal that the proposed architecture can classify the academic performance of students in each evaluation period with an accuracy of around 80%. It was also possible to identify factors related to the resolution of questions, such as the rate of correct answers and the mathematical steps to solve an exercise, classifying the level of difficulty of the questions proposed by the system, acting on the student's deficiencies, and the system. The results obtained from the data analysis can serve as a basis for decision-making, helping in the teaching process for teachers, tutors and managers to monitor academic performance, enabling the correction of problems and identifying the individual and collective difficulties present in each class. In summary, these results also provide a general roadmap on the performance of using EDM techniques in a given context.

Index Terms—Educational Data Mining, Intelligent Tutoring System.

I. INTRODUCTION

With the increasing use of technologies to support the teaching and learning process in the educational field, a large volume of data has been generated based on the different types of interaction with the system, covering mainly students, tutors, and teachers. However, much of this data has not been analyzed, establishing a significant gap, given the amount of information that can potentially be extracted from such data [1].

Consequently, data mining techniques are gaining more and more important in the education sector. Educational Data Mining (EDM) has attracted significant interest in recent years in an attempt to personalize and improve students' learning process [2]. Some other areas, teaching is discovering the potential impact of these techniques on the learning process. Certainly, these techniques can provide information to support the development of educational models that increase the efficiency and quality of [3] teaching and learning.

According to Organization for Economic Co-operation and Development (OECD), through the International Student Assessment Program (PISA), 68.1% of 15-year-old Brazilians students have a lower mathematics learning degree than that classified as basic. [4].

Taking into account the current moment in public high schools, it becomes increasingly complex for the teacher to detect individual or collective difficulties, given the abundance of variables. Often, this recognition is only noticed through the results of the final grades that are part of the evaluation method that makes up the [5] educational process.

The environment of an educational institution involves three types of actors: teacher, student, and the environment. The interaction between these three actors generates voluminous data that can be systematically grouped to extract valuable information. These data analyses allow the prediction of student performance, associate learning styles of different profiles and their behaviors, and collectively improve institutional performance [6].

Given this scenario and considering the amount of data generated by the support systems for the teaching and learning process of students [7], the project proposes to apply mining techniques in data generated by a web-based Intelligent Tutor System (ITS) called LEIA (LEarnIng Algebra), which helps students develop foundational math skills [8].

The rest of the paper is organized as follows. Section 2 presents and discusses related work. Section 3 presents the research method used while Section 4 shows the results and discussions. Section 5 presents conclusions and future work.

II. THEORETICAL FOUNDATION

This Section exposes the fundamental concepts for understanding this work. The concepts of Educational Data Mining and the main works related to the research are presented.

A. Educational Data Mining

EDM is a subarea of Data Mining, that is, this line is composed of models, tasks, methods and algorithms that support the exploration of data from educational environments, to discover patterns and predictions that characterize the behaviors, practices, contents of domain, assessments and educational activities of students [9]. EDM is concerned with

Apêndice B - Segundo artigo publicado: Applying Educational Data Mining to Classify Students in an Intelligent Tutoring System for Algebra Instruction.

Applying Educational Data Mining to Classify Students in an Intelligent Tutoring System for Algebra Instruction

Matheus Freitas de Menezes
 Institute of Computing (IComp)
 Federal University of Amazonas
 Manaus, AM, Brazil
 matheus.menezes@icomp.ufam.edu.br

José Francisco de M. Netto
 Institute of Computing (IComp)
 Federal University of Amazonas
 Manaus, AM, Brazil
 jnetto@icomp.ufam.edu.br

Arcanjo Miguel Mota Lopes
 Institute of Computing (IComp)
 Federal University of Amazonas
 Manaus, AM, Brazil
 amml@icomp.ufam.edu.br

Fabiann Mattheus Dantas Barbosa
 Institute of Computing (IComp)
 Federal University of Amazonas
 Manaus, AM, Brazil
 fabiann.barbosa@icomp.ufam.edu.br

Abstract—This complete article of the research category presents proposes the application of Educational Data Mining (EDM) techniques using data generated by a web-based Intelligent Tutoring System (ITS) developed for teaching algebra. The aim is to classify students' academic performance, seeking evidence to assist teachers in the decision-making process. To achieve this goal, we utilized the Knowledge Discovery in Databases (KDD) methodology to analyze students' academic data and identify patterns, trends, and relationships that can be used to enhance the teaching and learning process and make informed decisions in the educational context. Practical results revealed that the Random Forest algorithm classified students' academic performance in each evaluation period with an accuracy of approximately 94%. These findings derived from data analysis can serve as a basis for decision-making, aiding the teaching process for teachers, tutors, and administrators, as well as monitoring academic performance and identifying challenges present in each class. In summary, these results provide a comprehensive guide to the use of EDM techniques in a specific context.

Index Terms—Educational Data Mining, Intelligent Tutoring System.

I. INTRODUCTION

Since the 1970s, Intelligent Tutoring Systems (ITS) have been achieving the main goal of assisting students and teachers in technology-mediated teaching and learning processes. One of the highlights for ITS is related to how it assesses students' performance during one-on-one tutoring. This allows combining tutoring with assessment to support pedagogical decision-making and provide an adaptive instructional model [1]. However, for adaptive teaching to occur, ITS needs to understand students and the difficulties that arise, mainly from knowledge gaps [2].

With advancements in technology and AI, these systems have become more effective with instant feedback and instruc-

tions without any human intervention [3], and much of this improvement comes from the use of Educational Data Mining (EDM) techniques that collect data such as time, sequence of information, and student context [4]. This has helped identify strengths and weaknesses and predict and filter good and poor students [5]. Additionally, it has assisted educators in making the right decisions about choosing learning environments to better meet students' needs and educational objectives [6], [7].

Based on the preceding information, this research aims to apply EDM techniques for student classification using an ITS focused on step-by-step algebraic equation solving. The application of EDM aims to enhance the instructional model so that students can progress independently, overcoming their zones of proximal development and developing mathematical skills. Furthermore, the results will generate reports to allow educators to perform, in the context of technology-mediated teaching, the didactic transposition of content according to individual students' difficulties.

The rest of the paper is organized as follows. Section 2 presents and discusses related work. Section 3 presents the research method used while Section 4 shows the results and discussions. Section 5 presents conclusions and future work.

II. THEORETICAL FOUNDATION

This Section exposes the fundamental concepts for understanding this work. The concepts of EDM and the main works related to the research are presented.

A. Educational Data Mining

EDM is used to discover significant phenomena and resolve educational issues occurring in the context of teaching and learning [9]. Methods from the psychometrics literature are

Apêndice C - Terceiro artigo publicado: The Proposal of a Dashboard for Analysis and Visualization of Educational Data of Intelligent Tutoring Systems.

The Proposal of a Dashboard for Analysis and Visualization of Educational Data of Intelligent Tutoring Systems

Matheus Freitas de Menezes
 Institute of Computing (IComp)
 Federal University of Amazonas
 Manaus, AM, Brazil
 matheus.menezes@icomp.ufam.edu.br

José Francisco de M. Netto
 Institute of Computing (IComp)
 Federal University of Amazonas
 Manaus, AM, Brazil
 jnetto@icomp.ufam.edu.br

Arcanjo Miguel Mota Lopes
 Institute of Computing (IComp)
 Federal University of Amazonas
 Manaus, AM, Brazil
 amml@icomp.ufam.edu.br

Ronilson Cavalcante Silva
 Institute of Computing (IComp)
 Federal University of Amazonas
 Manaus, AM, Brazil
 ronilson.silva@icomp.ufam.edu.br

Ramayana A. Menezes Júnior
 University of the State of Amazonas
 Higher School of Technology (EST)
 Manaus, AM, Brazil
 rajunior@uea.edu.br

Abstract—This research paper discusses the impact of teaching and learning technologies on generating extensive data in education. Despite this data, a significant portion remains unanalyzed, creating a gap in utilizing its potential insights. Educational dashboards address this issue by presenting student information and interactions in online learning environments, utilizing Educational Data Mining (EDM) techniques. EDM aims to understand student behavior, enhance teaching processes, identify learning patterns, and personalize content. It aids in performance analysis, helping struggling students and guiding teaching adjustments. The study proposes an educational dashboard for analyzing Intelligent Tutoring System (ITS) data, specifically in algebra teaching. Using decision tree algorithm to extract insights from student interactions. The developed web dashboard organizes and visualizes student performance data, aiding monitoring, decision-making, communication, and transparency. Educational dashboards enhance decision-making, data analysis, and communication, benefiting both students and the educational system's performance.

Index Terms—Educational Data Mining, Intelligent Tutoring System, Educational Dashboard.

I. INTRODUCTION

Considering the current situation in public high schools, it is becoming increasingly complex for teachers to identify individual or collective difficulties, given the abundance of variables. Often, this recognition is only evident through the results of final grades, which are part of the evaluative method that constitutes the educational process [1]. Teachers face several challenges in monitoring student performance. Some of these challenges include a large number of students in the classroom, lack of unified and outdated educational data, and the difficulty of analysis, preventing data interpretation. In summary, teachers encounter various obstacles to efficiently, accurately, and promptly track student performance.

To enhance information visualization for the teacher, it is important to use technologies that allow dynamic and simplified presentation of this information in their activities [2]. With this purpose, many institutions employ the concept of dashboards, created to compile key metrics in a simple and easy-to-understand interface, enabling students, teachers, and administrators to quickly visualize information in the most practical way. These systems are automatically updated based on data stored in educational databases [3], [4].

The lack of adequate analysis of educational data can result in inefficient teaching and student failure. Without access to precise information about student performance, it is difficult to identify weaknesses and develop improvement plans. Additionally, the absence of data analysis restricts teachers' ability to assess the impact of their teaching strategies and make necessary adjustments. Many educational systems do not store data correctly or do not invest in the analysis of available data.

As an example, Intelligent Tutoring Systems (ITS) typically generate an abundance of data about student progress and learning. The analyses that can be derived from this data include, for example, estimates of student knowledge, categorized by skills and misconceptions in a specific subject of interest, as well as tracking time and progress in various activities within the ITS. While it seems highly likely that human teachers could benefit from accessing this information, ITSs do not usually have dashboards for teachers, leaving educators without a view of students' needs in various contexts. However, ITS analyses are used for research purposes. Even though some ITSs present statistical information that can be useful for teachers, these are not normally designed specifically to meet all needs [5].

Considering dashboards as important tools in the learning process, enabling students to make modifications during their

Apêndice D - Quarto artigo publicado: ERPLab: Remote Laboratory for Teaching Robotics and Programming.

ERPLab: Remote Laboratory for Teaching Robotics and Programming

Ronilson Cavalcante da Silva
 Institute of Computing (IComp)
 Federal University of Amazonas
 Manaus, AM, Brazil
 ronilson.silva@icomp.ufam.edu.br

José Francisco de Magalhães Netto
 Institute of Computing (IComp)
 Federal University of Amazonas
 Manaus, AM, Brazil
 jnetto@icomp.ufam.edu.br

Arcanjo Miguel Mota Lopes
 Institute of Computing (IComp)
 Federal University of Amazonas
 Manaus, AM, Brazil
 amml@icomp.ufam.edu.br

Matheus Freitas de Menezes
 Institute of Computing (IComp)
 Federal University of Amazonas
 Manaus, AM, Brazil
 matheus.menezes@icomp.ufam.edu.br

Ramayana Assunção Menezes Júnior
 UEA - Universidade do Estado do Amazonas
 Manaus, AM, Brazil
 rajunior@uea.edu.br

Abstract—This full paper of the research category describes the ERPLab as an innovative educational platform that aims to promote inclusivity and diversity in programming and robotics. It provides a unique learning experience, particularly in remote or resource-limited areas, making scientific education more accessible to a broader audience. ERPLab offers students an understanding and integration of scientific knowledge by combining biology, engineering, and programming. ERPLab plays a significant role in educational technology research by exploring new possibilities for remote teaching and providing valuable data and insights to evaluate the effectiveness of this approach. One of the standout features of ERPLab is its use of an automated greenhouse controlled by a 3D printer-produced robotic arm equipped with sensors, and two Wi-Fi cameras, esp32cam, facilitating observation from multiple angles, allowing students to explore new avenues of scientific learning. This system enables students to learn about plant cultivation while developing their programming and technological skills. ERPLab combines advanced technology with natural repellent plants to promote innovative teaching and research methods. The plant selection within the ERPLab includes natural repellents like rosemary (*Rosmarinus officinalis*), rue (*Ruta graveolens*), and basil (*Ocimum basilicum*). These essential programming skills are directly applicable to the rapidly advancing INDUSTRY 4.0. Some preliminary work was carried out in a school environment, and the results of feedback from some students and teachers were promising regarding the proposed approach. For future work, we will carry out qualitative and quantitative research to observe the interest and development of users for the program of the platform in the context of regular classes.

Index Terms—Remote Laboratory, Educational Robotics, Educational Technologies, Programming Learning, Virtual Laboratory

Robotics and Programming Teaching Laboratory called ERPLab.

ERPLab (Environment-Robotic-Programmin-Laboratory) is an educational platform that promotes inclusion and diversity in programming and robotics. The platform offers learning opportunities in remote areas or with limited resources, making scientific education more accessible and democratized [8]–[10]. ERPLab combines biology, engineering, and programming, promoting a broader and integrated understanding of scientific knowledge.

ERPLab was an automated greenhouse controlled by a 3D-printed robotic arm with a sensor that allowed students to learn about plant cultivation and care while developing programming and technology skills. They are encouraged to understand the relationship between soil, water, and nutrients, identify planting, fertilization, and harvesting processes, and develop essential programming skills for INDUSTRY 4.0 [11]–[15]. The plants used in ERPLab are natural repellents, including rosemary (*Rosmarinus officinalis*), rue (*Ruta graveolens*), and basil (*Ocimum basilicum*).

ERPLab contributes to research in educational technology by exploring new possibilities for remote teaching and providing data and information to evaluate the effectiveness of this approach. With the addition of a 3D-printed robotic arm, it can represent a significant technological contribution to ERPLab and its future projects [8], [16]–[22]. Using two Wi-Fi ESP32CAM cameras that facilitate observation from various angles, students can explore new learning possibilities in the scientific field.

Although there are challenges in fields of study that require students to learn through practical experiences [23]–[25], ERPLab has shown promising results in overcoming challenges such as lack of equipment and physical space in schools and limited access to learning opportunities in remote areas. The combination of cutting-edge technology and natural repellents in plants cultivated in ERPLab is a way to encourage the

Apêndice E - E-mail de aceite do quinto artigo: Visualização de Dados Educacionais: Desenvolvendo uma Ferramenta para Análise de Dados de Sistemas Tutores Inteligentes.



Universidade
Federal do
Amazonas

Matheus Menezes <matheus.menezes@icomp.ufam.edu.br>

Your RENOTE-2024-1 paper 241042

1 mensagem

JEMS <jems@sbc.org.br>

27 de junho de 2024 às 11:36

Responder a: liane@penta.ufrgs.br

Para: matheus.menezes@icomp.ufam.edu.br

Cc: José Francisco Netto <jfmnetto@gmail.com>, "Arcanjo Miguel M. Lopes" <arcanjolopes@gmail.com>

Prezado(a) Matheus de Menezes:

Parabéns! Seu artigo "Visualização de Dados Educacionais: Desenvolvendo uma Ferramenta para Análise de Dados de Sistemas Tutores Inteligentes" foi aprovado para publicação na revista Revista Novas Tecnologias na Educação 2024-1 V22 N1. Nesta edição foram submetidos 176 artigos e a taxa de aprovação foi de 36%.

As revisões encontram-se abaixo ou podem ser encontradas em
<https://submissoes.sbc.org.br/PaperShow.cgi?m=241042>

A versão final, devidamente identificada, e com as correções indicadas no processo de revisão efetuadas, deve ser feita até a data de 15 de julho de 2024. O mesmo processo de submissão, via JEMS, será usado para envio do arquivo pdf com a versão final.

A versão final deverá conter os dados dos autores, sua afiliação e ORCID ID (por extenso) bem como atentar para as orientações dos revisores e para a correta formatação do artigo. Neste sentido é imprescindível atentar para as Orientações para os Autores contidas em
<http://penta3.ufrgs.br/RENOTE/RENOTE-2022-2/normaspalestras.html>.

Artigos que não estiverem de acordo com esta orientação em sua versão final não serão publicados.

Cordiais saudações,
Eliseo Reatgeui e Liane Tarouco
Editores
RENOTE - Revista Novas Tecnologias na Educação
CINTED/UFRGS

B

ANEXOS

Termo de Consentimento Livre e Esclarecido aplicado na condução do teste de usabilidade.

Termo de Consentimento Livre e Esclarecido

Prezado (a) Professor (a),

Você está sendo convidado(a) a participar da Pesquisa de Avaliação de Usabilidade de um dashboard educacional voltado para sistemas tutores inteligentes, coordenada pelo pesquisador Matheus Freitas de Menezes. Este estudo visa compreender a interação e percepção dos usuários em relação à interface do dashboard, visando melhorias para a comunidade educacional.

Sua participação é voluntária e será realizada por meio de um teste de usabilidade conduzido via questionário online. Suas respostas poderão contribuir significativamente para o desenvolvimento da ferramenta educacional. Você tem o direito de retirar seu consentimento a qualquer momento, sem consequências. Não haverá custos ou remuneração associados a esta pesquisa.

Seus dados serão tratados com confidencialidade e sua identidade será mantida em sigilo. Para mais informações, entre em contato com o pesquisador pelo e-mail: matheus.menezes@icomp.ufam.edu.br ou pelo Whatsapp: 92991721293. Selezionando o botão próxima, estará concordando em participar da pesquisa e sabendo que poderá sair quando quiser.