

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
CENTRO INTERDISCIPLINAR DE NOVAS TECNOLOGIAS NA EDUCAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA NA EDUCAÇÃO

IGOR KÜHN

**Identificação de Trajetórias de Aprendizagem com o uso de  
Grafos Direcionados e Técnicas de Mineração de Dados  
Visando a Detecção de Evasão em Cursos EAD**

Tese apresentada ao Programa de Pós-Graduação em Informática na Educação (PPGIE/UFRGS) como requisito parcial para a obtenção do grau de Doutor em Informática na Educação.

Orientador: Prof. Dr. Leandro Krug Wives

Porto Alegre, Janeiro 2023.

## CIP - Catalogação na Publicação

kühn, Igor

Identificação de Trajetórias de Aprendizagem com o uso de Grafos Direcionados e Técnicas de Mineração de Dados Visando a Detecção de Evasão em Cursos EAD / Igor kühn. -- 2023.

163 f.

Orientador: Leandro Krug Wives.

Tese (Doutorado) -- Universidade Federal do Rio Grande do Sul, Centro de Estudos Interdisciplinares em Novas Tecnologias na Educação, Programa de Pós-Graduação em Informática na Educação, Porto Alegre, BR-RS, 2023.

1. Trajetórias de Aprendizagem. 2. Mineração de Dados. 3. Representação em Grafos. 4. Educação a Distância. 5. Ambientes Virtuais de Ensino e Aprendizagem. I. Krug Wives, Leandro, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos André Bulhões

Vice-Reitora: Profa. Patricia Pranke

Pró-Reitor de Pós-Graduação: Prof. Júlio Otávio Jardim Barcellos

Coordenador PPGIE: Prof. Dante Augusto Couto Barone

Diretor Cinted: Marcus Vinicius de Azevedo Basso

*“Nossas melhores ideias são geralmente aquelas que unem dois mundos distintos.”*

Marvin Minsky

*“Escrever é fácil. Você começa com uma letra maiúscula e termina com um ponto final, no meio você coloca as ideias.”*

Pablo Neruda

## **Imagem**

*Substantivo feminino , representação, reprodução ou imitação de uma pessoa ou objeto.  
Aspecto particular pela qual um ser ou um objeto é percebido; cena, quadro.*

Oxford Languages

## RESUMO

A expansão das Tecnologias de Informação e Comunicação (TICs) mudou as formas de interação e socialização entre as pessoas e o ambiente em que estão inseridas. As instituições de ensino também receberam forte influência de tais mudanças, necessitando adequar-se a essa nova realidade, em especial no que diz respeito à Educação a Distância. O recente crescimento dessa modalidade está baseado na intensificação do uso de TICs e tem levado pesquisadores a buscarem novas soluções para lacunas relativas ao ensino. Neste estudo são propostas soluções para tais hiatos que podem levar à evasão escolar. Para tanto, sugere-se o uso de Mineração de Dados e representações de Trajetórias de Aprendizagem (TAs), em especial através de representações visuais conhecidas como Grafos. A Mineração de Dados (MD) é empregada para que, com o uso de dados gerados pelas interações na plataforma Moodle, seja possível obter informações úteis que levem à identificação do perfil de alunos ou grupo de alunos, de acordo com as características apresentadas ao longo das TAs percorridas. As representações visuais foram utilizadas com o objetivo de identificar comportamentos que possam levar à evasão de alunos em cursos na modalidade a distância. Em especial, por meio de uma ferramenta de visualização desenvolvida especificamente para esta tese, foi possível visualizar as interações de um conjunto de alunos com os recursos presentes no Moodle. A ferramenta foi testada e avaliada por um conjunto de professores por meio de um questionário modelo TAM. Como resultado foi possível observar padrões específicos de interações de acordo com cada grupo de alunos agrupados de acordo com suas notas. Por último, houve um retorno positivo quanto a percepção dos participantes da avaliação da ferramenta e de seus resultados quanto a sua utilidade em tarefas docentes. O presente trabalho possui contribuições para as áreas da Educação e Computação, além de contribuição social, propondo uma nova ferramenta computacional que apresenta facilidades quanto ao seu uso por parte de professores e tutores de cursos EAD, possibilitando a visualização das interações dos alunos com recursos educacionais no ambiente Moodle, podendo ser utilizada para detecção precoce de alunos com propensão a evasão, contribuindo assim para um melhor acompanhamento dos estudantes quando utilizam a plataforma Moodle.

**Palavras-chave:** Trajetórias de Aprendizagem. Mineração de Dados. Representação em Grafos. Educação a Distância. Ambientes Virtuais de Ensino e Aprendizagem.

# **Analysis of Different Learning Trajectories using Data Mining Techniques and Directed Graphs to Support the Identification of Evasion on Distance Learning Courses**

## **ABSTRACT**

The expansion of Information and Communication Technologies (ICTs) has changed the forms of interaction and socialization between people and the environment in which they are inserted. Teaching institutions were also strongly influenced by such changes, needing to adapt to this new reality, especially regarding Distance Education. The recent growth of this modality is based on the intensification of the use of ICTs and has led researchers to seek new solutions for gaps related to teaching. This study proposes solutions for such gaps that can lead to school dropout. For that, it is suggested the use of Data Mining and representations of Learning Paths (LPs), especially through visual representations known as Graphs. Data Mining (DM) is used so that, with the use of data generated by interactions on the Moodle platform, it is possible to obtain useful information that leads to the identification of the profile of students or groups of students, according to the characteristics presented throughout of the LPs traversed. The visual representations were used with the aim of identifying behaviors that could lead to dropout of students in distance learning courses. In particular, through a visualization tool developed specifically for this thesis, it was possible to visualize the interactions of a group of students with the resources present in Moodle. The tool was tested and evaluated by a group of teachers through a TAM model questionnaire. As a result, it was possible to observe specific patterns of interactions according to each group of students grouped according to their grades. Finally, there was positive feedback regarding the perception of the participants in the evaluation of the tool, and its results, regarding its usefulness in their teaching tasks. This work has contributions to the areas of Education and Computing, besides Social contributions, proposing a new computational tool that is of easy use by teachers, tutors of distance learning courses, enabling the visualization of student interactions with educational resources in the Moodle environment. Moreover, the tool can be used for early detection of students prone to dropout, contributing to a better monitoring of students when using the Moodle platform.

**Keywords:** Learning Trajectories. Data Mining. Directed Graphs. Distance Learning. Virtual Teaching and Learning Environments.

## LISTA DE FIGURAS

Figura 1 – Número de Matrículas no EAD entre 2009 - 2018.....	14
Figura 2 – Número de Ingressos em Cursos de Graduação por Modalidade de Ensino entre 2011-2021 .....	15
Figura 3 – Taxa de Evasão cursos Superiores EAD-Brasil .....	16
Figura 4 – Gerações da Educação a Distância .....	22
Figura 5 – Diagrama Efeitos da Reprovação .....	25
Figura 6 – Interações Educacionais .....	32
Figura 7 – Níveis de Análise das TAs .....	<del>37</del> <sup>36</sup>
Figura 8 – Modelo de Trajetória .....	<del>38</del> <sup>37</sup>
Figura 9 – Modelo de Trajetória de Aprendizagem Conceitual.....	<del>39</del> <sup>38</sup>
Figura 10 – Interação entre aluno e recursos da TA .....	<del>40</del> <sup>39</sup>
Figura 11 – Hierarquia dos Elementos de uma TA.....	<del>41</del> <sup>40</sup>
Figura 12 – Modelo de Trilha de Aprendizagem.....	<del>41</del> <sup>40</sup>
Figura 13 – Estruturação Visual de Dados.....	<del>43</del> <sup>42</sup>
Figura 14 – Exemplos de Grafos .....	<del>45</del> <sup>44</sup>
Figura 15 – Exemplo de grafo bipartido .....	<del>49</del> <sup>48</sup>
Figura 16 – Exemplo de grafo árvore.....	<del>50</del> <sup>49</sup>
Figura 17– Pirâmide DIKW .....	<del>52</del> <sup>51</sup>
Figura 18 – Processo de DCBD .....	<del>54</del> <sup>53</sup>
Figura 19 – Exemplo de Validação Cruzada .....	<del>58</del> <sup>57</sup>
Figura 20 – Etapas da Revisão Sistemática.....	<del>61</del> <sup>60</sup>
Figura 21 – Número de trabalhos por Biblioteca .....	<del>62</del> <sup>61</sup>
Figura 22 – Número de trabalhos por ano.....	<del>63</del> <sup>62</sup>
Figura 23 – Nuvem de PalavrascChave 01 .....	<del>67</del> <sup>66</sup>
Figura 24 – Nuvem de Palavras-Chave 02 .....	<del>68</del> <sup>67</sup>
Figura 25 – Sequência de OA.....	<del>70</del> <sup>69</sup>
Figura 26 – Representação dos caminhos percorridos pelos alunos .....	<del>70</del> <sup>69</sup>
Figura 27 – Representação dos acessos às atividades .....	<del>72</del> <sup>71</sup>
Figura 28 – Painel de controle das atividades.....	<del>73</del> <sup>72</sup>
Figura 29 – Estruturação da Pesquisa .....	74
Figura 30 – Log de atividade do Curso.....	79
Figura 31 – Notas dos alunos ao longo do semestre.....	80
Figura 32 – Tela geral MDV.....	82
Figura 33 – Modo de visualização Árvore.....	83
Figura 34 – Previsão de desempenho.....	84

Figura 35 – Gráfico diagrama de caixa.....	87
Figura 36 – Gráfico de dispersão 01.....	88
Figura 37 – Clusters notas da disciplina Derad 109.....	9294
Figura 38 – Interações alunos extrato superior Derad 109.....	89
Figura 39 – Interações alunos extrato médio Derad 109.....	90
Figura 40 – Interações alunos extrato inferior Derad 109.....	91
Figura 41 – Gráfico de dispersão 02.....	92
Figura 42 – Clusters notas da disciplina Derad 202.....	93
Figura 43 – Interações alunos extrato superior Derad 202.....	93
Figura 44 – Interações alunos extrato médio Derad 202.....	94
Figura 45 – Interações alunos extrato inferior Derad.....	95
Figura 46 – Gráfico de dispersão 03.....	96
Figura 47 – Clusters notas da disciplina Derad 501.....	96
Figura 48 – Interações alunos extrato superior Derad 501.....	97
Figura 49 – Interações alunos extrato médio Derad 501.....	97
Figura 50 – Interações alunos extrato inferior Derad 501.....	98
Figura 51 – Gráfico dispersão 04.....	99
Figura 52 – Clusters notas da disciplina Derad 502.....	99
Figura 53 – Interações alunos extrato superior Derad 502.....	100
Figura 54 – Interações alunos extrato médio Derad 502.....	101
Figura 55 – Interações alunos extrato inferior Derad 502.....	102
Figura 56 – Modelo TAM.....	104
Figura 57 – Ferramenta comparada a outras ferramentas.....	106
Figura 58 – Ferramenta em relação ao benefício.....	107
Figura 59 – Ferramenta facilidade para o trabalho.....	108
Figura 60 – Utilidade da ferramenta no trabalho.....	108
Figura 61 – Facilidade de operação da ferramenta.....	109
Figura 62 – Facilidade de execução.....	110
Figura 63 – Interação com a ferramenta.....	110
Figura 64 – Flexibilidade e interatividade da ferramenta.....	111
Figura 65 – Habilidade para utilizar a ferramenta.....	112
Figura 66 – Elementos visuais da ferramenta.....	116
Figura 67 – Análise e compreensão das interações.....	113
Figura 68 – Seleção de turmas.....	114
Figura 69 – Função predição.....	115
Figura 70 – Opções para visualização.....	115
Figura 71 – Seleção de notas.....	116

## LISTA DE QUADROS

Quadro 1 – Evolução Histórica da EAD no Brasil.....	23
Quadro 2 – Trabalhos Seleccionados.....	<del>64</del> 63
Quadro 3 – Trabalhos por Biblioteca Pós Filtro.....	<del>68</del> 67
Quadro 4 – Síntese de sugestões e críticas a ferramenta.....	<del>122</del> 124
Quadro 5 – Síntese de sugestões e críticas ao manual .....	<del>125</del> 124



## LISTA DE ABREVIATURAS E SIGLAS

AVA	Ambiente Virtual de Aprendizagem
AVEAs	Ambientes Virtuais de Ensino e Aprendizagem
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
DCBD	Descoberta de Conhecimento em Base de Dados
DDBMS	Sistema de Gerenciamento de Banco de Dados Distribuídos
EAD	Educação a Distância
GPS	Sistema de Posicionamento Global
WEKA	<i>Waikato Environment for Knowledge Analysis</i>
MD	Mineração de Dados
MDE	Mineração de Dados Educacionais
MOODLE	<i>Modular Object-Oriented Dynamic Learning</i>
MDV	<i>Moodle Data Visualization</i>
PLAGEDER	Curso Superior em Desenvolvimento Rural
TA	Trajatória de Aprendizagem
TAM	<i>Technology Acceptance Model</i>
TAs	Trajórias de Aprendizagem
TDT	Teoria da Distância Transacional
TI	Tecnologia da Informação
TICs	Tecnologias da Informação e Comunicação
TRA	Teoria da Ação Raciocinada
UAB	Universidade Aberta do Brasil
UCI	<i>Machine Learning Repository</i>
UFRGS	Universidade Federal do Rio Grande do Sul

## SUMÁRIO

<b>RESUMO.....</b>	<b><u>32</u></b>
<b>ABSTRACT .....</b>	<b><u>43</u></b>
<b>LISTA DE FIGURAS .....</b>	<b><u>54</u></b>
<b>LISTA DE QUADROS .....</b>	<b><u>76</u></b>
<b>LISTA DE ABREVIATURAS E SIGLAS.....</b>	<b><u>87</u></b>
<b>1 INTRODUÇÃO.....</b>	<b>11</b>
1.1 QUESTÃO DE PESQUISA .....	18
1.2 OBJETIVO GERAL.....	18
1.3 OBJETIVOS ESPECÍFICOS.....	18
1.4 ESTRUTURA DO DOCUMENTO .....	19
<b>2. FUNDAMENTAÇÃO .....</b>	<b>21</b>
2.1 EDUCAÇÃO A DISTÂNCIA.....	21
2.2 EVASÃO NA MODALIDADE EAD.....	24
2.3 APRENDIZAGEM .....	26
2.4 TEORIA DA DISTÂNCIA TRANSACIONAL .....	28
2.5 TRAJETÓRIAS DE APRENDIZAGEM .....	33
2.5.1 Níveis de Análise de Trajetórias de Aprendizagem .....	<u>37</u> <u>36</u>
<b>3. REPRESENTAÇÕES VISUAIS .....</b>	<b><u>42</u><u>41</u></b>
3.1 TEORIA DOS GRAFOS .....	<u>43</u> <u>42</u>
3.1.1 Classificação dos Grafos .....	<u>45</u> <u>44</u>
3.1.2 Classificação quanto à Centralidade .....	<u>47</u> <u>46</u>
3.2 Visualização em grafos .....	<u>49</u> <u>48</u>
<b>4. MINERAÇÃO DE DADOS.....</b>	<b><u>52</u><u>51</u></b>
4.1 PROCESSO DE DESCOBERTA DE CONHECIMENTO EM BASE DE DADOS .....	<u>54</u> <u>53</u>
4.1.1 Seleção de dados .....	<u>54</u> <u>53</u>
4.1.2 Pré- processamento .....	<u>55</u> <u>54</u>
4.1.3 Transformação .....	<u>55</u> <u>54</u>
4.1.4 Mineração.....	<u>55</u> <u>54</u>
4.1.5 Análise/Avaliação.....	<u>55</u> <u>54</u>

4.2 TAREFAS DE MINERAÇÃO DE DADOS .....	<u>5655</u>
4.2.1 A Validação dos Dados.....	<u>5756</u>
4.3 MINERAÇÃO DE DADOS EDUCACIONAIS .....	<u>5857</u>
<b>5. TRABALHOS RELACIONADOS .....</b>	<b><u>6059</u></b>
5.1 REVISÃO SISTEMÁTICA .....	<u>6059</u>
<b>6. METODOLOGIA.....</b>	<b><u>7574</u></b>
6.1 ETAPA 1 .....	<u>7675</u>
6.2 ETAPA 2 .....	<u>7776</u>
6.3 ETAPA 3 .....	<u>7776</u>
6.4 ETAPA 4 .....	<u>7776</u>
6.5 ETAPA 5 .....	<u>7877</u>
6.6 RESULTADOS PARCIAIS.....	<u>7877</u>
6.7 PLATAFORMA MOODLE .....	80
<b>7 A FERRAMENTA DATA MOODLE VISUALIZATION (MDV) .....</b>	<b>84</b>
7.1 DESCRIÇÃO DA FERRAMENTA .....	84
<b>8 EXPERIMENTOS DE AVALIAÇÃO E VALIDAÇÃO.....</b>	<b><u>8887</u></b>
8.1 UTILIZAÇÃO DA MD E DA FERRAMENTA MDV PARA A ANÁLISE DAS DISCIPLINAS SELECIONADAS .....	<u>8988</u>
8.2 AVALIAÇÃO DA FERRAMENTA UTILIZANDO O MODELO TAM .....	<u>105104</u>
8.2.1 Análise e discussão dos dados do questionário.....	<u>108107</u>
8.2.2 Questões fechadas .....	<u>109108</u>
8.2.2.1 Percepção do usuário quanto a aplicabilidade da ferramenta.....	<u>109108</u>
8.2.2.2 Percepção do usuário quanto a utilização da ferramenta.....	<u>112111</u>
8.2.2.3 Percepção do usuário quanto as funcionalidades específicas .....	<u>116115</u>
8.3 RESPOSTAS RELATIVAS ÀS QUESTÕES ABERTAS .....	<u>119118</u>
<b>9 RESULTADOS E DISCUSSÕES .....</b>	<b><u>127126</u></b>
<b>10 CONCLUSÃO .....</b>	<b>126</b>
<b>REFERÊNCIAS.....</b>	<b><u>135134</u></b>
<b>ANEXO A - IMAGEM DA TELA DA FERRAMENTA NEO4J.....</b>	<b><u>145144</u></b>
<b>ANEXO B - IMAGEM DA TELA DA FERRAMENTA NOTEBOOK AZURE MICROSOFT.....</b>	<b><u>146145</u></b>

<b>ANEXO C - IMAGEM FERRAMENTA WEKA.....</b>	<b><u>147146</u></b>
<b>ANEXO D- REPORTAGEM JORNAL CORREIO DO POVO E SITE G1 .....</b>	<b><u>148147</u></b>
<b>ANEXO E - RUN INFORMATION .....</b>	<b>144</b>
<b>ANEXO F – MANUAL USUÁRIO – MOODLE DATA VISUALIZATION.....</b>	<b><u>149148</u></b>
<b>APÊNDICE 1 – QUESTIONÁRIO TAM APLICADO COM DOCENTES .....</b>	<b><u>154153</u></b>

## 1 INTRODUÇÃO

A Humanidade passa por constantes transformações. Em especial, as décadas mais recentes foram marcadas por profundas mudanças, entre as quais se destaca o desenvolvimento tecnológico das comunicações. Como fruto dessas mudanças, é possível verificar que o processo de construção de instrumentos tecnológicos amplia a capacidade de interação e socialização entre as pessoas e o ambiente em que estão inseridas, sejam estes físicos ou virtuais. Segundo Castells (2013, p.58), “As redes interativas de computadores estão crescendo exponencialmente, criando novas formas e canais de comunicação, moldando a vida e, ao mesmo tempo, sendo moldadas por elas”.

Nas universidades, escolas e no mundo corporativo, o acesso facilitado e instantâneo a um conjunto cada vez maior de informações altera significativamente as relações sociais. A partir da década de 1970, as Tecnologias da Informação (TIs), identificada atualmente como Tecnologias de Informação e Comunicação (TICs), correspondem a uma das áreas tecnológicas que mais evoluíram, tendo em vista expansão e popularização da Internet. Ainda sobre o tema, Castells (2013, p.58) afirma que:

[...] o primeiro computador programável e o transistor, fonte da microeletrônica, o verdadeiro cerne da Revolução da Tecnologia da Informação no século XX. Porém, defendo que de fato, só na década de 70 as novas tecnologias da informação difundiram-se amplamente, acelerando seu desenvolvimento sinérgico e convergindo em um novo paradigma.

Essa mudança de paradigma continua afetando de modo intenso a vida em sociedade, de forma mais específica, dos indivíduos que passaram de uma longa era analógica para um novo universo digital e, em decorrência da intensificação do uso das TICs, mudanças sociais, culturais e econômicas passaram a ocorrer de forma mais rápida e intensa. Os processos de ensino e aprendizado não estão à parte dessa mudança, tais processos também estão sendo modificados ao serem introduzidos em ambientes escolares elementos tecnológicos como computadores, *tablets* e *smartphones*, os quais até algumas décadas não faziam parte desse cenário. Em decorrência da pandemia, causada pelo Coronavírus Covid-19, tais tecnologias se mostram fundamentais, pois, escolas, universidades e empresas passaram a utilizar a forma de Ensino Remoto Emergencial (ERE) e o trabalho remoto como possível

solução para que de alguma forma as aulas e o trabalho nas instituições não fossem completamente suspensos.

Segundo Perkins e Solomon (2005, p.71), os efeitos do uso da tecnologia podem ser observados a partir de três aspectos: efeitos **com** a tecnologia; efeitos **da** tecnologia e efeitos **através** da tecnologia. Os efeitos cognitivos com o uso da tecnologia ocorrem quando certas funções intelectuais são transferidas para a tecnologia, configurando uma união com o usuário. Tal relação é caracterizada por oferecer um suporte cognitivo às atividades como ortografia, cálculo e medida. Tais atividades poderiam ser realizadas sem a o uso da tecnologia, mas seu uso melhora de forma significativa a *performance cognitiva*<sup>1</sup>.

Ainda segundo os mesmos autores, existem efeitos da tecnologia que permanecem quando se interrompe o seu uso após um período de uso contínuo. A ideia em relação a esse aspecto é que o uso da tecnologia se assemelha ao uso de uma ferramenta, e os efeitos em relação a esse aspecto sugerem processos cognitivos de internalização. Exemplos desse tipo de uso são os diferentes simuladores como os de voo e bonecos utilizados para aulas de primeiros socorros.

As novas tecnologias de forma qualitativa muitas vezes remodelam profundamente e reorganizam sistemas, ou seja, mudam as atividades de modo que sem a tecnologia tais atividades não são mais possíveis de serem realizadas. Os recursos digitais para criação de vídeos, os editores de textos e o uso dos dispositivos móveis são exemplos desse tipo de efeito.

Os três tipos de efeitos causados pelo uso da tecnologia apresentados podem ser encontrados quando analisamos os Ambientes Virtuais de Ensino e Aprendizado (AVEAs), pois as tarefas realizadas nesse tipo de plataforma são realizadas com o auxílio da tecnologia e o uso desses ambientes também pode gerar habilidades que permanecem com os usuários mesmo após seu uso. Por último, a forma como se ensina e aprende nesse tipo de sala de aula é algo inovador e que muda de forma significativa as relações, seja entre os sujeitos envolvidos, seja entre os sujeitos e os recursos educacionais. Ainda sobre tais mudanças, acredita-se que a mais significativa delas seja a relação entre espaço e tempo em que as relações professor-alunos são realizadas, pois podem ser ampliadas quando comparadas a aulas

---

<sup>1</sup> Entendemos a performance cognitiva ou desempenho cognitivo como o funcionamento ideal e eficaz das capacidades das funções executivas e capacidades cognitivas de cada sujeito.

tradicionais.

Em termos de disponibilização dos recursos educacionais, a forma como esses são apresentados aos alunos também está sendo modificada. Tal mudança apresenta diversas possibilidades, dentre elas uma consiste em adaptar materiais didáticos já existentes para o formato digital e a forma mais atualizada, sendo a criação de materiais didáticos digitais que pode ser realizada por redes, como por exemplo a Rede Interativa Virtual de Educação (RIVED). Essa rede que é sediada na Secretaria de Educação a Distância (SEED), sendo composta por um conjunto de instituições que criam e disponibilizam seus materiais didáticos.

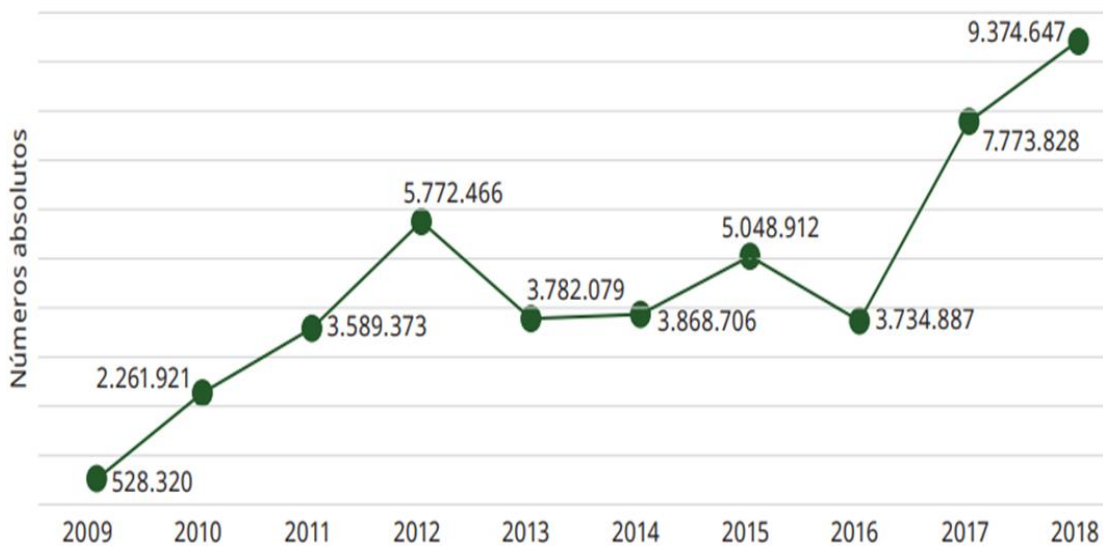
Essa transposição apresenta desafios não somente em sua forma, mas principalmente na maneira como o conteúdo é apresentado, acessado e apropriado pelo aluno. Essa diversidade de recursos digitais, por sua vez, oportuniza uma maior possibilidade do uso personalizado do conteúdo didático, oferecendo assim experiências únicas de percurso em ambientes digitais de aprendizado. O uso dos AVEAs amplia as possibilidades de interações, sejam elas professor-aluno, aluno-aluno ou aluno-conteúdos, possibilitando adaptações dos conteúdos em relação ao perfil de aprendizado dos alunos. O uso de computadores, bem como, o de dispositivos móveis como ferramentas de acesso e uso dos AVEAs têm possibilitado a obtenção em grande escala de dados gerados a partir desses ambientes. Segundo Mostow e Beck (2005, p.15), tais sistemas podem acumular extensa quantidade de dados que são extremamente valiosos para a análise do comportamento dos alunos. Ainda de acordo com Zorrilla et al. (2005), essa nova fonte de informação, onde professores/tutores, passam a poder orientar de forma mais eficiente, levando em consideração o ponto de vista de um conjunto de alunos ou de um determinado aluno.

Com a possível utilização de tais dados, torna-se viável identificar Trajetórias de Aprendizagem (TAs), que, devido às diferentes possibilidades de interação com recursos, tarefas, colegas, professores e tutores, tornam-se únicas ou com características comuns para grupos de alunos. A Educação a Distância (EAD) pode fazer uso de ferramentas de suporte para o ensino e aprendizado, e como mencionado anteriormente os AVEAs são parte fundamental para a viabilização dessa modalidade de ensino.

Não apenas no Brasil, mas em diversos países, com o desenvolvimento tecnológico e a disponibilização da Internet de banda larga, é possível identificar o crescimento de cursos e do processo de promoção de habilidades e competências

humanas por meio da EAD. A EAD é uma modalidade em expansão, principalmente na última década, conforme gráfico disponível na Figura 1, onde é possível verificar um crescimento de mais de 100% entre os anos 2016-2018<sup>2</sup>.

Figura 1– Número de Matrículas no EAD entre 2009 - 2018



Fonte: Censo EAD, Abed (2019).

Um dos motivos do crescimento dessa modalidade são suas características específicas quando comparadas com a educação presencial, sendo algumas delas: uma maior flexibilidade de horários, conciliação entre estudo e trabalho, a não necessidade de deslocamento até o curso, acesso a diferentes tipos de cursos além da possibilidade do desenvolvimento de um aprendizado autônomo (FIRAT, 2016, p. 191).

Como afirma Pontes (2012), a modalidade EAD vem ganhando cada vez mais espaço no cenário nacional como uma alternativa para evitar a desqualificação da força de estudo e a exclusão social de grande parcela da população.

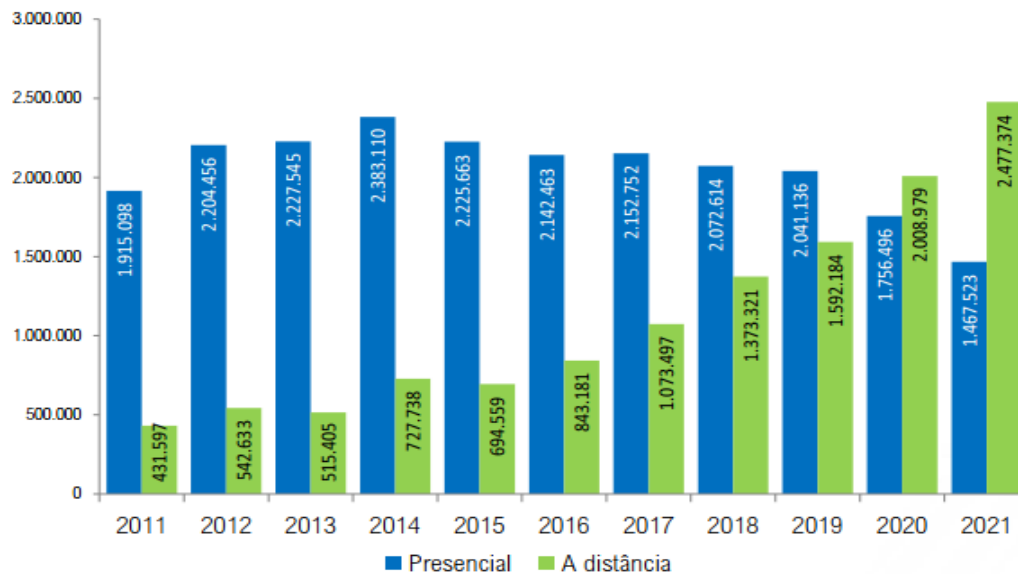
De acordo com o Censo da Educação Superior de 2019 (INEP, 2019), entre os anos de 2009 e 2019, a modalidade EAD apresentou variação percentual de ingressantes de 14,3% positivos. O ensino presencial por sua vez, a partir do ano de 2015, apresentou um decréscimo no número de matrículas realizadas. Dados do Censo da Educação Superior de 2021 (INEP, 2021) quando comparado aos cursos

<sup>2</sup> Abed a partir de 2019 modificou sua metodologia para a apresentação visual dos dados relativos ao censo.



presenciais entre os anos 2011-2021, a modalidade EAD apresentou crescimentos percentuais significativos (Figura 2), o que nos leva a acreditar que a modalidade de ensino selecionada para este estudo está em consonância com uma realidade de crescimento dos cursos ofertados no Brasil.

Figura 2 – Número de Ingressos em Cursos de Graduação por Modalidade de Ensino entre 2011-2021



Fonte: Censo INEP (2021).

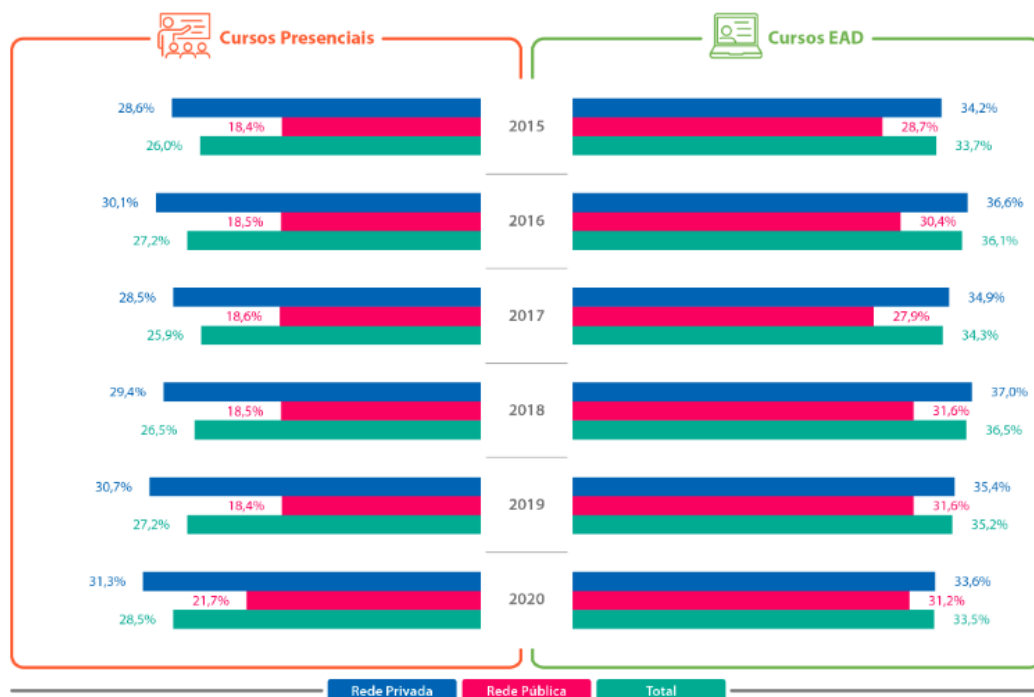
De forma similar à modalidade presencial, essa também apresenta fragilidades em relação ao ensino e aprendizado, dentre as fragilidades, podemos destacar os índices de desempenho nos cursos EAD. Conforme aponta Fornari (2010, p.4), baixo desempenho nas disciplinas, alto grau de reprovação e retenção são fatores significativos que podem levar à evasão em cursos EAD. Identificados tais problemas, o que se busca, por parte das instituições de ensino, além da expansão de seus cursos, é a melhora dos índices de desempenho dos alunos que revertam em um aumento dos níveis de permanência dos alunos em seus cursos.

Dessa forma, o acompanhamento e a análise contínuos do desempenho dos alunos são estratégias cada vez mais empregadas por instituições de ensino na busca pela garantia de bons níveis de aprendizagem, conforme os estudos de Detoni, Cechinel e Araújo (2015), Ramos (2017) e Queiroga (2018). Sendo assim, acredita-se que tais acompanhamentos e análises, com o uso de informações de qualidade possam ser uma possível solução para identificação de problemas como retenção,

abandono e evasão em cursos.

Em relação a evasão, elevados índices podem ser observados em cursos ofertados em universidades no Brasil e fora do país, por sua vez, a expansão do número de matrículas é acompanhada também pela elevação dos índices de alunos evadidos. Tais valores são apresentados pelo Mapa do Ensino Superior no Brasil (2022) e revelam que a evasão no ensino superior permanece em patamares elevados. Os percentuais mantêm uma tendência de crescimento que pode ser observado já no primeiro ano (2015) e chegando a valores percentuais de 33,6% (ensino privado) e 31,2% (ensino público) no ano de 2020, sendo a média percentual de 35,3% para o ensino privado e 30,2% para o ensino privado e ensino público, respectivamente. Isso nos leva a constatar que aproximadamente um a cada três alunos matriculados abandonará o curso.

Figura 3– Taxa de Evasão cursos Superiores EAD-Brasil



Fonte: Mapa do ensino superior do Brasil (2022, p.36).

O problema da evasão em cursos superiores não é algo recente e vem despertando o interesse de pesquisadores ao longo dos anos, como é possível verificar na subseção 2.2 do presente estudo. Em decorrência da relevância do problema, esse recebeu uma definição feita por Woodley e Simpson (2014, p.5) como sendo um “elefante invisível dentro da sala de aula”, visto que é um grande problema, significativo e muitas vezes as instituições fingem não enxergar.

Uma dentre as possíveis alternativas para a identificação de elementos que avaliem a probabilidade de evasão em cursos EAD é o uso da Mineração de Dados (MD), pois essa pode fornecer a partir de dados interacionais gerados em AVEAs, informações úteis e precisas para tomada de decisão por parte de professores, tutores e gestores de cursos.

Buscando-se cada vez mais o aproveitamento dos recursos tecnológicos para o combate ao “elefante invisível”, dessa forma é realizado neste estudo o uso de um conjunto de ferramentas de MD, bem como uma ferramenta de visualização de TAs por meio de grafos, tendo como origem dados da plataforma Moodle. Como conceito inicial entendemos um grafo como sendo uma representação visual constituída pelo conjunto de vértices e arestas onde é possível verificar as relações existentes entre os diferentes vértices por meio das suas conexões as arestas.

O conjunto de dados referido no parágrafo anterior foi utilizado para auxiliar professores e tutores usuários de AVEAs a identificar TAs de sucesso e não sucesso<sup>3</sup> para que essas sejam usadas como referência de percursos de aprendizagem a serem trilhados em tais ambientes. A identificação de tais TAs são realizadas por meio do uso de visualizações gráficas, sendo assim possível mapear e identificar as TAs na forma de grafos (grafos são melhor descritos na subseção 3.1). Entende-se que essa forma de representação possibilita uma análise mais adequada e completa das TAs, podendo ser realizado um melhor reconhecimento das diferenças e similaridades entre elas.

Em relação a modalidade de ensino utilizada neste estudo, é possível identificar que as instituições de ensino percebem a EAD como uma modalidade de ensino fundamental no processo educacional. Tal importância é demonstrada com a crescente oferta de cursos de graduação e especialização, fruto de investimento contínuo e substancial nessa modalidade de ensino, como pode ser verificado na Figura 2. Por sua vez, desde os anos 1990, a Universidade Federal do Rio Grande do Sul (UFRGS) empenha-se em promover um ambiente institucional adequado ao desenvolvimento de cursos EAD, seja por meio de recursos próprios ou de financiamento de projetos por meio de órgãos de fomento como a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Nesse sentido, foi possível identificar ambientes para a aplicação da pesquisa proposta neste documento. Dentre

---

<sup>3</sup> As trajetórias de sucesso entendemos como as trajetórias dos alunos que ao final da disciplina foram aprovados, e de insucesso as trajetórias relativas ao aluno não aprovados.

os cursos ofertados, encontram-se: cursos de graduação em Desenvolvimento Rural (PLAGEDER) em sua quarta edição, segunda edição em formato de Bacharelado e de pós-graduação totalmente EAD. O curso mencionado foi o curso selecionado para os experimentos realizados neste trabalho. A Universidade também proporciona o acesso universal a Cursos Abertos Online e Massivos, i.e., os MOOCS (*Massive Open On-Line Courses*), em seu repositório de cursos online gratuitos – a plataforma Lúmina.

### 1.1 QUESTÃO DE PESQUISA

Tendo com cenário os elementos mencionados anteriormente, delineou-se a seguinte questão de pesquisa: **Como detectar por meio de ferramentas tecnológicas a presença de fatores que determinem sucesso e insucesso nas trajetórias de aprendizagem, bem como elementos comuns e incomuns entre elas, e se tais fatores e elementos podem indicar alunos com provável propensão à evasão?**

### 1.2 OBJETIVO GERAL

O objetivo principal desta pesquisa é analisar, por meio de técnicas de Mineração de Dados e do uso de Grafos, diferentes Trajetórias de Aprendizagem, identificando as trajetórias de sucesso e de insucesso, possibilitando assim a detecção de elementos que indiquem probabilidade de evasão em um curso EAD.

### 1.3 OBJETIVOS ESPECÍFICOS

Tendo em vista o objetivo principal, objetivos específicos foram ponderados, seguem a seguir os objetivos específicos deste trabalho:

1. Buscar, em diferentes disciplinas ofertadas na modalidade EAD, *data sets* e a partir desses dados, utilizando técnicas de Mineração de Dados, identificar Trajetórias de Aprendizagem;
2. Identificar e agrupar, utilizando a técnicas de MD, conjuntos de alunos de acordo com seus desempenhos nas disciplinas;
3. Desenvolver em conjunto com alunos do curso de Ciências da Computação uma ferramenta para representação visual das trajetórias;
4. Construir visualmente, por meio do uso de grafos, relações entre os diferentes elementos (alunos e recursos) que constituem um Trajetória de Aprendizagem, indicando diferenças ou similaridades entre elas, bem como, a possível relação com o sucesso e insucesso.
5. Analisar e comparar Trajetórias de Aprendizagem de alunos em diferentes disciplinas do curso Plageder<sup>4</sup>, buscando identificar a existência de um ou mais perfis de alunos propensos à evasão de acordo com suas interações com os recursos educacionais disponibilizados.

#### **1.4 ESTRUTURA DO DOCUMENTO**

Para atingir os objetivos e responder à questão de pesquisa, a Tese está organizada da seguinte forma: o primeiro capítulo apresenta a introdução à tese inicialmente seguido pelo Capítulo 2 constituído pela fundamentação teórica, contemplando temas como educação a distância, evasão na EAD, aprendizagem, teoria da distância transacional, trajetórias de aprendizagem e níveis de análise em uma TA. No Capítulo 3, o tema desenvolvido são as representações visuais, teoria dos grafos e visualização em grafos. O Capítulo 4 trata sobre mineração de dados, processo de descoberta de conhecimento em base de dados, tarefas na MD e mineração de dados educacionais. O Capítulo 5 é destinado aos estudos relacionados a tese e apresenta a revisão bibliográfica realizada, bem como, os trabalhos resultantes da revisão que serviram como referência para a tese. No Capítulo 6, é apresentada a estrutura metodológica do trabalho e as etapas realizadas ao longo da construção da tese. O Capítulo 7 é destinado a apresentação da ferramenta MDV,

---

<sup>4</sup> Curso Superior em Desenvolvimento Rural/ UFRGS

bem como, testes e avaliações da ferramenta. O Capítulo 8 é destinado a apresentação dos experimentos de avaliação e validação. No Capítulo 9 por sua vez, são realizadas as apresentações e discussões dos resultados e o Capítulo 10 exhibe as conclusões do presente estudo.

## 2 FUNDAMENTAÇÃO

No desenvolvimento desta pesquisa, alguns conceitos serão utilizados visando estabelecer o embasamento teórico que contempla esta tese. Tais conceitos serão apresentados e discutidos neste capítulo.

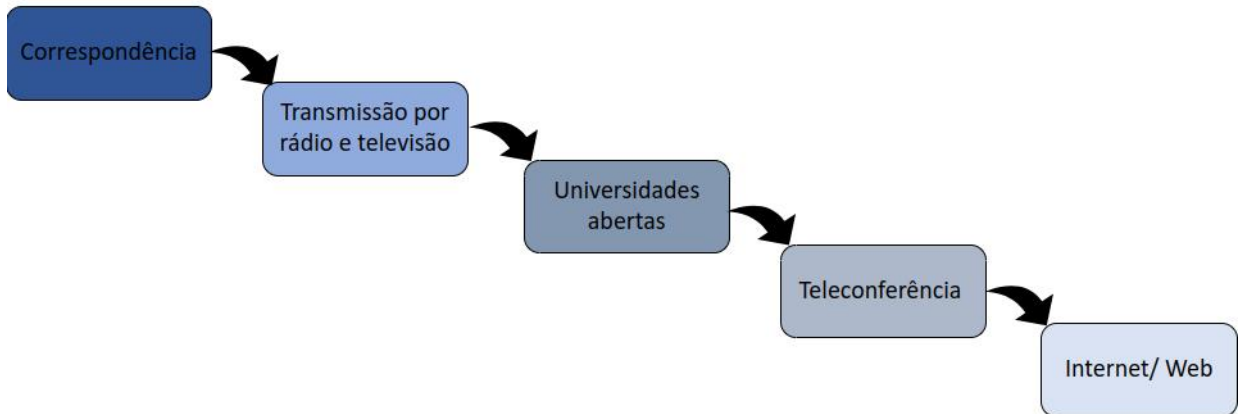
### 2.1 EDUCAÇÃO A DISTÂNCIA

A Educação a distância teve como origem os cursos por correspondência, sendo essa uma opção ou complemento à educação tradicional, como afirma Moore:

Educação a distância é o aprendizado planejado que ocorre normalmente em um lugar diferente do ensino, o que requer comunicação por meio de tecnologias e uma organização institucional especial (MOORE, 2007, p.25).

Moore identifica cinco diferentes gerações quanto ao desenvolvimento dessa modalidade e sua evolução, sempre apresentando uma relação estreita com a tecnologia (Figura 4). É possível ser verificado que tal relação é apresentada em seu primeiro momento, tendo o sistema de correspondência como o elemento tecnológico mediador entre aluno e curso, ou seja, a troca de informação por meio de cartas. Com o avanço dos anos e com o uso mais intenso das TICs, os processos de comunicação incorporaram mudanças passando da forma analógica para uma forma onde o caráter digital passou a responder por grande parte dos processos, a possibilidade da passagem de relações majoritariamente assíncronas para relações síncronas entre os sujeitos usuários da EAD.

Figura 4– Gerações da Educação a Distância



Fonte: Adaptado de Moore (2007).

Com o avanço e popularização das TICs surgiram os cursos à distância com o uso da Web, sendo esses ofertados em diferentes níveis de ensino. Por meio da convergência e potencialização das mídias utilizadas até então, se ampliou a possibilidade de interação entre os agentes envolvidos no processo de ensino e aprendizado.

Com as novas ferramentas e com o uso da Internet, um número maior de indivíduos pode compartilhar suas aprendizagens em tempo real, mantendo interações com o professor ou com outro estudante, independentemente do local onde estiver (UNESCO, 2002).

No Brasil, a educação a distância, segundo o Instituto de Pesquisas Avançadas em Educação (IPAE, 2007), data dos primórdios dos anos 1900, onde já eram ofertados cursos por correspondência. Essa modalidade em âmbito nacional apresenta etapas geracionais semelhantes ao quadro internacional: ensino por correspondência, transmissão por rádio e televisão, universidades abertas, teleconferência e internet/web, de acordo com a sequência observada na Figura 4. O Quadro 1 apresenta um resumo dos principais fatos relativos à evolução histórica da EAD no Brasil, a partir do ano de 1900.



Quadro 1– Evolução Histórica da EAD no Brasil

1900	No Rio de Janeiro Professores já ofereciam cursos profissionalizantes por correspondência para formação em datilografia.
1923	Henrique Morize e Roquete-Pinto lideraram um grupo que fundou a Rádio Sociedade do Rio de Janeiro, uma iniciativa privada que, então, deu início à educação pelo rádio.
1937	Teve origem o serviço de Radiodifusão Educativa do Ministério da Educação.
1939	A Marinha e o Exército do Brasil utilizavam a educação a distância para formar oficiais na Escola de Comando do Estado Maior.
1941	Foi criado o Instituto Universal Brasileiro, que foi o segundo instituto brasileiro a oferecer também cursos profissionalizantes.
1947	O Serviço Social do Comércio (SESC), o Serviço Nacional de Aprendizagem Comercial (SENAC) e as emissoras associadas criaram no Rio de Janeiro e em São Paulo, a Universidade do Ar oferecendo cursos radiofônicos comerciais.
1960	Primeira iniciativa governamental de apoio à Educação a Distância no Brasil com a criação, pelo Ministério da Educação e Cultura, do Programa Nacional de Teleducação (PRONTEL).
1967	Início das atividades do Instituto Brasileiro de Administração Municipal (IBAM), que oferecia cursos por correspondência com foco em políticas públicas e em instrumentos gerenciais.
1967	Criação do núcleo EAD da Fundação Padre Landell de Moura, ofertando ensino por correspondência e via rádio.
1967	Início do Projeto Mobral, vinculado ao governo federal, alfabetização de adultos com abrangência nacional com o uso do rádio.
1967	Projeto Satélite Avançado de Comunicações Interdisciplinares (SACI), criado pelo Instituto Nacional de Pesquisa Espacial (INPE) com o objetivo de criar um sistema nacional de telecomunicações via satélite.
1970	Projeto Minerva, programa de elaborado pelo governo federal e que teve por finalidade educar pessoas adultas.
1970	A Fundação Roberto Marinho lançou um programa de educação supletiva a distância, o programa evoluiu e tornando-se o Telecurso 2000(ainda em atividade).
1996	A EAD ascendeu a um novo <i>status</i> passando a ser reconhecida como uma modalidade de ensino oficial. Lei n. 9.394, de 20 de dezembro de 1996.
1998	São oficializados os cursos de graduação e pós-graduação. Decreto 2.561, de 10 de fevereiro de 1998, pelo decreto 2.561, de 27 de abril de 1998, e pela Portaria Ministerial 301 do Ministério da Educação (MEC), de 7 de abril de 1998.
2005	Publicação do Decreto Nº 5.622, de 19 de dezembro de 2005, que regulamentou o artigo 80 da Lei 9394/96 de Diretrizes e Bases da Educação Nacional (LDBEN), oficializando os cursos a distância no Brasil.
2006	Criação do sistema Universidade Aberta do Brasil (UAB) e oficialização pelo Ministério da Educação.
2007	Ministério da Educação elaborou uma nova proposta de EAD para ensino médio, o Sistema Escola Técnica Aberta do Brasil (E-TEC).
2008	É criada pelo Decreto nº 53.536, a Univesp (Universidade Virtual do Estado de São Paulo).

Fonte: Adaptado de Mattar (2011).

Como mais recente marco legal, podemos destacar o decreto Nº 9.057, de 25 de maio de 2017, que define EAD no Brasil:

Art. 1º Para os fins deste Decreto, considera-se educação a distância a modalidade educacional na qual a mediação didático-pedagógica nos processos de ensino e aprendizagem ocorra com a utilização de meios e tecnologias de informação e comunicação, com pessoal qualificado, com políticas de acesso, com acompanhamento e avaliação compatíveis, entre outros, e desenvolva atividades educativas por estudantes e profissionais da educação que estejam em lugares e tempos diversos (BRASIL, 2017).

Tais marcos relativos a EAD, como o citado, ajudam a reforçar essa modalidade de ensino, pois desde sua origem ela produziu certa desconfiança em termos de qualidade, quando comparada ao ensino tradicional presencial. Na subseção a seguir, o tema evasão é abordado de forma mais específica na modalidade de ensino EAD, sendo essa a modalidade de ensino uma realidade presente no meio educacional e como visto anteriormente neste trabalho cada vez mais as pessoas buscam sua formação por meio dela.

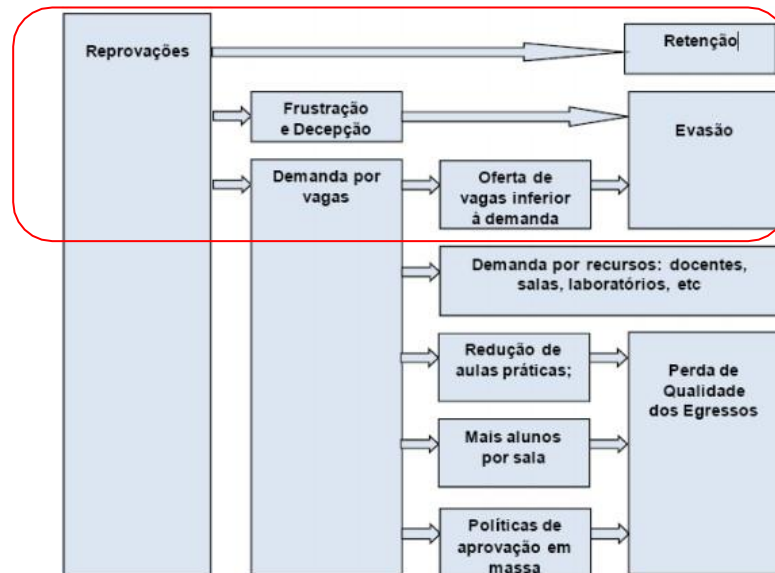
## **2.2 EVASÃO NA MODALIDADE EAD**

A evasão na modalidade de EAD, foi e continua sendo um problema para as instituições de ensino, e os números chamam a atenção devido aos seus valores elevados. Este estudo tem entre seus objetivos identificar elementos na modalidade EAD que contribuam para o aumento da evasão. A afirmação de Maluf (2012, p.1) chama a atenção para tal fenômeno: “[...] por mais atraente que seja o ambiente virtual, por mais agradáveis e motivadoras que sejam as aulas, os grupos de estudos e a equipe de tutores, ainda assim, com todos os estímulos considerados, encontraremos um significativo número de alunos evadidos”. Canto (2012) alerta para esse tipo de fenômeno em turmas do Curso de Engenharia Elétrica ofertadas na UFRGS, tendo as reprovações como fator diretamente relacionado à evasão dos alunos. Segundo o autor:

Elevados índices de reprovação podem implicar diversos resultados indesejáveis, tais como: o aumento da evasão, o aumento de demanda por recursos (docentes, salas, laboratórios, etc.), ou a adoção de políticas de aprovação em massa, que levaria à perda de qualidade dos egressos (CANTO,2012).

Tais relações são representadas pelo diagrama a seguir (Figura 5).

Figura 5– Diagrama Efeitos da Reprovação



Fonte: Canto (2012).

De acordo com estudos realizados na *Open University*<sup>5</sup> do Reino Unido por Woodley e Simpson (2014), o fenômeno da evasão foi identificado como “o elefante invisível dentro da sala de aula” devido a seu tamanho e a forma como é escondido e negligenciado. Também são evidenciados dados recentes que, em seus cursos de graduação na modalidade EAD, apenas 22% dos seus alunos os concluem. Em outra afirmação dos pesquisadores da mesma instituição, o problema é evidenciado:

[...] as instituições com essa modalidade (EAD) precisam reconhecer e enfrentar o problema, cutucando o elefante de qualquer maneira. É essencial compreender o perfil do aluno que evade, assim como os motivos da evasão (WOODLLEY; SIMPSON, 2014, p.12).

Ainda sobre o mesmo tema, Patto (1996) alerta que a evasão escolar é um problema antigo na educação, se estende ao longo dos tempos e prejudica o desenvolvimento de uma política educacional no país, traz perdas para as instituições e para os alunos, engloba aspectos sociais, comportamentais e financeiros.

Em relação a essa questão, Baggi e Lopes afirmam que:

[...] a evasão é um fenômeno social complexo, definido como interrupção no ciclo de estudos. É um problema que vem preocupando as instituições de

<sup>5</sup> O acesso ao material citado pode ser feito através de <http://www.open.ac.uk/>

ensino em geral, sejam públicas ou particulares, pois a saída de alunos provoca graves consequências sociais, acadêmicas e econômicas (BAGGI; LOPES, 2011, p.1).

Tendo como pano de fundo a realidade apresentada, somos instigados a nos dedicar a uma pesquisa mais aprofundada sobre o tema, a partir de afirmações como a seguinte:

Claramente, precisamos de muito mais pesquisas sobre o que acontece com alunos a distância. Mas, no momento, parece seguro afirmar que a porcentagem média de graduações em educação a distância é consideravelmente mais baixa do que na educação convencional (WOODLEY; SIMPSON, 2014).

Dois aspectos sobre o fenômeno da evasão são identificados por Rossi (2008) que, em seus estudos, aponta como principais causas da evasão em cursos no ensino superior ofertados no convênio com o sistema UAB: a falta de tempo e disposição para estudar após um dia de estudo. Ainda segundo Biazus (2004), na modalidade EAD alguns motivos da evasão são: a insatisfação com o tutor, as dificuldades em acessar o Ambiente Virtual de Aprendizagem (AVA), a falta de habilidade em utilizar a plataforma ou a escolha inadequada de ferramentas tecnológicas que pautam a EAD, elaboração incorreta do curso e a falta de tempo.

De posse de estudos já realizados e tendo conhecimento que as causas da evasão são multifatoriais nesse estudo é apresentada, testada e avaliada uma ferramenta que pode auxiliar na identificação de comportamento(s) que possam levar a evasão de alunos. Sendo proposto o uso da MD e do uso de grafos para que, a partir de um conjunto de dados, seja possível gerar informações úteis para identificação e visualização de TAs. No capítulo seguinte, a aprendizagem é melhor descrita, bem como, a maneira em que se relaciona com o estudo.

### **2.3 APRENDIZAGEM**

A aprendizagem se dá no decorrer de nossa existência por meio de processos que se dão de forma internalizada, bem como, nas interações dos indivíduos com o meio em que vivem. Segundo Illeris (2013, p.35):

A aprendizagem humana é a combinação de processos ao longo da vida, pelos quais a pessoa inteira – corpo (genético, físico e biológico) e mente (conhecimento, habilidades, atitudes, valores, emoções, crenças e sentidos) – experiência as situações sociais, cujo conteúdo percebido é transformado no sentido cognitivo, emotivo ou prático (ou por qualquer combinação) e integrado à biografia individual da pessoa, resultando em uma pessoa continuamente em mudança (ou mais experienciada).

Ainda, segundo Moran (2000, p.23), aprendemos quando interagimos com os outros e o mundo, e, depois, quando interiorizamos, voltamos para dentro, fazendo nossa própria síntese, nosso reencontro do mundo exterior com a nossa reelaboração pessoal.

Ao longo do tempo, diferentes teorias de aprendizagem foram construídas, elas são tentativas de sistematizar uma área de conhecimento humano que é a aprendizagem ou os processos educacionais. As teorias de aprendizagem, segundo Moreira (1999, p.13), são divididas em três grandes áreas ou correntes de pensamento: a Comportamentalista (behaviorismo), a Humanista e a Cognitivista (construtivismo).

Segundo Moreira (1999) e Ostermann (2011), a Teoria Comportamentalista, também conhecida como Teoria Behaviorista, está baseada nos comportamentos observáveis e mensuráveis dos sujeitos e suas respostas aos estímulos. Tal teoria surgiu no início do século XX na América do Norte em contraposição à Teoria Mentalista de origem europeia. Os estudos atuais em relação ao comportamentalismo estão direcionados a medir e avaliar as consequências posteriores aos estímulos, ou seja, é realizado um processo em ordem inversa quando se parte da resposta para se buscar a fonte do estímulo. A principal ideia desse modelo de estudo é dar continuidade a um estímulo, a partir de uma resposta positiva do sujeito. Os principais representantes do comportamentalismo são: Pavlov, Watson, Guthrie, Thorndike e Skinner.

Os mesmos autores, Moreira (1999) e Ostermann (2011), afirmam que o Humanismo por sua vez é baseado no indivíduo em sua totalidade, na pessoa em sua autorrealização e crescimento pessoal. O aprendizado não está baseado somente no aumento de conhecimentos, mas também nos sentimentos, ações e pensamentos de forma integrada. Representantes dessa escola de pensamento são: Carl Rogers, que criou o chamado “ensino centrado no aluno”; George Kelly, que baseia sua teoria na ideia de construtos mentais que são criados e modificados de acordo com suas experiências de vida.

Por sua vez, Joseph Novak, a partir da aprendizagem significativa de Ausubel (1980,p. 685), lança a ideia de que o aprendiz<sup>6</sup> é um sujeito que sente, age, pensa e que ao fim busca o crescimento pessoal e a autorrealização.

E, por fim, Moreira (1999) e Ostermann (2011) concluem afirmando que o Cognitivismo está fundamentado na análise de como ocorrem os processos mentais, ou seja, a cognição humana. Essa corrente teórica tem como objeto de seus estudos as variáveis intervenientes que estão localizadas entre os estímulos e respostas, bem como nos processos mentais superiores<sup>7</sup> e na própria cognição humana. A Teoria Construtivista encontra-se em consonância com a Teoria Cognitivista, que, segundo Moreira (1999, p.15), está baseada na premissa de que o processo cognitivo se desenvolve a partir de uma construção de conhecimento, sendo então o construtivismo uma posição filosófica cognitivista interpretacionista.

Essa teoria por sua vez vê o aluno como sujeito na construção do conhecimento a partir de sua estrutura cognitiva, e não compreende o aluno como um mero receptor de conhecimento. Dentre seus representantes encontram-se Piaget, Bruner, Vygotsky, Johnson Laird, Ausubel e Kelly.

Na próxima subseção é apresentada de forma mais particularizada a teoria da Distância Transacional relativa à aprendizagem e suas relações com a aprendizagem em AVEAs.

## **2.4 TEORIA DA DISTÂNCIA TRANSACIONAL**

A Teoria da Distância Transacional (TDT) foi anunciada por Michael Moore, no ano de 1972, com a intenção de descrever o universo de relações entre aluno e professor quando estão separados no espaço ou tempo, tendo como principais influências para formulação dessa teoria os trabalhos de Carl Rogers, Abraham Maslow e Charlotte Buhler, psicólogos da corrente denominada humanista, possuindo seus preceitos teóricos baseados no humanismo. Tal corrente de pensamento parte

---

<sup>6</sup> O termo “aprendiz” é usado aqui no sentido geral de “o ser que aprende”, e não no sentido específico de “aquele que aprende arte ou ofício”.

<sup>7</sup> Tais processos são: processamento de informação, resolução de problemas, tomada de decisões, percepção e compreensão.

de princípios baseados na particularidade do ser humano, sendo ele um ser único com sua complexidade e singularidade. O surgimento dessa corrente de pensamento ocorreu em contraponto ao behaviorismo, tendo como discordância a ideia de não aceitar o ser humano como máquina ou animal sujeito a condicionamentos. Moore, tendo como base então os conceitos ancorados na visão de um ser humano singular e com características específicas relativas ao aprendizado, os aplicou de forma mais específica na modalidade EAD.

O conceito de Distância Transacional possui como elemento principal o espaço cognitivo entre professor e aluno em um ambiente educacional, mais especificamente na EAD. Segundo Moore (2007, p.240), a distância entre professor e aluno não é meramente geográfica e temporal, mas educacional e psicológica. Ainda segundo Moore, “a distância pedagógica, cognitiva e social que existe entre professor e aluno influencia a eficácia e eficiência da aprendizagem”.

Ainda de acordo com Moore “[...] a Teoria da Distância Transacional permite a geração de números quase infinitos de hipóteses para pesquisas sobre a interação entre professores e alunos” (2013, p.80). O autor propõe como elemento de mensuração a “Interação da Distância”, em que quanto maior o distanciamento aluno-professor maior o valor da variável. Segundo o autor, o uso das TICs colabora para a diminuição dessa distância, podendo ele ser reduzida a valores mínimos para que não afete de forma significativa o processo de ensino e aprendizado. Ainda sobre a TDT, pode-se dizer que ela parte do princípio de que as relações em ambientes EAD são ordenadas conforme os seguintes componentes: estrutura dos programas educacionais, interação entre alunos e professores e o grau de autonomia do aluno.

Em relação ao componente estrutura dos programas, Moore e Kearsley (2007), entendem que esse componente compreende o conjunto de elementos usados na concepção do curso, tais como: objetivos de aprendizado, temas do conteúdo, apresentações de informações, estudos de caso, ilustrações, exercícios e testes. Este componente recebe então forte influência da filosofia da instituição de ensino, dos professores e do nível acadêmico dos alunos, bem como os aspectos organizacionais da instituição.

Sobre as diferentes formas de interação, dos mesmos autores, o termo diálogo é usado para descrever interações de professor e aluno com uma determinada finalidade, sendo construtivo e valorizado por cada participante. Sua abrangência e natureza são determinadas pela filosofia educacional dos responsáveis pela

elaboração de um curso, pela matéria envolvida e por fatores ambientais (linguagem, meios de comunicação). Podemos aqui fazer uma breve observação, os alunos em um AVEA podem realizar diferentes formas de interações, isto é, ocorrem não somente interações diretas com professores e outros alunos, mas também interações com os diferentes recursos presentes em tais ambientes como por exemplo fóruns, webconferências, participação em questionários, acesso a materiais disponibilizados pelo professor, etc. Como terceiro componente, é apresentado o grau de autonomia do aluno. Esse componente, para os autores, compreende a capacidade para tomar decisões a respeito de seu próprio aprendizado, como: desenvolver um plano pessoal de estudo, encontrar condições em ambiente de aprendizado ou de trabalho e decidir quando o progresso está satisfatório; assim se aceita a independência do aluno como um recurso valioso no processo ensino-aprendizagem e não como uma perturbação que precise ser controlada. Ainda sobre a autonomia do aluno, a mesma poderá se modificar de acordo com estrutura que os cursos são oferecidos, por exemplo se os cursos são: construídos de forma em que seus recursos são disponibilizados de forma sequencial ao longo do tempo sendo liberado o acesso a esses recursos somente após a conclusão do anterior, ou se o aluno em primeiro momento inicial tem acesso a todo conteúdo do curso ou disciplina, sendo este último um modelo menos usual. Essas definições na construção didática do curso serão também determinantes para definir em maior ou menor grau a autonomia do aluno. O grau de autonomia se relaciona também com o momento de escolha com quais os recursos interacionais o aluno irá interagir.

No presente estudo, as disciplinas analisadas foram construídas de maneira que os alunos têm acesso aos recursos educacionais e estes são disponibilizados de acordo com o avançar da disciplina, ou seja, o grau de autonomia do aluno é limitado de acordo com a possibilidade de acesso novos recursos. Esta forma de disponibilização está diretamente relacionada com o avanço dos alunos quanto à apropriação dos conteúdos\conceitos e obedecendo critérios de liberação de acordo com o planejamento didático do professor, ou seja, existe uma Trajetória de Aprendizagem Planejada (trajetória em que o professor construiu a sequência didática da disciplina) e a Trajetória de Aprendizagem Realizada (a trajetória efetivamente realizada pelo aluno). Em termos de análise, foram estudadas TAs geradas a partir do avanço dos alunos nas disciplinas tendo seu grau de liberdade de escolha (autonomia) entre os diversos recursos oferecidos dependendo então da disponibilização de tais



recursos pelo docente.

Ainda segundo Moore (1989, p.2) na EAD existem três tipos de interação: aluno-conteúdo, aluno-professor e aluno-aluno.

**Interação aluno-conteúdo** - o primeiro tipo de interação é a interação entre o aluno e o conteúdo ou tema de estudo. Esta é uma característica definidora da educação, sem ela não pode haver educação, uma vez que é o processo de interagir intelectualmente com o conteúdo que resulta em mudanças na compreensão do aluno. É este tipo de interação que se acredita menos envolvido no que Holmberg (1986) chama a "conversa didática interna", ocorre quando os alunos "falam a si próprios" sobre as informações e ideias que encontram num texto, programa de televisão, palestra, ou em outro lugar. Porém sem esse tipo de interação o processo de apropriação dos conteúdos por parte dos alunos se mostra limitado e em alguns casos a apropriação se torna inviável.

**Interação aluno-professor** - o segundo tipo de interação é a interação entre o aluno e o especialista que preparou o material, ou algum especialista que atua como professor. Nesta interação, os professores à distância tentam alcançar objetivos comuns com todos os outros educadores. Primeiro tendo planejado ou tendo recebido um currículo, um programa de conteúdo a ser ensinado e procuram estimular ou pelo menos manter o interesse do estudante pelo que deve ser ensinado, para motivar o estudante a aprender, melhorar e manter o interesse do aluno incluindo a autodireção e a automotivação. Os instrutores organizam avaliações para verificar se os alunos estão progredindo, e para ajudar a decidir se devem ou não mudar as estratégias pedagógicas. A ferramenta proposta nesse trabalho tem como objetivo auxiliar os professores a realizarem uma avaliação de suas disciplinas, a partir de uma análise quanto ao desempenho dos alunos e aos recursos acessados, podendo então reformular o conteúdo de sua disciplina enfatizando conteúdos que sejam mais relevantes ao aprendizado.

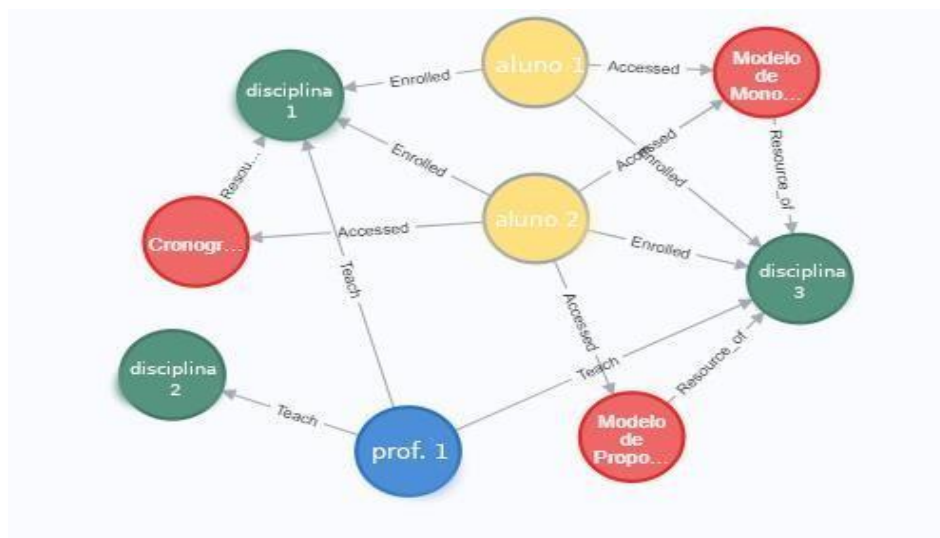
**Interação aluno-aluno** - é a terceira forma de interação, esta dimensão da educação à distância é proposta por Moore nos anos 1990. É uma interação de aprendizagem, entre um aluno e outros alunos, sozinho ou em grupo, com ou sem a presença em tempo real de um instrutor. Nesse tipo de interação fatores como: idade, experiência e nível de autonomia do aluno podem determinar o grau e tipo de interação.

De acordo com o autor, a relação aluno-conteúdo constitui a base do conceito de educação e remete para a integração dos conteúdos na estrutura cognitiva dos alunos. Nesta tese, foram selecionadas para estudo as interações aluno-conteúdo, de forma mais específica as interações com os diferentes recursos presentes em um AVEA. Como exemplo de recursos podemos citar: fóruns, webconferências, participação em questionários, acesso a materiais disponibilizados pelo professor, etc. Indo além, é também realizado neste estudo a identificação e mensuração da quantidade de interações dos alunos com os recursos digitais presentes nas disciplinas apresentadas nas TAs.

Segundo Goel *et al.* (2012, p.733), a área problemática predominante na Teoria da Distância Transacional tem sido a sua medição, bem como, a de seus componentes. Alguns trabalhos apresentam mecanismos de medição, mas nenhum deles conseguiu, até então, contemplar de maneira abrangente todos os conceitos e propostas da teoria (Huang *et al.*, 2016, p.734). Alguns exemplos de ferramentas utilizadas para medir tais componentes são: questionários com escalas do tipo *Likert* (HUANG *et al.*, 2016) e entrevistas abertas (KASSANDRINO *et al.*, 2014).

Nesse estudo, entende-se que as interações são possíveis de serem mensuradas e representadas visualmente por meio da frequência de acessos entre os alunos e os recursos educacionais como é possível verificar na Figura 6 abaixo.

Figura 6 – Interações Educacionais



Fonte: Elaboração própria.

Ainda sobre as interações e a TDT, de acordo com Cabau e Costa (2017, p.431):

Salientamos que quando o tema é interação, os autores buscam investigar, basicamente o processo de interação entre aluno e professor, mas também levantam questões relativas ao papel do professor como gestor da interação, a natureza interativa das tarefas e a interação na comunicação mediada por tecnologias como computador e os ambientes 3D online.

Como parte deste trabalho, a identificação dos caminhos percorridos pelos alunos é fundamental para que se possa entender os comportamentos dos estudantes nos AVEAs. Diante disso, na seção seguinte são conceituadas algumas dessas trajetórias.

## 2.5 TRAJETÓRIAS DE APRENDIZAGEM

Existem diversos termos que definem o percurso ou caminho de aprendizagem de um aluno, alguns deles são: *knowledge tracing* (PIECH et al,2015) , *learning paths* (KLAHR e NIGAM,2004) *learning route* (BARRASH,1994), *learning itineraries* (AGUDELO e IBÁÑEZ), *learning journey* (SMITH e ENG,2013) , *trajectories of cognitive* (FELDMAN e EIDELMAN,2009), trilhas de aprendizagem(RAMOS et al ,2015), progressões de aprendizagem(MONTEIRO et al,2021) e trajetórias de aprendizagem(SIMBINE, 2017).

Para definir esses percursos realizados pelos alunos em ambientes de aprendizagem, neste estudo o termo utilizado são Trajetórias de Aprendizagem (TAs). Sendo a aprendizagem um processo que ocorre ao longo da vida da pessoa, como afirma Illeris (2013), o termo Trajetória de Aprendizagem aplica-se de forma a exemplificar o percurso de aprendizagem de forma análoga ao que ocorre durante a existência de um indivíduo, com um ponto inicial e final e diversas etapas a serem percorridas entre eles ao longo do tempo.

Não existe uma definição única e consolidada para o termo Trajetória de Aprendizagem. Para Serrazina e Oliveira (2010), as TAs são construtos pedagógicos

úteis, bem como construtos teóricos.

A expressão TA é uma metáfora que considera o objeto de aprendizagem similar a um caminho que conecta conceitos: o ponto de partida é um conjunto de conceitos e proposições presumidamente conhecidos pelo estudante e o ponto de chegada é o conjunto de conceitos e proposições relacionados aos objetivos educacionais (CANTO et al, 2014a, p.167).

Para Confrey et al. (2009, p.2), a TA

[...] é a rede ordenada de constructos desenvolvidos pelo estudante em um processo instrucional, uma concepção de pesquisa desenvolvida empiricamente que descreve uma (isto é, atividades, tarefas, ferramentas, formas, de interação e métodos de avaliação) com o objetivo de, através de sucessivos refinamentos de representações, articulações e reflexões, movê-lo de um contexto de ideias informais em direção a um contexto de conceitos progressivamente mais complexos.

Figueiredo et al. (2015) definem TAs como os caminhos que o estudante deve seguir *a priori* para que seja capaz de se apropriar ou construir os conhecimentos pretendidos pelo professor. Para Sarama e Clements (2009), as TAs são descrições do pensamento dos estudantes quando eles aprendem visando a atingir objetivos específicos de um tópico, associados a um percurso hipotético específico, por meio de um conjunto de tarefas de ensino organizado para ativar aqueles processos ou ações mentais dos estudantes.

Segundo os mesmos autores, as TAs ainda podem ser conceituadas como uma sequência de trajetos, onde trajeto é a apropriação de competências por meio do processo de ensino aprendido, tendo como ponto de partida competências ou conceitos previamente conhecidos.

Sobre às TAs, neste estudo buscou-se mapear e identificar em diferentes teses e artigos quais tipos de TAs e seus significados. A seguir estão descritas algumas:

- **Trajectoria Hipotética** - é o caminho previsto pelo professor, em tempo planejamento pedagógico, que assume o caráter de uma tomada de decisão, baseado em conjecturas apoiadas na expertise docente. Ainda, segundo Simonl (1995), uma trajetória hipotética de aprendizagem é uma construção de ensino - algo que professores pressupõem como uma forma de dar sentido para onde os alunos e professor podem ir. É hipotética porque uma trajetória de aprendizagem real não é conhecida com antecedência.

- **Trajectoria de Aprendizagem Conceitual** - é um processo de

aprendizagem que ocorre através do estabelecimento progressivo de relacionamentos entre conceitos previamente conhecidos e novos conceitos ou proposições (ARGÔLO, 2016).

- **Trajectoria Emergente**- é a trajetória que o aluno, o estudante realiza (SIMBINE, 2017).

- **Trajectoria Situacional** - refere-se à trajetória efetivamente realizada (SAINT- GEORGES, FILLIETTAZ, 2008).

- **Trajectoria de Aprendizagem Planejada** (ou Trajetória Pedagógica) - é um conjunto de atividades planejadas com o intuito de alcançar um determinado objetivo educacional (CANTO et al., 2014a).

- **Trajectoria de Aprendizagem Realizada** - é o conjunto de atividades realizadas por um indivíduo com a finalidade de alcançar um determinado objetivo educacional (CANTO et al., 2014a).

- **Trajectorias Conceituais de Aprendizagem** - são o resultado de um processo de agregação pedagogicamente orientada de conceitos elementares (ARGÔLO, 2016).

- **Trajectorias Conceituais de Aprendizagem Intencionais** - podem ser entendidas como projeções de pontos de vista do responsável pedagógico sobre o processo de ensino que deverá se desenrolar através de uma trajetória de aprendizagem intencional (ARGÔLO, 2016).

- **Trajectoria Conceitual de Aprendizagem Referencial** - é a projeção de pontos de vista docente na construção de trajetórias conceituais de aprendizagem. Implica na concepção de uma trajetória de aprendizagem a partir da qual as atividades discentes tomarão lugar. Pode ser vista uma agregação de conceitos e situações estrategicamente elaboradas que proporcionam um Espaço Conceitual Sistêmico (ARGÔLO, 2016).

Ainda, segundo Canto et al. (2014a), o conceito de trajetória de aprendizagem traz consigo uma inerente necessidade de monitoramento e mecanismos de avaliação da trajetória realizada.

Para este estudo o conceito utilizado de TA são os caminhos resultantes das interações dos alunos com diferentes recursos educacionais presentes em um AVEA.

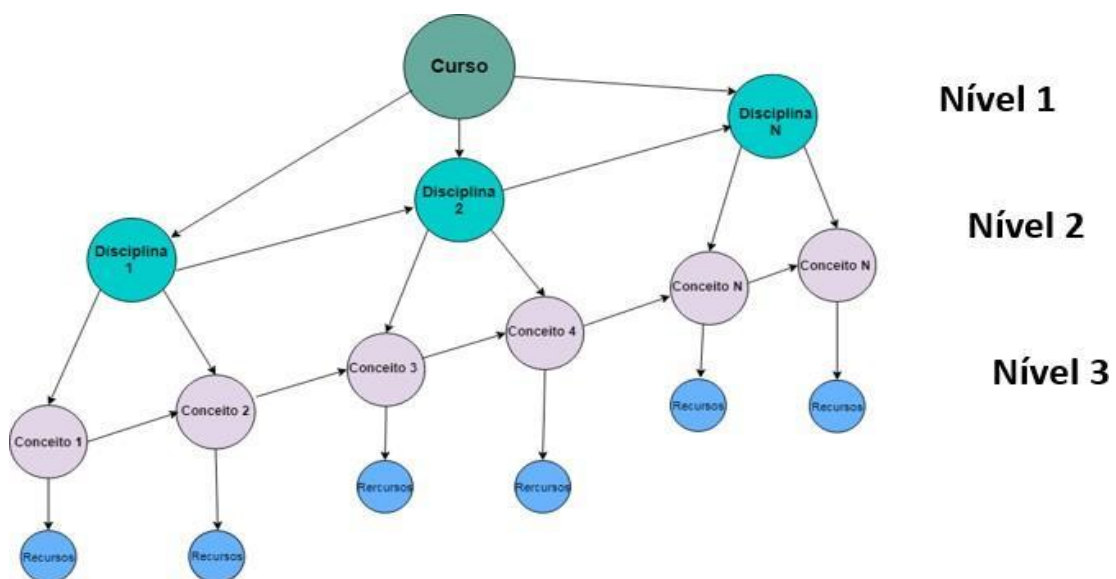
As TAs podem apresentar diferentes níveis de construção de acordo com o que se busca analisar, na subseção seguinte são apresentados opções de análises,

a partir de diferentes níveis.

### 2.5.1 NÍVEIS DE ANÁLISE DE TRAJETÓRIAS DE APRENDIZAGEM

Quando tentamos entender melhor o que é uma TA, como verificado anteriormente, existem diversos termos e conceituações que podem ser utilizados para essa identificação. De forma análoga, quando nos propomos a analisar os caminhos percorridos pelos alunos, tal análise pode ser realizada em diferentes níveis.

Figura 7 – Níveis de Análise das Tas



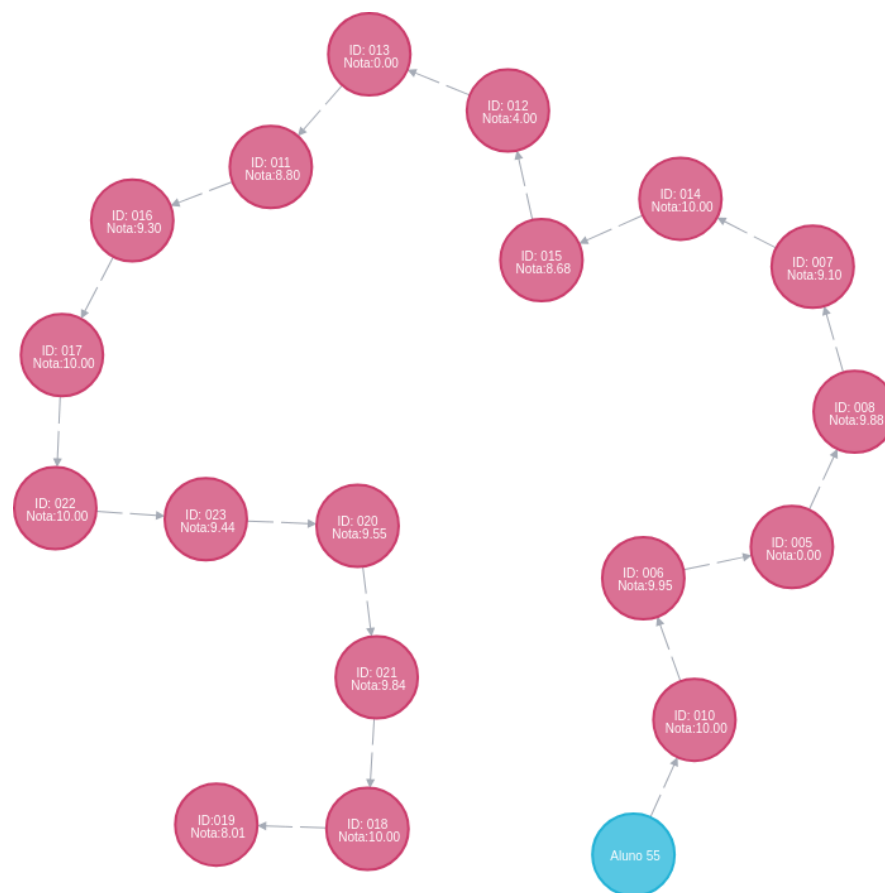
Fonte: Elaboração própria.

Com auxílio da Figura 7, propomos aqui, de forma teórica, 3 níveis de análise, onde em um primeiro nível (Nível 1) são analisadas as relações existentes entre as disciplinas cursadas pelos alunos sendo essa análise caracterizada por um nível mais alto identificando uma TA a partir de suas disciplinas. Neste trabalho são realizadas análise relativas aos recursos, ou seja, são observadas as relações dos alunos com diferentes recursos disponibilizados na plataforma Moodle, sendo assim são análises em nível 3.

A Figura 8 representa uma TA construída a partir de dados reais de um aluno e demonstra a sequência didática percorrida. Nessa imagem é representada uma primeira tentativa de exibir por meio de dados interacionais e com utilização de grafos uma visualização que busque por meio de nodos, onde estão presentes a identificação

da disciplina bem como, a nota do aluno. Também a partir de um nodo inicial onde ocorre a identificação do aluno(nodo azul) e as arestas fazem as relações entre os nodos indicando a sequência realizada. A imagem possui elementos que a caracterizam como uma TA como: um início e um fim, uma sequência lógica e uma interconexão entre seus elementos. Esse modelo, bem como, os próximos que estão presentes nessa seção, foram utilizados como referência para o modelo final proposto neste trabalho.

Figura 8 – Modelo de Trajetória

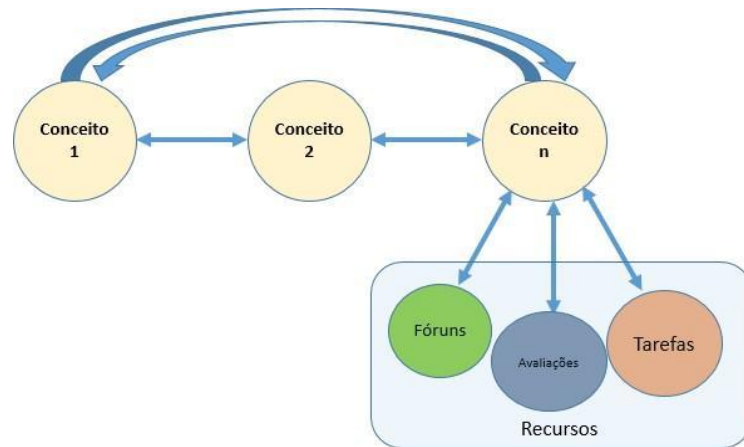


Fonte: Elaboração própria.

Em um exercício analítico, é possível então baixar o nível de análise (Nível 2), tendo como elemento de conexão os diferentes conceitos presentes em uma disciplina (Figura 9). Esse modelo de TA está presente no estudo de Argôlo (2016), que define esse tipo de trajetória como sendo uma TA Conceitual.



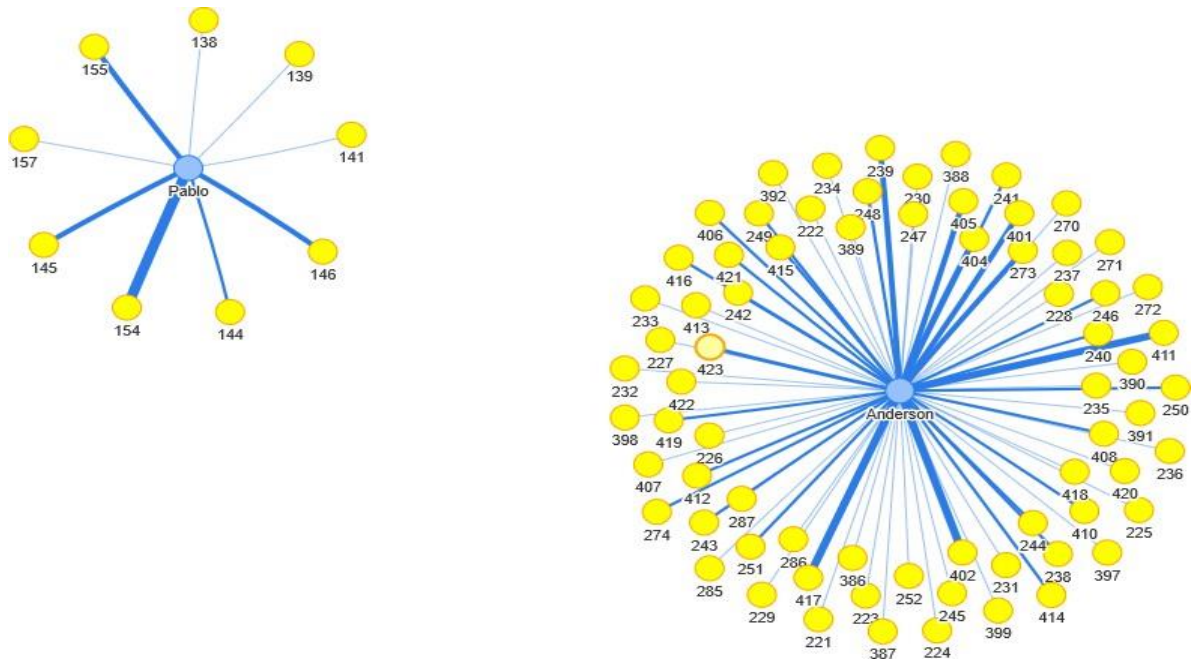
Figura 9– Modelo de Trajetória de Aprendizagem Conceitual



Fonte: Elaboração própria.

No último nível de análise (Nível 3), encontram-se as relações entre atividades em um AVEA, podendo essas serem mensuradas, ou seja, o que se busca neste estudo é a análise das interações dos alunos com diferentes recursos presentes em uma TA, sejam eles fóruns, chats, postagens de tarefas e demais recursos educacionais. A representação das interações entre alunos e os recursos pode ser visualizada na Figura 10, onde dois alunos são identificados pelos nodos azuis e os nodos amarelos representam os diferentes recursos presentes em uma disciplina. As interações são exibidas por meio das arestas, as diferentes espessuras das arestas ocorrem em virtude de uma maior ou menor frequência de acesso aos recursos.

Figura 10 – Interação entre aluno e recursos da TA



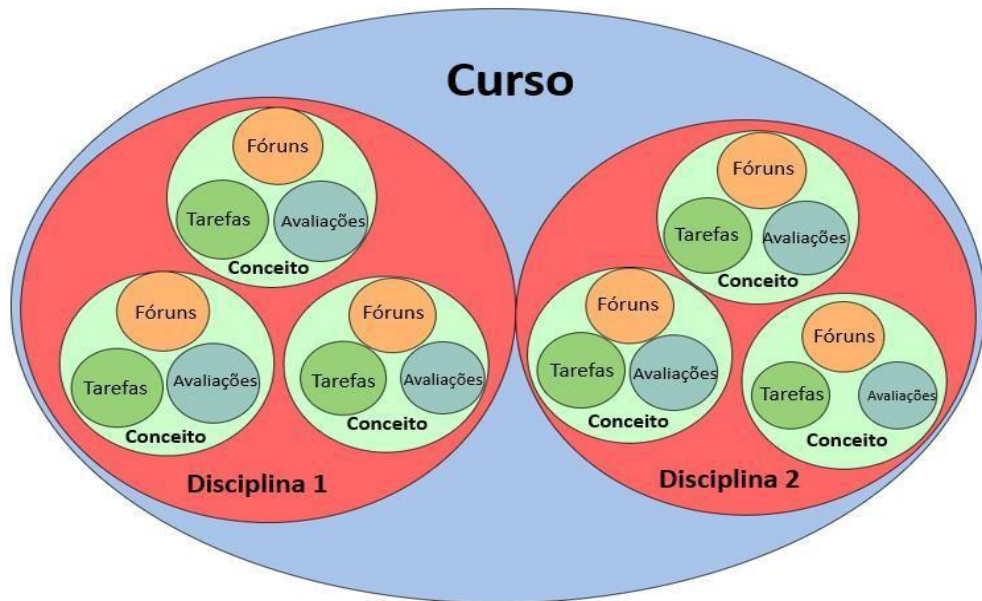
Fonte: Elaboração própria.

Nesse nível de análise, é importante mencionar que também é possível medir a distância transacional, analisando as interações dos alunos com os recursos disponibilizados pelo professor. Por meio de uma identificação visual, podemos observar os recursos mais acessados percebendo as arestas mais densas como de maior frequência de acesso entre alunos e recursos.

De acordo com a Teoria da Distância Transacional de Moore (2007), essa distância é um componente fundamental para um maior ou menor desempenho do aluno e por consequência o seu sucesso e insucesso.

Na Figura 11, é possível identificar a ordem hierárquica dos elementos presentes em uma TA. Dessa forma é viável a identificação das relações e do nível de análise que se busca, a nível de curso, disciplinas, conceitos ou recursos.

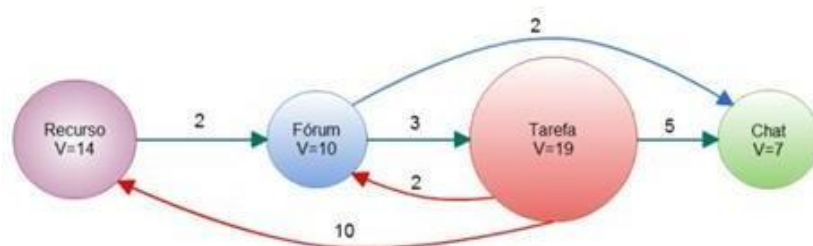
Figura 11– Hierarquia dos Elementos de uma TA



Fonte: Elaboração própria.

Como trabalho similar, quanto ao uso de grafos e referência para o presente estudo, o trabalho de Ramos (2017) propõe a criação de trilhas de aprendizagem (Figura 12) tendo como ponto de partida e chegada os diferentes elementos da TA como: fóruns, tarefas, chats. O modelo é representado em forma de linha que avança no tempo em que estão presentes as interações do aluno e os acessos que o mesmo realizou, bem como, o número de vezes que esses acessos são foram realizados.

Figura 12– Modelo de Trilha de Aprendizagem



Fonte: Ramos (2017, p.4).

No capítulo a seguir, são apresentados conceitos sobre representações visuais, grafos, bem como uma metodologia para a estruturação visual de dados.

### 3- REPRESENTAÇÕES VISUAIS

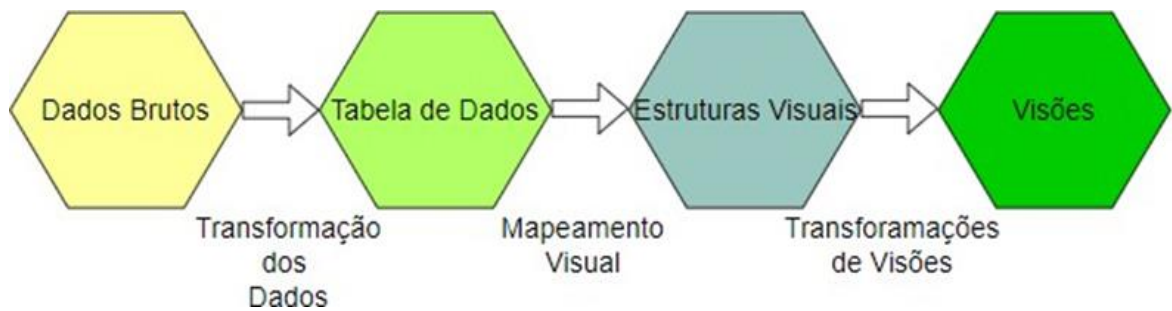
O processo de construção visual das trajetórias e interações tem por objetivo a melhor compreensão dos fenômenos que ocorrem em AVEAs no âmbito do ensino e aprendizado, podendo contribuir para o processo analítico. A visualização refere-se ao processo de transformar dados, informação e conhecimento em uma forma de apresentação visual, passível de ser interpretada com facilidade pelas pessoas.

Segundo Simbine (2017), visualização é o processo de mapear informações e dados em uma forma compacta de apresentação visual (imagética ou gráfica) permitindo uma captação mais eficiente das informações e dados apresentados.

A área de pesquisa relativa à Visualização de Informações é um campo de trabalho emergente e está preocupada com a construção de representações visuais de dados abstratos (NASCIMENTO e FERREIRA, 2005). Se por um lado o grande número de dados interacionais resultantes de plataformas como o Moodle podem ser uma fonte importante sobre o comportamento dos alunos em tais ambientes, o uso correto desses dados para gerar informações úteis não é algo trivial, pois a grande disponibilidade de dados acaba por exigir a escolha precisa do que realmente será significativo para a posterior análise. A escolha também deve ser bastante ajustada pois se não utilizarmos uma quantidade suficiente, a análise pode ser prejudicada. No uso de dados para construir imagens, por exemplo, a quantidade excessiva de dados pode gerar imagens não sejam inteligíveis ocasionando a incompreensão e dificuldade para se obter informações úteis. Em relação ao uso de imagens, Branco (2003) afirma que: “Esta é motivada não só pelo grande acúmulo de informações na atualidade, mas também pela complexidade intrínseca ao processo de se alcançar um mapeamento visual eficaz”.

Quanto à metodologia para estruturação visual de dados, Card, Mackinlay e Shneiderman (1999) apresentam um modelo de referência para a visualização de informações (Figura 13).

Figura 13 – Estruturação Visual de Dados



Fonte: Adaptado de Card *et al.* (1999, p.17).

A estruturação visual de dados apresenta três etapas distintas sendo elas:

1º) **Transformação dos Dados**: é caracterizada pela organização e processamento de um conjunto de dados brutos e sua transposição em forma de uma representação lógica mais estruturada. Essa etapa ainda compreende a filtragem e o agrupamento dos dados relevantes, eliminando os dados redundantes, errados ou incompletos.

2º) **Mapeamento Visual**: caracteriza-se por construir e disponibilizar uma estrutura que pode representar visualmente, por exemplo, dados de uma tabela.

3º) **Transformações Visuais**: onde são criados mecanismos de interação que permitem ao usuário explorar diferentes cenários para melhor entender os dados visualizados.

No presente trabalho, as representações visuais utilizadas são figuras conhecidas como grafos. Na seção seguinte, é apresentada de forma mais específica a teoria utilizada para constituição de tais representações.

### 3.1 – Teoria dos grafos

A teoria dos grafos tem origem na matemática e data do Séc. XVIII, onde foi utilizada na resolução de problemas. A palavra “grafo” é um neologismo derivado da palavra *graph*, e foi utilizada pela primeira vez com os trabalhos de Euler, Kirchhoff e Cayley (SANTOS; MELLO; MURAMI, 2007, p.297). De acordo com Dunham (2007), o primeiro problema, chamado o *problema das pontes de Königsberg*, foi anunciado por Euler, no ano de 1736, em que consistia na possibilidade de atravessar sete pontes sem repetir nenhuma. Euler então provou que não existia caminho que

possibilitasse tais restrições, usando um raciocínio muito simples, transformou os caminhos em linhas e as intersecções em pontos, criando provavelmente o primeiro grafo da história.

O uso de grafos apresentou uma maior intensificação nas décadas mais recentes em decorrência do seu uso para análise de redes sociais, bem como a identificação de comunidades dentro dessas redes. Os grafos também permitem análises mais refinadas dos dados, reconhecimento de padrões e análise de medidas de centralidade.

De uma forma simplificada, o conceito proposto neste trabalho indica que um grafo é uma representação visual constituída de vértices e arestas onde é possível verificar as relações existentes entre os diferentes vértices por meio de suas conexões, as arestas. Dois vértices ligados por uma aresta são sempre chamados de adjacentes.

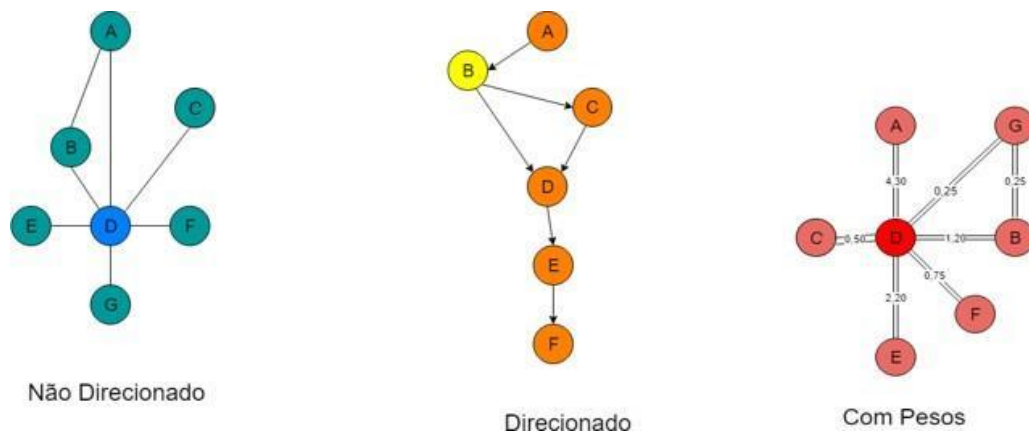
Como uma aresta conecta sempre dois vértices, a soma do grau de todos os vértices é sempre o dobro do número de arestas. Matematicamente, um grafo pode ser representado por um conjunto  $V$ , e por  $V^2$  um conjunto de todos os pares não ordenados de elementos de  $V$ . Tendo  $V$   $n$  elementos, então  $V^2$  possui  $(n/2) = n(n-1)/2$  elementos. Elementos  $V^2$  são identificados como subconjuntos de  $V$  que têm cardinalidade 2. Desse modo, cada componente de  $V^2$  terá a forma  $\{v, w\}$ , sendo  $v$  e  $w$  dois elementos diferentes de  $V$ .<sup>8</sup>

Em relação a direção, os grafos podem ser direcionados ou não (Figura 14), quando não direcionados, não existe um sentido pré-estabelecido para fluxo entre os nodos. As arestas dos grafos podem possuir valor ou não, e quando atribuído valor às arestas, essas servem para determinar o peso das relações entre os nodos. Por exemplo, quando utilizado um grafo em um modelo probabilístico, o valor das arestas pode servir para demonstrar a probabilidade de ocorrência de um conceito contido em um nodo quando se parte de um nodo anterior.

---

<sup>8</sup>As definições citadas podem ser acessadas em: <https://www.inf.ufsc.br/grafos/definicoes/definicao.html>

Figura 14 – Exemplos de Grafos



Fonte: Elaboração própria.

### 3.1.1 Classificação dos Grafos

Grafos podem ser utilizados para os mais diversos objetivos de acordo com sua classificação, de acordo com Souza (2013), os grafos podem ser classificados da seguinte forma:

- Grafo de Relacionamento** - são usados para representar diversas relações entre pessoas, como por exemplo, um gráfico simples representa duas pessoas que se conhecem, ou seja, possuem uma relação entre elas. Cada pessoa é representada por um vértice e uma aresta não orientada. No presente estudo, este tipo de grafo é utilizado para representar a relação dos alunos com diferentes recursos.
- Grafo de Influência** - esse modelo de grafo é utilizado quando após a observação do comportamento de um grupo é possível identificar a influência de um indivíduo em relação aos outros do grupo. Por meio do processo de modelagem, cada indivíduo é representado por um vértice e uma aresta orientada exibe quando um indivíduo representado pelo vértice *a* influencia outro indivíduo representado pelo vértice *b*, pois o vértice *a* aponta para vértice *b*.
- Grafo da Web** - é possível modelar um grafo orientado representando a

*World Wide Web*, sendo que cada página é representada por um vértice e uma aresta que inicia na página  $a$  da web e finaliza na página  $b$ , se existir um link em  $a$  que direciona para  $b$ . Devido à grande velocidade e quantidade de páginas criadas e retiradas da web, o grafo se modifica constantemente e possui mais de três milhões de vértices e vinte bilhões de arestas.

- **PageRank** - é um mecanismo de votação, em que sua maior representação do seu uso é na ferramenta de buscas Google. Quando se digita uma palavra-chave de algo que se deseja obter informações. Cada site é representado por um vértice de um grafo orientado e cada link em que uma página aponta para outra. Um “voto” representado por uma aresta  $(u,v)$ , onde  $u$  indica a página  $v$ . Quanto mais elevado for o valor PageRank da página de onde o link tem origem, maior o valor do voto dado à página de destino. A ferramenta de busca Google possui milhares de páginas cadastradas tornando o cálculo dos valores das páginas cada vez mais complexo.

- **Grafo de Torneios Round-Robin** – em um torneio um time joga contra outro time exatamente uma vez, o jogo pode ser modelado usando grafos orientados e cada time é representado por um vértice. É um grafo orientado simples que não contém nem laços nem arestas múltiplas.

- **Grafos de Mapas Rodoviários**- grafos podem ser usados para modelar mapas rodoviários, os vértices representam as intersecções e as arestas representam estradas. Arestas orientadas representam estradas de mão única, e as arestas não orientadas representam estradas de mão dupla. Arestas não orientadas múltiplas representam estradas de mão dupla múltiplas que conectam as mesmas duas intersecções, por sua vez, as arestas orientadas múltiplas representam estradas de mão única múltiplas que começam em uma intersecção e terminam em outra intersecção. Grafos mistos são necessários para representar mapas rodoviários que incluem tanto estradas de mão única como de mão dupla.

Grafos podem também ser classificados quanto a sua centralidade como pode ser visto subseção seguinte.



### 3.1.2 Classificação quanto à Centralidade

As medidas de centralidade surgiram da análise em redes sociais, que podem ser caracterizadas por um conjunto de indivíduos conectados por relações sociais. Em redes, como as mencionadas anteriormente, cada posição corresponde a um vértice no grafo e cada possível relação entre as posições corresponde a uma aresta que conecta um par de vértices.

Intuitivamente, em uma rede, os vértices mais centrais são aqueles que a partir dos quais podemos atingir qualquer outro com mais facilidade ou rapidez.

Nesta subseção são descritas diferentes medidas de centralidade que avaliam a importância dos nós em uma rede de acordo com sua posição estrutural.

Segundo Freitas (2010), os grafos podem ser analisados por meio de sua medida em relação à centralidade, a seguir são identificadas algumas dessas medidas.

- **Centralidade de Grau** - é assim chamada por ser uma função do grau de um vértice, conta o número de arestas incidentes a um vértice do grafo. Essa medida de centralidade é considerada a mais simples intuitiva no que diz respeito à centralidade de um vértice e o número de contatos diretos que ele possui. No presente trabalho, essa medida pode ser verificada por meio do número de acessos em relação ao nodo, ou seja, a frequência de acesso dos alunos aos recursos.

- **Centralidade de Proximidade** - esta medida de centralidade, por sua vez, está relacionada com a distância total de um vértice a todos os demais vértices do grafo.

Segundo Sabidussi (1966), a centralidade de proximidade é baseada na soma das distâncias de um vértice em relação aos demais vértices do grafo.

- **Centralidade de Eficiência** - em relação a esta medida de centralidade, a mesma minimiza as distâncias de um vértice para alcançar qualquer outro vértice no grafo.

De acordo com Harary (1991), a medida chamada centralidade de eficiência está baseada no conceito de excentricidade de um vértice. Tal medida pode ser aplicada em problemas de localização como estabelecer um local de modo que minimize o tempo de viagem entre o ponto e as demais localizações. De forma aplicada, por exemplo, pode ser utilizado quanto a determinação do posicionamento de um hospital, em que um dos fatores a ser considerado é a redução do tempo

máximo de atendimento de uma ambulância a uma possível emergência.

- **Centralidade de Intermediação** - mede quantas geodésicas<sup>9</sup> entre todos os pares de vértices do grafo que passam através de um determinado vértice.

As redes sociais no mundo real são sempre muito extensas e geralmente desconexas, limitando a utilização das medidas de centralidade conhecidas. Freeman (1977) propôs medidas que mesmo sendo baseadas em interligações entre pares de vértices, poderiam ser aplicadas a grafos desconexos. Surgindo primeiramente o conceito de intermediação parcial de um vértice em uma dada rede para então chegar a um valor que pudesse medir a centralidade de tal vértice. Significando então que este valor deveria ser capaz de expressar a influência que o indivíduo (representado pelo vértice) poderia exercer sobre os seus pares em uma rede de comunicação no mundo real.

- **Centralidade de Intermediação de Percursos Aleatórios**- a centralidade de intermediação pode ser vista como uma medida da influência que um nó tem sobre a propagação do fluxo de informação (ou de qualquer conteúdo) através da rede. Esse tipo de medida de centralidade considera todos os caminhos, geodésicos e não-geodésicos, passando através de um vértice, possíveis de serem aleatórios.

Para exemplificar o uso de grafos na construção de trajetórias, no estudo de Argôlo (2016), grafos são utilizados para representar um panorama imagético da quantidade de registros básicos realizados em função da modelagem tempóreo-espacial. Tal estudo objetivou organizar e demonstrar as trajetórias conceituais de aprendizagem, tendo como elemento principal a relação entre os núcleos conceituais.

No estudo aqui proposto, grafos e arestas são utilizados para representar a visualização das relações dos alunos com diferentes recursos presentes nas TAs.

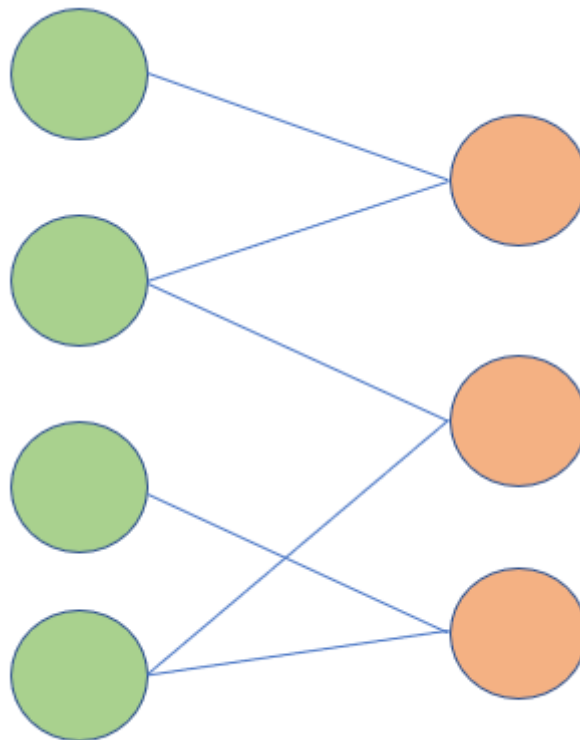
---

<sup>9</sup> É o caminho entre dois pontos em um espaço tridimensional.

### 3.2 VISUALIZAÇÃO EM GRAFOS

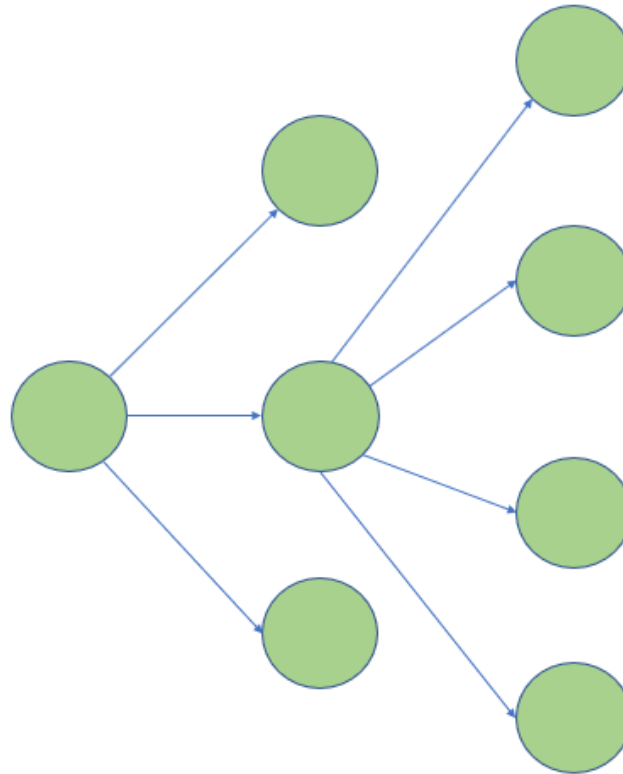
Para este estudo, as informações são modeladas em forma de grafos, mais especificamente duas estruturas de grafos são utilizadas: grafos bipartidos, aqui chamados de bigrafos, e árvores. Um grafo bipartido é aquele que é possível separar os nodos em dois conjuntos U e V de forma que toda aresta do grafo esteja ligada a um nodo de U e outro de V. Uma árvore é um grafo que pode ser organizado de forma hierárquica onde cada nodo tem um único antecessor, exceto o nodo raiz que não possui antecessor. Exemplo dos dois tipos são apresentados nas Figuras 15 e 16.

Figura 15 - Exemplo de grafo bipartido



Fonte: Elaboração própria.

Figura 16 – Exemplo de grafo árvore



Fonte: Elaboração própria.

No trabalho de Bennett et al. (2007), é apresentado um modelo para os fatores envolvidos na interpretação de visualização em grafos. Em um primeiro momento, o leitor observa a imagem e o cérebro capta rapidamente os elementos principais da visualização, como linhas e cores fortes e a distribuição dos elementos no espaço. Assim visto, são enunciados alguns critérios que formam uma boa visualização (SUN e WONG, 2005)

- Simplicidade da figura.
- Elementos com características similares, como forma ou cor, aparentam estar agrupados.
- Elementos colineares aparentam estar agrupados.
- Elementos próximos tendem a formar um grupo ou compartilhar características entre si.
- Elementos conectados aparentam estar agrupados.
- Elementos com formatos familiares tendem a ter alguma relação entre si.

- Objetos devem ser mais evidentes que a imagem de fundo da visualização.
- Áreas demonstrando simetria tendem a serem vistas como uma figura separada do resto da visualização.
- Itens alinhados vertical ou horizontalmente tendem a ser vistos como uma figura.
- Contornos auxiliam na definição ou percepção das figuras.

Neste trabalho, tais critérios foram usados para formar uma visualização que seja intuitiva ao usuário. Por ser uma visualização interativa, a manipulação do grafo permite diminuir ou aumentar a complexidade das informações exibidas, facilitando a percepção dos elementos à medida que a visualização é alterada.

No Capítulo seguinte é melhor descrita a etapa de Mineração de Dados e como foi aplicada nesse trabalho.

## 4 MINERAÇÃO DE DADOS

Atualmente são gerados grandes volumes de dados tendo as mais diversas origens, como por exemplo, redes sociais, comunidades virtuais, dispositivos médicos, sensores em roupas, casas, aviões, trens e carros, além de equipamentos como câmeras de vigilância, leitores de código de barra, *smartphones*, *tablets* e computadores. Esses dados são gerados em diferentes formatos, velocidades e volumes.

Segundo Palazzo (2018), a palavra *Dado* tem como origem o termo do latim *datum*, que significa algo oferecido, dado. Em seu estudo, Han (2006) refere-se a esse momento em que vivemos como um período "rico em dados, pobre em informação".

Tendo como referência a pirâmide de hierarquia da sabedoria de Rowley (2007), e fazendo uso de um conjunto de dados, neste estudo serão realizadas três etapas presentes na referida pirâmide (Figura 17), pois o objetivo é, a partir de um conjunto de dados, extrair informações que gerem conhecimento para que sirva como subsídio para a identificação de trajetórias de aprendizagem.

Figura 17 – Pirâmide DIKW



Fonte: Adaptado de Rowley (2007)

Segundo Fayyad (1996), o modelo tradicional para transformação de dados em

informação consiste em um processamento manual de todas essas informações por especialistas que, então, produzem relatórios que deverão ser analisados. Na maioria das situações, devido ao grande volume de dados, o processo de forma manual torna-se impraticável. Para tornar esse processo viável, uma alternativa é utilizar a MD.

Por sua vez, um AVEA fornece uma diversidade de dados interacionais que quando recebem um tratamento adequado por meio do processo de MD, por exemplo, podem sinalizar comportamentos e padrões. Alguns exemplos desse tipo de interação são: número de acessos, número de exercícios respondidos, frequência de interação com os recursos educacionais, interação via fóruns e interações via chat.

O uso da MD em cursos de educação a distância é fonte de estudos com distintos objetivos como análise de evasão escolar (PORTAL, 2016), colaboração entre estudantes (Vasconcelos et al.; 2018) e desempenho acadêmico (Rodrigues et al.; 2015).

A MD, segundo *Microsoft Developer Network* (2006), “É o processo de descoberta de informações acionáveis em grandes conjuntos de dados” e que “[...] usa análise matemática para derivar padrões e tendências que existem nos dados”. Finalmente, “[...] esses padrões não podem ser descobertos com a exploração de dados tradicional pelo fato de as relações serem muito complexas ou por haver muitos dados”.

Configura-se então como a aplicação de técnicas, implementadas por meio de algoritmos computacionais capazes de receber como entrada um conjunto de dados e devolver como saída um padrão comportamental que pode ser demonstrado, por exemplo, como uma modelagem de um perfil, uma função de mapeamento ou uma regra de associação (técnica descrita mais adiante nesse documento).

A MD é uma etapa de um processo mais amplo, denominado de KDD (*Knowledge Discovery in Databases* ou Descoberta de Conhecimento em Bases de Dados), que será melhor discutido na subseção seguinte.

## 4.1 PROCESSO DE DESCOBERTA DE CONHECIMENTO EM BASE DE DADOS

O processo de Descoberta de Conhecimento em Base de Dados (DCBD), conforme pode ser visualizado na Figura 18, é constituído por cinco etapas distintas que são: a seleção dos dados, o pré-processamento, a transformação, a mineração e a análise e Avaliação. Esse, segundo Fayyad (1996), "É um processo não trivial de identificação de novos padrões válidos, úteis e compreensíveis". O processo pode ser interativo ou iterativo, ou seja, em cada etapa pode ser executada mais de uma vez de forma sequencial ou a partir de uma etapa específica. Tal descoberta está associada a um processo analítico, sistemático, se possível automatizado.

Figura 18 – Processo de DCBD



Fonte: Adaptado de FAYYAD (1996).

### 4.1.1 Seleção de dados

Na primeira etapa, é realizada a seleção e coleta dos dados, tendo como principal cuidado garantir a qualidade da fonte do material. Os dados que são utilizados neste estudo têm como origem a plataforma Moodle em um curso de



graduação na modalidade EAD, ofertado pela Universidade Federal do Rio Grande do Sul (UFRGS).

#### **4.1.2 Pré-processamento**

Na segunda etapa são realizadas tarefas com objetivo a eliminação de dados repetidos ou dados inconclusivos. Também é possível nesse momento do processo recuperar dados incompletos. Nessa etapa, por vezes, devido ao grande número de dados utilizados, é necessária a redução do tamanho da amostra com a finalidade de também diminuir o número de variáveis que compõem o algoritmo.

#### **4.1.3 Transformação**

Na terceira etapa, após serem selecionados, limpos e pré-processados, os dados necessitam ser armazenados e formatados adequadamente para que os algoritmos possam ser aplicados. A transformação e a combinação também são utilizadas para obter “dados derivados” que são os dados faltantes.

#### **4.1.4 Mineração**

Na quarta etapa do processo, é realizada a MD, que segundo alguns especialistas é a principal fase do processo. Cabe lembrar que se as etapas anteriores não forem realizadas com um rigor técnico elevado, a qualidade da mineração acaba sendo comprometida levando a resultados equivocados ou inconclusivos. Algumas possibilidades de mineração são descritas na Seção 4.2.

#### **4.1.5 Análise/Avaliação**

Como quinta e última etapa, é realizada a análise e avaliação dos dados já transformados em informação. Ainda sobre esta etapa, podemos destacar que em alguns casos existe a necessidade de um especialista para que este possa realizar a correta leitura e análise das informações obtidas.

Quando os resultados apresentados ao final do processo não forem satisfatórios e não sendo possível identificar em qual etapa houve alguma falha, é recomendado retornar à primeira etapa do processo de MD e realizar novamente o

processo. Na subseção a seguir são apresentadas diferentes tarefas de MD.

## 4.2 TAREFAS DE MINERAÇÃO DE DADOS

A MD possibilita a realização de diferentes tarefas de acordo com o objetivo final do processo, algumas delas são descritas a seguir.

A tarefa de **Agrupamento** (*clustering*), segundo Ward (2005), é uma tarefa não supervisionada e tem como objetivo identificar e aproximar os registros similares. Um agrupamento é um conjunto de registros similares entre si, porém sendo diferentes dos outros registros nos demais agrupamentos. Essa tarefa se diferencia da classificação por não necessitar que os registros sejam previamente categorizados (o aprendizado é dito “não-supervisionado”). Ela apenas identifica os grupos de dados similares e não tem a pretensão de classificar, estimar ou prever o valor de uma variável. No presente estudo, a técnica de agrupamento foi utilizada para a identificação de diferentes grupos de alunos em quatro turmas selecionadas. Os alunos foram agrupados utilizando como fator de agregação o seu desempenho nas disciplinas, tendo como indicador numérico a nota final.

**Associação**, de acordo com Novaes (2002), é uma tarefa que consiste em identificar quais atributos se relacionam. Possuem a seguinte forma: SE atributo X ENTÃO atributo Y. Devido aos bons resultados obtidos é uma das tarefas mais conhecidas, principalmente nas análises da “Cestas de Compras” (*Market Basket*), onde identificamos quais produtos são adquiridos em conjunto pelos consumidores.

A **Classificação** é uma tarefa supervisionada e tem por objetivo identificar a qual classe pertence um determinado registro. De acordo com Silva (2005), nessa tarefa, o modelo analisa o conjunto de registros, com cada registro já contendo a indicação à qual classe pertence, a fim de “aprender” como classificar um novo registro (o aprendizado é dito “supervisionado”). Nesse tipo de tarefa, faz-se uso de classificadores, que são algoritmos utilizados para as tarefas automáticas de classificação de dados, podendo esses gerar informações para tomada de decisões. Alguns tipos de classificadores são: *Gaussian NB*, *Logistic Regression*, *MLP Classifier*, *Random Forest Classifier*, *Gradient Boosting Classifier*, *Decision Tree Classifier*. Eles não serão detalhados aqui, mas encontram-se disponíveis na maioria das ferramentas e bibliotecas de MD.

**Descrição** é um tipo de tarefa utilizada para especificar padrões e tendências

revelados pelos dados. A descrição normalmente oferece uma possível interpretação para os resultados obtidos. A tarefa é comumente utilizada em conjunto com as técnicas de análise exploratória de dados para comprovar a influência de certas variáveis no resultado obtido.

**Regressão** ou Estimação são similares à classificação também são supervisionadas, porém são usadas quando o registro é identificado por um valor numérico e não um categórico. Assim, pode-se estimar o valor de uma determinada variável analisando-se os valores das demais.

A tarefa de **Predição** é uma tarefa que segundo Han *et al.* (2006) possui similaridade às tarefas de classificação e estimação, porém ela visa descobrir o valor futuro de um determinado atributo. Um exemplo desse tipo de tarefa é predizer se um aluno atingirá ou não um conceito em uma disciplina baseado em suas notas anteriores.

Em testes e simulações realizadas no decorrer da escrita deste estudo, foram utilizadas as tarefas de predição, classificação e associação utilizando um conjunto de dados gerados a partir da plataforma Moodle. A tarefa de predição também está presente como uma funcionalidade na ferramenta de visualização (MDV).

#### 4.2.1 A Validação dos Dados

A validação de dados é a forma de avaliar a qualidade dos modelos de MD para que a escolha do modelo seja a mais adequada possível e atenda corretamente a expectativa quanto à tarefa a ser executada.

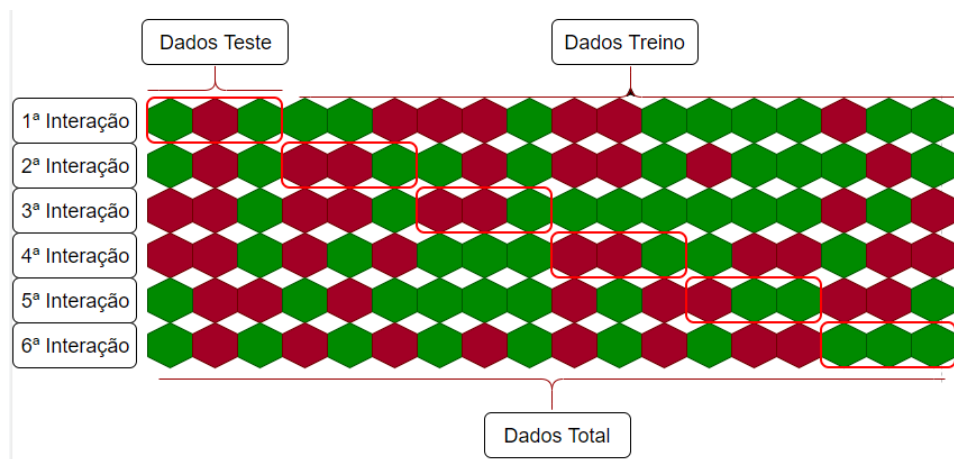
O grande volume de dados gerados impossibilita o uso de todos os registros do repositório para a construção do modelo de MD, tornando essa opção inviável. Assim, utiliza-se uma amostra (mais representativa possível) que é dividida em dois conjuntos:

1. Conjunto de Treinamento (*Training Set*): conjunto de registros usados no qual o modelo é desenvolvido;
2. Conjunto de Testes (*Test Set*): conjunto de registros usados para testar o modelo construído.

Neste estudo, o modelo de validação dos dados utilizado é o modelo cruzado,

onde é utilizado todo o conjunto de dados, sendo esse dividido em diversas partes. Os subconjuntos de dados gerados serão, em algum momento do processo, usados como conjunto de dados de teste, bem como, conjunto de dados de treino. Isso é possível pois os conjuntos passarão por um número significativo de interações, como exemplificado na Figura 19.

Figura 19 – Exemplo de Validação Cruzada



Fonte: Elaboração própria.

A metodologia proposta por meio da validação cruzada tem como objetivo evitar o risco que, em algum momento do processo, a amostra selecionada não seja representativa em relação ao conjunto de dados (população) de onde foi retirada. A validação cruzada permite verificar a acurácia de um modelo gerado a partir da análise de dados de treinamento e teste sem necessidade de selecionar um conjunto de dados em separado para posterior validação do modelo.

Neste estudo a validação é utilizada em uma funcionalidade presente na ferramenta de análise visual MDV, onde o usuário pode realizar a predição quanto a aprovação ou não de um aluno.

### 4.3 MINERAÇÃO DE DADOS EDUCACIONAIS

Para identificar relações entre os dados, e como consequência informações que possam ser posteriormente analisadas, a Mineração de Dados Educacionais (MDE) mostra-se uma alternativa.

A MDE é definida por Baker e Carvalho (2011) como a área de pesquisa que possui como principal foco o desenvolvimento de métodos para explorar conjuntos de dados coletados em ambientes educacionais. Esse tipo de ferramenta proporciona ao seu usuário filtrar e selecionar dados que possam fornecer alguma informação sobre o aluno ou grupo de alunos.

Como exemplo, segundo os autores, é possível identificar em que situação um tipo de abordagem instrucional (aprendizagem individual ou colaborativa) proporciona melhores benefícios educacionais ao aluno. Também é possível verificar se o aluno está desmotivado ou confuso e, assim, personalizar o ambiente e os métodos de ensino para oferecer melhores condições de aprendizagem. MDE está focada em duas linhas:

(1) A análise de dados e a criação de modelos para melhor compreender os processos de aprendizagem; e

(2) O desenvolvimento de métodos mais eficazes para dar suporte à aprendizagem quando o aluno estuda utilizando softwares educacionais.

O presente estudo encontra-se em consonância com a primeira linha apresentada, sendo o objetivo a análise de dados e a busca de padrões de interação nas trajetórias dos alunos em AVEAs.

A MDE tem sido utilizada em diversas áreas, sendo suas principais áreas de aplicação, de acordo com Baker e Yacef (2009), são as que seguem: modelagem do estudante, modelagem do domínio, suporte pedagógico e descoberta científica. Os modelos do estudante armazenam informação sobre características dos alunos, tais como conhecimento, motivação, atitudes, personalidade, além de questões sociais. As técnicas de MDE podem ser utilizadas para dar uma maior precisão no modelo de estudante e proporcionar uma maior personalização e adaptação de um AVEA. Modelar as diferenças existentes entre os estudantes possibilita acompanhar o aprendizado de forma individualizada, melhorando significativamente o seu aprendizado. Utilizando métodos de MDE é possível modelar atributos do estudante em sistemas de tempo real, podendo ser selecionados os que realmente terão uma influência positiva no aprendizado.

Como etapa desta pesquisa foi realizada uma revisão sistemática em que se buscou em bibliotecas digitais trabalhos relacionados diretamente com esta tese. A seguir é descrito o processo de pesquisa relativo a tais trabalhos.

## 5 TRABALHOS RELACIONADOS

Neste capítulo, a fim de contextualizar brevemente a área de pesquisa de modo a discutir alguns de seus processos e contextos, são apresentados trabalhos identificados a partir de uma revisão sistemática focada em estudos em que foram utilizadas ferramentas para análise de alunos em AVEAs. A presente revisão sistemática de literatura foi realizada entre o segundo semestre do ano de 2017 ao primeiro semestre de 2019.

### 5.1 REVISÃO SISTEMÁTICA

Nesta etapa do estudo, foi realizada uma revisão sistemática que se compreende como sendo um método de pesquisa que fornece direções para realizar revisões de literatura. De acordo com Kitchenham e Charters (2007), ela permite que evidências sobre um determinado domínio sejam reveladas em um alto nível de granularidade e que sejam identificados estudos agrupados de acordo com uma série de características, definidas pelos revisores. Ainda segundo Sampaio e Mancini (2006), “[...] nesse sentido, revisões sistemáticas e metanálise são os métodos mais adequados e atuais para resumir e sintetizar evidências sobre a eficácia e os efeitos de intervenções”.

A revisão apresentada neste estudo tem como embasamento teórico a metodologia sugerida por Petersen *et al.* (2008). A presente metodologia tem como principais etapas da revisão sistemática: definição de questões de pesquisa, busca por trabalhos relevantes, seleção e filtragem de trabalhos (critérios de inclusão e exclusão) e como etapa final a extração de dados e mapeamento (Figura 20).

Figura 20 – Etapas da Revisão Sistemática



Fonte: Adaptado de Petersen *et al.* (2008, p.18).

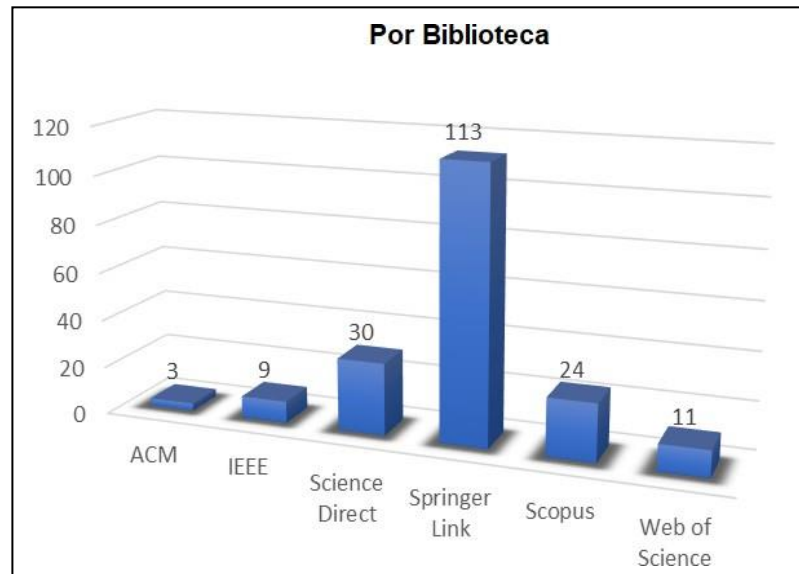
Como primeira etapa deste estudo foi definida a seguinte questão de pesquisa:  
**“Como a utilização de técnicas, sejam elas de mineração de dados ou outras podem identificar padrões de comportamento de alunos em trajetórias de aprendizagem, caminhos de aprendizagem ou trilhas de aprendizagem?”**

Dessa forma, com a utilização de termos selecionados foi realizada nas principais bibliotecas digitais uma pesquisa utilizando palavras chaves que foram elencadas como as principais do trabalho. Mesmo com a informação de que as bibliotecas digitais pesquisadas são constituídas de documentos em sua maioria absoluta na língua inglesa, foram mantidos os termos em português com a intenção de verificar entre os resultados da pesquisa a existência de trabalhos em Língua Portuguesa. Abaixo segue a *string* utilizada para a pesquisa:

("Educational Data Mining" OR "EDM" OR "mineração de dados educacionais") AND ("Learning Path" OR "learning trajectory" OR "trajectory of Learning" OR "learning trail" OR "Trilha de aprendizado" OR "trilha de aprendizagem" OR "trajetória de aprendizado" OR

As fontes de pesquisa, ou seja, as bibliotecas digitais definidas para o estudo foram as seguintes: *ACM, IEEE, Science Direct, Springer Link, Scopus e Web of Science*, entendendo que sejam fontes de consulta bem estabelecidas e com um grau de confiança perante a comunidade acadêmica. Também como definição quanto a utilização das bibliotecas, sua escolha foi realizada tendo como referência teses anteriores relativas às trajetórias de aprendizagem que realizaram revisões semelhantes. Como resultado, foi obtido o retorno de 190 documentos entre artigos completos, artigos curtos, trabalhos de conclusão de graduação, dissertações e teses.

Figura 21 – Número de trabalhos por Biblioteca



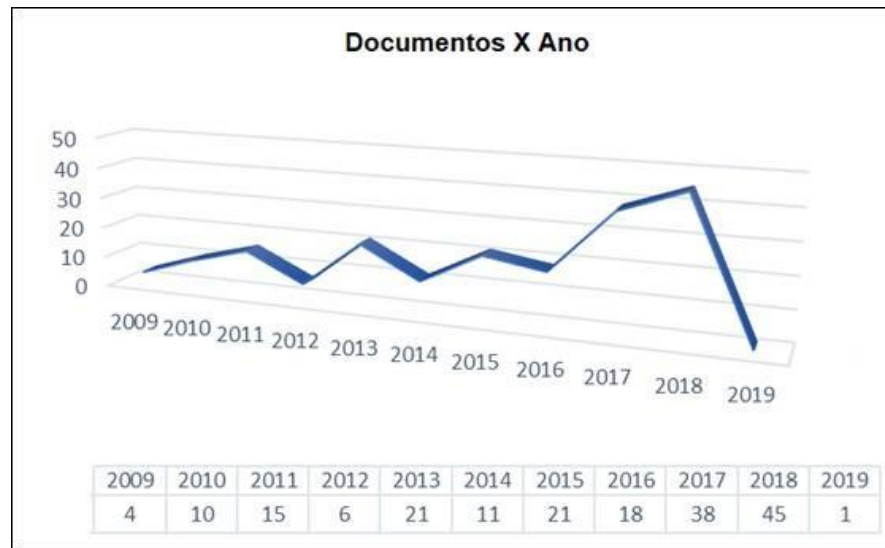
Fonte: Elaboração própria.

Buscou-se neste estudo obter o número de trabalhos encontrados por biblioteca, como é possível verificar na Figura 21. Esse tipo de levantamento possui relevância, pois pode ser utilizado como referência para trabalhos futuros em que se busque termos usados nesse trabalho ou termos afins que possam ser utilizados. Tal informação pode concentrar a busca em bibliotecas que retornem um maior número de trabalhos, poupando assim o tempo dos pesquisadores.

Foi também realizada uma pesquisa ano a ano, para o período de 2009-2019, houve uma predominância maior de produção de material entre os anos 2017 e 2018 como é possível visualizar na Figura 22. Cabe aqui uma nota em relação ao ano de 2019 que apresentou a ocorrência de apenas um trabalho em virtude da coleta de dados ter seu início no mês janeiro do referido ano.



Figura 22 – Número de trabalhos por ano



Fonte: Elaboração própria.

Como primeiro critério de exclusão foram descartados documentos duplicados resultando em 154 documentos, em um segundo momento foi estabelecido como critério de exclusão documentos que foram publicados há mais de 10 anos, trabalhos de conclusão de cursos de graduação (TCCs). Como critério de inclusão foram incorporados artigos que apresentaram os termos: evasão, retenção, reprovação e desistência, bem como grafos e seus termos similares, resultando em 52 artigos. A partir daí, foi estabelecido como critério para seleção a leitura dos títulos, abstract e palavras-chave.

Posteriormente, foram identificados os artigos considerados diretamente relacionados à questão de pesquisa sendo resultante 13 trabalhos que foram lidos em sua integralidade (Quadro 2).

Quadro 2 – Trabalhos Selecionados

Título	Autores	Palavras-chave	Evasão	Base de Dados	Visualização das Trajetórias	Ferramenta	Algoritmo	País	Ano
Temporal predication of dropouts in MOOCs: Reaching the low hanging fruit through stacking generalization.	Wanli Xing, Xin Chen, Jared Stein, Michael Marcinkowski	Mooc, Dropout, Prediction, Algorithim, Stacking, Learning analytics	Sim	Mooc	Não	Canvas API	General Bayesian Network, (GBN), decision tree (C4.5)	Estados Unidos da América	2016
*Visualization of student learning model in serious games.	Miroslav Minovic, Miloš Milovanovic*, Uroš Šošević, Miguel Ángel Conde González	Learning analytics, Data visualization, Serious games	Não	Não, tarefas e conteúdos produzidos pelos educadores	Sim, três níveis de análise baseado na Taxonomia de Anderson	UML tool, EduGameLO (Educational Game Learning Object) and EduGameAO (Educational Game Assessment Object)	Não	Espanha e Sérvia	2015
Visual learning analytics of educational data: A systematic literature review and research agenda	Camilo Vieira, Paul Parsons, Vetrica Byrd	Visual analytics, Learning analytics, Educational data mining, Literature review	Sim	Não	Não	Não	Não	Estados Unidos da América	2018
Predictive Analytics: Another Vision of the Learning Process	David Bañeres, Montse Serra	Predictive analytics, Learning analytics, Awareness system, e- learning, Machine Learning	Sim	Open University of Catalonia	Não	WEKA, CMI-5for xAPI, JSON	J48, C4.5, EM, Naïve Bayes, SVM(Support Vector Machines)	Espanha	2018

Título	Autores	Palavras-chave	Evasão	Base de Dados	Visualização das Trajetórias	Ferramenta	Algoritmo	País	Ano
Curriculum Assessment of Higher Educational Institution Using Aggregate Profile Clustering	Satrio Adi Priyambada, Mahendrawathi ER, Bernardo Nugroho Yahya	Educational Data Mining; Clustering; Students' Learning Path; Curriculum	Sim	Mooc	Não	Weka, J48, C4.5	k-means	Indonésia e Coreia do Sul	2017
Visualizing Educational Data from Logic Tutors	Matthew Johnson, Tiffany Barnes	Educational Data Visualization	Não	Não	Sim	EDM Vis	Não	Estados Unidos da América	2010
Log Analysis in a Virtual Learning Environment for Engineering Students	Gilles Lavigne, Genoveva Gutiérrez Ruiz, McAnally-Salas, Javier Organista Sandoval	Logs, Educational technology, educational data mining, Virtual learning environment, Moodle	Não	Moodle	Sim	SPSS, Organization Risk Analyzer (ORA)	Não	México	2015
Detection of Student Modelling Anomalies	Sergey Sosnovsky, Laurens Müter, Marc Valkenier, Matthieu Brinkhuis, Abe Hofman	Student modelling, Adaptive educational system, Educational data mining, Student modelling anomaly	Não	AES Math Garden	Não	Elo Rating System (ERS)	Elo-rating algorithm	Reino Unido	2018
*A Learning Analytics Approach to Correlate the Academic Achievements of Students with Interaction Data from an Educational Simulator	Mehrnoosh Vahdat, Luca Oneto, Davide Anguita, Mathias Funk, Matthias Rauterberg	Learning analytics, Educational data mining, Technology Enhanced Learning, Process mining, Complexity, Interaction data, Educational simulator	Não	LA, Data Collector (LADC)	Sim	Disco tool	fuzzy miner algorithm	Itália e Reino Unido	2015

Título	Autores	Palavras-chave	Evasão	Base de Dados	Visualização das Trajetórias	Ferramenta	Algoritmo	País	Ano
Cluster Analysis in Personalized E-Learning Systems	Danuta Zakrzewska	Personalized e-learning system, Teaching paths, Student clustering	Não	Institute of Computer Science Technical University of Lodz dataset	Não	Spatial Clustering of Applications with Noise (DBSCAN)	Fuzzy clustering algorithm, K-means, Single Layer Clustering, Hierarchical Agglomerative Clustering	Polônia	2009
Studying situated learning in a constructionist programming camp: A multimethod microgenetic analysis of one girl's learning pathway	Katarina Pantic, Deborah A. Fields, Lisa Quirke	Computer science education, do-it-yourself media, learning trajectories, microgenetic methods, big data, Scratch	Não	Scratch Camps, MOOSE Crossing	Sim	Não	Não	Estados Unidos da América e Canadá	2016
*Student Drop- out Modelling Using Virtual Learning Environment Behaviour D	Jakub Kuzilek, Jonas Vaclave, Viktor Fuglik, Zdenek Zdrahal	Student Drop- out · Modelling Virtual learning environment Markov chains	Sim	Open University Learning Analytics dataset (OULAD)	Não	Linguagem R	Markov chain	República Tcheca	2018
*Informing learning design through analytics: Applying network graph analysis	Dirk Ifenthaler, David Gibson, Eva Dobozy	Learning experience, digital learning, Navigation sequence, Learning design	Sim	Moocs Curtin, University	Não	Curtin Challenge digital learning platform, Blackboard	Não	Alemanha e Austrália	2018
*Graph theory based model for learning path recommenda	Guillaume Durand, Nabil Belacel, François, LaPlante	Learning object recommendation system, Learnin path, Graph theory, Soft computation	Não	Apache Cassandra	Sim	Linguagem R	Não	Canadá	2013

Fonte: Elaboração própria.

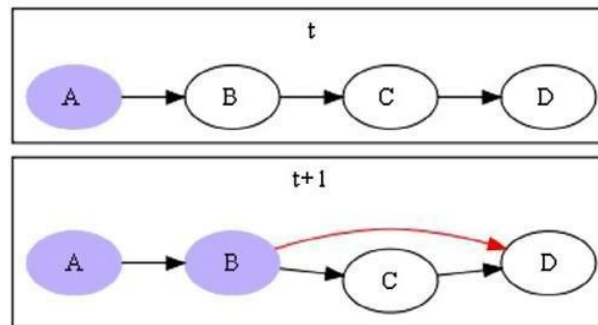




Como etapa final, foram selecionados 5 artigos que possuíam informações que se relacionam fortemente à tese (trabalhos grifados com um asterisco no Quadro 2). A relação dos trabalhos com a tese é, a forma como os dados são representados, a fonte de dados ser alguma plataforma digital que auxilia professores, o monitoramento do percurso dos alunos quando utilizam a plataforma, bem como, se existe o uso da MD conjuntamente com ao uso da plataforma ou até mesmo como funcionalidade da plataforma. Os trabalhos serão melhor descritos a seguir.

No primeiro estudo intitulado “**Graph theory based model for learning path recommendation**” de Durand, Belacel e La Plante (2013) é proposto um modelo de recomendação de objetos de aprendizagem (OA) em que o aluno projeta o seu caminho de aprendizado. Trata-se de um sistema de recomendação de aprendizagem que auxilia os alunos a encontrarem objetos de aprendizado apropriados, bem como, criar caminhos de aprendizado eficientes durante o período de aprendizado com o objetivo de adquirir competências necessárias. O modelo leva em consideração não somente a indicação de OA, mas também a sequência mais adequada ao aluno fazendo uso de grafos para representar a sequência dos recursos recomendados. O conjunto de dados e a ferramenta utilizada foi o Apache Cassandra que é um sistema de gerenciamento de banco de dados distribuído (DDBMS) de código aberto altamente escalável. Não fica claro no artigo a aplicação do modelo em algum grupo de alunos. A pesquisa apresenta relação com a tese em relação ao uso de grafos para a representação de trajetórias de aprendizagem, porém o modelo possui limitações quanto ao número de vezes em que os objetos são acessados (Figura 25), pois não existe a demonstração da informação nas imagens. O modelo não apresenta nenhum tipo de recurso em que o usuário possa acessar esse tipo de informação, como por exemplo uma aba ou caixa de texto. A ausência desse tipo de informação motivou a implementar essa funcionalidade no modelo que propomos.

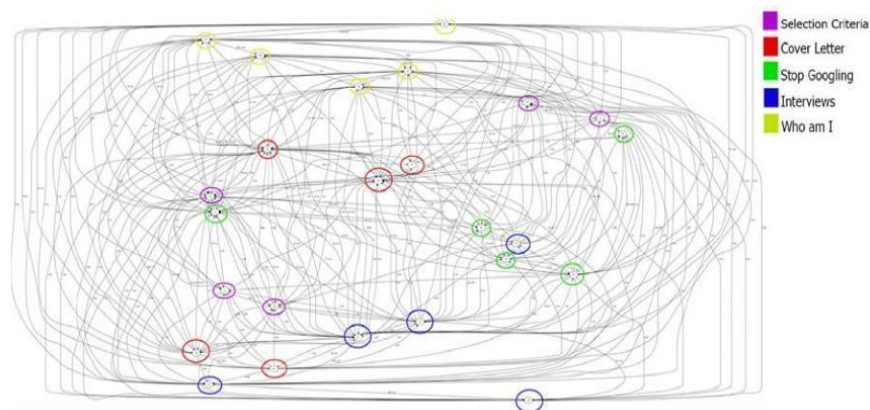
Figura 25 – Sequência de OA



Fonte: Durand *et al.* (2013).

No estudo “**Informing learning design through analytics: Applying network graph analysis.**” de Fenthaler, Gibson e Dobozy (2018) é proposto o uso do desenho dos caminhos percorridos e análise dos mesmos, utilizando um conjunto de dados da Universidade de Curtin, com auxílio da Blackboard e da plataforma edX MOOCs. Foram implementados os experimentos em 12 módulos do curso chamado Careers Challenge. Também são utilizados grafos para análise de redes e representação dos caminhos percorridos. Sendo os elementos em comum com a tese: o uso de um banco de dados de alunos, o uso de grafos, bem como, a busca de padrões para indicar caminhos de aprendizagem realizados em um AVEAs. Nesse trabalho relacionado, um fator negativo presente é a representação dos caminhos percorridos (Figura 26) ser bastante confusa quando um conjunto maior de dados é utilizado.

Figura 26 – Representação dos caminhos percorridos pelos alunos



Fonte: Fenthaler *et al* (2018).

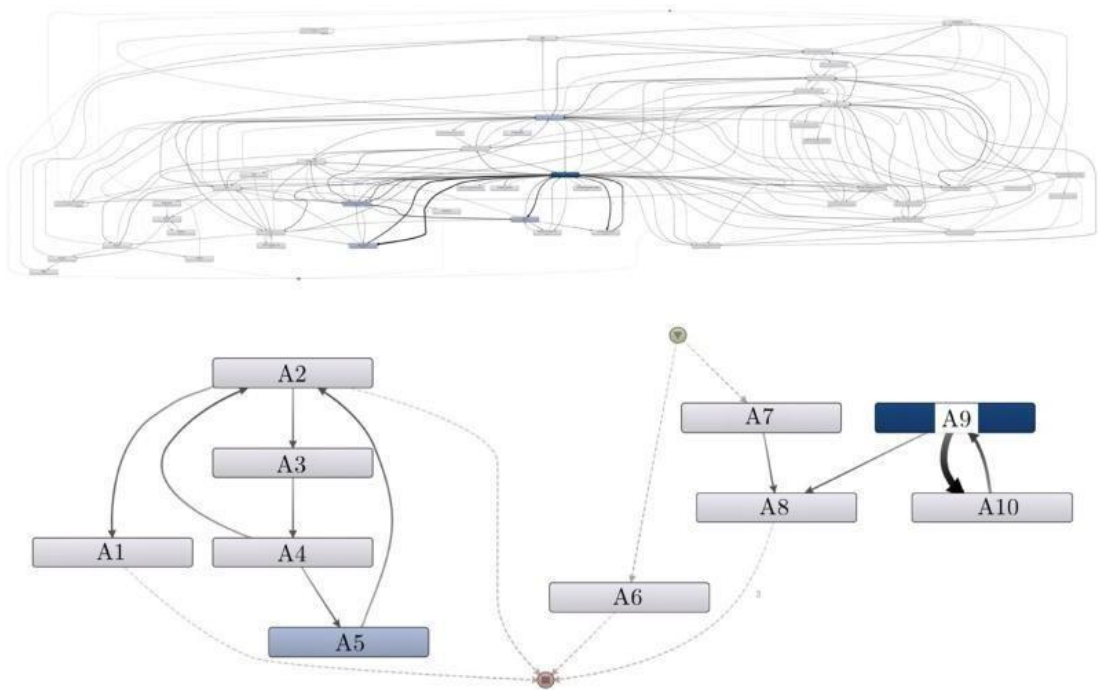


O terceiro trabalho analisado intitulado “***Student Drop-out Modelling Using Virtual Learning Environment Behaviour Data***” de Kuzilek *et al.* (2018) utiliza de dados interacionais em AVEAs para identificar padrões comportamentais de alunos com probabilidade alta de evasão. Para isso, foi utilizado um modelo de cadeia de Markov desenvolvido especificamente para o experimento.

A base de dados utilizada foi a OULAD, pertencente à Open University, localizada no Reino Unido. Os dados foram coletados por meio da Plataforma Moodle e a variável utilizada foram as notas dos 32.593 alunos entre os anos de 2013 e 2014, sendo esses dados disponibilizados ao público. A relação do estudo com a tese encontra-se no âmbito do uso de informações interacionais de alunos em AVEAs, de forma mais específica o Moodle, com o objetivo de identificar perfis de alunos propensos a evadir. O estudo não apresenta nenhum tipo de análise de trajetória de aprendizado, bem como, não faz uso de nenhum tipo de técnica de mineração de dados para detecção de evasão.

O quarto trabalho intitulado “***A Learning Analytics Approach to Correlate the Academic Achievements of Students with Interaction Data from na Educational .Simulator***” de Vahdat *et al.* (2015) teve como objetivo a criação de modelos de comportamento de alunos baseados em dados coletados por ferramentas específicas. O uso de tais dados possui como objetivo a personalização e adaptação do aprendizado para que o mesmo seja otimizado. Para que tais práticas fossem implementadas, foram utilizadas técnicas de mineração de dados educacionais como algoritmo de mineração difuso. Com o uso da ferramenta é possível gerar um modelo de processo realizado pelo aluno e quanto mais escuro o retângulo mais o aluno acessou a atividade.

Figura 27 – Representação dos acessos às atividades



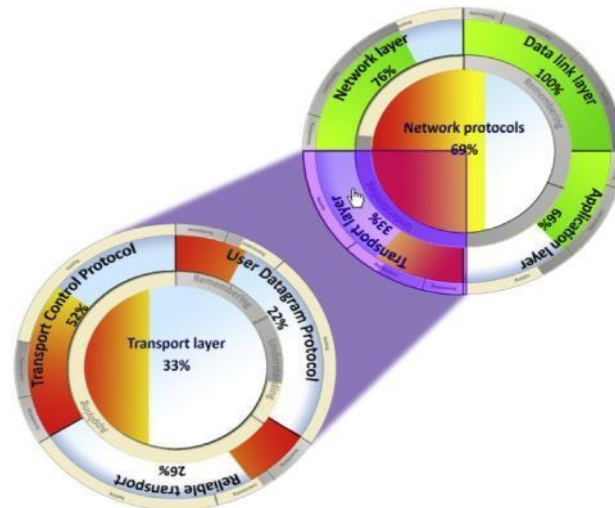
Fonte: Vahdat *et al.* (2015).

O banco de dados utilizado foi gerado a partir de seis sessões de uso de um laboratório, onde estudantes cursando o primeiro ano de Engenharia da Computação da Universidade de Gênova, utilizavam um simulador de eletrônica digital. O trabalho apresenta um avanço na análise de caminhos percorridos AVEAs, porém como fatores negativos podemos destacar o uso de um algoritmo pouco comum na mineração de dados e o uso de uma representação de trajetórias que quando usado um conjunto maior de dados a visualização não se torna tão clara. A relação com a tese se apresenta a partir de alguns elementos comuns como o uso de imagens para representação dos caminhos percorridos pelos alunos e o uso de técnicas de mineração de dados, a partir de um banco de dados originário das interações dos alunos com recursos educacionais.

O quinto e último trabalho intitulado “**Visualization of student learning model in serious games**” de Minovic *et al.* (2015) possui como objetivo o mapeamento dos caminhos de aprendizagem de alunos em um jogo sério chamado EduGame. Para a mensuração do nível de conhecimento do aluno foi construído um dashboard

chamado “escala de Anderson” baseado na Taxonomia de Bloom, com elementos relativos ao domínio cognitivo como, conhecimento, compreensão, aplicação, análise, síntese e avaliação sendo medidos.

Figura 28 – Painel de controle das atividades



Fonte: Minović *et al.* (2015).

O experimento foi realizado por um grupo de 6 educadores monitorando *online* a performance de 20 estudantes que passaram por duas etapas distintas em uma disciplina de Redes de Computadores. A ferramenta utilizada foi criada usando a mesma estrutura e dados geralmente fornecidos em ferramentas populares do LMS como o Moodle e a simulação fornece aos educadores uma visão específica sobre a dinâmica de desenvolvimento do modelo de aluno. A relação do estudo com a tese se dá por meio do uso de uma plataforma de aprendizado, onde os conhecimentos dos alunos são medidos por meio das interações com os recursos educacionais. O estudo utiliza uma representação visual em relação ao caminho percorrido, porém o acesso aos recursos não fica claro, somente o percentual de avanço em relação ao conteúdo.

Nessa perspectiva, tendo como referência os trabalhos acima elencados, onde foram apresentadas pontos positivos e negativos dos mesmos, a tese buscou apresentar os resultados de uma ferramenta de análise construída que ao mesmo tempo utilizasse as soluções propostas pelos trabalhos, bem como, tentasse de alguma forma incorporar elementos, funcionalidades e elementos que julgamos importantes e se mostraram ausentes nos estudos.

A revisão sistemática foi realizada com o auxílio da ferramenta Mendeley<sup>10</sup>, trata-se de um software acadêmico gratuito utilizado para organizar citações e referências bibliográficas. As principais funções do software são: organizar as fontes pesquisadas, editar informações de cada referência, gerar citações de forma mais rápida, formatar as referências bibliográficas conforme a norma selecionada e ainda possibilita a criação de grupos de pesquisa semelhante a uma rede social.

---

<sup>10</sup> O acesso às informações pode ser obtido através de <https://www.mendeley.com/>

## 6 METODOLOGIA

De acordo com Alves (2003, p.59), a metodologia é um instrumento do pesquisador, uma vez que é por meio da especificação dos caminhos a serem adotados que se torna possível delimitar a criatividade e definir o **como, onde, com quem, com quê, quanto e de que maneira** se pretende captar a realidade e seus fenômenos. Essa delimitação também objetiva a possibilidade de reprodução dos experimentos realizados durante o processo de construção do estudo.

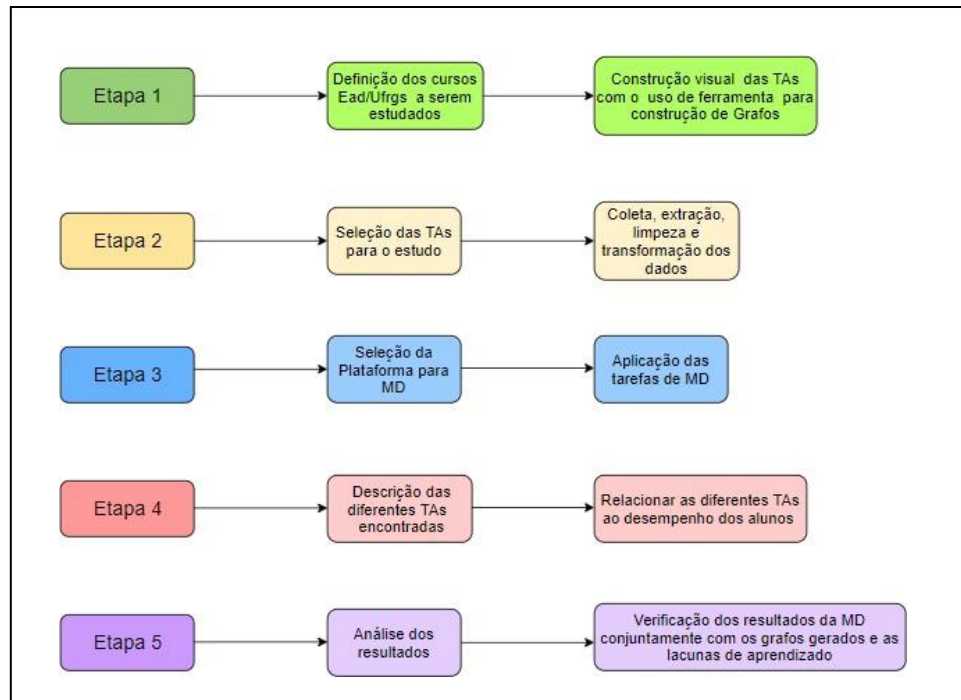
Considerando as características deste estudo, o mesmo encontra-se inserido no modelo de pesquisa descritiva que, segundo Vergara (2000, p.47), expõe as características de determinada população ou fenômeno, estabelece correlações entre variáveis e define sua natureza. De forma mais específica, busca-se neste trabalho, a partir de um conjunto de dados, descobrir e descrever padrões relativos às TAs.

Ainda segundo Gil (2010, p.184), as pesquisas descritivas têm como objetivo primordial a descrição de determinada população ou fenômeno ou, então, o estabelecimento de relações ou associações entre variáveis. Ainda segundo o autor, entre as pesquisas descritivas, salientam-se aquelas que têm por objetivo estudar as características de um grupo: sua distribuição por idade, sexo, procedência, nível de escolaridade, estado de saúde física e mental.

Assim, para se alcançar o sucesso em uma pesquisa desse porte, torna-se importante a definição de um método científico, o qual propicie um conjunto de ações sistemáticas para que seja aferida segurança aos resultados alcançados (LAKATOS e MARCONI, 2007).

Para que os melhores resultados sejam alcançados, a pesquisa foi dividida em cinco diferentes etapas (apresentadas na Figura 29, a seguir), as quais são detalhadas nas subseções seguintes.

Figura 29 – Estruturação da Pesquisa



Fonte: Elaboração própria.

## 6.1 ETAPA 1

Nesta etapa, foi realizado um mapeamento de possíveis cursos de graduação na modalidade EAD e foi realizado um contato com seus coordenadores, a fim de firmar convênios com os cursos, com o objetivo de ter acesso aos bancos de dados dos alunos, presentes no Moodle. Posteriormente, de posse dos dados, as TAs foram construídas visualmente com uso da ferramenta Moodle Data Visualizer (MDV) desenvolvida para esta tese por Paim (2022). Também nesse estudo foi utilizada a ferramenta Neo4j, que oferece escalabilidade e suporte a transações com forte consistência para processamento distribuído e paralelo em rede de comunicação de dados. As Figuras 6, 8 e 10 deste estudo são exemplos de grafos gerados com o uso da ferramenta Neo4j<sup>11</sup>.

<sup>11</sup> O Neo4J é um sistema de gerenciamento de banco de dados gráfico. Disponível em <https://neo4j.com/>

## 6.2 ETAPA 2

A partir de um conjunto de TAs construídas na etapa anterior, foram analisadas as diferentes TAs realizadas pelos alunos. Elas foram agrupadas de acordo com as suas similaridades. Foram coletados dados desses conjuntos de percursos via banco de dados do Moodle, e os dados receberam um primeiro tratamento. Dentre as ações de pré-processamento, que incluem tratamento para eliminação das instâncias repetidas e/ou valores discrepantes, normalização dos dados, foi realizada a seleção efetiva dos dados que utilizados. Essa etapa do processo tomou tempo significativo da realização do trabalho, pois dependendo da qualidade dos dados, o processo necessita ser realizado inteiramente de forma manual, não podendo ser automatizado.

## 6.3 ETAPA 3

Com os dados já ajustados, buscou-se nessa etapa selecionar a melhor plataforma. Considerando os testes já realizados por meio de simulação, a ferramenta *Microsoft Azure Notebooks*<sup>12</sup> apresentou vantagens em relação a outras ferramentas, entre elas a grande quantidade de bibliotecas disponíveis e a versatilidade da linguagem (Python). A mesma também possui a característica de ser Web, livre, de código aberto e utiliza a modalidade de computação em nuvem. Nesta etapa, foi selecionada como ferramenta principal a plataforma para MD Weka que apresentou um bom desempenho, sendo utilizada a tarefa de agrupamento nas turmas selecionadas para o estudo.

Foram realizadas nessa etapa as tarefas de MD, sendo: classificação, descrição, agrupamento e associação. A tarefa de predição foi incorporada a MDV como uma funcionalidade da ferramenta.

## 6.4 ETAPA 4

Com resultado das tarefas de MD e das TAs, construídas e selecionadas de acordo com suas similaridades, nessa etapa foi realizada a descrição detalhada das TAs, principalmente a relação da trajetória com o desempenho do aluno, bem como a

---

<sup>12</sup> A ferramenta pode ser acessada através de <http://notebooks.azure.com/>

ocorrência de desistência de alguma disciplina e o abandono do curso.

## 6.5 ETAPA 5

A análise nessa etapa consistiu em encontrar, nas TAs realizadas, possíveis lacunas em relação ao aprendizado que indiquem elementos que possam levar alunos a abandonarem as disciplinas, bem como, evadirem do curso. Também se busca a existência de um perfil comum desses alunos nas diferentes disciplinas.

## 6.6 RESULTADOS PARCIAIS

Após a definição do tema de pesquisa, foram realizados experimentos com dois *datasets* e duas ferramentas para MD (Quadro 4). Tais experimentos são melhor descritos ao longo desta subseção.

**Quadro 4 – Resultados Parciais**

Dataset	Ferramenta\ Plataforma	Tarefa Realizada	Quantidade de Alunos
UCI Machine Learning Repository	Azure Notebook	Comparação entre classificadores para predição	173.739
Especialização em Informática/PPGIE	Azure Notebook	Comparação entre classificadores para predição	385
Especialização em Informática/PPGIE	Azure Notebook	Classificadores de predição para indicação de TAs	385
Especialização em Informática/PPGIE	Weka	Agrupamento e Associação com utilização dos algoritmos <i>Apriori</i> e <i>SimpleKMeans</i> .	385

Fonte: Elaboração própria.

O primeiro experimento realizado por mim teve como objetivo realizar testes com a ferramenta *Azure notebook Microsoft* com a intenção de explorar os recursos que a mesma apresenta quanto à utilização da mesma para tarefas de MD. Para isso foi utilizado um *dataset* com origem no repositório UCI (*Machine Learning Repository*) pertencente à *Open University*, que disponibiliza de forma gratuita o uso de seus dados. As tarefas realizadas foram: manipulação de arquivos de dados, classificação



de dados, realização de tarefas supervisionadas e não supervisionadas, bem como, o uso de diferentes classificadores.

O segundo experimento foi realizado com o uso da ferramenta citada anteriormente, onde buscou-se analisar a acurácia para dados do *dataset* ainda não ajustados para MD, onde os dados foram utilizados para realizar ajustes e as calibrações dos algoritmos de forma análoga ao primeiro experimento. Foram simulados dados semelhantes ao desempenho dos alunos para que os resultados do uso dos algoritmos tivessem maior proximidade com a realidade. Cumpriu-se todas as etapas do processo DCBD necessárias para que fosse possível executar as tarefas relativas a MD. Os atributos utilizados foram os registros “notas” de 385 alunos. No terceiro experimento, foram realizados testes semelhantes ao experimento anterior, porém aumentando-se o número de atributos utilizados. Esse aumento de atributos teve como objetivo testar a precisão do modelo quanto a capacidade de predição.

Na etapa seguinte, optou-se por utilizar o mesmo conjunto de dados já ajustados fazendo uso de outra ferramenta. Com o uso do Weka (*Waikato Environment for Knowledge Analysis*)<sup>13</sup>, executou-se tarefas diferentes as executadas em testes anteriores. A ferramenta também apresentou desempenho esperado de acordo com relatos de experimentos mencionados em artigos referenciados no capítulo sobre trabalhos relacionados (Capítulo 5).

## 6.7 PLATAFORMA MOODLE

A plataforma Moodle é uma ferramenta de ensino e aprendizado amplamente utilizada por diversas instituições, que permite a professores disponibilizarem conteúdos e tarefas para os alunos, bem como a organização de fóruns que possibilitam discussões sobre as atividades. Por ser gratuito e de código livre, é uma ferramenta bastante difundida, contando com mais de 180 mil sites e aproximadamente 39 milhões de cursos ativos, espalhados em mais de 240 países, segundo seu site oficial (Moodle, 2021).

A apresentação e as funcionalidades da plataforma Moodle foram desenvolvidos baseados em uma abordagem pedagógica conhecida como socioconstrutivismo, que define que o processo de desenvolvimento e aprendizagem

---

<sup>13</sup>A ferramenta pode ser acessada através de [https://waikato.github.io/weka-wiki/downloading\\_weka/](https://waikato.github.io/weka-wiki/downloading_weka/)

humano é construído, principalmente, a partir de interações com outros indivíduos (HQ, 2022b). Devido a isso, um elevado número de componentes em uma página de curso do Moodle envolve interações com outros usuários ou a possibilidade de escrever *feedback* e comentários adicionais.

Neste trabalho, os componentes de uma página de curso do Moodle são elementos centrais na visualização e representam o segundo conjunto de elementos com os quais os estudantes interagem. Portanto, são listados a seguir os principais componentes do Moodle utilizados por cursos da UFRGS:

- **Página:** é a página inicial de uma disciplina, que usualmente contém uma seção de avisos conjuntamente com links relevantes sobre o funcionamento das aulas, como plano de ensino, cronogramas e links para FAQs. Podendo ser dividido em módulos que representam segmentações do conteúdo ou do cronograma do semestre, por exemplo;

- **Arquivo:** esse componente permite o docente fazer *upload* de arquivos que ficarão disponíveis na página para acesso dos estudantes, são identificados por um link. Esta é uma das formas mais comuns de disponibilizar conteúdos na plataforma;

- **Fórum:** é um espaço criado para a disciplina, que concentra dúvidas dos estudantes relativas ao conteúdo ou andamento das aulas. Professores, tutores e alunos podem responder aos tópicos criados;

- **Webconferência Mconf:** Mconf é um software *open source* de webconferência desenvolvido em um projeto de pesquisa na UFRGS, em 2010, que foca na integração desse software com outros sistemas utilizados (INF/UFRGS, 2022). O software é disponibilizado como um serviço pelo Centro de Processamento de Dados (CPD) da UFRGS e é integrado ao Moodle na forma de *plugin* que estende as funcionalidades base da plataforma.

- **Tarefa:** é uma seção especial do curso que inclui um espaço para descrever ao que está sendo solicitado, um espaço para adicionar arquivos ou escrever um texto, um espaço para notas, bem como, uma seção de comentários sobre o envio que pode ser utilizado como uma forma de *feedback* para o aluno;

- **Pasta:** é um agrupador de componentes para melhorar a organização da página do curso;

- **Questionário:** é uma forma de criar tarefas pelo Moodle, permite a elaboração de questões objetivas e dissertativas e a possibilidade de utilizar uma

correção automática.

- **Relatório de Notas:** espaço reservado ao desempenho do aluno, concentrando as notas do estudante relativas as tarefas e avaliações ao longo do curso.
- **Sistema:** componente que compreende várias funcionalidades secundárias para os atores do ensino, como atribuição de notas automática feita pelo Moodle, visualização de relatórios, adicionar e remover usuários à disciplina e outras atividades de caráter administrativo. Por exemplo, quando um usuário visualiza a página do curso essa ação é logada pelo sistema.

A partir da plataforma, um usuário com permissão de administrador sobre a página do curso pode realizar o *download* dos logs pelo caminho *Administration > Course administration > Reports > Logs*. Um exemplo de log pode ser verificado na Figura 30.

Figura 30 - Log de atividade do Curso

```
Hora,"Nome completo","Usuário afetado","Contexto do Evento",Componente,"Nome do evento",Descrição,Origem,"endereço IP"
"11/08/2021 16:22","ANA ALMEIDA",-,"Curso: DERAD109 - Métodos De Pesquisa Científica - Turma A (2020/2)","Sistema","Curso
visto","The user with id '157839' viewed the course with id '76611'","web,170.43.86.1
```

Fonte: Paim (2022).

Um *log* é composto por:

- **Hora:** data e hora que o evento ocorreu.
- **Nome Completo:** usuário que iniciou o evento.
- **Usuário Afetado:** usuário que recebe o evento, no caso de não ter afetado um usuário a coluna é preenchida com '-'.
  - **Contexto do Evento:** curso onde o evento ocorreu.
  - **Componente:** componente onde o evento foi executado.
  - **Nome do Evento:** nome do evento ocorrido. Por exemplo, "Curso visto" é uma visualização da página inicial do curso.
  - **Descrição do Evento:** breve descrição do ocorrido, incluindo os IDs dos usuários envolvidos.
  - **Origem:** "web", indicando que se originou da interface web; ou "cli", indicando que veio da interface por linha de comando.
  - **Endereço IP:** IP de origem do usuário que iniciou o evento.

Além dos logs, o outro dado necessário para compor a visualização é a nota final do aluno na disciplina. Ela serve como parâmetro para a identificação das diferentes trajetórias. A Figura 31 mostra um exemplo da planilha de notas em um curso no Moodle exportada como arquivo CSV. Cada coluna é a nota de alguma atividade parcial ao longo do semestre. A última coluna é o *timestamp* que indica quando foi a última vez que a planilha de notas foi exportada, e a penúltima coluna indica o Total do Curso, que é a nota em que estamos interessados<sup>14</sup>.

Figura 31 - Notas dos alunos ao longo do semestre

ANA, ALMEIDA	100	98	30	60	50	85	-	90	"70,55"	1631734635
BERNARDO, SANTOS	-	88	75	-	50	70	-	75	"66,55"	1631734635
CARLOS, SILVA	-	90	80	85	85	90	100	90	"87,55"	1631734635

Fonte: Paim (2022).

Os dados disponibilizados para a realização desse trabalho foram os *logs* e as notas de quatro turmas: DERAD109 - Métodos de Pesquisa Científica, DERAD202 - Questão Ambiental e Legislação, DERAD501 - Segurança Alimentar e Nutricional e DERAD502 - Cooperativismo e Economia Solidária. As disciplinas foram ministradas entre setembro de 2020 e fevereiro de 2021. Em relação aos usuários, as disciplinas envolveram 528 alunos, excluindo os professores, tutores e palestrantes.

<sup>14</sup> O Total do curso é o valor que aparece entre aspas na imagem.

## **7 A FERRAMENTA DATA MOODLE VISUALIZATION (MDV)**

Como parte da presente tese, foi desenvolvida a ferramenta MDV<sup>15</sup> com o objetivo de servir como apoio a professores, tutores e técnicos que venham a atuar em um curso que use a plataforma Moodle em suas disciplinas. O desenvolvimento da ferramenta teve como objetivos: representar de forma visual as interações de alunos com recursos presentes na plataforma Moodle; buscar por meio da MDV incorporar funcionalidades ausentes nas ferramentas identificadas na revisão sistemática; ser uma ferramenta útil para profissionais da educação; propor uma forma clara de representação das TAs e a possibilidade futura de integrar outras funções presentes em trabalhos de colegas de orientação. Por fim, a ferramenta deveria ser capaz de processar dados tabulares exportados da plataforma Moodle, modelar os dados em grafos e disponibilizar esses modelos ao usuário em uma interface interativa que permita a manipulação da visualização para que possa realizar uma análise da forma como os discentes utilizam a plataforma.

Para melhor compreensão e utilização da ferramenta foi organizado um guia do usuário presente no Anexo F deste documento.

### **7.1 DESCRIÇÃO DA FERRAMENTA**

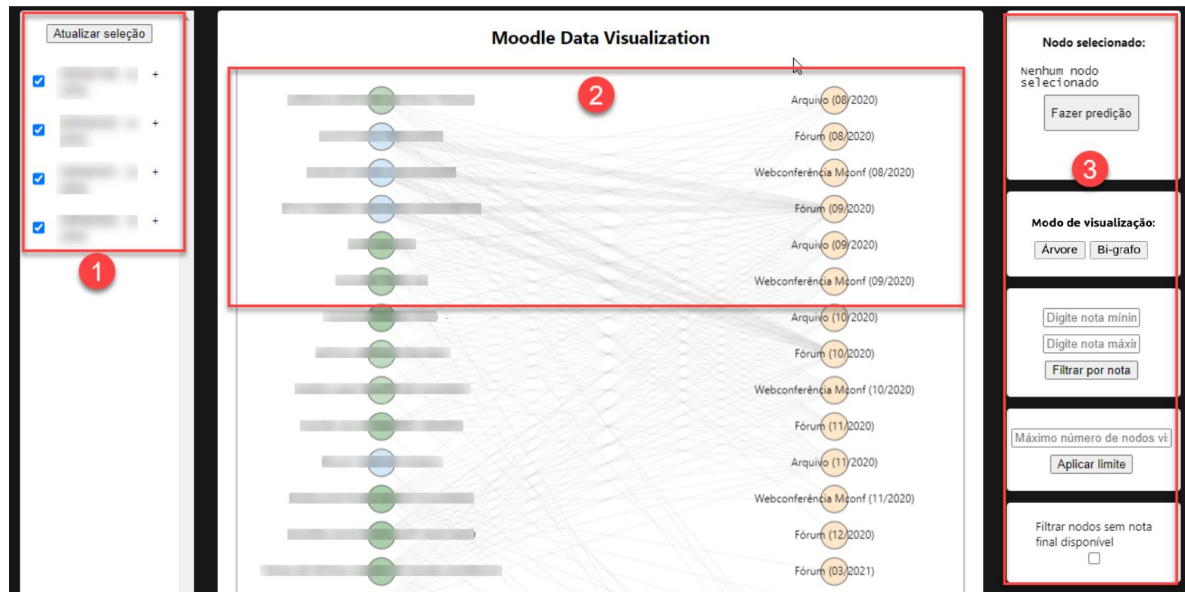
A tela principal da MDV é constituída de 3 áreas com funções específicas, sendo elas, área de seleção, área de visualização e análise e área de ajustes, ações e parametrização (Figura 32).

A área de seleção possibilita o usuário selecionar as turmas, uma turma, um conjunto de alunos ou um aluno. A área de visualização e análise estrategicamente ocupa maior área da tela principal sendo este espaço destinado a visualização dos gráficos de análise os quais representam os indivíduos (alunos, tutores e professores) e os recursos que acessaram na(s) turma(s) escolhida(s). A visualização pode ser realizada de duas formas distintas: modo bigrafo (Figura 32) e modo árvore (Figura 33), onde cada um possui sua própria estrutura criada a partir dos dados brutos, mas apenas um é exibido por vez ao usuário.

---

<sup>15</sup> O acesso à ferramenta pode ser obtido através de <http://mdv.inf.ufrgs.br/>

Figura 32 – Tela Geral MDV

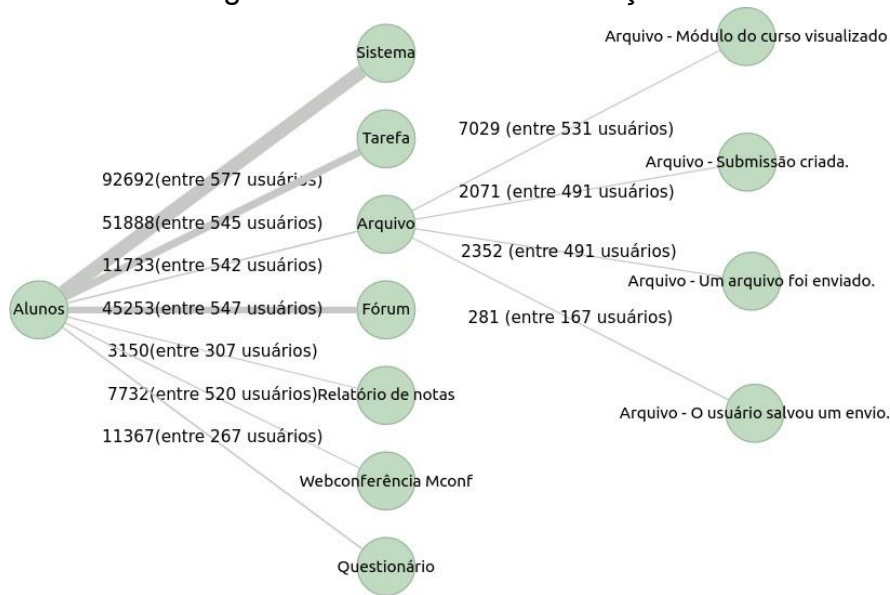


Fonte: Paim (2022).

O modo bigrafo exibe as interações de indivíduos com os recursos do Moodle, sendo que a coluna da esquerda representa os usuários selecionados. Nodos azuis indicam que o indivíduo não é um aluno e não possui nota final (professores, tutores e técnicos), os nodos verdes indicam alunos que aprovaram na disciplina. E por sua vez os nodos vermelhos indicam alunos que reprovaram. A coluna da direita indica o recurso e o mês que foi acessado. As arestas de cada nodo indicam a quantidade de interações realizadas pelos usuários com o recurso.

O modo árvore exibe informações agregadas sobre o conjunto de alunos selecionados evidenciando quais recursos foram mais acessados. Na Figura 33, o primeiro nodo representa uma centralização das interações do grupo de estudante, e as suas arestas indicam o número total de interações seguido pelo número de usuários que realizaram as interações. É possível visualizar, além do sistema, os recursos do Moodle que os alunos dessas disciplinas mais acessaram. Neste caso, são eles: Tarefas, com 51.888 interações, realizadas por 545 usuários, e o Fórum, com 45.253 interações realizadas por 547 usuários. Além dos componentes principais, é possível analisar também quais ações dentro de cada componente foram realizadas pelos usuários.

Figura 33 – Modo de visualização Árvore



Fonte: Paim (2022).

O terceiro componente da tela principal é a área de ajustes, ações e parametrização onde o usuário pode escolher a melhor forma de visualizar os dados, utilizando para isso filtros e seleções específicos. Nesta área, existe uma funcionalidade em que podem ser realizadas predições usando o módulo de IA, o módulo utiliza uma *graph neural network* para fazer predições a partir de um grafo que representa as interações de um aluno com o Moodle. A função predição ou previsão serve como um auxílio ao(s) professores, tutores e gestores para que possam identificar alunos com desempenho abaixo do esperado e, a partir dessa informação, identificar aluno(s) com propensão ao abandono do curso. Esse modelo foi treinado com os logs de interação e com as notas finais dos alunos contidos nos dados disponíveis. Como o objetivo é predizer o desempenho acadêmico, a saída poderia ser uma nota final no intervalo de 0 a 100, no entanto, o baixo volume de dados de treinamento não permite uma saída complexa com acurácia alta. Por esse motivo a saída foi simplificada com um resultado de predição "0", que representa não aprovado, ou "1" que representa aprovado (Figura 34).

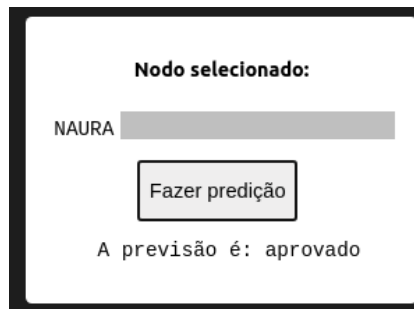
Diferentemente das demais funcionalidades da ferramenta, que funcionam sem requisições adicionais para o servidor, o modelo de predição não é transferido para o usuário conjuntamente com os arquivos estáticos. O treinamento do modelo foi realizado com os dados das quatro turmas disponíveis, totalizando 528 grafos de

estudantes disponíveis para o treinamento. No entanto, apenas 21 desses estudantes reprovaram na disciplina, o que dá uma porcentagem de 4% de reprovação. Uma distribuição como essa pode enviesar o treinamento: um modelo que prediz sempre que o aluno irá aprovar, por exemplo, terá uma acurácia de 96%. Tendo isso em vista, foi selecionado um subconjunto do grupo de estudantes que contém 60 alunos que aprovaram e os 21 que reprovaram.

Para a seleção de hiperparâmetros foi rodado o algoritmo de *grid search* contendo os seguintes valores: para *learning rate*, foram testados 0.0001, 0.001 e 0.01; para número de episódios foi testado 50, 100, 200 e 400 episódios; para número de camadas foi testado 2 e 3; para número de nodos nas camadas ocultas foi testado 6, 12, 48 e 128. O modelo escolhido foi salvo e adicionado ao servidor.

Quando o módulo de IA recebe uma requisição para prever uma entrada, a função de transformação é chamada para adaptar o grafo para o padrão da biblioteca NetworkX e em seguida é feita a transformação do grafo em uma entrada válida para o modelo de predição, com a biblioteca deepSNAP. O modelo é então executado e o resultado é enviado para o usuário.

Figura 34 – Previsão de desempenho



Fonte: Paim (2022).

O processo de desenvolvimento computacional e implementação foi realizado como tema de Trabalho de Conclusão de Curso (TCC) por Paim (2022), aluno do curso de Ciência da Computação da UFRGS, conjuntamente com seu orientador e coorientador (o autor desse trabalho). O trabalho foi apresentado à banca ao final do ano de 2022, obtendo nota máxima e informações mais detalhadas sobre a arquitetura e implementação podem ser obtidas no referido TCC<sup>16</sup>.

<sup>16</sup> O TCC pode ser acessado no seguinte endereço: <https://lume.ufrgs.br/handle/10183/251780>





## 8 EXPERIMENTOS DE AVALIAÇÃO E VALIDAÇÃO

A UFRGS, visando à ampliação do acesso ao Ensino Superior por meio do oferecimento de cursos de graduação, com base no princípio do ensino gratuito e de qualidade, criou em 2006, o Programa Especial de Graduação – PEG (Resolução nº 37/2006 do Conselho de Ensino, Pesquisa e Extensão – CEPE/UFRGS). Tal programa tem como objetivo o desenvolvimento da atuação da instituição na área da graduação, por meio de cursos que não tenham o caráter de oferta permanente, atendendo necessidades emergenciais e temporárias da comunidade, consideradas e analisadas pelas instâncias pertinentes da Universidade, podendo se dar por meio da EAD. A partir dessa perspectiva surgiu o PLAGEDER (PLAGEDER, 2018).

O currículo do curso foi construído com vistas à elaboração e avaliação de projetos, estágios e monografia considerando a realidade local nas áreas em torno de cada polo de apoio presencial. Apesar de ser um curso a distância, os alunos precisam ter disponibilidade para frequentar as atividades presenciais obrigatórias (encontros presenciais no polo com professor, tutores, estudantes e coordenador do polo) (PLAGEDER, 2018).

O PLAGEDER tem como objetivos específicos:

1) Compreender e interpretar, de maneira crítica e autônoma, o desenvolvimento agrário brasileiro, em geral, e o rio-grandense, em especial, no período contemporâneo;

2) Proporcionar instrumental teórico e aplicado para o diagnóstico e análise de realidades agrárias complexas visando a formulação de ações e políticas públicas em prol do desenvolvimento rural sustentável;

3) Disponibilizar métodos e técnicas para a realização das atividades de avaliação, planejamento, gestão de projetos e programas de desenvolvimento rural em unidades de produção agrícolas;

4) Disponibilizar técnicas e procedimentos para a avaliação, planejamento e gestão de projetos agroindustriais de pequeno e médio porte (PLAGEDER, 2018).

Tendo como cenário o curso acima descrito, foram selecionadas para o experimento, quatro disciplinas (Derad 109, Derad 202, Derad 501, Derad 502), totalizando aproximadamente 528 alunos. A partir da seleção do curso e das disciplinas, foi realizada a extração dos dados interacionais tendo como origem o banco de dados da plataforma Moodle do referido curso.

## 8.1 UTILIZAÇÃO DA MD E DA FERRAMENTA MDV PARA A ANÁLISE DAS DISCIPLINAS SELECIONADAS

Como estratégia para a análise foi utilizada a técnica de agrupamento, e por meio da ferramenta de MD (Weka) foi possível identificar três grupos distintos de acordo com suas notas. A técnica citada tem como objetivo agrupar automaticamente, por aprendizado não supervisionado, indivíduos com característica(s) semelhantes, ou seja, otimizar a homogeneidade dentro de cada grupo e otimizar as diferenças entre os grupos. Tal processo busca, a partir de tais grupos, uma realização da análise mais apropriada possível. A técnica também é conhecida por Clusterização, *Clustering, Typology, Classification Analysis ou Numerical Taxonomy*.

Esta análise inicial tem como objetivo compreender, por meio de uma análise exploratória, os perfis das turmas utilizadas para o experimento.

Como primeiro grupo foi identificado o conjunto de alunos que foram aprovados, ou seja, obtiveram sucesso<sup>17</sup> durante as disciplinas, esse conjunto de alunos se caracteriza por apresentar desempenho acima da nota 80 e com notas máximas no valor de 100. Um segundo grupo identificado foi o grupo de alunos aprovados com nota entre 60 e 80, que por sua vez também obtiveram sucesso na disciplina. Foi também identificado um terceiro grupo caracterizado pelos alunos que não foram aprovados, ou seja, não obtiveram sucesso durante as disciplinas. Sobre o terceiro grupo foram identificados dois subgrupos, o primeiro trata-se dos alunos que não foram aprovados e obtiveram nota zero e o segundo compreendem os alunos que obtiveram notas acima de zero e menor que 60.<sup>18</sup>

Como análise exploratória quantitativa, podemos verificar nas turmas elementos presentes em gráficos tipo diagrama de caixa (*boxplot*) tais como: o limite superior, limite inferior, a mediana, os valores extremos (*outliers*), a dispersão, a simetria, o terceiro e primeiro quartil (Figura 35)<sup>19</sup>. Tais métricas possibilitam, de forma visual, identificar em cada turma estudada elementos que possam identificar os grupos especificados.

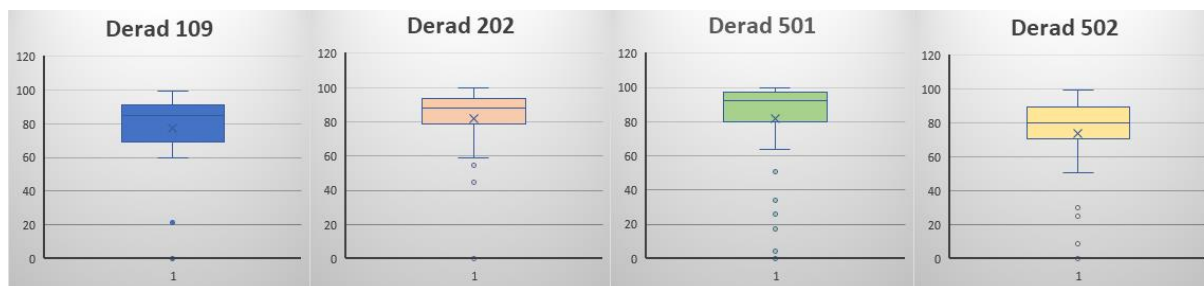
---

<sup>17</sup> Sucesso e insucesso foram os termos utilizados para caracterizar os alunos baseados em sua aprovação ou não ao final da disciplina.

<sup>18</sup> A nota mínima para aprovação na disciplina é 60.

<sup>19</sup> Os gráficos foram construídos com auxílio da ferramenta Excel Microsoft.

Figura 35 – Gráfico diagrama de caixa



Fonte: Elaboração própria.

Sobre o gráfico da Figura 35, é possível observar que existe pouca variação dos valores limites, tanto inferior como superior<sup>20</sup>, havendo uma variação maior no limite inferior na Derad 502, ou seja, as notas (linha vertical da Figura 35) dessa turma possuem uma maior variação. Existe por sua vez, nessa disciplina, uma distribuição simétrica identificada pela linha central da caixa (mediana), as demais turmas apresentam uma distribuição assimétrica. Ainda podemos verificar que os “X” presentes nos gráficos representam as médias de cada turma.

Em relação a análise das interações por meio da ferramenta MDV, foi utilizada a divisão sugerida pelo processo de MD e analisadas as 4 disciplinas, que por sua vez, foram divididas em 3 grupos de acordo com suas notas. A escolha das disciplinas também teve como critério o maior número de alunos matriculados. As duas primeiras disciplinas são ofertadas no início do curso, as outras duas ao final do curso selecionado. O que se buscou foi observar aspectos como as notas, frequência de acesso e com quais recursos os alunos tiveram uma maior interação. Também tentamos verificar a existência de padrões quanto aos acessos ao longo da disciplina. Por último, buscou-se relacionar as notas obtidas pelos alunos nas disciplinas e a frequência de acesso aos recursos.

O conjunto de notas da primeira disciplina (Derad 109) e sua distribuição é demonstrada por meio de dois gráficos. A Figura 36 representa por meio de um gráfico de dispersão o conjunto de notas (pontos azuis), onde é possível visualizar a distribuição em uma imagem de duas dimensões em que o eixo horizontal apresenta o número de alunos e o eixo vertical o valor das notas<sup>21</sup>. A utilização deste modelo de gráfico auxilia para que seja possível visualizar a distribuição das notas e a quantidade

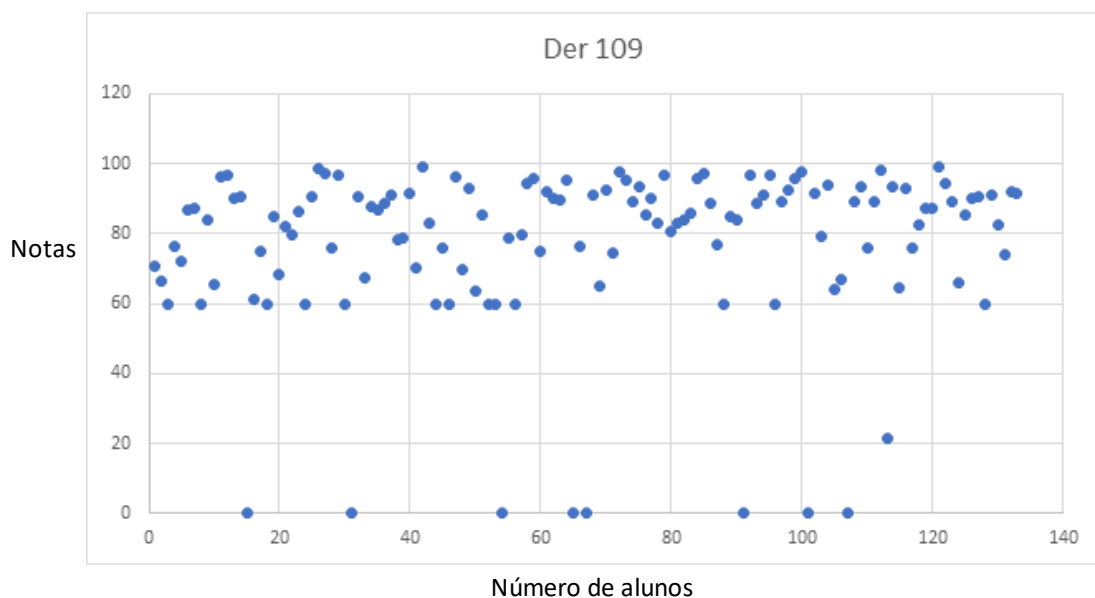
<sup>20</sup> Os valores superiores não considerando os *outliers*, representados no gráfico pelos pontos abaixo do limite inferior.

<sup>21</sup> Os gráficos foram construídos com auxílio da ferramenta Excel Microsoft.

de alunos que foram aprovados, bem como os alunos que obtiveram notas abaixo da média (60) e não foram aprovados.

A nota máxima atingida pelos alunos foi de 99,2 e a nota mínima 0, sendo a nota média 77,3 e o desvio padrão 23,5. A utilização de informações, como as anteriormente mencionadas, podem auxiliar na análise das disciplinas quanto à existência ou não de similaridade entre elas quando o aspecto observado é o desempenho dos alunos.

Figura 36 – Gráfico dispersão 01

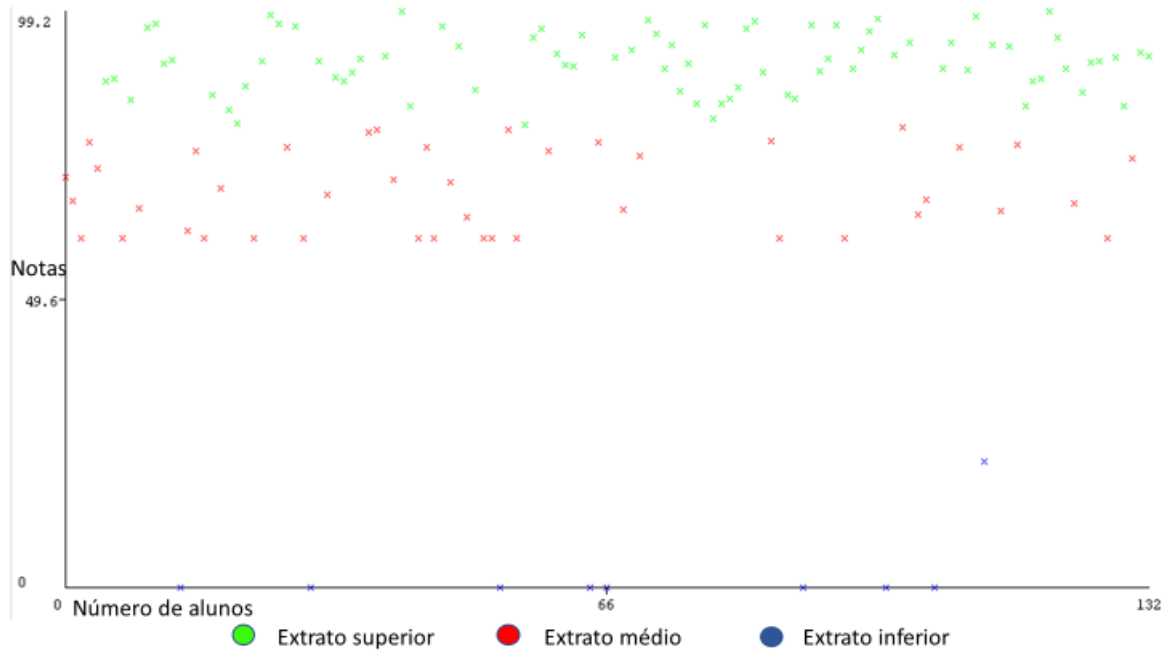


Fonte: Elaboração própria.

Em relação ao processo de agrupamento, na Figura 37 podemos identificar três grupos de alunos onde podemos observar que os alunos com notas acima de 60 pontos são a maioria, ou seja, existe uma concentração de pontos(alunos) na área superior da figura. Os grupos podem ser identificados pela variação de cores entre eles: em verdes estão os alunos pertencentes ao extrato superior, em vermelho estão os alunos pertencentes ao extrato médio e em azul os alunos do extrato inferior. Gráficos como os apresentados nessa figura podem auxiliar dando um panorama geral de uma turma e a forma como estão distribuídos de acordo com sua nota<sup>22</sup>.

<sup>22</sup> Gráfico construído com auxílio da ferramenta Weka.

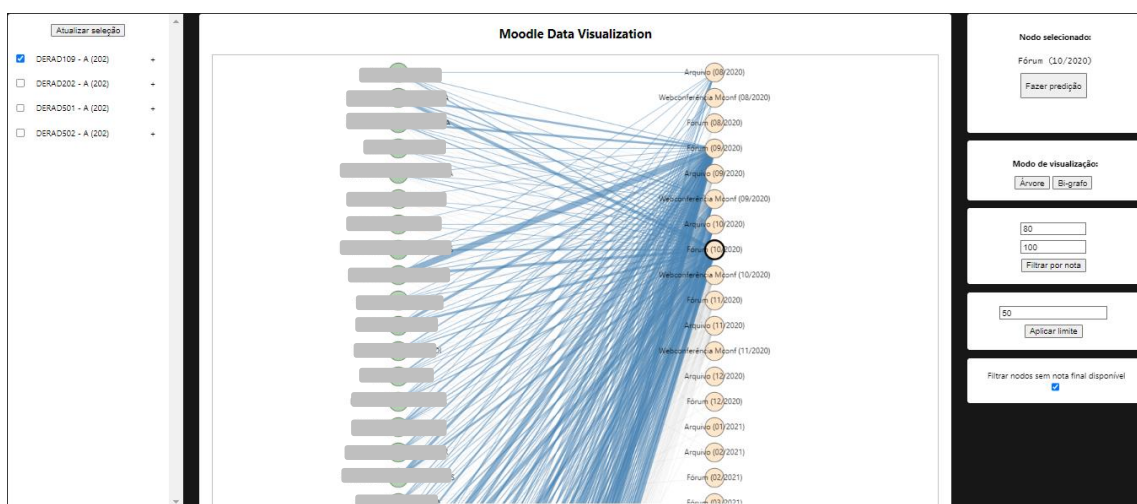
Figura 37 – Clusters notas da disciplina Derad 109



Fonte: Elaboração própria.

Por meio da ferramenta MDV, foi realizada a análise dos alunos pertencentes a esta edição da disciplina (Figura 38).

Figura 38 – Interações alunos extrato superior Derad 109

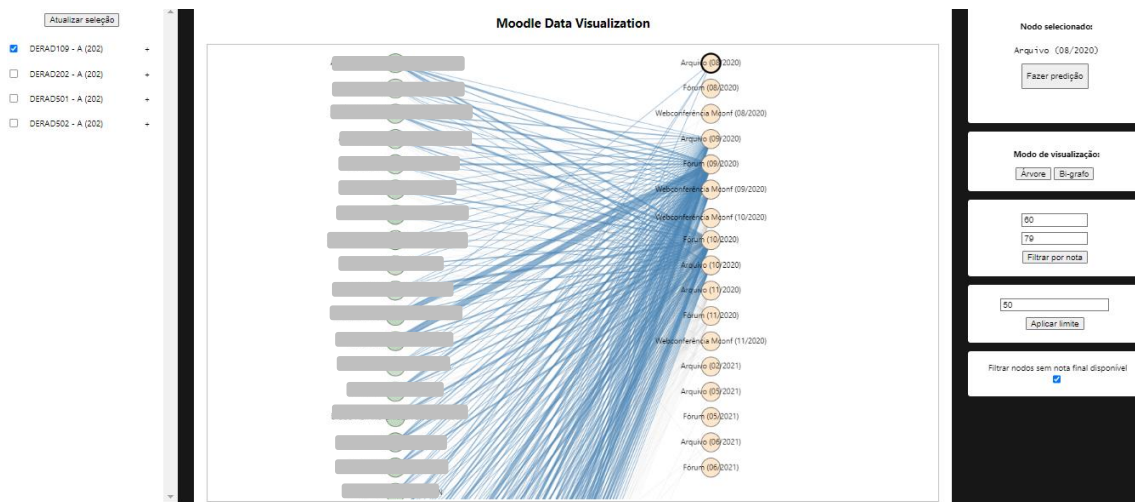


Fonte: Elaboração própria.

Em relação aos alunos, que passamos a tratar aqui como os alunos pertencentes ao extrato superior, existe uma maior interação com os recursos nos

momentos iniciais do curso, ou seja, nos três primeiros meses houve uma maior frequência de acesso como pode ser verificado na Figura 38, que correlaciona os alunos com os recursos e atividades da disciplina, na plataforma Moodle. Por essa razão, na referida imagem buscou-se dar maior evidência às interações iniciais. Podemos observar que os recursos mais acessados foram os fóruns, arquivo e webconferências.

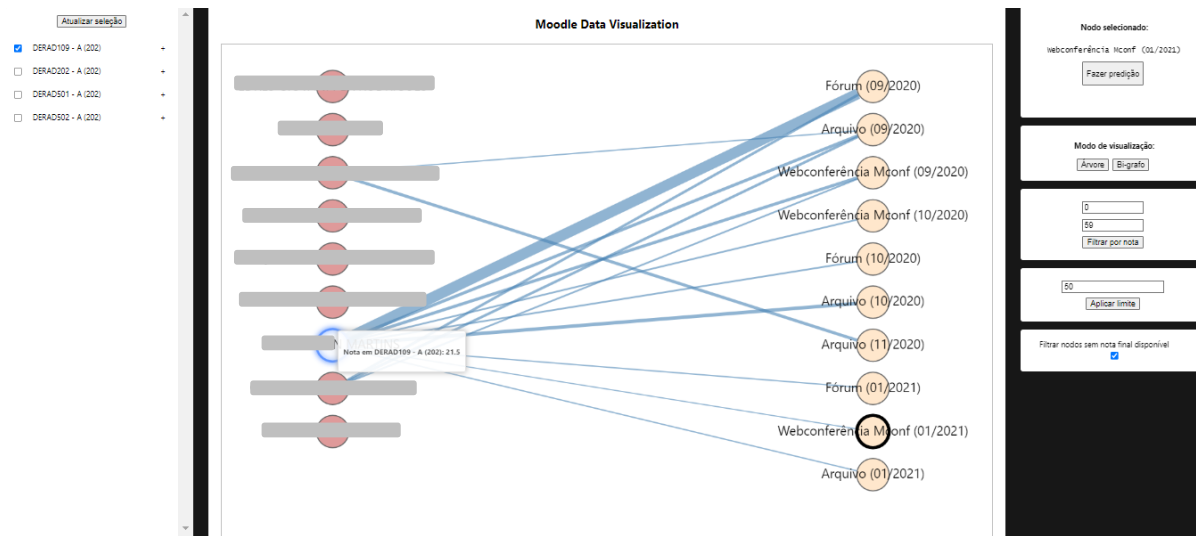
Figura 39 – Interações alunos extrato médio Derad 109



Fonte: Elaboração própria.

Em relação aos alunos do grupo extrato médio, esses apresentaram um padrão de frequência de acesso semelhante ao grupo anterior, ou seja, nos momentos iniciais da disciplina houve um maior interesse dos alunos em interagir com os recursos. Essa similaridade entre os grupos pode indicar que para alunos que obtiveram sucesso na disciplina existe um comportamento comum independente da nota que obtiveram, ou seja, alunos aprovados possuem um comportamento similar entre si quanto ao acesso aos recursos ao longo do tempo.

Figura 40 – Interações alunos extrato inferior Derad 109



Fonte: Elaboração própria.

Para a análise de prováveis reprovações, retenções e evasões este é o grupo onde direcionamos o nosso olhar e ao longo das análises tentamos compreender o seu comportamento interacional e os possíveis padrões de comportamento.

Os alunos que não obtiveram sucesso na disciplina podem ser categorizados em dois subgrupos, que são: os alunos reprovados, mas que possuem alguma nota diferente de zero e os alunos reprovados com nota zero.

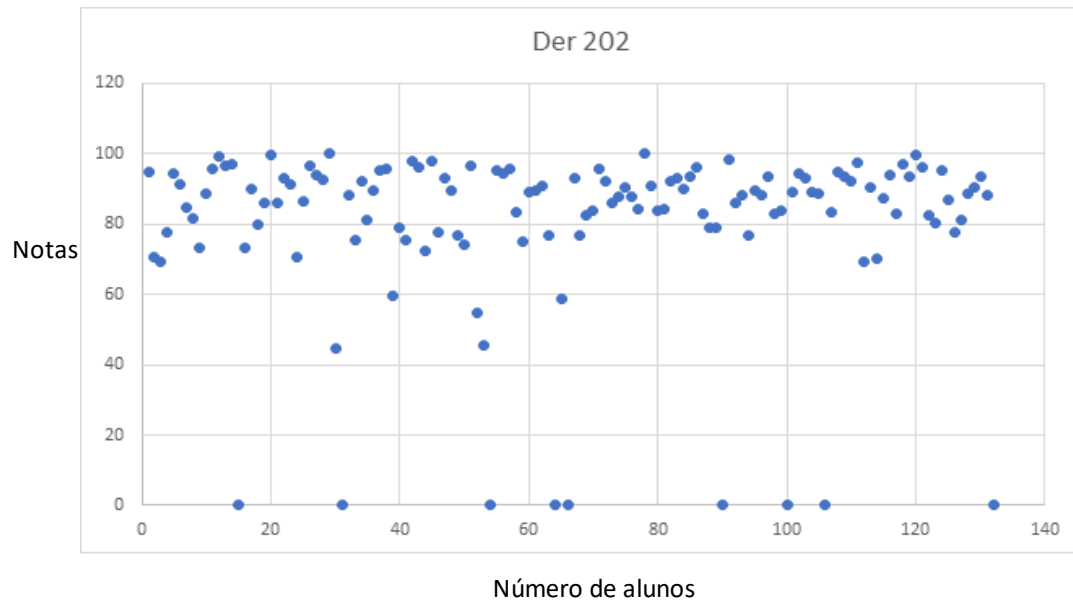
Em relação a esse segundo subgrupo, infelizmente temos poucas informações sobre suas interações por esse motivo a análise fica parcialmente prejudicada, necessitando uma entrevista com os alunos para que se possa compreender o motivo do seu baixo desempenho. Cabe aqui uma rápida observação, sobre o grupo existe a presença de três alunos que chamam a atenção. O primeiro deles foi reprovado com uma nota baixa, o segundo e terceiro foram reprovados com nota zero, sendo o ponto em comum entre eles a realização de interações. O aluno que obteve nota diferente de zero interagiu de forma mais frequente nos primeiros meses do curso, os outros dois alunos tiveram comportamentos diferentes. Um dos alunos obteve uma baixa interação no início e voltando a acessar aos recursos somente dois meses depois. Outro aluno realizou mais acessos no primeiro mês do curso e não interagindo mais ao longo da disciplina. De acordo com a literatura, este último comportamento identificado demonstra um comportamento padrão de alunos com forte propensão a desistência e evasão.

Na segunda disciplina analisada (Derad 202), de forma similar à disciplina anterior, foram utilizadas as notas dos alunos para que por meio de gráficos



podéssemos ter uma ideia inicial em relação a sua distribuição. As notas se concentram no extrato superior, ou seja, a maioria dos alunos obteve sucesso sendo, portanto, aprovados com notas acima de 60.

Figura 41 – Gráfico dispersão 02

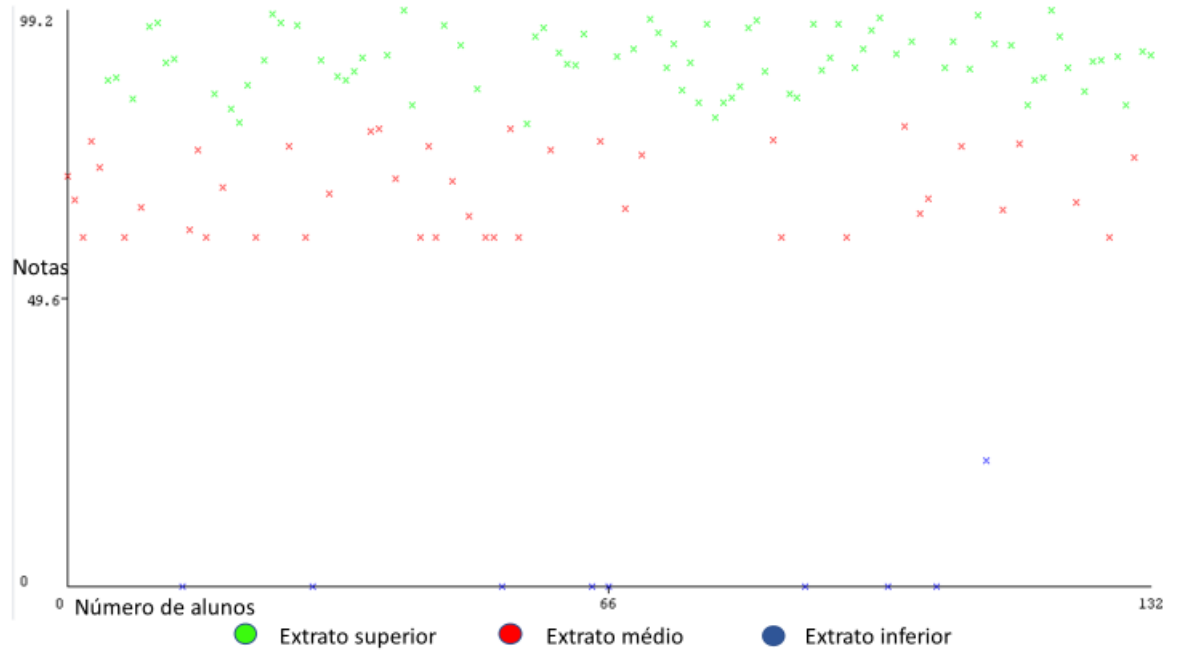


Fonte: Elaboração própria.

A nota máxima obtida foi 100, a nota mínima 0 sendo a média 88,5 e o desvio padrão 24,2.

Também foi gerado por meio da MD um gráfico onde pode ser visualizados os distintos grupos de acordo com suas notas, bem como, a forma como foram agrupados (Figura 42).

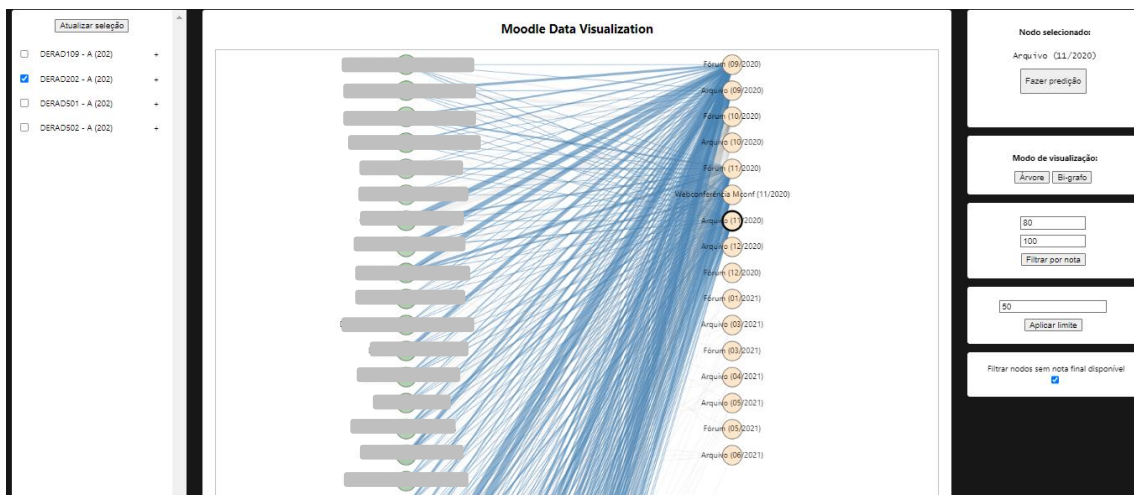
Figura 42 – Clusters notas da disciplina Derad 202



Fonte: Elaboração própria.

Em relação à disciplina Derad 202, os alunos identificados como sendo do extrato superior também apresentam um comportamento interacional bastante intenso no primeiro mês da disciplina. Os nodos localizados ao lado direito (recursos) estão organizados por data, por meio dessa ordem temporal é possível verificar que os primeiros nodos na parte superior representam os primeiros recursos disponibilizados no início da disciplina.

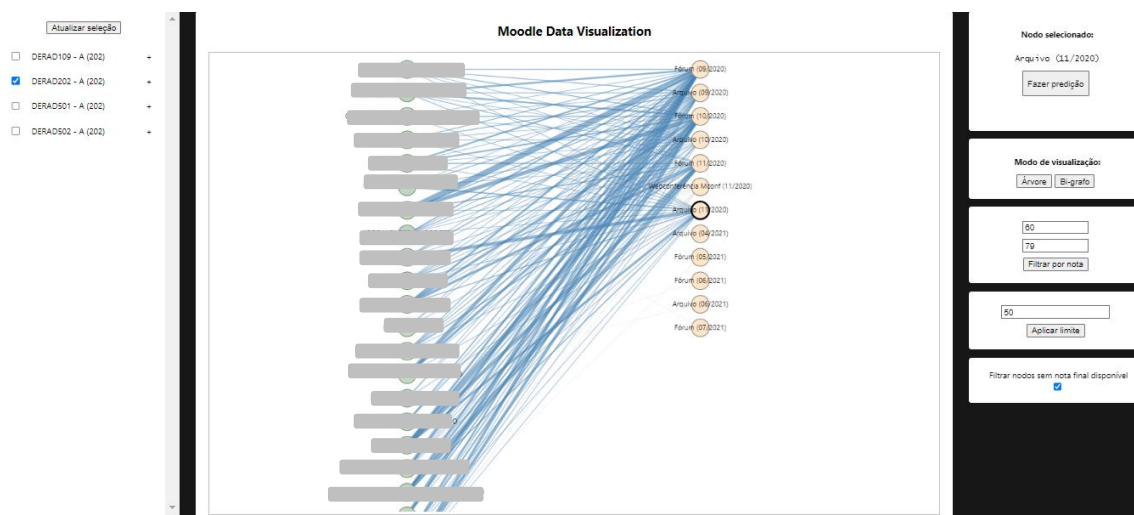
Figura 43 – Interações alunos extrato superior Derad 202



Fonte: Elaboração própria.

Todos os alunos, em algum momento, interagiram com algum recurso, para uma minoria não existe uma relação direta entre a quantidade de interações com o seu desempenho. Porém, a maioria dos alunos desse grupo apresenta uma significativa frequência de acesso a recursos, sendo assim podemos afirmar, com auxílio da Figura 43, que existem evidências que apontam a interação como algo importante para esse grupo de alunos. De forma semelhante ao mesmo extrato de alunos da disciplina analisado anteriormente, o grupo de alunos se mostra mais propenso a interagir com os recursos no início do curso<sup>23</sup> e apresentando um crescente desinteresse. Tal desinteresse segundo os dados obtidos não afeta de forma relevante seu desempenho ao longo da disciplina.

Figura 44 – Interações alunos extrato médio Derad 202



Fonte: Elaboração própria.

De forma semelhante ao grupo anterior, também é notado aqui uma alta frequência de acesso aos recursos logo no início da disciplina ocorrendo a sua redução ao longo do tempo.

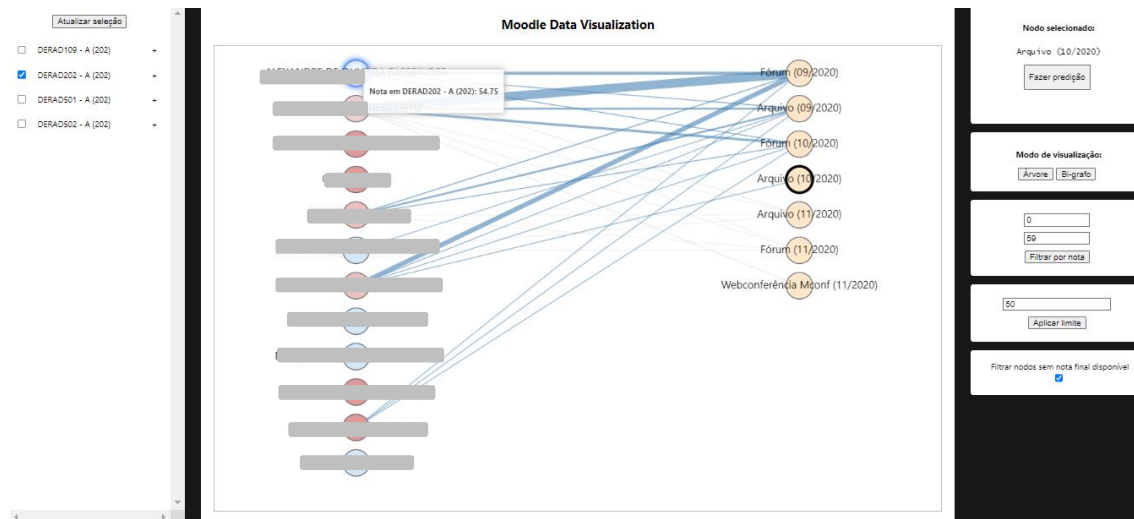
Nessa turma, para o grupo identificado no extrato inferior, algumas observações podem ser realizadas. A primeira delas é uma limitação técnica da ferramenta em não filtrar os nodos finais sem nota, esta foi uma falha detectada já com a ferramenta construída e que será corrigida para a próxima versão. Em relação aos alunos, novamente identificamos dois grupos distintos, os alunos com nota

<sup>23</sup> Os nodos dos recursos são apresentados de acordo com a ordem cronológica, sendo os nodos no topo da imagem os com data mais antiga.

diferente de zero e alunos nota zero. Para os alunos com nota zero, notamos que somente um deles realizou algumas interações ao início da disciplina e não voltou a interagir. Outro grupo com nota zero, não realizou nenhuma interação com os recursos. Por sua vez, o grupo que possui nota realizou diversas interações e algumas com uma frequência significativa. Um fenômeno que podemos observar, em relação a este último grupo, é a proporcionalidade das notas dos alunos em relação as suas frequências de acesso, ou seja, mesmo os alunos não sendo aprovados os que obtiveram notas maiores realizaram um maior número de interações. Essas interações além de serem proporcionais às notas, para os alunos com notas maiores ocorreram interações não somente no início da disciplina, mas ao longo dos meses.

A possibilidade de verificação destes comportamentos dos alunos pode servir como uma informação valiosa na tentativa de recuperação dos alunos. No exemplo desta turma, os alunos com notas não ficaram com suas notas muito longe da média mínima para aprovação. Havendo então interesse por parte do curso, professores e alunos, algumas ações corretivas poderiam ter sido implementadas para evitar a reprovação dos alunos. Informações como estas, obtidas por meio da ferramenta MDV é o principal objetivo pedagógico e reconhece a sua utilidade nas tarefas docentes.

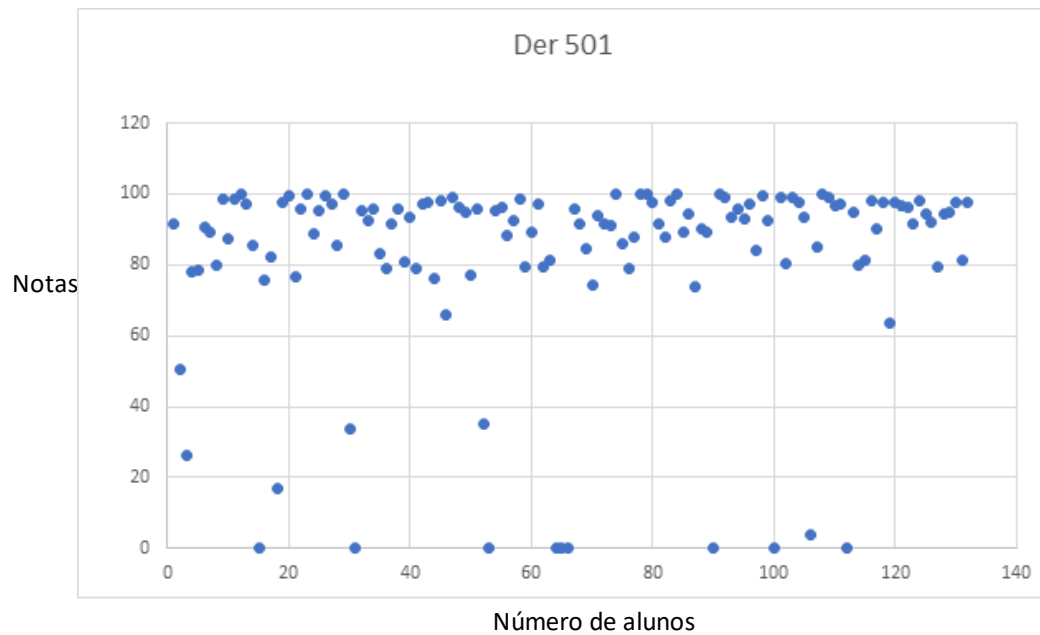
Figura 45 – Interações alunos extrato inferior Derad 202



Fonte: Elaboração própria.

A terceira disciplina analisada (Derad 501) também apresentou uma maior concentração de notas no extrato superior.

Figura 46 – Gráfico de dispersão 03

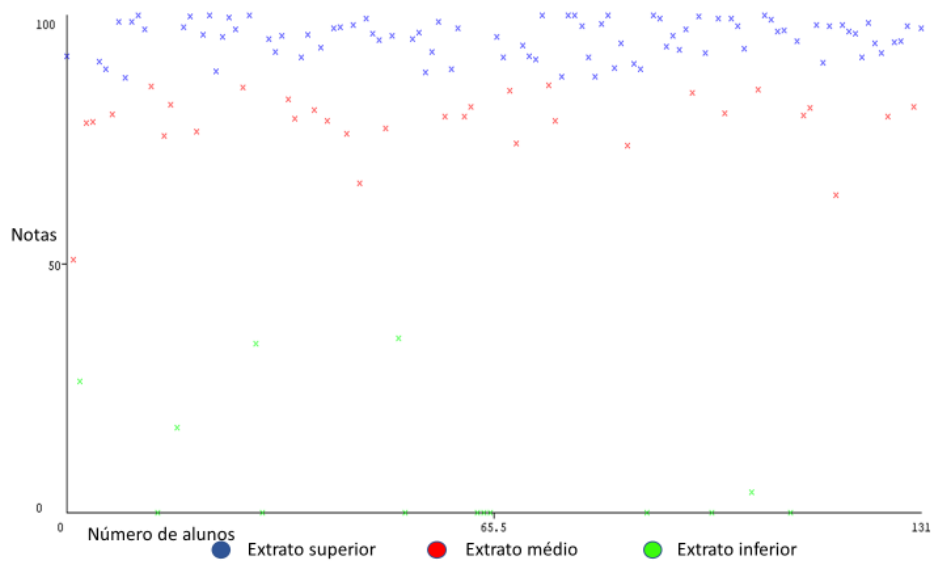


Fonte: Elaboração própria.

A maior nota atingida foi 100, por sua vez a menor foi 0. A nota média 82 e o desvio padrão 27,2.

Com auxílio da Figura 47 é possível verificar como estão divididos os 3 grupos de alunos de acordo com suas notas.

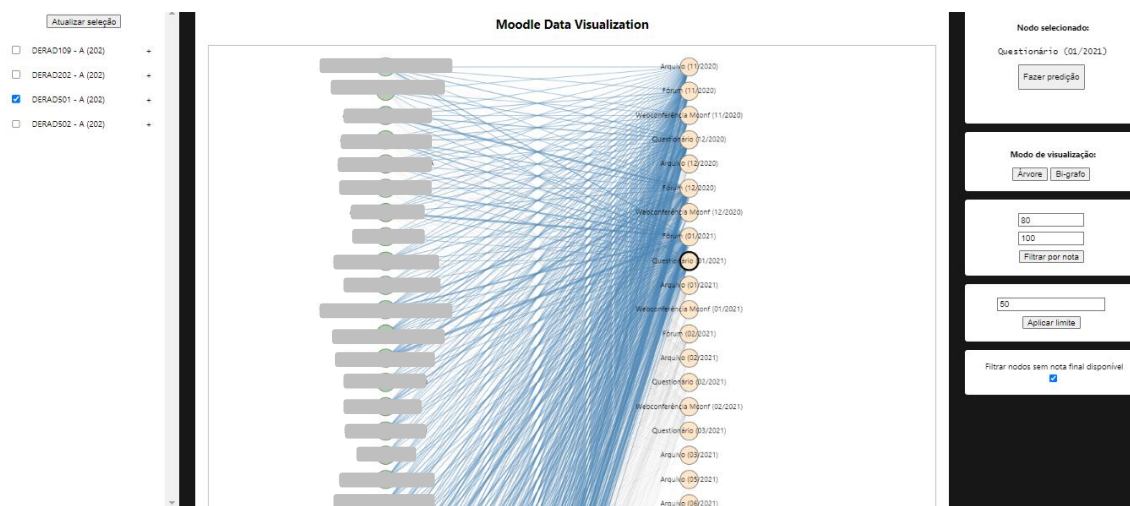
Figura 47 – Clusters notas da disciplina Derad 501



Fonte: Elaboração própria.

Para esta disciplina também foram analisadas as interações dos alunos com os diferentes recursos. A seguir são melhor apresentadas tais interações.

Figura 48 – Interações alunos extrato superior Derad 501

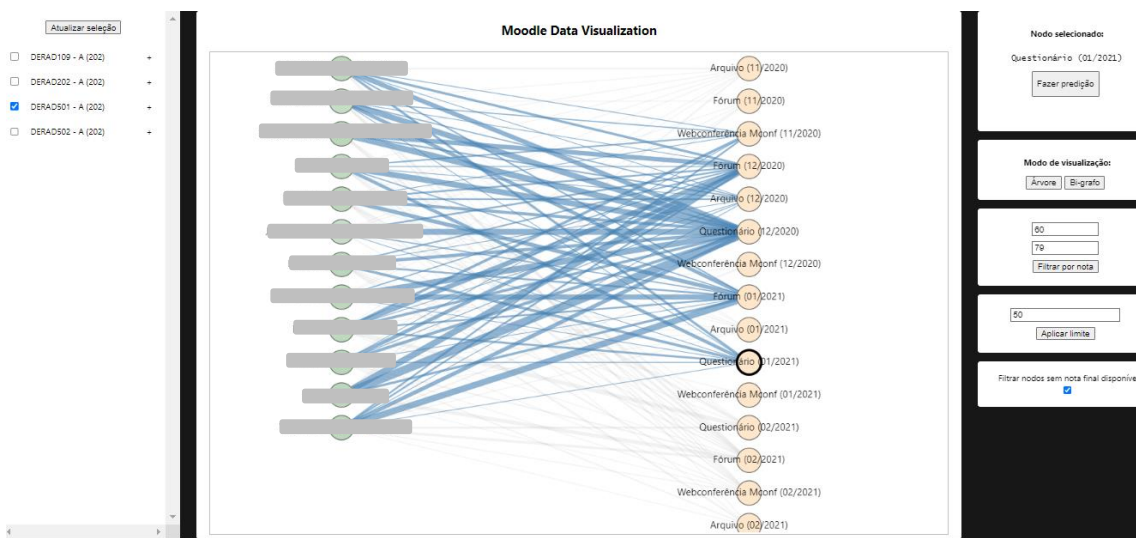


Fonte: Elaboração própria.

Novamente, podemos observar que para os alunos com bom desempenho, notas entre 80 e 100, existe um nível de interação significativo já nos primeiros meses da disciplina e reduzindo os acessos ao longo do período.

O grupo considerado intermediário dessa disciplina apresenta uma característica interessante, em relação as turmas analisadas este é o extrato que apresenta o menor número de representantes (Figura 48).

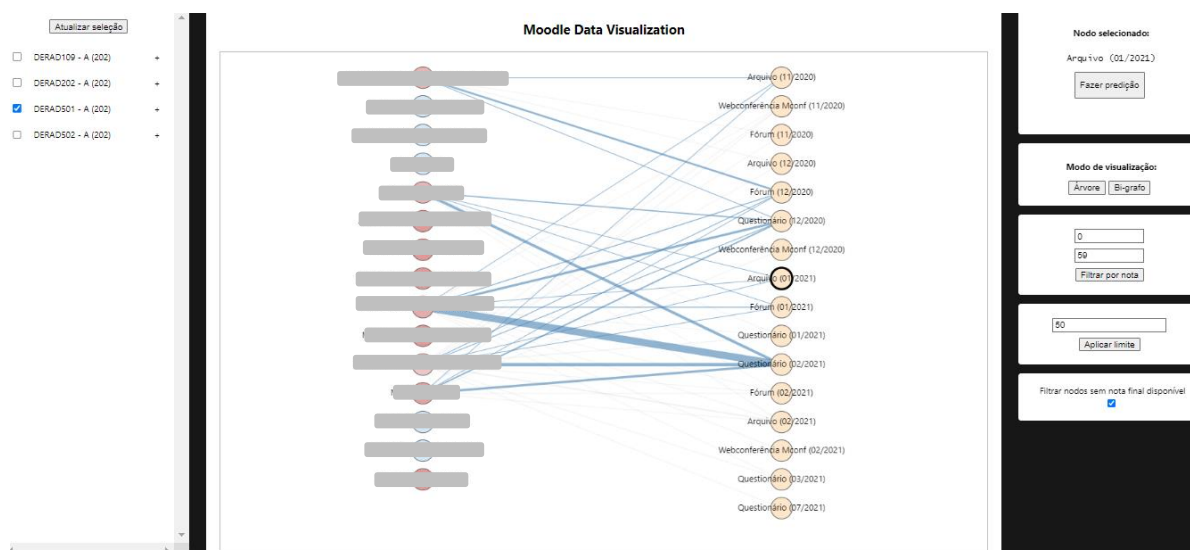
Figura 49 – Interações alunos extrato médio Derad 501



Fonte: Elaboração própria.

Outra característica que podemos observar é em relação a distribuição dos acessos ao longo do tempo apresentando interações distribuídas de forma mais equilibrada ao longo da disciplina com uma frequência de acesso significativa.

Figura 50 – Interações alunos extrato inferior Derad 501



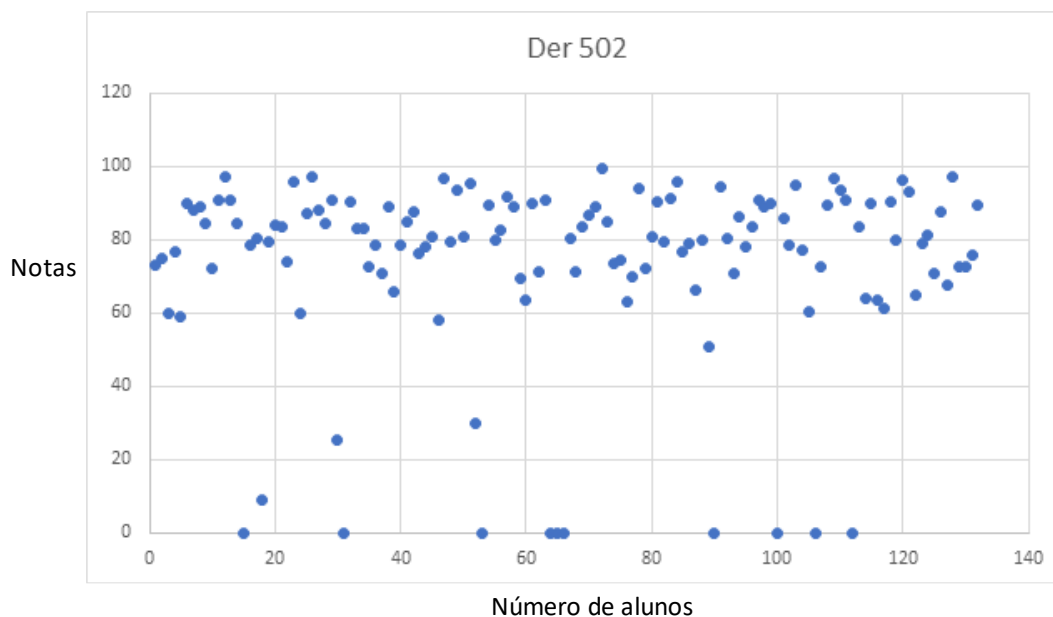
Fonte: Elaboração própria.

Os alunos da disciplina identificados como do extrato inferior também apresentam dois subgrupos, com nota zero e com notas abaixo de 60 diferente de zero. Em relação aos alunos com notas diferentes de zero, os mesmos apresentaram em seu conjunto notas inferiores aos alunos do mesmo grupo em outras disciplinas analisadas.

Em relação à forma como foram acessados os recursos, não apresentam um padrão específico, podemos identificar alunos que acessaram recursos presentes ao longo da disciplina. Em relação à frequência de acesso, um aluno apresentou uma quantidade significativa a um recurso (questionário), não resultando em uma maior nota quando comparado ao desempenho dos demais alunos presentes neste subgrupo.

Como última disciplina analisada (Derad 502), a partir do gráfico de dispersão podemos ter uma ideia geral de como estão distribuídas as notas da turma. A nota máxima da turma foi 99,5 sendo nota mínima 0. A média da turma foi 84,8 e o desvio padrão 25,2.

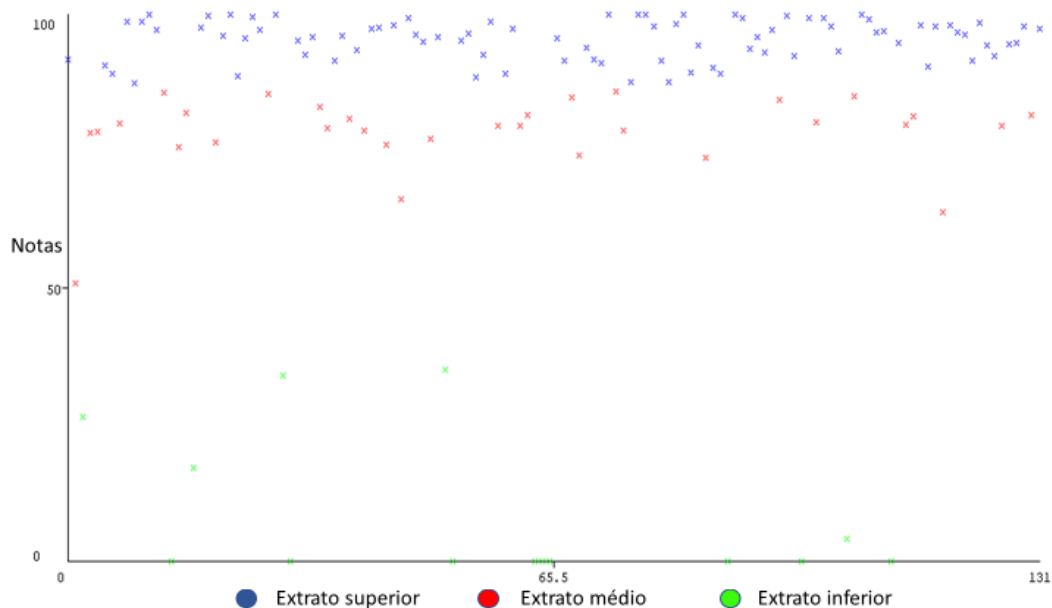
Figura 51 – Gráfico de dispersão 04



Fonte: Elaboração própria.

A identificação por meio de agrupamento nos auxilia novamente a identificar quais os grupos presentes e a forma como estão distribuídos de acordo com o seu desempenho.

Figura 52 – Clusters notas da disciplina Derad 502



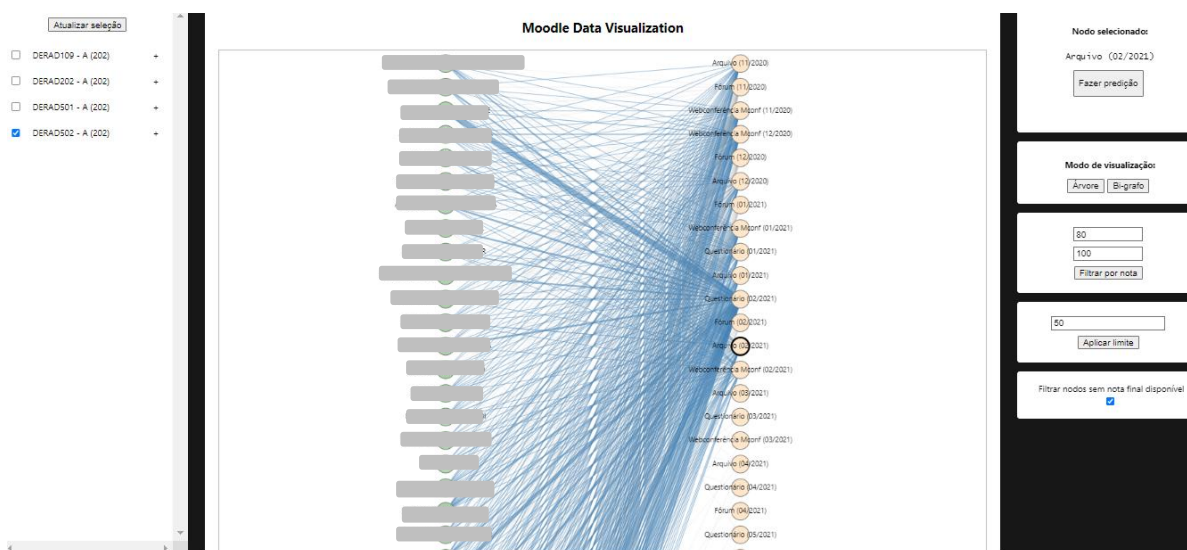
Fonte: Elaboração própria.

A ferramenta MDV possibilita por meio de dados coletados de AVEAs, gerar



informações visuais. Como realizado com as disciplinas anteriores, a ferramenta também foi utilizada e se buscou a análise e obtenção de informações relevantes quanto a turma.

Figura 53 – Interações alunos extrato superior Derad 502



Fonte: Elaboração própria.

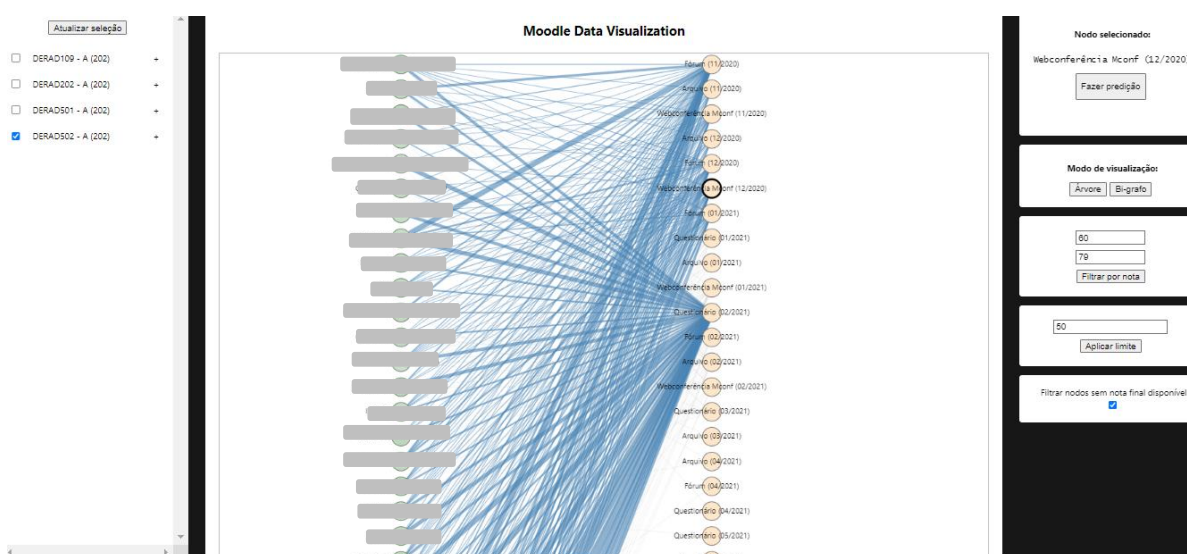
A última disciplina a ser analisada apresenta algumas peculiaridades que podem ser observadas. O comportamento interacional apresenta similaridade as turmas já vistas, porém podemos destacar que para o(s) aluno(s) com as maiores notas a proporcionalidade em relação a notas fica bastante evidente. Este dado é algo importante pois, se essa afirmação é verdadeira, uma representação visual em que o número de arestas é mais evidente pode nos levar a acreditar que o desempenho deste grupo de alunos é melhor. Ou seja, de uma maneira rápida, e sem a necessidade de consultar tabelas com as notas, essa informação é disponibilizada para os agentes responsáveis pelo acompanhamento dos alunos. Servindo, então, essa informação como subsídio para as futuras decisões por parte dos envolvidos com o curso.

Os alunos do extrato médio também apresentam algumas características interacionais interessantes. Os acessos se deram com maior frequência em dois recursos, como podemos verificar na Figura 53. O primeiro trata-se de um fórum no primeiro mês da disciplina e o segundo um questionário no mês de fevereiro. Tais atividades contaram com a participação da maioria dos alunos, senão com a sua

totalidade, variando somente a frequência de acesso entre os alunos. Cabe aqui uma especulação em relação a este fenômeno, talvez as duas atividades tenham um caráter avaliativo, decorrendo deste condicionamento os significativos acessos aos referidos recursos.

Para esse grupo não é possível fazer uma relação direta da nota dos alunos com a quantidade de acessos aos recursos, necessitando de mais dados para tirarmos conclusões mais definitivas.

Figura 54 – Interações alunos extrato médio Derad 502



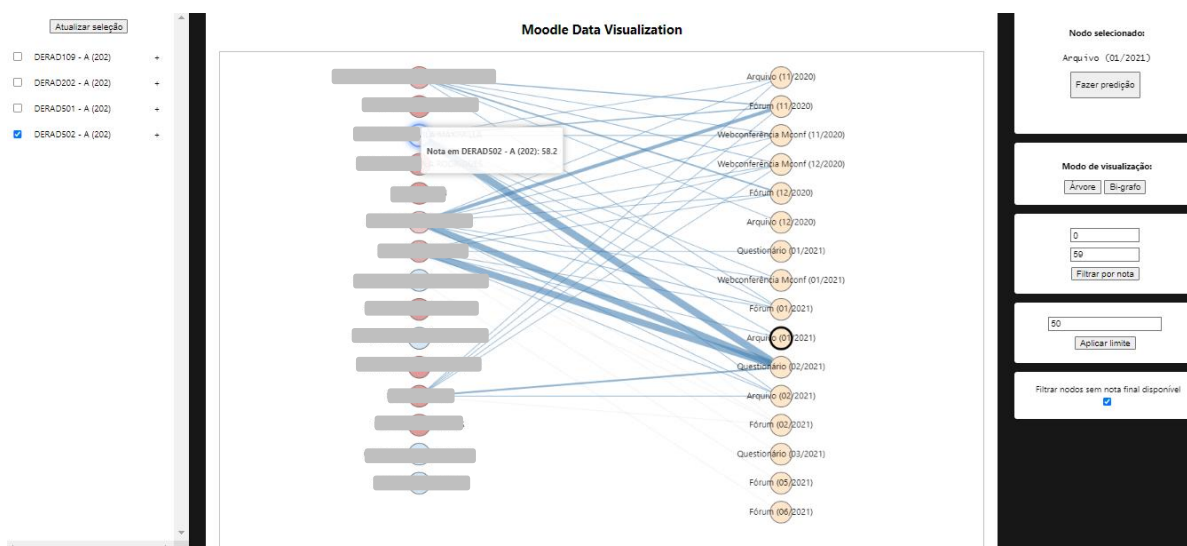
Fonte: Elaboração própria.

Para essa disciplina, e de forma mais específica os alunos que obtiveram um menor desempenho, foi possível observar a impossibilidade de uma melhor análise em relação aos alunos com nota zero. Para esse grupo, em um primeiro momento, a ferramenta não apresenta uma utilidade relevante, pois a mesma com os dados selecionados não gera informações úteis. Para não descartar completamente o uso da MDV em relação a esse grupo de alunos, a mesma pode identificar os alunos com nota zero e a ausência de interações por parte dos mesmos.

Notamos aqui um comportamento análogo ao da disciplina analisada anteriormente em que mesmo com o uma frequência de acesso significativa (2 alunos) esta não foi suficiente para a aprovação dos alunos. Percebemos uma relação proporcional positiva entre a quantidade de interações e as notas observadas, ou seja, para esse grupo é possível observar que mesmo com desempenho baixo, os alunos quando comparados entre si, indicam que os que possuem melhores notas

apresentam uma maior frequência de acesso aos recursos, como pode ser visualizado na Figura 55. A ferramenta apresentou uma limitação na filtragem dos nodos não sendo possível selecionar somente os alunos.

Figura 55 – Interações alunos extrato inferior Derad 502



Fonte: Elaboração própria.

Ao fim deste capítulo, após a experimentação da ferramenta com as disciplinas e turmas selecionadas, buscamos trazer algumas considerações.

As turmas selecionadas pertenciam a um mesmo curso de graduação variando em relação ao período escolhido para a seleção das turmas. As duas primeiras turmas fazem parte dos semestres iniciais e as outras duas turmas do meio do curso para o final. O critério para escolha das turmas em diferentes períodos do curso teve como intenção observar a existência de diferenças e similaridades entre as turmas.

Em relação aos alunos foi possível observar as interações de 528 alunos com diferentes recursos.

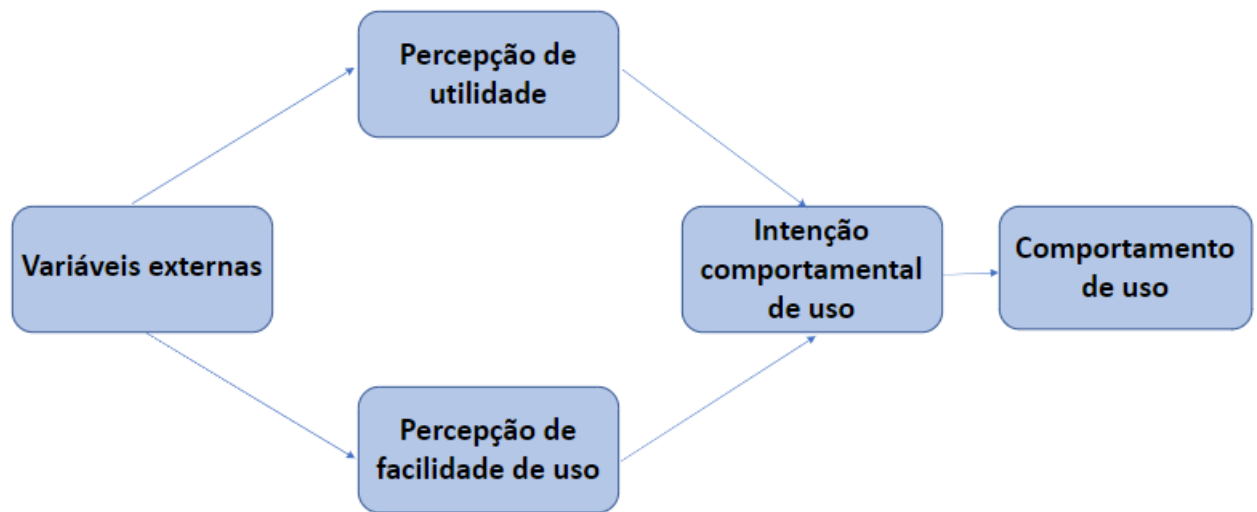
## 8.2 AVALIAÇÃO DA FERRAMENTA UTILIZANDO O MODELO TAM

Como forma de avaliação da MDV, foi aplicada uma ferramenta de consulta em forma de questionário com os(as) professores(as) docentes do curso em estudo, em

que se buscou a percepção quanto à utilidade e a facilidade de uso. O modelo de questionário está fundamentado na Teoria de Aceitação da Tecnologia, também conhecido como TAM (*Technology Acceptance Model*). O modelo foi proposto por Davis (1989, p. 319) sendo uma derivação do modelo TRA (*Theory of Reasoned Action*) de Icek *et al.* (1967). A Teoria da Ação Raciocinada é uma teoria geral sociopsicológica/comportamental utilizada para explicar um conjunto de comportamentos humanos, como o voto, exercícios e o uso de preservativos. A teoria sustenta que o comportamento de uma pessoa é uma função da sua intenção comportamental que é determinada pela sua atitude e pelas suas crenças sobre as expectativas de outra pessoa, ou seja, as crenças sociais normativas. Davis (1989) propôs o TAM para focar no porquê dos usuários aceitarem ou rejeitarem a tecnologia da informação e como melhorar a aceitação, oferecendo, desse modo, um suporte para prever e explicar a aceitação.

Dessa forma, o modelo está fundamentado em dois construtos: a **percepção de utilidade** e a **percepção de facilidade de uso**. A Figura 56 apresenta o fluxo do processo representado no modelo de instrumento de pesquisa proposto. Pode-se perceber que ambos construtos mediam completamente os efeitos das variáveis externas, como características do sistema, processo de desenvolvimento, treinamento, na intenção de uso (DAVIS, 1989, pp. 319). Segundo o autor, o objetivo do modelo TAM é focar no porquê dos usuários aceitarem ou rejeitarem a tecnologia da informação, bem como, como melhorar a aceitação, oferecendo, desse modo, um suporte para prever e explicar a aceitação. Ainda segundo o mesmo autor (DAVIS, 1989), o modelo é útil para detectar o porquê da não aceitação de um sistema ou tecnologia pelos usuários e, conseqüentemente, implementar os ajustes necessários.

Figura 56 – Modelo TAM



Fonte: Adaptado de Davis (1989).

**Percepção de utilidade (*Perceived Usefulness*)** – representa o grau em que uma pessoa acredita que o uso de um sistema particular pode melhorar o seu desempenho.

**Percepção de facilidade de uso (*Perceived Easy of Use*)** - é o grau em que uma pessoa acredita que o uso de um sistema de informação não necessite de esforços significativos para sua utilização.

**Intenção comportamental de uso (*Behavioral intention of use*)** - as variáveis externas<sup>24</sup> formam a intenção comportamental de uso ou de não uso da tecnologia por meio do seu impacto entre a percepção de utilidade e a percepção de facilidade de uso, por isso um dos principais efeitos do TAM é a de fornecer uma base para o rastreamento dessas variáveis (Yarborough e Smith, 2007). Em outras palavras, poderíamos dizer que as variáveis externas fornecem uma melhor compreensão do que influencia a utilidade percebida e a facilidade de uso percebida. No presente estudo, podemos considerar como variável externa, por exemplo, o Manual do Usuário MDV.

**Comportamento de uso (*Usage behavior*)** - é a medida relacionada diretamente ao uso da tecnologia durante um tempo determinado. Existe uma estreita

<sup>24</sup> Podem ser consideradas variáveis externas por exemplo: o design do sistema, características funcionais do sistema, a estrutura do sistema, manuais de treinamento e suporte ao usuário.

relação entre intenção comportamental de uso e o comportamento de uso, busca-se então validar ou não a primeira, a partir das observações da segunda variável.

### 8.2.1 ANÁLISE E DISCUSSÃO DOS DADOS

Como instrumento de coleta de dados foi aplicado um questionário utilizando o Google Forms<sup>25</sup>, que teve como público-alvo docentes do curso Plageder, foram contatados 48 professores por meio eletrônico com link de acesso a ferramenta MDV, bem como, senha de acesso e manual de usuário da ferramenta. A dinâmica de testes se deu na seguinte ordem: os participantes deveriam ler o manual do usuário e posteriormente por meio do link disponibilizado acessar a ferramenta. Acessando a ferramenta realizariam de forma exploratória o uso das diferentes funcionalidades, posteriormente deveriam responder a um questionário relativo as funcionalidades da ferramenta e ao manual do usuário. Foram também agendados encontros coletivos e individuais via vídeoconferência, para que os participantes da pesquisa pudessem tirar dúvidas quanto ao acesso e a utilização da ferramenta e manual. A escolha de perfil do primeiro público se deu em decorrência da familiaridade dos docentes com o curso, com os alunos e com a plataforma Moodle. Houve uma baixa adesão por parte dos docentes do curso em realizar testes e responder ao questionário, como consequência foi decidido modificar o perfil dos usuários ampliando também a docentes de outros cursos de graduação e pós-graduação. A mudança de público ocorreu então ampliando à doutorandos do Curso de Pós-graduação em Informática na Educação – PPGIE, este novo público é constituído por docentes do ensino superior e tecnológico e pesquisadores de áreas diversas do conhecimento.

O questionário (Apêndice 1) é constituído por 20 perguntas fechadas e 2 perguntas abertas. Como forma de resposta as perguntas fechadas, cada questão pode ser respondida em uma Escala Likert com 7 pontos que variam de *discordo completamente* até *concordo completamente*. As duas perguntas abertas compreendiam aspectos relativos a sugestões de melhorias em relação a ferramenta e o manual de usuário. Como número final foram obtidas 20 respostas às questões

---

<sup>25</sup> Google Forms é um aplicativo de gerenciamento de pesquisas. Os usuários podem usar o Google Forms para pesquisar e coletar informações sobre outras pessoas e também podem ser usados para questionários formulários de registro.

fechadas e 11 em relação às questões abertas.

A seguir, são disponibilizados os principais resultados e comentários fornecidos pelos respondentes.

## 8.2.2 Questões fechadas

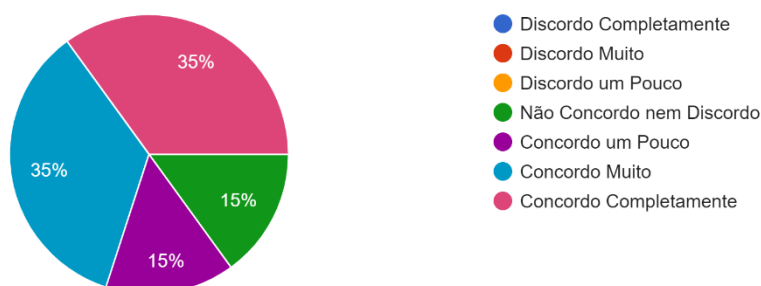
Para melhor compreensão, as perguntas fechadas foram separadas em 4 blocos, sendo eles: a percepção do usuário quanto a aplicabilidade da ferramenta em seu trabalho como docente, a percepção quanto à utilização da ferramenta, a percepção quanto às funcionalidades específicas e a utilidade da ferramenta quanto à identificação de alunos e suas interações.

### 8.2.2.1 Percepção do usuário quanto a aplicabilidade da ferramenta

A primeira pergunta questiona o usuário sobre a redução do tempo para executar tarefas quando utiliza a ferramenta.

Figura 57 – Ferramenta comparada a outras ferramentas

1- A utilização desta ferramenta no meu trabalho permite-me realizar as tarefas mais rapidamente do que se as executasse em outros produtos semelhantes.  
20 respostas



Fonte: Elaboração própria.

O percentual preponderante de respostas, ou seja, 85% apontaram como “concordo” quanto ao ganho de tempo proporcionada pela ferramenta durante o seu dia a dia como docente. Tal resultado representa um retorno positivo por parte dos usuários, pois um dos objetivos do desenvolvimento de uma ferramenta como a MDV

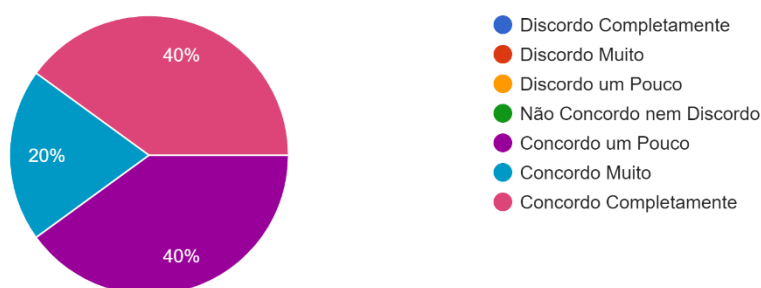
é também permitir ganhos de tempo aos docentes tendo em vista que a alta carga de trabalho dos docentes. Essa realidade acaba por impedir a utilização deste tipo de ferramenta, pois em uma relação custo-benefício quanto ao tempo, o docente acaba abrindo mão do uso da tecnologia em suas tarefas docentes, caso perceba que ela toma muito seu tempo. Ainda em relação a este aspecto da ferramenta, o percentual elevado de respostas positivas é corroborado pela resposta do Sujeito 11 que aponta o ganho de tempo como uma vantagem quanto a utilização da ferramenta.

A questão seguinte solicita ao usuário se posicionar quanto ao benefício da ferramenta em relação ao seu trabalho.

Figura 58 – Ferramenta em relação ao benefício

2- A utilização desta ferramenta no meu trabalho aumenta o meu benefício, interesses ou vantagens.

20 respostas



Fonte: Elaboração própria.

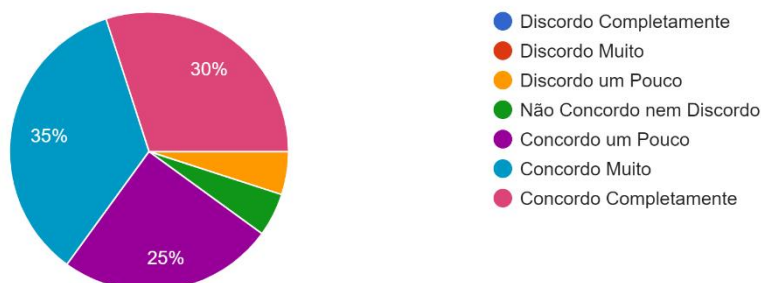
A respostas obtidas também apresentaram como características uma boa aceitação por parte dos usuários. O total de respostas fica localizado entre “concordo um pouco” e “concordo plenamente”, não ocorrendo nenhuma resposta negativa em relação ao questionamento das características da ferramenta. Podemos concluir então que que ela aumenta benefícios, interesses ou vantagens por parte dos docentes.

De forma semelhante, a pergunta anterior o que se buscou com a pergunta é relacionar a ferramenta com o trabalho docente.



Figura 59 – Ferramenta facilidade para o trabalho

3- Utilizar esta ferramenta torna o meu trabalho mais fácil.  
20 respostas

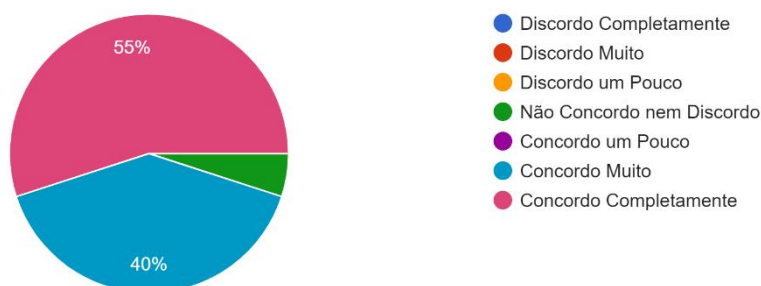


Fonte: Elaboração própria.

O maior percentual das repostas aprovou a ferramenta em relação a mesma tornar o trabalho docente mais fácil, onde 90% das respostas apontaram a facilitação do trabalho como uma função importante. O resultado quantitativo pode ser reforçado pelo depoimento do Sujeito 7 que afirma que a funcionalidade da ferramenta torna o trabalho de professores e tutores mais fácil.

Figura 60 – Utilidade da ferramenta no trabalho

4- Considero esta ferramenta útil no meu trabalho.  
20 respostas



Fonte: Elaboração própria.

As respostas da questão 4 apontam de forma significativa a utilidade da ferramenta no trabalho como docente, a partir desse resultado positivo verificamos que um dos objetivos do desenvolvimento da ferramenta foi atingido, sendo esse objetivo da mesma, ser algo útil para o trabalho docente.

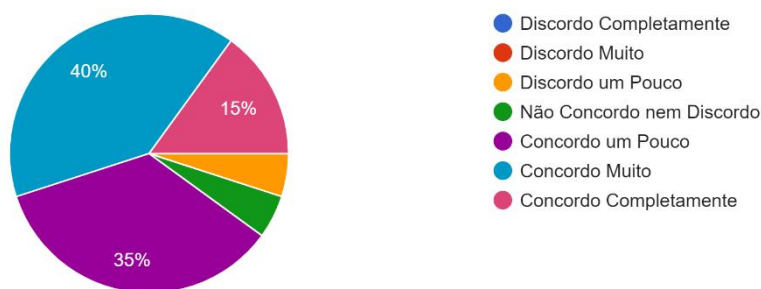
### 8.2.2.2 Percepção do usuário quanto a utilização da ferramenta

Nesta subseção é possível verificar quanto os usuários da ferramenta atribuem em grau de importância a utilização da mesma.

Figura 61 – Facilidade de operação da ferramenta

5- Foi fácil para mim operar esta ferramenta.

20 respostas

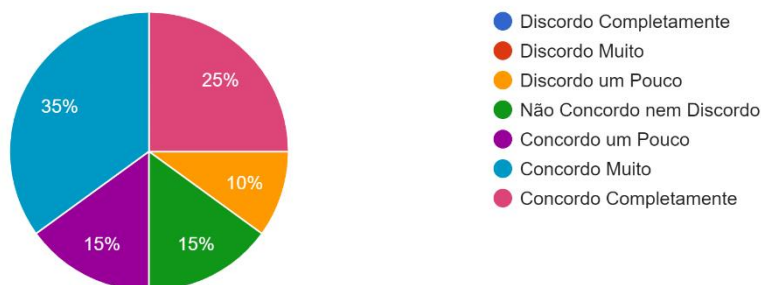


Fonte: Elaboração própria.

A maior parte dos usuários respondeu que quanto a operacionalização da ferramenta, essa se mostrou fácil. Obteve-se um percentual de 90% de aprovação quanto a esta característica, havendo somente 10% de respostas relativas a “discordo um pouco”. Então, é possível afirmar com base nesse percentual que a ferramenta apresenta facilidade quanto ao seu uso mesmo para docentes que não estejam familiarizados com a utilização deste tipo de ferramenta.

Figura 62 – Facilidade de execução

6- Acho fácil que esta ferramenta faça o que eu preciso.  
20 respostas

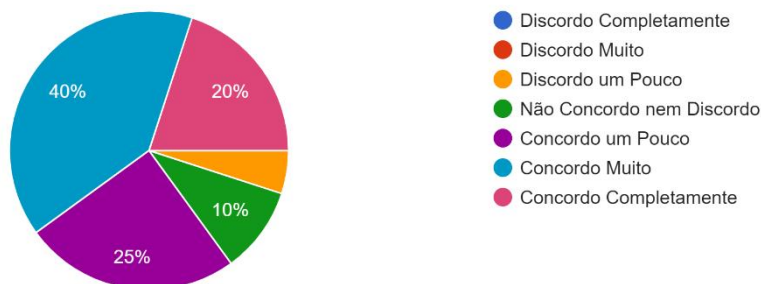


Fonte: Elaboração própria.

A questão 7 busca identificar como foi a interação com a ferramenta.

Figura 63 – Interação com a ferramenta

7- A minha interação com esta ferramenta foi clara e compreensível.  
20 respostas



Fonte: Elaboração própria.

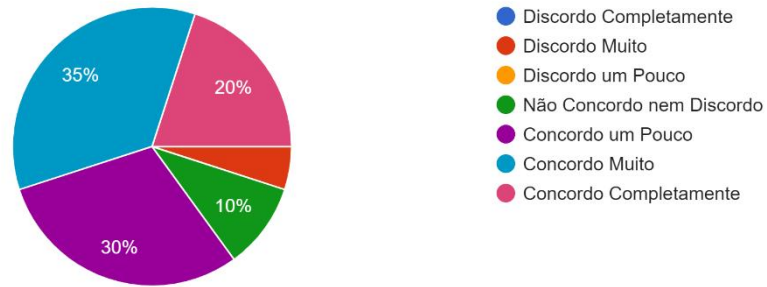
As respostas expressam, em sua maioria, que a interação com a ferramenta é clara e compreensível. A partir dos resultados, é possível afirmar que a ferramenta apresenta facilidade de interação com o usuário. Chama a atenção o percentual de 10% de respostas não concordando nem discordando, bem como, 5% discordando um pouco quanto a clareza e a facilidade de compreensão quanto a interação. O percentual de 80% de respostas favorável aponta algo importante, pois quando foi discutida a construção da ferramenta esses eram aspectos identificados como

fundamentais para que a utilização da ferramenta fosse algo intuitivo. A questão 8 também possui relação com a avaliação em relação a interação do usuário com a ferramenta.

Figura 64 – Flexibilidade e interatividade da ferramenta

8- Acho que esta ferramenta é suficientemente flexível e interativa.

20 respostas



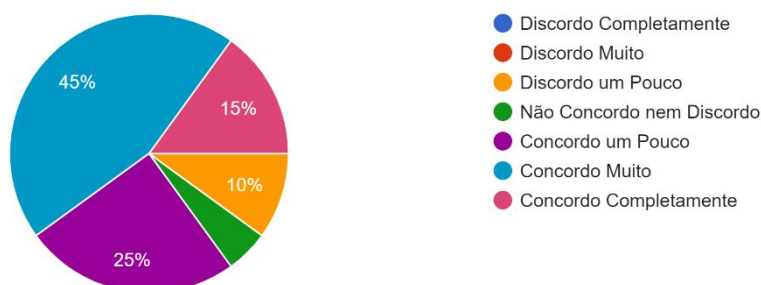
Fonte: Elaboração própria.

De forma semelhante às respostas relativas à questão 7, quanto à flexibilidade e interatividade, a ferramenta obteve 85% das respostas concordando que em relação a esses atributos a ferramenta se mostrou sólida. Despertou atenção, o percentual de 5% em relação a opção “discordo muito” se possível seria interessante localizar esse usuário(s) para tentar entender o motivo de sua resposta.

Ainda sobre a facilidade de uso, a questão 9 busca identificar a percepção de facilidade de uso do usuário a partir das suas habilidades quanto ao uso de ferramentas digitais. Obteve-se um percentual significativo de aprovação quanto a essa facilidade de uso, compreendendo então que com o conhecimento que os usuários já possuem, é possível utilizar a ferramenta não sendo necessários conhecimentos técnicos específicos, além do manual disponibilizado.

Figura 65 – Habilidade para utilizar a ferramenta

9- Com as minhas habilidades, foi fácil utilizar esta ferramenta.  
20 respostas

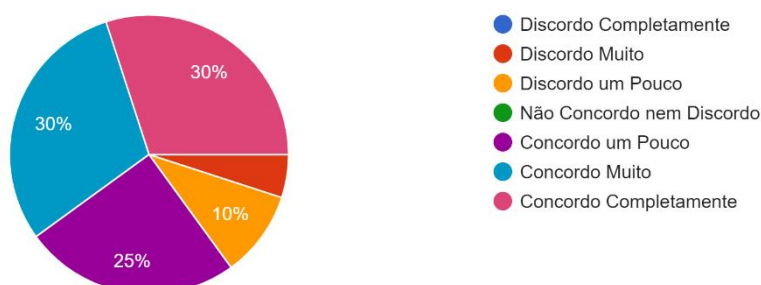


Fonte: Elaboração própria.

As questões 10 e 12 buscam identificar a percepção quanto aos aspectos visuais apresentados pela ferramenta.

Figura 66 – Elementos visuais da ferramenta

10- Considero que a ferramenta possui elementos visuais agradáveis e intuitivos.  
20 respostas



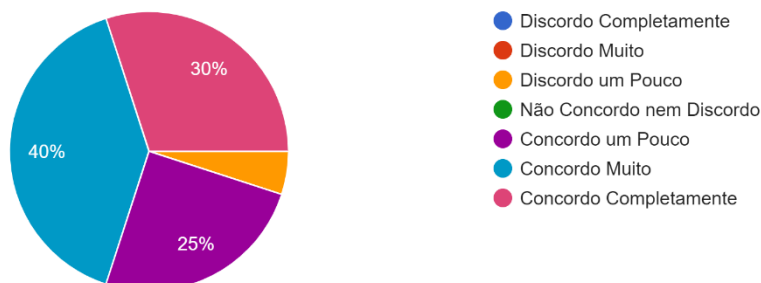
Fonte: Elaboração própria.

Quanto à construção visual da ferramenta, desde as primeiras reuniões este aspecto era uma preocupação. Quando realizada a revisão bibliográfica, foram localizados trabalhos semelhantes a esta tese em que a apresentação dos dados sempre foi uma dificuldade. Então em relação a esta lacuna, existia uma especial preocupação em desenvolver uma ferramenta em que fosse possível demonstrar em uma tela uma quantidade de dados para que o usuário possa compreender e utilizar tais informações, evitando sempre a poluição visual.

Figura 67 – Análise e compreensão das interações

12- O modo como podemos visualizar recursos, alunos e suas relações e iterações na área de visualização e análise é claro e compreensível.

20 respostas



Fonte: Elaboração própria.

Outro aspecto que se buscou contemplar foi o visual e intuitivo dos botões e ícones utilizados para selecionar as funcionalidades da ferramenta. Tendo em vista os resultados apresentados, podemos afirmar que o objetivo de construir uma ferramenta em que o aspecto visual seja um ponto positivo foi também atingido.

A resposta do Sujeito 5, sugere a utilização de cores mais fortes na composição dos grafos, deixando as cores mais visíveis e atraentes. Tais informações são importantes pois quando a próxima versão da ferramenta for desenvolvida, essas percepções serão incorporadas.

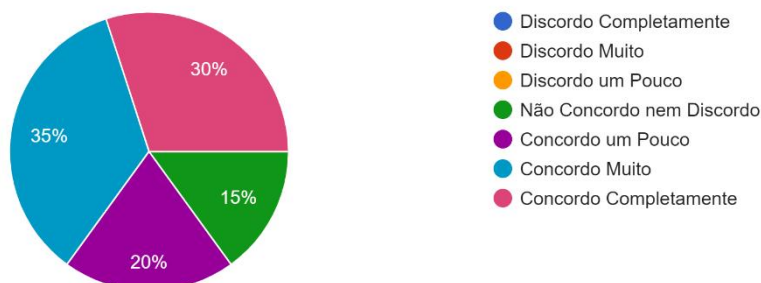
### 8.2.2.3 Percepção do usuário quanto as funcionalidades específicas

Nesta subseção é possível verificar a percepção dos usuários em relação a funcionalidades da ferramenta.

Em relação a forma como são selecionadas turmas e alunos, o modo de seleção teve um bom aceito por parte dos usuários. O *feedback* positivo indicou que 85% dos usuários concordam com o modo como são realizadas as seleções, é um modo adequado. 15% optaram pela opção “não concordam nem discordam”, o que nos leva a acreditar que para esses usuários essas funcionalidades não são algo fundamental na ferramenta.

Figura 68 – Seleção de turmas

11- Considero adequada a forma de seleção de turmas e alunos disponível na área de seleção.  
20 respostas



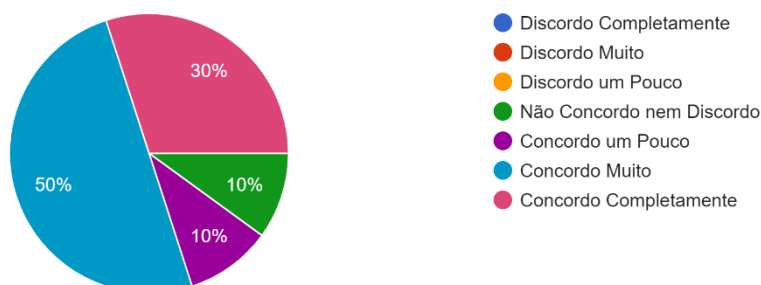
Fonte: Elaboração própria.

Sobre a questão 13 é possível salientar que foi uma agradável “surpresa” o percentual de 50% dos usuários responderem que “concorda muito”, bem como, um total de 90% responderem que concordam que essa função tem utilidade no seu trabalho docente. A utilização da predição está presente nesta tese seja por meio do uso da MD nos estudos parciais, onde foi identificada essa função como algo importante para o trabalho docente. A incorporação da função predição com o uso do módulo de IA a ferramenta MDV, demandou um número elevado de horas de trabalho por parte da equipe desenvolvedora. Como pode ser visto na subseção 6.1.1, o resultado binário sendo “0” reprovado e “1” aprovado talvez não represente a complexidade que foi pensar, desenvolver, testar e calibrar esta funcionalidade. O retorno positivo por parte dos usuários em relação a este item, nos indica que identificação da predição, desde o início do doutorado, como algo importante para o trabalho dos docentes, tutores e gestores, indicando também que a decisão de incorporar esta funcionalidade a ferramenta foi uma decisão correta.

Figura 69 – Função previsão

13- Em relação a área de ajustes, ações e parametrização, a função previsão apresenta utilidade no meu trabalho como docente.

20 respostas



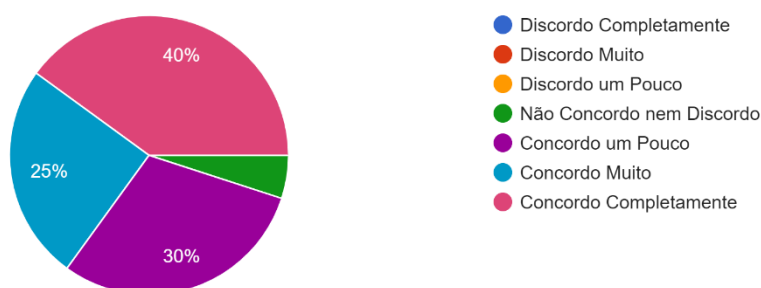
Fonte: Elaboração própria.

A questão 14, por sua vez, a partir das respostas dadas, aponta que as duas opções de visualização bigrafo e árvore podem gerar informações úteis ao seu usuário. A possibilidade em visualizar dados de alunos em formatos diferentes proporciona a ferramenta uma funcionalidade que durante a revisão bibliográfica não foi localizada em ferramentas presentes em trabalhos relacionados e é indicada como importante pelos usuários.

Figura 70 – Opções para visualização

14- A funcionalidade modo de visualização apresenta duas opções para visualização das informações. As duas formas de apresentação se m... realização da análise das turmas e dos alunos.

20 respostas



Fonte: Elaboração própria.

De acordo com a resposta, a questão 21 do Sujeito 11, indica que a opção bigrafo e árvore se mostram eficientes para a visualização das relações dos alunos com os recursos.

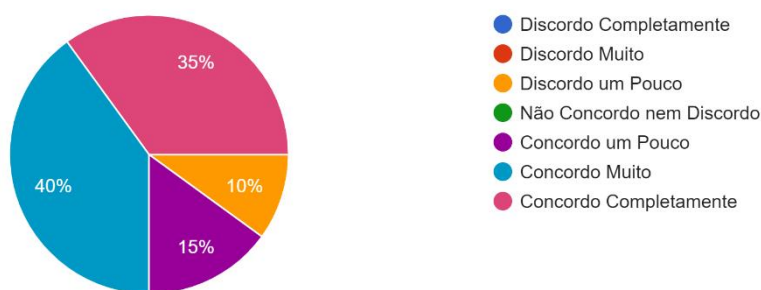


A questão 15 evidencia que filtrar o conjunto de alunos por nota máxima e mínima possibilita uma maior acurácia em relação à análise. Pode-se separar ou agrupar alunos e posteriormente visualizar suas interações com diferentes recursos.

O percentual de 95% de respostas favoráveis a importância da possibilidade de escolha tendo como critério de seleção as notas, indica que esta funcionalidade se mostra útil.

Figura 71 – Seleção de notas

15- O modo como selecionamos as notas na função filtrar por nota se mostra adequado.  
20 respostas



Fonte: Elaboração própria.

A seguir são apresentadas as respostas dos participantes nas questões abertas.

### 8.3 RESPOSTAS RELATIVAS ÀS QUESTÕES ABERTAS<sup>26</sup>

**21- Em relação à ferramenta, gostaria de sugerir as seguintes melhorias (opcional):**

#### **Sujeito 1:**

1. Adicionar barra de rolagem na visualização de dados do Moodle, pois quando usamos a ferramenta em um notebook, sem mouse externo, a função zoom fica prejudicada;
2. Quando uma função for acionada, abrir uma caixa explicativa sobre essa função;

<sup>26</sup> As respostas a essas questões não eram obrigatórias.

**Sujeito 2:**

“A ferramenta carregou corretamente no Chrome e todas as funcionalidades atenderam seus objetivos.

A minha sugestão seria adicionar no manual ou até mesmo na ferramenta uma explicação mais minuciosa sobre a área de ajustes, ações e parametrização, com objetivo de cada função.

Sobre este questionário, não ficou clara a questão 6”.

**Sujeito 3:**

“Na primeira tentativa, tive algumas dificuldades na interação. Não ficou claro para mim o que precisava fazer. Porém, ao interagir mais vezes com a ferramenta, pude descobrir as formas de interação disponíveis.

Uma sugestão é trazer um tutorial animado dentro da própria aplicação, indicando onde o usuário deve interagir e o que será possível realizar com aquela interação.”

“Durante a minha interação, fiquei em dúvida sobre o que eu deveria extrair de informações a partir da interação com a ferramenta.

Eu não tenho muita familiaridade com mineração de dados, árvores e grafos, e isso dificultou um pouco a compreensão/leitura dos dados apresentados na ferramenta.”

**Sujeito 4:**

Ter um botão de limpar seleção (de aluno e recursos - nodos selecionados); Apagar a predição (previsão de aprovado ou reprovado) ao clicar em qualquer outro nodo (fica aparecendo a predição do último aluno escolhido, mesmo após clicar em outro nome e isso pode transmitir uma informação incorreta).

O arquivo com os dados (exampledata.js) é carregado juntamente com a página e é muito grande (quase 40MB).

Sugiro que o mesmo seja carregado de modo assíncrono (com AJAX, através da chamada de um webservice) e só com os dados da disciplina selecionada. Além disso, ele já poderia enviar um arquivo JSON normalizado e otimizado para reduzir o seu tamanho e sem repetições de dados, como o nome do aluno e disciplina.

Na tela inicial, ele só mostrou os dados de 15 alunos da turma, por padrão. O padrão deveria ser mostrar todos ou ser mais intuitiva essa opção no aplicar limite, onde poderia ter ali, um checkbok "Mostrar todos".

A ferramenta é ótima! Isso não é uma crítica, são somente sugestões para deixá-la ainda melhor. Parabéns!!!

**Sujeito 5:**

A minha sugestão seria deixar as cores mais visíveis e atraentes ao usuário (usar tons mais fortes).

**Sujeito 6:**

Talvez pensar um perfil de usuário coordenação e outro docente e os docente só teriam acesso as suas turmas.

**Sujeito 7:**

Pelos conhecimentos que tenho sobre Moodle e no quanto eu uso e já usei, acredito que a ferramenta esteja muito boa.

Seja funcional a ponto de facilitar em muito o trabalho do professor e dos tutores.

No momento atual, não tenho sugestões de melhorias, talvez com o tempo de uso estar poderiam aparecer, por hora, apenas parabênzimo pelo excelente trabalho.

**Sujeito 8:**

A implementação de um botão para "retornar" o zoom inicial da visualização do gráfico. Várias vezes enquanto eu manipulava tive dificuldades em retornar ao enquadramento inicial.

Nesse ponto, relato que tive dificuldades na navegação pelo gráfico e precisei recarregar a página algumas vezes para retomar a exploração da ferramenta.

**Sujeito 9:**

Sugestão: incluir uma legenda das cores dos nodos.

**Sujeito 10:**

Senti a falta de um campo que faz o reinício da função ou funções que indique algum erro na execução de alguma análise.

Talvez seria interessante um campo que explique os resultados das análises feitas.

**Sujeito 11:**

Poder visualizar o que ocorre quando os alunos utilizam o Moodle é muito importante.

Achei muito boa a forma como ela foi organizada, gostei muito do ambiente a aparência é funcional e agradável.

O modo de visualização bigrafo e modo árvore, os filtros se mostraram eficientes.

A ferramenta otimiza o tempo do docente porque não precisa buscar as notas dos alunos em tabelas e possibilita visualizar de forma mais rápida grupos de alunos de acordo com suas notas.

É possível verificar se os alunos estão realizando as tarefas e se existe uma relação

entre a execução das tarefas com o seu desempenho.

A ferramenta também possibilita verificar não só alunos com propensão a evasão, mas também alunos com baixo rendimento e propensos a reprovação.

Sugiro um uma opção de salvar as pesquisas em forma de documento, tabela ou pdf.

### Sujeito 12:

Parabenizo pela ferramenta, tenho certeza que será de grande utilidade para os professores que usam o Moodle em suas aulas.

Quadro 4: Síntese de sugestões e críticas à ferramenta

Resumo dos testes com a ferramenta, realizados pelos usuários	Sugestões	Críticas
Sujeito 1	<ul style="list-style-type: none"> <li>➤ Adicionar barra de rolagem na visualização de dados do Moodle</li> <li>➤ Quando uma função for acionada, abrir uma caixa explicativa sobre essa função</li> </ul>	
Sujeito 2	<ul style="list-style-type: none"> <li>➤ Adicionar no manual ou até mesmo na ferramenta uma explicação mais minuciosa sobre a área de ajustes, ações e parametrização, com objetivo de cada função.</li> </ul>	<ul style="list-style-type: none"> <li>➤ Sobre este questionário, não ficou clara a questão 6</li> </ul>
Sujeito 3	<ul style="list-style-type: none"> <li>➤ Trazer um tutorial animado dentro da própria aplicação, indicando onde o usuário deve interagir e o que será possível realizar com aquela interação.</li> </ul>	<ul style="list-style-type: none"> <li>➤ Fiquei em dúvida sobre o que eu deveria extrair de informações a partir da interação com a ferramenta.</li> </ul>
Sujeito 4	<ul style="list-style-type: none"> <li>➤ Ter um botão de limpar seleção (de aluno e recursos - nodos selecionados); Apagar a predição (previsão de aprovado ou reprovado).</li> <li>➤ Ao clicar em qualquer outro nodo (fica aparecendo a predição do último aluno escolhido, mesmo após clicar em outro nome e isso pode transmitir uma informação incorreta).</li> <li>➤ Sugiro que o arquivo seja carregado de modo assíncrono (com AJAX, através da chamada de um webservice) e só com os dados da disciplina selecionada.</li> <li>➤ Além disso, ele já poderia enviar um arquivo JSON normalizado e otimizado para reduzir o seu tamanho e sem</li> </ul>	<ul style="list-style-type: none"> <li>➤ Na tela inicial, ele só mostrou os dados de 15 alunos da turma, por padrão. O padrão deveria ser mostrar todos ou ser mais intuitiva essa opção no aplicar limite, onde poderia ter ali, um <i>checkbox</i> "Mostrar todos".</li> <li>➤ O arquivo com os dados (exampledata.js) é carregado juntamente com a página e é muito grande (quase 40MB).</li> </ul>

	repetições de dados, como o nome do aluno e disciplina.	
<b>Sujeito 5</b>	➤ A minha sugestão seria deixar as cores mais visíveis e atraentes ao usuário (usar tons mais fortes).	
<b>Sujeito 6</b>	➤ Talvez pensar um perfil de usuário coordenação e outro docente e os docentes só teriam acesso as suas turmas.	
<b>Sujeito 8</b>	➤ A implementação de um botão para "retornar" o zoom inicial da visualização do gráfico.	➤ Relato que tive dificuldades na navegação pelo gráfico e precisei recarregar a página algumas vezes para retomar a exploração da ferramenta.
<b>Sujeito 9</b>	➤ Incluir uma legenda das cores dos nodos.	
<b>Sujeito 10</b>	➤ Senti a falta de um campo que faz o reinício da função ou funções que indique algum erro na execução de alguma análise.	➤ Talvez seria interessante um campo que explique os resultados das análises feitas.
<b>Sujeito 11</b>	➤ Sugiro uma opção de salvar as pesquisas em forma de documento, tabela ou pdf.	

Fonte: Elaboração própria.

## 22- Em relação ao manual do usuário, gostaria de sugerir as seguintes melhorias (opcional):

### Sujeito 1:

Nos títulos de cada seção, destaque em negrito e caixa alta; Usar mesmo destaque (cor e fonte) das imagens no texto explicativo, a fim de de levar o usuário a relacionar imagem-explicação, de forma direta.

### Sujeito 2:

Sugestão de um vídeo para esclarecer o objetivo das funcionalidades da ferramenta.

### Sujeito 3:

Eu usei o manual do usuário, mas preferia que ele fosse animado, dentro da própria aplicação.

Assim eu poderia visualizar por onde o autor considera melhor seguir com a interação.

### Sujeito 4:

Tive dificuldade só em descobrir como mostrar todos os alunos de uma disciplina e

como desmarcar os que já tinha clicado.

O "aplicar limite" poderia ser mais bem explicado para isso.

**Sujeito 5:**

O manual está bem claro e auxilia na compreensão da ferramenta.

**Sujeito 6:**

Quanto ao manual, muito bom também, deixou muito clara a ideia e de entender tudo o que e como pode ser feito usando a ferramenta, mesmo sem eu ter trabalhado com ela.

Somente com a leitura e visualização das figuras foi possível abstrair/imaginar como as coisas vão acontecendo.

Eu só aconselharia um melhor/maior detalhamento no item 3 do manual "Área de ajustes, ações e parametrizações" porque sinto que tem muito mais potencial do que ficou claro na escrita. Mais uma vez, parabéns.

**Sujeito 7:**

Uma breve descrição sobre a natureza dos dados, se são de um curso ou uma disciplina específica.

**Sujeito 8:**

Sugestão: o manual do usuário poderia trazer (logo no início do documento) mais detalhes sobre o objetivo da ferramenta.

O texto inicial informa que a ferramenta tem como objetivo auxiliar e apoiar os professores, tutores e técnicos, entretanto, não fica claro de que forma ela pode fazer isso.

Questões como quais informações podem ser visualizadas/analizadas e quais insights os profissionais da educação podem obter com a utilização da ferramenta seriam importantes estar explicitadas no início do manual.

**Sujeito 9:**

Deixar mais claro como os campos devem ser preenchidos. ex: formato das notas. Deve esclarecer se a quando seleção de algum campo algum outro deve ser selecionado em simultâneo, isto é, se existe dependência entre os campos.

**Sujeito 10:**

Revisar alguns termos e a gramática de uma forma mais geral, dado que o manual da ferramenta se tornará algo permanente em relação ao uso da ferramenta.

**Sujeito 11:**

No manual do usuário, poderia ser acrescentado o que representa as cores das

circunferências. Eu estou vendo várias verdes e uma azul. Imagino que sejam os alunos e a azul o professor, mas poderia ser também o tutor? Quais as cores para tutor e técnico? Li o manual e não encontrei essa informação.

Quadro 5: Síntese de sugestões e críticas ao manual

Resumo da avaliação do manual de usuário	Sugestões	Críticas
<b>Sujeito 1</b>	<ul style="list-style-type: none"> <li>➤ Nos títulos de cada seção, destaque em negrito e caixa alta.</li> <li>➤ Usar mesmo destaque (cor e fonte) das imagens no texto explicativo, a fim de levar o usuário a relacionar imagem-explicação, de forma direta.</li> </ul>	
<b>Sujeito 2</b>	<ul style="list-style-type: none"> <li>➤ Sugestão de um vídeo para esclarecer o objetivo das funcionalidades da ferramenta.</li> </ul>	
<b>Sujeito 3</b>	<ul style="list-style-type: none"> <li>➤ Eu usei o manual do usuário, mas preferia que ele fosse animado, dentro da própria aplicação. Assim eu poderia visualizar por onde o autor considera melhor seguir com a interação.</li> </ul>	
<b>Sujeito 4</b>	<ul style="list-style-type: none"> <li>➤ O "aplicar limite" poderia ser mais bem explicado para isso.</li> </ul>	
<b>Sujeito 6</b>	<ul style="list-style-type: none"> <li>➤ Aconselharia um melhor/maior detalhamento no item 3 do manual "Área de ajustes, ações e parametrizações"</li> </ul>	
<b>Sujeito 8</b>	<ul style="list-style-type: none"> <li>➤ O manual do usuário poderia trazer (logo no início do documento) mais detalhes sobre o objetivo da ferramenta.</li> <li>➤ Questões como quais informações podem ser visualizadas/analizadas e quais insights os profissionais da educação podem obter com a utilização da ferramenta seriam importantes estar explicitadas no início do manual.</li> </ul>	<ul style="list-style-type: none"> <li>➤ O texto inicial informa que a ferramenta tem como objetivo auxiliar e apoiar os professores, tutores e técnicos, entretanto, não fica claro de que forma ela pode fazer isso.</li> </ul>
<b>Sujeito 9</b>	<ul style="list-style-type: none"> <li>➤ Deixar mais claro como os campos devem ser preenchidos. ex: formato das notas.</li> <li>➤ Deve esclarecer se a quando seleção de algum campo algum outro deve ser selecionado em simultâneo isto é, se existe dependência entre os campos.</li> </ul>	
<b>Sujeito 10</b>	<ul style="list-style-type: none"> <li>➤ Revisar alguns termos e a gramática de uma forma mais geral, dado que o manual da ferramenta se tornará algo permanente em relação ao uso da ferramenta.</li> </ul>	

Fonte: Elaboração própria.

Os apontamentos realizados pelos participantes são fundamentais para que ajustes e alterações possam ser realizados, tendo em vista a melhoria das funcionalidades e para que a ferramenta apresente um caráter prático para seu uso no dia a dia dos docentes e pesquisadores que utilizam a plataforma Moodle, seja como ferramenta no EAD e mesmo no ensino presencial ou híbrido.

No capítulo a seguir são apresentados resultados e conclusões a partir dos objetivos específicos da tese.



## 9 RESULTADOS E DISCUSSÕES

Tendo como referência os objetivos específicos construídos para esse trabalho, nesse capítulo estes são melhor discutidos, bem como, como foi possível atingir cada um deles.

- **Buscar, em diferentes disciplinas ofertadas na modalidade EAD, *data sets* e a partir desses dados, utilizando técnicas de Mineração de Dados, identificar Trajetórias de Aprendizagem;**

Para esta tese foram utilizadas diversas bases de dados, sendo algumas delas: o repositório UCI (*Machine Learning Repository*) pertencente à *Open University* que disponibiliza de forma gratuita o uso de seus dados, dados relativos a alunos do curso de Especialização em Informática/PPGIE e dados relativos a logs e as notas de quatro turmas do Curso Superior em Desenvolvimento Rural - PLAGEDER. Com a MD foi possível identificar diferentes perfis de alunos de acordo com seus desempenhos para posterior construção das TAs por meio da MDV.

- **Identificar e agrupar, utilizando a técnicas de MD, conjuntos de alunos de acordo com seus desempenhos nas disciplinas;**

Por meio da ferramenta de MD Weka foram identificados 3 grupos de alunos (extrato superior, extrato médio e extrato inferior) dentro das 4 turmas do curso selecionado, contando com um universo de 528 alunos. Foram agrupados utilizando como critério as notas obtidas nas disciplinas.

- **Desenvolver em conjunto com alunos do curso de Ciências da Computação uma ferramenta para representação visual das trajetórias;**

O processo de construção da ferramenta se deu em diferentes etapas contando sempre com a participação de alunos, alunos bolsistas, do Instituto de Informática da Universidade Federal do Rio Grande do Sul, seu orientador e por mim doutorando do PPGIE. A partir da revisão sistemática, foram elencados trabalhos que possuíam como objetivo de alguma forma representar conjuntos de dados em forma de imagem. Dentre os trabalhos selecionados buscou-se identificar pontos positivos e lacunas em relação ao processo de representação de imagens tendo como origem base de dados. Outro elemento fundamental na construção foi incorporar a ideia de TA, dado a grande

quantidade de conceitos e formas de representação das TAs, diversas versões de representação de modelos foram construídas. A maior dificuldade sempre foi obter um modelo que pudesse utilizar um conjunto de significativos de dados e ao mesmo tempo gerar imagens em que o nível de poluição visual fosse baixo. As etapas realizadas durante a construção da ferramenta sempre tiveram um caráter incremental, pois a cada novo modelo proposto foram realizadas reuniões com as equipes de desenvolvedores para avaliação das funcionalidades que deveriam ser incluídas, a melhor linguagem computacional a ser utilizada, bem como, a melhor ferramenta para gerar grafos. Durante o processo de desenvolvimento 4 alunos do Curso de Ciências da Computação/UFRGS fizeram parte da equipe e como resultado foram apresentados 2 TCCs obtendo conceito A nas respectivas bancas avaliadoras, artigos publicados em periódicos e apresentação de trabalhos em eventos.

- **Construir visualmente, por meio do uso de grafos, relações entre os diferentes elementos (alunos e recursos) que constituem um Trajetória de Aprendizagem, indicando diferenças ou similaridades entre elas, bem como, a possível relação com o sucesso e insucesso.**

A partir das turmas, disciplinas e dados selecionados, por meio da ferramenta MDV foram construídas visualmente as TAs dos alunos, compreendendo a ideia de trajetória deste trabalho que são o resultado das interações entre os alunos e os diferentes recursos presentes na plataforma Moodle. Na seção 6.1 são apresentadas as construções visuais resultantes das interações realizadas pelas 4 turmas e agrupadas de acordo com as notas dos alunos, somando um total de 12 telas.

- **Analisar e comparar Trajetórias de Aprendizagem de alunos em diferentes disciplinas do curso Plageder<sup>27</sup>, buscando identificar a existência de um ou mais perfis de alunos propensos à evasão.**

Como resultado desse objetivo específico foi possível observar que houve interações mais efetivas no início do curso, os alunos pertencentes ao extrato inferior realizam interações desordenadas (falta de padrão), houve pouca diferença entre o grupo do extrato superior e extrato médio, os dois grupos apresentaram um padrão em relação a frequência de acesso a recursos.

---

<sup>27</sup> Curso Superior em Desenvolvimento Rural/ UFRGS

## 10 CONCLUSÃO

Nas últimas décadas a tecnologia modificou profundamente a sociedade. A presença das TICs em nosso ambiente de lazer, trabalho e estudo já é algo que faz parte das nossas rotinas de forma constante. De acordo com Kenski (2004, p.3), “O homem transita culturalmente mediado pelas tecnologias que lhe são contemporâneas. Elas transformam suas maneiras de pensar, sentir e agir. Mudam também suas formas de se comunicar e de adquirir conhecimentos”.

Seja pelo uso de *smartphones*, *desktops* ou *tablets* a forma como passamos a interagir e quantidade de interações se ampliou de forma significativa. O uso de ferramentas, plataformas e aplicativos em que as pessoas interagem são uma grande fonte de dados que são utilizados para os mais variados fins como, por exemplo, os sistemas de recomendação. Tratando-se da área educacional diversas plataformas virtuais se mostraram fundamentais durante o período da pandemia causada pela Covid-19. O uso dessas plataformas chamadas AVEAs já vinha sendo utilizado pelas instituições de ensino mas seu uso se acentuou no período anteriormente mencionado. Por sua vez, as AVEAs também possuem como característica a possibilidade de armazenar em seus bancos de dados, dados resultantes das interações dos alunos, tutores e professores com recursos, entre alunos, bem como, com professores. Essa grande quantidade de material que possui como origem as chamadas pegadas digitais (*digital footsteps*) resultantes dos caminhos realizados ao navegar pelas plataformas digitais, pode gerar informações relevantes para a tomada de decisão por parte de professores, tutores e gestores de cursos EAD, híbridos ou cursos presenciais que utilizam tais plataformas como apoio ao ensino.

De acordo com Moore e Kearsley (2007), um bom sistema de monitoramento e avaliação é essencial para o sucesso no desenvolvimento de cursos a distância. Tendo então como pano de fundo tal cenário, e constata-se como a construção de ferramentas de monitoramento e análise em AVEAs tem se tornado essenciais para que se possa reduzir fenômenos como o de retenção, abandono de cursos e evasão.

Este trabalho buscou, então por meio de dados de um curso EAD coletados de uma plataforma (Moodle) e com o uso de ferramentas de MD e do uso de grafos, identificar trajetórias de alunos por meio de suas interações. Como um dos objetivos específicos foi desenvolvida uma ferramenta (MDV) para que se pudesse por meio de

dados de um AVEA construir representações visuais. O desenvolvimento de tal ferramenta mostra a contribuição desta tese tanto para Educação como para a Computação trazendo inovações quanto a forma de visualização de trajetórias de aprendizagem em forma de interações representadas por grafos. O desenvolvimento da ferramenta buscou suprir algumas lacunas identificadas em trabalhos presentes na revisão bibliográfica como: o uso de uma representação simplificada e compreensível de como os alunos percorrem um AVEA, a facilidade de operação da ferramenta deixando a sua interface final a mais intuitiva possível, a possibilidade do usuário realizar seus próprios agrupamentos na forma de seleção das funcionalidades possibilitando a posterior análise dos dados algo bastante específico e ao mesmo tempo se mostrando algo útil no dia a dia dos professores.

Com o uso da MDV foi possível, então, perceber diferentes trajetórias de acordo com os grupos identificados e o seu desempenho, foi possível identificar alunos que, de acordo com seu desempenho (alunos do extrato inferior), apresentaram um padrão de interações que sinalizam características que podem indicar um perfil de aluno com propensão a evadir, sendo a detecção da evasão na EAD também um objetivo desta tese. Tais características são: uma forma desordenada e não sequencial de acesso a recursos, diferença significativa de frequência de acesso a recursos, não acesso ou baixíssimo acesso a recursos pelos alunos com nota 0.

Diversos são os motivos que levam os alunos a evadir de um curso, pode-se, por exemplo, apontar três deles: acesso a plataforma, participação em fóruns e desempenho nas avaliações, esses indicadores são possíveis de serem mensurados por meio da MDV. Sendo possível, então, perceber diferentes trajetórias de acordo com os grupos identificados de acordo com o seu desempenho, foi possível identificar alunos que, de acordo com seu desempenho (alunos do extrato inferior), apresentaram um padrão de interações que sinalizam características que podem indicar um perfil de aluno com propensão a evadir. Tais características são: uma forma desordenada e não sequencial de acesso a recursos, diferença significativa de frequência de acesso a recursos, não acesso ou baixíssimo acesso a recursos pelos alunos com nota 0.

Ainda de acordo com Kampff (2009, p. 94), “Situações de baixa frequência ou ausência no ambiente virtual por um longo período, falta de participação ou contato com professores e colegas, pouco acesso a materiais disponibilizados, perda de prazo

para entrega de atividades ou desempenho insatisfatório nas tarefas são alguns elementos que podem ser monitorados com periodicidade”.

O acesso ao AVEAs é um indicador quantitativo que se pode medir pelo número de logins realizado pelo aluno e o tempo médio de permanência. Outro elemento possível de mensuração, é a participação em fóruns, a baixa participação nesses espaços destinados a discussões pode indicar desinteresse por parte dos alunos.

Por último, outro elemento que pode indicar uma sinalização de abandono é o baixo desempenho nas avaliações. Fazemos aqui a ressalva que a nota do aluno é parte do conjunto avaliativo, uma nota baixa é um indicador clássico de que o aluno é um forte candidato a reprovação.

No trabalho de Lara (2016, p.11), é ressaltado que quando alunos não realizam acesso frequente ao AVEA, eles tornam subutilizados todos os recursos do ambiente, bem como, a estrutura de apoio oferecido pela instituição. Ainda do mesmo autor é explicitado que “o mapeamento do comportamento dos alunos dentro do ambiente virtual, pode auxiliar na concepção e estrutura dos cursos ofertados nestas plataformas, bem como, auxiliar os professores a planejarem a execução de sua disciplina,”

Sendo a evasão a consequência final de um conjunto de comportamentos indicativos ao cursar as disciplinas, para Lobo (2012), a evasão é um dos maiores problemas de qualquer nível de ensino. O abandono do aluno sem a finalização dos seus estudos representa uma perda social, de recursos e de tempo de todos os envolvidos no processo de ensino, pois perdeu o aluno, seus professores, a instituição de ensino, o sistema de educação e toda a sociedade (ou seja, o País).

Para os alunos, quando estes não se diplomam vem à tona diversos sentimentos como, a desmotivação, medo do futuro, insegurança, frustração, conformismo, passividade e o sentimento de fracasso, intitulado de abalo emocional, que se reverterá em sentimento de incapacidade intelectual para concluir qualquer curso (PPPI/UFBA, 2011; BAGGI, 2010; BRUNS, 1985; MCCUBBIN, 2003)

Para evitar as consequências negativas acima destacadas, podemos sugerir, a a partir dos resultados deste trabalho, que o monitoramento e acompanhamento dos alunos nos primeiros sinais de não interação com os recursos seria, por parte de professores, tutores e gestores de cursos, algo a ser considerado.

Em consonância com este estudo o trabalho de Silva (2013, p.13) após a

utilização da função sobrevivência de Kaplan- Meier<sup>28</sup> foi verificado que existe alta uma probabilidade do aluno evadir já nos momentos iniciais dos cursos, sendo que este risco se reduz ao longo do tempo. Foi identificado um padrão de interação entre os alunos com bom desempenho, em que nos primeiros momentos das disciplinas interagiram com os recursos com uma maior frequência. A baixa interação já nos primeiros momentos pode servir como indício para identificar alunos com propensão a evadir. Sendo assim, a ferramenta MDV possibilita a identificação desses padrões, bem como, prever, a partir das primeiras notas, a probabilidade de reprovação.

De acordo com Moore e Kearsley (2007), um bom sistema de monitoramento e avaliação é essencial para o sucesso no desenvolvimento de cursos à distância. Por meio do acompanhamento das interações, se torna possível a detecção e redução da distância transacional a quem Moore (2013) atribui com um dos principais fatores para o mau desempenho dos alunos nos cursos EAD.

A ferramenta MDV também possibilita a reflexão do docente quando comparadas as trajetórias planejadas e as trajetórias realizadas pelos alunos. Seja analisando os alunos que obtiveram um bom desempenho e fazendo a relação com os recursos que tais alunos acessaram com maior frequência, seja realizando uma análise dos alunos que não obtiveram um rendimento satisfatório e qual o seu comportamento nas disciplinas. Ainda sobre a análise dos docentes, é possível, a partir das informações coletas, aprimorar constantemente os materiais e a forma como são disponibilizadas as informações na ferramenta, pois os docentes por meio da MDV podem visualizar quais são os recursos mais acessados e por qual grupo de alunos esse acesso é realizado. O presente trabalho possui contribuições para as áreas da Educação, Computação e Social propondo uma nova ferramenta computacional que apresenta facilidade quanto ao seu uso por parte de professores, tutores de cursos EAD possibilitando a visualização das interações dos alunos com recursos educacionais no ambiente Moodle. Podendo a ferramenta ser utilizada para detecção precoce de alunos com propensão a evasão contribuindo para um melhor acompanhamento dos estudantes quando utilizam a plataforma Moodle.

Como sugestão de trabalhos futuros foram elencados alguns pontos como:

---

<sup>28</sup> A função mede o risco de evasão do aluno ao longo do curso utilizando o semestre como unidade de medida.

- Realizar mais experimentos com outros bancos de dados, selecionando turmas de cursos que apresentem um maior número de alunos com baixo desempenho (alunos do extrato inferior).
- Analisar as interações do(s) mesmo(s) alunos que fizeram parte das 4 turmas para se possa comparar a forma como interagiram ao longo das 4 disciplinas.
- Agregar à ferramenta, a visualização de outros recursos do Moodle (como, por exemplo, a funcionalidade visualizar o tempo de permanência dos alunos em cada recurso).
- Implementar as sugestões viáveis elencadas pelos usuários respondentes da pesquisa, bem como, corrigir as falhas apresentadas pela ferramenta ao longo dos testes.
- Estabelecer um protocolo de desenvolvimento para as próximas etapas definindo, por exemplo, a linguagem computacional a ser utilizada.
- Transpor na forma de *plugin* para o Moodle ou App a ferramenta MDV.
- Implementar na estrutura do MDV a possibilidade de leitura de dados de forma *online* nos bancos de dados.
- Indicar o percentual de recursos acessados em cada disciplina e disponibilizados e se os alunos mudam a forma de acesso ao longo de um curso.
- Analisar outros tipos de interações, alunos-alunos e alunos-professor.
- Utilizar dados como data e hora dos acessos para uma análise mais refinada.
- Tornar o login MDV único para que o professor possa registrar e analisar suas interações de forma pessoal
- Incorporar conceitos de *learning analytics* relacionando os mesmos a MD.
- Incluir em uma futura revisão sistemática trabalhos com a data posterior ao ano de 2019.
- Verificar por meio de gráficos de correlação ou histograma a existência de uma distribuição padrão em relação as notas com objetivo de

verificar(identificar) grupos de alunos.

- Manter e ampliar parcerias com grupos de pesquisa de outras Universidades dentro e fora do país<sup>29</sup>.
- Verificar a acurácia da funcionalidade preditiva usando dados reais de aprovação dos alunos.

---

<sup>29</sup> Foi estabelecida a construção de um convênio com a Université de Lorraine, Laboratoire Loria em Nancy, em decorrência de tal aproximação alguns experimentos foram realizados. Com exemplos podemos destacar uma análise conjunta dos dados coletados para esta tese. Tal trabalho teve como objetivo identificar por meio de técnicas e observações de pesquisadores brasileiros e franceses a compreensão dos fenômenos que ocorrem nos AVEAs a partir do uso de tais ambientes.



## REFERÊNCIAS

- AGUDELO, Olga Lucía; IBÁÑEZ, Jesus Salinas. **Flexible learning itineraries based on conceptual maps**. *Journal of New Approaches in Educational Research*, v. 4, n. 2, p. 70-76, 2015.
- ALVES, Magda. **Como escrever teses e monografias: um roteiro passo a passo**. Rio de Janeiro: Elsevier, p. 59, 2003.
- ARGOLO, Eráclito. DE Souza. **Trajетórias Conceituais Intencionais de Ensino e Aprendizagem: Investigação Em Fluxo Temporal Em Espaços E Contextos Nos Processos Educacionais Em EaD**. [s.l.: s.n.].
- AUSUBEL, David P., NOVAK, Joseph D., HANESIAN, Helen. **Psicologia educacional**. Tradução Eva Nick. Rio de Janeiro: Interamericana, p. 685, 1980.
- BAGGI, Cristiane Aparecida dos Santos; LOPES, Doraci Alves. **Evasão e avaliação institucional no ensino superior: uma discussão bibliográfica**. *Avaliação*, v. 16, n. 2, Sorocaba, pp.1-20, 2011.
- BAKER, Ryan; ISOTONI, Seiji; CARVALHO, Adriana. **Mineração de Dados Educacionais: Oportunidades para o Brasil**. *Revista Brasileira de Informática na Educação*. v.19, n.2, 2011.
- BAKER, Ryan, YACEF, Kalyna. **The State of Educational Data Mining in 2009: A Review and Future Visions**. *Journal of Educational Data Mining*, pp. 3-17, 2009.
- BARRASH, Joseph. **Age-related decline in route learning ability**. *Developmental Neuropsychology*, v. 10, n. 3, p. 189-201, 1994.
- BENNETT, C. et al. **The aesthetics of graph visualization**. In: CAe. [S.l.: s.n.], 2007. p. 57–64.
- BIAZUS, Cleber Augusto. **Sistema de fatores que influenciam o aluno a evadir-se dos cursos de graduação na UFSM e na UFSC: um estudo no curso de Ciências Contábeis**. Tese –Programa em Engenharia de Produção, Universidade Federal de Santa Catarina, 2004.
- BLANCHARD, E. G., ROY, M., LAJOIE, S. P., FRASSON, C. **An evaluation of sociocultural data for predicting attitudinal tendencies**. In *Proceedings of the International Conference on Artificial Intelligence in Education*, pp. 399-406, 2009.
- BRANCO, Vinícius M. A. **Visualização como Suporte à Exploração de uma Base de Dados Pluviométricos**. 2003. Dissertação (Mestrado em Ciência da computação e Matemática Computacional). Instituto de Ciências Matemáticas e de Computação. Universidade de São Paulo. 2003.
- BRASIL. Decreto nº 5.622, de 19 de dezembro de 2005. Regulamenta o art. 80 da Lei n. 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional, 2005. Disponível em:

<[http://www.planalto.gov.br/ccivil\\_03/\\_ato2004-2006/2005/decreto/d5622.htm](http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2005/decreto/d5622.htm)>. Acesso em: 19 de junho 2019.

BRITO, A. C. U. (2013) **Práticas de mediação de uma professora de educação infantil**. São Paulo: Universidade de São Paulo, USP.

BRUNS, Maria Alves de Toledo. **Evasão Escolar: Causas e Efeitos Psicológicos e Sociais**.1985. 177 f. Dissertação (Mestrado em Educação) – Programa de Pós-Graduação em Educação em Psicologia Educacional. Universidade Estadual de Campinas. São Paulo. 1985.

CABAU, Nubia; COSTA, Maria. **A Teoria da Distância Transacional: um mapeamento de teses e dissertações brasileiras**. Revista Eletrônica de Educação. 12. 431-447,2017.

CANTO, Alberto Bastos.; DE LIMA, José Valdeni; TAROUCO, Liane. Projeto e uso de objetos de aprendizagem: explorando as dimensões afetiva e cognitiva. **Revista Iberoamericana de Educación en Tecnología y Tecnología en Educación**, La Plata, v. 14, p. 7-17, 2014.

CANTO, Alberto Bastos et al. **Evasão e retenção em cursos de engenharia**. In: OLIVEIRA, V. F.; CHAMBERLAIN, Z. M. (Eds.). Desafios da educação em engenharia: vocação, formação, exercício profissional, experiências metodológicas e proposições. Brasília: ABENGE, 2012. p. 18.

CAMILO, Cassio Oliveira; SILVA, João Carlos. **Mineração de Dados: conceitos, tarefas, métodos e ferramentas**. Relatório Técnico. Instituto de Informática – UFMG, 2009.

CARD, S. K.; MACKINLAY, J. D.; SHNEIDERMAN, B. **Readings in information visualization: using vision to think**. Morgan Kauffman, 1999.

CARMONA, C., MILLAN, E., PÉREZ-DE-LA-CRUZ, J.-L., TRELLA, M., CONEJO, R.: **“Introducing prerequisite relations in a multi-layered Bayesian student model.”** In: Ardissono, L., Brna, P., Mitrović, A. (eds.) UM 2005. LNCS (LNAI), v. 3538, pp. 347–356. Springer, Heidelberg, 2005.

CASTELLS, Manuel. **A sociedade em rede**. São Paulo: Paz e Terra, p. 58, 2013.

CENSO EAD.Br. Associação Brasileira de Educação a Distância – ABED (Org.). **Relatório analítico da aprendizagem a distância no Brasil**.2017.SãoPaulo: Pearson Education do Brasil, 2018.

CHEN, G.; LIU, C.; Ou, K.; LIU, B. B. **“Discovering Decision Knowledge from Web Log Portfolio for Managing Classroom Processes by Applying Decision Tree and Data Cube Technology.”** Journal of Educational Computing Research, 23(3), pp.305– 332. 2000.

CONFREY, Jere et al. **Equipartitioning/splitting as a foundation of rational number reasoning using learning trajectories**. Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education, v. 1, n. 1998,

pp. 1–8, 2009.

DAVIS, F. D. (1989). “**Perceived usefulness, perceived ease of use, and user acceptance of information technology.**” *MIS Quarterly*, 13, 3, 319–340.

DETONI, D, CECHINEL, C & ARAÚJO, R M. **Modelagem e Predição de Reprovação de Acadêmicos de Cursos de Educação a Distância a partir da Contagem de Interações.** *Revista Brasileira de Informática na Educação*, 2015.

DUNHAM, W. **The Genius of Euler: Reflections on His Life and Work: Mathematical.** Association of America. 2007.

DURAND, Guillaume e BELACEL, Nabil e LAPLANTE, François. **Graph theory based model for learning path recommendation.** 2013.

EDUCAÇÃO DO SÉCULO XXI: **Cognição, tecnologias e aprendizagens.** Petrópolis, RJ: Vozes; Rio de Janeiro: Editora PUC, 2016.

FAYYAD, Usama, M. GREGORY, Piatetsky Shapiro. SMYTH, Padhraic. **Advances In knowledge discovery and data mining.** Mit Press, 1996.

FELDMAN, Ruth; EIDELMAN, Arthur I. **Biological and environmental initial conditions shape the trajectories of cognitive and social-emotional development across the first years of life.** *Developmental science*, v. 12, n. 1, p. 194-200, 2009.

FIGUEIREDO, Sonner Arfux de. COSTA, Nielce Meneguelo Lobo da. LLINARES, Salvador. GONZÁLEZ Julia Valls. **Caracterização em uma Trajetória de Aprendizagem com Funções Trigonométricas.** In: XIV CIAEM-IACME, Chiapas, México, 2015.

FIRAT, Mehmet. **Measuring the e-learning autonomy of distance education students.** *Open Praxis*, v. 8, n. 3, pp. 191-201, 2016.

FREEMAN, L. C. (1977). **A set of measurements of centrality based on betweenness.** *Sociometry*, 40(1), 35-41.

FREITAS, Leandro Quintanilha. **Medidas de Centralidade em Grafos.** - Rio de Janeiro:UFRJ/COPPE, 2010.

FORNARI, Liamara Teresinha. **Reflexões acerca da reprovação e evasão escolar e os determinantes do capital.** *Revista Espaço Pedagógico*, v. 17, n. 1, p. 112-124, 2010.

SILVA, Glauco Peres da. **Análise de evasão no ensino superior: uma proposta de diagnóstico de seus determinantes.** *Avaliação: Revista da Avaliação da Educação Superior (Campinas)* [online]. 2013, v. 18, n. 2 [Acessado 20 Novembro 2022] , pp. 311-333.

GOEL, A. **A mixed integer programming formulation and effective cuts for minimising schedule durations of Australian truck drivers.** *Journal of Scheduling*, 15(6), 733–741. <https://doi.org/10.1007/s10951-012-0282-0>. 2012.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**. 5 ed. São Paulo: Atlas, p. 184, 2010.

HAN, Jiawei; Kamber, Micheline. **Data Mining Concepts and Techniques**. San Francisco, EUA, Morgan Kaufmann, p.550, 2001.

Harary, F. **Graph theory**. Addison-Wesley, 1991.

HERSEY, Paul; BLANCHARD, Kenneth Hartley. **Management of organizational behavior:**

Utilizing human resources. Prentice-Hall, Inc, 1993.

HQ, M. **About Moodle**. 2022. <[https://docs.moodle.org/400/en/About\\_Moodle](https://docs.moodle.org/400/en/About_Moodle)>. [Online; acessado Agosto-2022].

HUANG, X., CHANDRA, A., DePAOLO, C. A., SIMMONS, L. L. **Understanding transactional distance in web-based learning environments: An empirical study**. British Journal of Educational Technology. pp. 734–747. 2016.

ICEK, A; MARTIN, F. **The prediction of behavior from attitudinal and normative variables**. Journal of Experimental Social Psychology, Volume 6, Issue 4, pp. 466-487. 1970.

IFENTHALER, D., GIBSON, D., & DOBOZY, E. (2018). **Informing learning design through analytics: Applying network graph analysis**. Australasian Journal of Educational Technology, 34(2). <https://doi.org/10.14742/ajet.3767>

ILLERIS, K. **Teorias Contemporâneas da Aprendizagem**. Porto Alegre: Penso. pp.35-36. 2013.

INEP Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Microdados do Censo da Educação Superior**, Brasília, 2018. Disponível em <<http://inep.gov.br/censo-da-educacao-superior>>. Acesso em 20 de Março de 2019.

INEP Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Microdados do Censo da Educação Superior**, Brasília, 2022. Disponível em <<http://inep.gov.br/censo-da-educacao-superior>>. Acesso em 20 de Julho de 2022.

INF/UFRGS. **About Mconf**. 2022. <<http://mconf.org/about/>>. [Online; acessado Agosto-2022].

IPAE Instituto Nacional de Estudos e Pesquisas Anísio Teixeira. **A história da Educação a Distância no Brasil**. Carta Mensal Rducacional, Rio de Janeiro, 2007. Disponível em <[http://www.ipae.com.br/pub/pt/cme/cme\\_82/index.htm](http://www.ipae.com.br/pub/pt/cme/cme_82/index.htm)> . Acesso em 10 de Junho 2022.

KAMPFF, Adriana Justin Cerveira. **Mineração de dados educacionais para geração de alertas em ambientes virtuais de aprendizagem como apoio à prática docente**. Tese de Doutorado, Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Informática na Educação, 2009.

KASSANDRINO, Amanda & ANGELAKI, Christina & MAVROIDIS, I. **Transactional Distance among Open University Students: How Does it Affect the Learning Process?** European Journal of Open, Distance and E-Learning. 17. 10.2478/eurodl-2014-0002.2014, 2014.

KING, R. D; Sternberg, M. J. E., Srinivasan, A. **Probabilistic reasoning in intelligent systems- networks of plausible inference.** Morgan Kaufmann, pp 1– 552, 1989.

KIRNER, C., TORI, R. Fundamentos de Realidade Aumentada. In: Tori, R., Kirner, C., Siscoutto, R. ed. **Fundamentos e Tecnologia de Realidade Virtual e Aumentada.** Porto Alegre, SBC, p. 20-37. 2006.

KITCHENHAM, Barbara; CHARTERS, Stuart. **Guidelines for Performing Systematic Literature Reviews in Software Engineering.** Technical Report EBSE, Keele University and Durham University Joint Report, 2007.

KLahr, David; NIGAM, Milena. The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. **Psychological science**, v. 15, n. 10, p. 661-667, 2004.

KUZILEK J., VACLAVEK J., FUGLIK V., ZDRAHAL Z. **Student Drop-out Modelling Using Virtual Learning Environment Behaviour Data.** In: Pammer-Schindler V., Pérez- Sanagustín M., Drachsler H., Elferink R., Scheffel M. (eds) Lifelong Technology-Enhanced Learning. EC-TEL, 2018.

LAKATOS, Eva Maria; MARCONI, Maria de Andrade. **Metodologia do Estudo científico.** Atlas, 7ª ed, p. 228, 2007.

LARA, Daiany Francisca, Claudia Alves Peres, and Franciano Antunes. "**ANÁLISE DA INTERAÇÃO NO AMBIENTE VIRTUAL DE APRENDIZAGEM E SUA INFLUÊNCIA NO DESEMPENHO: UM ESTUDO DE CASO NA UNIVERSIDADE XXXXX.**" SIED: EnPED-Simpósio Internacional de Educação a Distância e Encontro de Pesquisadores em Educação a Distância (2016).

LOBO, Maria Beatriz de Carvalho Melo. **Panorama da evasão no ensino superior brasileiro: aspectos gerais das causas e soluções.** In: ABMES Cadernos nº 25, 2012.

MALUF, Rosângela. **A evasão escolar e o ensino a distância.** 2012. Disponível em: [www.abed.org.br/media/textoevasao.pdf](http://www.abed.org.br/media/textoevasao.pdf) .Acesso em 22/06/2019.

**Mapa do Ensino Superior no Brasil/SEMESP .(2019),** Disponível em:

<https://www.semesp.org.br/pesquisas/mapa-do-ensino-superior-no-brasil-2021/>>Acessado em Janeiro 2022.

MATTAR, João. **Guia de educação a distância.** São Paulo: Cengage Learning: Portal Educação,2011.

**Microsoft Developer Network** . Disponível através de <[https://msdn.microsoft.com/pt-br/library/ms174949\(v=sql.120\).aspx](https://msdn.microsoft.com/pt-br/library/ms174949(v=sql.120).aspx)>. Acesso em: 20 Maio 2018.

MIROSLAV, Minović; MILOŠ, Milovanović; UROŠ, Šošević & MIGUEL Ángel Conde González. **Visualisation of student learning model in serious games**. Computers in Human Behavior. Volume 47, pp 98-107. ISSN 0747-5632, 2015.

MONTEIRO, Luciana G.; GUERRA, Antonio CO; SILVA, Joaquim FM. **Estudo e validação de uma progressão de aprendizagem em cinética química de estudantes de nível superior**. Química Nova, v. 44, p. 366-371, 2021.

MORAN, José Manuel et al. **Novas tecnologias e mediação pedagógica**. 6. ed. Campinas: Papyrus, 2000.

MOORE, Michael Grahame. **The theory of transactional distance**. In: MOORE, M. G. (Ed.), Handbook of distance education, New York: Routledg, 2013.

MOORE, Michael; KEARSLEY, Greg. **A educação a distância: uma visão integrada**. Trad. Roberto Galman, São Paulo: Thomson Learning, 2007.

MOORE, Michael; KEARSLEY, Greg. **Educação a distância: sistemas de aprendizagem online**. 3.ed. SãoPaulo: Cengage Learning, 2013.

MOREIRA, Marco Antônio. **Mapas Conceituais e Aprendizagem Significativa**. I Workshop sobre Mapeamento Conceitual, Vol. 24, N. 6, do PPGE nFis/IF-UFRGS, Brasil, 2013.

MOREIRA, Marco Antônio. **Teorias de Aprendizagem**. São Paulo: EPU ,1999.

MOREIRA, Marco Antônio. **Aprendizagem significativa: a teoria e textos complementares**. São Paulo: Livraria Editora da Física, 179p, 2011.

MOSTOW, Jack et al. An educational data mining tool to browse tutor-student interactions: Time will tell. In: **Proceedings of the Workshop on Educational Data Mining, National Conference on Artificial Intelligence**. AAAI Press, 2005. pp. 15-22.

NASCIMENTO, H. A. do; FERREIRA, C. B. R. **Visualização de Informações - Uma Abordagem Prática**. In: UNISINOS, 2005. Rio Grande do Sul. In: XXV Congresso da Sociedade Brasileira de Computação. Anais... Rio Grande do Sul: São Leopoldo, 2005.

NETTO, Oscar Picchi. **Um filtro interativo de utilizando árvores de decisão**. Dissertação de Mestrado, Programa Inter unidades de Bioinformática, Universidade de São Paulo, 2013.

NEVES, Rita de Cassia das. **Pré-processamento no processo de descoberta de conhecimento em banco de dados**. Dissertação (Programa de Pós-Graduação em Computação) – Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2003.

NOVAES, URIAS DA ROSA. **Agrupamento de Dados Através de Algoritmos SWARM**. Tese apresentada a COPPE/UFRJ, 2002. Disponível em . Acesso em out. de 2004.

NOVAK, Joseph Donald. CAÑAS, Alberto. **A teoria subjacente aos mapas conceituais e como elaborá-los e usá-los.** Práxis Educativa, Ponta Grossa, v.5, n.1, p. 9-29, jan.-jun. 2010. Disponível <http://www.revistas2.uepg.br/index.php/praxiseducativa/article/view/1298> Acesso em: 20 Junho, 2019.

NOVAK, Joseph Donald; GOWIN, Bob D. **Aprender a Aprender.** Lisboa: Plátano Edições Técnicas, 1996.

OSTERMANN, Fernanda. **Teorias de Aprendizagem/** Fernanda e Cláudio José de Holanda Cavalcanti. Evangraf; UFRGS, 2011.

PAIM, Eduardo. Ferramenta para visualização de interações de estudantes em recursos do ambiente Moodle. UFRGS, 2022.

PATTO, Maria Helena Souza. **A produção do fracasso escolar histórias de submissão e rebeldia.** São Paulo. T. A. Queiroz, 1996.

PALAZZO.MdeOliveira.Acessadoem20demarço2019  
Disponível em <<https://www.palazzo.pro.br/Wordpress/?p=156>> Acesso em 12 de Abril 2019.

PETERSEN, Kai et al. **Systematic mapping studies in software engineering.** In: 12th International Conference on Evaluation and Assessment in Software Engineering. sn, 2008.

PIECH, Chris et al. Deep knowledge tracing. **Advances in neural information processing systems**, v. 28, 2015.

PLAGEDER. Porto Alegre: UFRGS. Site:  
<http://www.ufrgs.br/cursopgdr/apresentacao.htm> - Acesso em Setembro 2022.

PONTES, Aldo. **O processo educativo na modalidade EAD: contornos, caminhos e mediações.**In: Oliani, Gilberto e Moura, Rogério (org.). Educação a distância: gestão e docência. Editora CRV, Curitiba, 2012.

PROGRAMA PENSE, PESQUISE E INOVE A UFBA – PROUFBA. Edital Propci – Proext–Prop/UFBA 01/2011 – PROUFBA. Coordenação de Pesquisa e Criação, versão 2. Universidade Federal da Bahia. Disponível em: <http://www.moodle.ufba.br/> acesso em 11 de Outubro de 2022.

QUEIROGA, Emanuel Marques. **Modelo de Predição da Evasão de Estudantes em Cursos Técnicos a Distância a partir da Contagem de Interações.** *Revista Thema*, pp. 425-438, 2018.

RAMOS, David et al. **Trilhas de aprendizagem em ambientes virtuais de ensino-aprendizagem: Uma revisão sistemática da literatura.** In: Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE). 2015. p. 338.

RAMOS, D. B. et al. **Um modelo para trilhas de aprendizagem em um ambiente virtual de aprendizagem.** In: Simpósio Brasileiro de Informática na Educação, 28.,

2017, Recife. **Anais**. Recife: SBIE, 2017. p. 1407-1416. Disponível em: <http://www.br-ie.org/pub/index.php/sbie/article/view/7669>. Acesso em: 22. ago, 2019.

RAMOS, Jorge Cavalcanti. **Um Modelo Preditivo da Evasão dos Alunos na EAD a Partir dos Construtos da Teoria da Distância Transacional**. *Simpósio Brasileiro de Informática na Educação-SBIE*, 2017.

RICHIT, A. **Implicações da teoria de vygotsky aos processos de aprendizagem e desenvolvimento em ambientes mediados pelo computador**. Disponível em [http://www.rc.unesp.br/igce/demac/maltempi/cursos/curso3/Artigos/Artigos\\_arquivos/Artigo%20Vygotsky%20-2004.doc](http://www.rc.unesp.br/igce/demac/maltempi/cursos/curso3/Artigos/Artigos_arquivos/Artigo%20Vygotsky%20-2004.doc): Acesso em: 11 de julho de 2022.

RIGO, Sandro José; CAMBRUZZI, Walter; BARBOSA, Jorge; CAZELLA, Silvio Cesar. **Aplicações de Mineração de Dados Educacionais e Learning Analytics com Foco na Evasão Escolar: Oportunidades e Desafios**. *Revista Brasileira de Informática na Educação*, v.22, n.1, 2014.

RIVERED, Rede Interativa Virtual de Educação. [http://rived.mec.gov.br/site\\_objeto\\_lis.php](http://rived.mec.gov.br/site_objeto_lis.php)

RODRIGUES, Jean Carlos; SIQUEIRA, Rafael Fernandes. **Mineração de dados do desempenho acadêmico na educação a distância**. 89 f. Estudo de Conclusão de Curso (Graduação) - Universidade Tecnológica Federal do Paraná, Ponta Grossa, 2015.

ROWLEY, Jennifer. **The wisdom hierarchy: representations of the DIKW hierarchy**. *Journal of Information Science*, 33, 163-180, 2007.

ROSSI, L. **Causas da evasão em curso superior a distância do consórcio da universidade aberta do Brasil**, 2008.

SABIDUSSI, Gert. **The centrality index of a graph**. *Psychometrika*, v. 31, n. 4, p. 581-603, 1966.

SALOMON, Gavriel; PERKINS, David. **Do Technologies Make Us Smarter? Intellectual Amplification With, Of, and Through Technology**. In R. J. Sternberg & D. D. Preiss (Eds.), *The educational psychology series. Intelligence and technology: The impact of tools on the nature and development of human abilities* (pp. 71-86). Lawrence Erlbaum Associates Publishers. Mahwah, NJ, US, 2005.

SAMPAIO, R.F, MANCINI MC, FONSECA ST. **Prática baseada em evidência: buscando informação para fundamentar a prática clínica do fisioterapeuta e do terapeuta ocupacional**. *Rev. Bras. Fisioter.* 2002;6(3):113-8.

SANTOS, J. P. O; MELLO, M. P; MURAMI, I. T. C. **Introdução à Análise Combinatória**. Rio de Janeiro: Ciência Moderna, 2007.

SARAMA, Julie; CLEMENTS, Douglas H. **Teaching Math in the Primary Grades: The Learning Trajectories Approach**. *Beyond the Journal, Young Children on the Web* March ,2009.

SAINT-GEORGES, I.; FILLIETTAZ, L. **Situated trajectories of learning in vocational**



training interactions. *European Journal of Psychology of Education*, 2008.

SERRAZINA, Lurdes; OLIVEIRA, Isolina. **Trajetórias de aprendizagem e ensinar para a compreensão.** Mapa do Ensino Superior no Brasil (2018), <<https://www.semesp.org.br/pesquisas/mapa-do-ensino-superior-no-brasil-2018/>>

SMITH, Becky; ENG, Min. **MOOCs: A learning journey.** In: International Conference on Hybrid Learning and Continuing Education. Springer, Berlin, Heidelberg, 2013. p. 244-255.

SILVA, Alice. **Análise Classificatória**, Universidade Nova de Lisboa - Faculdade de Ciências e Tecnologia.  
Disponível em < <http://ferrari.dmat.fct.unl.pt/services/AnaliseDados/Cluster.pdf> > . Acesso em julho. de 2020.

SILVA, Glauco Peres da. **Análise de evasão no ensino superior: uma proposta de diagnóstico de seus determinantes.** Avaliação: Revista da Avaliação da Educação Superior (Campinas) [online]. 2013, v. 18, n. 2 [Acessado 20 Novembro 2022] , pp. 311-333.

SIMON, Martin A. **Reconstructing mathematics pedagogy from a constructivist perspective.** *Journal for Research in Mathematics Education*, v. 26, n. 2, p. 114–145, 1995.

SIMBINE, F. B. **Comparando Roteiros Guiado e Livre nas Trajetórias de Aprendizagem.** Tese. PGIE-UFRGS, 2017.

SOARES, Silviane L. **Aplicação de técnicas de Mineração de Dados na Gestão de Sistemas de Energia Elétrica.** 106 f. Dissertação (Mestrado em Engenharia Elétrica) – Programa de Pós- Graduação em Engenharia Elétrica, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2005.

SOUZA, Audemir Lima de. **Teoria dos grafos e aplicações.** 2013. 78 f. Dissertação (Mestrado em Matemática) - Universidade Federal do Amazonas, Manaus, 2013.

SUN, D.; WONG, K. **On evaluating the layout of uml class diagrams for program comprehension.** In: 13th International Workshop on Program Comprehension (IWPC'05). [S.l.: s.n.], 2005. p. 317–326.

WARD, JOE H. **Hierarchical Grouping to optimize an objective function.** *Journal of American Statistical Association*, 58(301), 236-244, 1963. Disponível em . Acesso em ago. de 2005.

WEKA. Waikato Environment for Knowledge Analysis.  
Disponível em < <https://www.cs.waikato.ac.nz/ml/weka/> > Acessado em Junho 19,2020.

WOODLEY, Alan; SIMPSON, Ormond. **Student dropout:** The elephant in the room. On-line distance education: Towards a research agenda, 2014.

UNESCO. **Open and Distance Learning:** Trends, Policy and Strategy Considerations. Division of Higher Education. Unesco, 2002.

VAHDAT, MEHRNOOSH & ONETO, LUCA & ANGUIA, DAVIDE & FUNK, MATHIAS & RAUTERBERG, MATTHIAS. **A Learning Analytics Approach to Correlate the Academic Achievements of Students with Interaction Data from an Educational Simulator.** 9307. 352-366. 2015.

VASCONCELOS, Raimundo Claudio; NETO, Antonio Justiano Moraes; TELES, Lúcio. **Proposta de um modelo de mineração de dados educacionais para identificar a colaboração entre estudantes da EAD.** CIET:EnPED, [S.l.], maio 2018. ISSN 2316-8722. Disponível em: <<http://cietenped.ufscar.br/submissao/index.php/2018/article/view/241>>. Acesso em: 23 jun. 2019.

VERGARA, Sylvia C. **Projetos e relatórios de pesquisa em administração.** 3.ed. Rio de Janeiro: Atlas, p.47, 2000.

Yarbrough AK, Smith TB. **Technology acceptance among physicians: a new take on TAM.** Med Care Res Rev. 2007 Dec;64(6):650-72. doi: 10.1177/1077558707305942. Epub 2007 Aug 23. PMID: 17717378.

YIN, Robert K. **Estudo de Caso: planejamento e métodos.** 4ª Edição. Porto Alegre: Bookman, 2010.

ZORRILLA, Pérez, MARÍA, Jesús et al. **Evaluación de la comprensión lectora: dificultades y limitaciones.** Revista de educación, 2005.

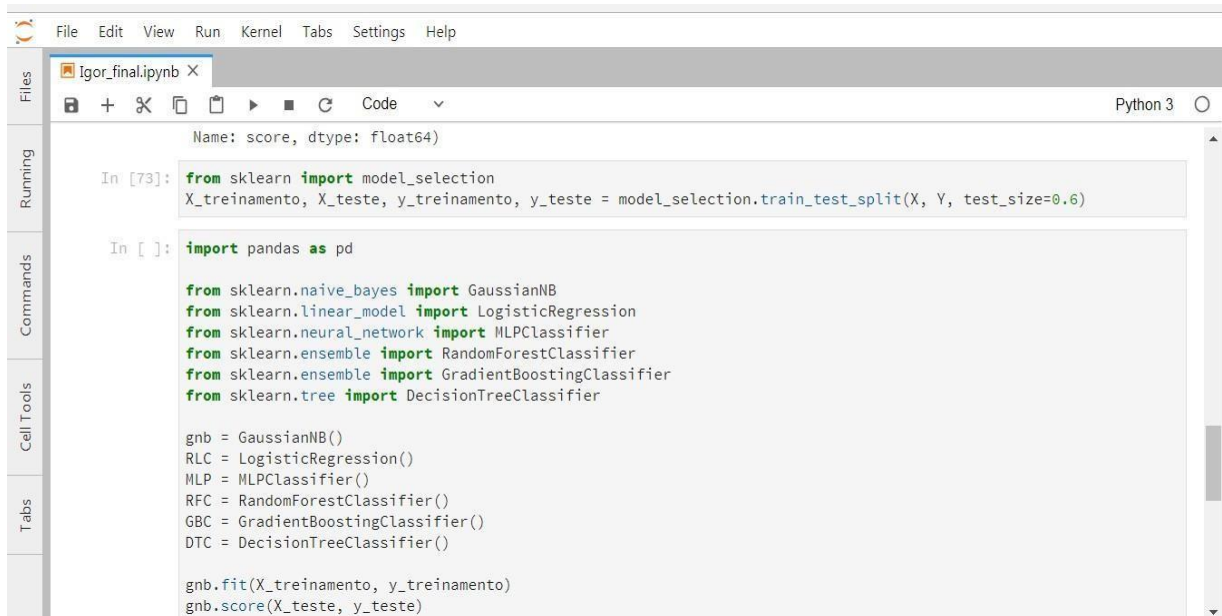
# ANEXO A

Figura 29 – IMAGEM DA TELA DA FERRAMENTA NEO4J

The screenshot displays the Neo4j web interface. On the left, a sidebar contains 'Database Information', 'Node Labels' (Hashtag, Source, Troll, Tweet, URL, User), 'Relationship Types' (HAS LINK, HAS\_TAG, IN\_REPLY\_TO, MENTIONS, POSTED, POSTED\_VIA, RETWEETED), and 'Property Keys' (created\_at, created\_str, description, expanded\_url, favorite\_count, favourites\_count, followers\_count, friends\_count, id, lang, listed\_count, location, name, retweet\_count). The main area shows a Cypher query: `1 MATCH (u:User)-[:POSTED]->(t:Tweet)-[:HAS_TAG]->(ht:Hashtag {tag: "thanksobama"})` and `2 RETURN *`. Below the query, a graph visualization shows 76 nodes and 100 relationships. The nodes are color-coded: blue for '#thanksobama', red for 'User', and green for 'Hashtag'. The graph shows a central green node connected to many blue nodes, which are in turn connected to red nodes. The interface also includes a 'Code' editor and a 'Graph' view selector.

## ANEXO B

Figura 30 – IMAGEM DA TELA DA FERRAMENTA NOTEBOOK AZURE MICROSOFT



The image shows a screenshot of the Microsoft Azure Notebook interface. The notebook is titled "Igor\_final.ipynb" and is running Python 3. The code is as follows:

```
Name: score, dtype: float64)

In [73]: from sklearn import model_selection
X_treinamento, X_teste, y_treinamento, y_teste = model_selection.train_test_split(X, Y, test_size=0.6)

In [ ]: import pandas as pd

from sklearn.naive_bayes import GaussianNB
from sklearn.linear_model import LogisticRegression
from sklearn.neural_network import MLPClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.ensemble import GradientBoostingClassifier
from sklearn.tree import DecisionTreeClassifier

gnb = GaussianNB()
RLC = LogisticRegression()
MLP = MLPClassifier()
RFC = RandomForestClassifier()
GBC = GradientBoostingClassifier()
DTC = DecisionTreeClassifier()

gnb.fit(X_treinamento, y_treinamento)
gnb.score(X_teste, y_teste)
```

## ANEXO C IMAGEM FERRAMENTA WEKA

**Weka Explorer**

Preprocess   Classify   Cluster   Associate   Select attributes   Visualize

Open file...   Open URL...   Open DB...   Generate...   Undo   Edit...   Save...

**Filter**  
Choose **None**   Apply

**Current relation**  
Relation: iris   Attributes: 5  
Instances: 150   Sum of weights: 150

**Selected attribute**  
Name: sepallength   Type: Numeric  
Missing: 0 (0%)   Distinct: 35   Unique: 9 (6%)

Statistic	Value
Minimum	4.3
Maximum	7.9
Mean	5.843
StdDev	0.828

Class: class (Nom)   Visualize All

**Attributes**  
All   None   Invert   Pattern

No.	Name
<input checked="" type="checkbox"/>	1 sepallength
<input type="checkbox"/>	2 sepalwidth
<input type="checkbox"/>	3 petallength
<input type="checkbox"/>	4 petalwidth
<input type="checkbox"/>	5 class

Remove

**Status**  
OK   Log   x 0

## ANEXO D REPORTAGEM JORNAL CORREIO DO POVO E SITE G1

CENSO DA EDUCAÇÃO SUPERIOR NO BRASIL

# Evasão acadêmica é de 56,8% no país

Dados do Inep/MEC também revelam que 24,6% dos acadêmicos estão em IES públicas; e 75,4%, nas particulares

Brasil apresenta alto índice de evasão no Ensino Superior: cerca de 56,8% dos estudantes com ingresso em 2010 desistiram dos estudos, 37,9% concluíram o curso, e os outros 5,3% permaneceram na graduação após 6 anos. Já entre os matriculados, 24,6% dos acadêmicos (2,1 milhões) estudam em instituição pública; e 75,4% (6,3 milhões) frequentam instituição privada.

Os dados são do Censo da Educação Superior 2018, realizado pelo Instituto Nacional de Estudos e Pesquisas Educacionais (Inep), que recolheu informações de 2.537 instituições de Ensino Superior (IES) brasileiras, totalizando 8,4 milhões de estudantes de graduação. O levantamento foi apresentado ontem pelo Ministério da Educação (MEC).

Outros indicadores do Censo apontam que, comparado a outros países da América Latina, o Brasil tem o menor número de ingressos no Ensino Superior na faixa etária entre 25 e 34 anos. Fica atrás de países como México, Colômbia, Chile e Argentina. O percentual de alunos que conclui a graduação dentro do esperado é de 33% no Brasil, abaixo da média da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), que é 41%. E cerca de 50% dos acadêmicos finalizam a graduação com 3 anos ou mais acima do esperado.

O ministro da Educação, Abraham Weintraub, disse que a menor taxa de desistência e a maior taxa de conclusão correspondem a matriculados de jovens pobres, que ingressaram no Ensino Superior privado através de financiamento. Segundo ele, ao optar pelo financiamento, o jovem "escolhe direito o curso que vai fazer e não desiste". E, mesmo com os problemas do Fundo de Financiamento Estudantil (Fies), como os altos índices de inadimplência, revela que "a taxa de sucesso desse jovem é melhor que de outros estudantes" que têm o curso custeado pelos pais.

O estudo mostra que, em 2016, dos alunos da rede privada que concluíram a graduação, em torno de 53,3% tinham Fies, enquanto 35,1% não possuíam o financiamento. Alunos com bolsas do ProUni também correspondem

ao maior número de conclusões de curso, atingindo 56%.

O diretor de Estatísticas Educacionais do Inep, Carlos Sampaio, acrescentou que, quanto a média

de anos de estudo entre alunos de 18 a 29 anos no país: é de 13,2 anos, em população com maior renda; 10,9 anos, negros ou pardos; e 9,7 anos, com menor renda.

**DADOS**

- **Rio Grande do Sul:** Existem 121 IES, que totalizam quase 520 mil estudantes. Destes, apenas 12% corresponde a alunos negros, pardos, amarelos ou indígenas. E são 10 instituições públicas, com cerca de 111 mil estudantes, e 111 instituições privadas, que contam com 407 mil acadêmicos. O Censo da Educação Superior 2018 pode ser acessado em: [bit.ly/2kqhBp4](http://bit.ly/2kqhBp4). Já as Sinopses Estatísticas da Educação Superior, que contêm os dados exposto os níveis estaduais, está em: [bit.ly/2gmr8v1](http://bit.ly/2gmr8v1).
- **Educação a Distância:** A pesquisa destaca que, pela primeira vez, o número de vagas oferecidas em cursos de graduação a distância (EAD), por volta de 7,1 milhões, é maior que as ofertas do ensino presencial, que representam 6,3 milhões de vagas. Entretanto, mesmo com uma maior oferta, o percentual de estudantes no Ensino Superior a distância corresponde a 40% do total, enquanto a modalidade presencial conta com 60%. E três alunos, em cada quatro, preferem estudos em sala de aula.

MENU

g1

EDUCAÇÃO

## Quase 3,5 milhões de alunos evadiram de universidades privadas no Brasil em 2021

Taxa de evasão do ensino superior privado chegou a 36,6% e é a segunda maior de toda a série histórica, ficando atrás apenas do ano de 2020. Inadimplência também cresceu nos últimos dois anos.

Por Amanda Lüder, GloboNews

02/01/2022 19h38 · Atualizado há 2 horas



## ANEXO E

=== Run information ===					
Scheme: weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 3 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10					
Relation: Derad 109					
Instances: 133 # número de alunos					
Attributes: 1					
Der 109					
Test mode: evaluate on training data					
=== Clustering model (full training set) ===					
Means					
=====					
Number of clusters selected by cross validation: 3					
Number of iterations: 10					
Within cluster sum of squared errors: 0.4599391345500558 # valor do erro com 3 clusters					
Initial starting points (random):					
Cluster 0: 65.25					
Cluster 1: 90.9					
Cluster 2: 97.55					
Missing values globally replaced with mean/mode					
Final cluster centroids:					
	Cluster#				
Attribute	Full Data	0	1	2	# número de clusters (3)
	(133.0)	(9.0)	(43.0)	(81.0)	# número de alunos em cada
cluster					
=====					
Der 109	77.288	2.3889	68.0895	90.4932	# média da nota em cada cluster
Time taken to build model (full training data) : 0 seconds					
=== Model and evaluation on training set ===					
Clustered Instances					
0	9 ( 7%)				
1	43 ( 32%)				
2	81 ( 61%)				

== Run information ==	
Scheme: weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 3 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10	
Relation: Derad 202	
Instances: 132	
Attributes: 1	

Der 202  
 Test mode: evaluate on training data  
 === Clustering model (full training set) ===

**kMeans**

=====

Number of clusters selected by cross validation: 3  
 Number of iterations: 7  
 Within cluster sum of squared errors: 0.49186397707761004

Initial starting points (random):

Cluster 0: 76.8  
 Cluster 1: 77.6  
 Cluster 2: 0

Missing values globally replaced with mean/mode

Final cluster centroids:

Attribute	Cluster Full Data	0	1	2
	(132.0)	(32.0)	(91.0)	(9.0)
Der 202	80.5284	72.2891	91.3901	0

=====

Time taken to build model (full training data) : 0.09 seconds

=== Model and evaluation on training set ===

Clustered Instances

0 32 ( 24%)  
 1 91 ( 69%)  
 2 9 ( 7%)

=== Run information ===

Scheme: weka.clusterers.SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 3 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10  
 Relation: Derad 501  
 Instances: 132  
 Attributes: 1

Test mode: evaluate on training data  
 === Clustering model (full training set) ===

**kMeans**

=====

Number of clusters selected by cross validation: 3  
 Number of iterations: 5  
 Within cluster sum of squared errors: 0.5072014315396596

Initial starting points (random):



Cluster 0: 95.85 Cluster 1: 66.1 Cluster 2: 0  Missing values globally replaced with mean/mode																													
Final cluster centroids: <table border="1"> <thead> <tr> <th>Attribute</th> <th>Cluster#</th> <th>0</th> <th>1</th> <th>2</th> </tr> </thead> <tbody> <tr> <td></td> <td>Full Data</td> <td></td> <td></td> <td></td> </tr> <tr> <td></td> <td>(132.0)</td> <td>(86.0)</td> <td>(32.0)</td> <td>(14.0)</td> </tr> <tr> <td colspan="5">=====</td> </tr> <tr> <td>Der 501</td> <td>81.9927</td> <td>95.3566</td> <td>78.3216</td> <td>8.2914</td> </tr> </tbody> </table> Time taken to build model (full training data) : 0.28 seconds					Attribute	Cluster#	0	1	2		Full Data					(132.0)	(86.0)	(32.0)	(14.0)	=====					Der 501	81.9927	95.3566	78.3216	8.2914
Attribute	Cluster#	0	1	2																									
	Full Data																												
	(132.0)	(86.0)	(32.0)	(14.0)																									
=====																													
Der 501	81.9927	95.3566	78.3216	8.2914																									
=== Model and evaluation on training set ===  Clustered Instances  0      86 ( 65%) 1      32 ( 24%) 2      14 ( 11%)																													

=== Run information ===																													
Scheme:    weka.clusterers.EM -I 100 -N -1 -X 10 -max -1 -ll-cv 1.0E-6 -ll-iter 1.0E-6 -M 1.0E-6 -K 10 -num-slots 1 -S 100 Relation:   Derad 502 Instances:   131 Attributes:   1																													
Der 502 Test mode:    evaluate on training data === Clustering model (full training set) ===																													
Initial starting points (random):  Cluster 0: 95.85 Cluster 1: 66.1 Cluster 2: 0  Missing values globally replaced with mean/mode																													
kMeans =====																													
Number of clusters selected by cross validation: 3 Number of iterations performed: 6 Within cluster sum of squared errors: 0.5160214314393595																													
<table border="1"> <thead> <tr> <th>Attribute</th> <th>Cluster</th> <th>0</th> <th>1</th> <th>2</th> </tr> </thead> <tbody> <tr> <td></td> <td>Full Data</td> <td></td> <td></td> <td></td> </tr> <tr> <td></td> <td>(131.0)</td> <td>(0.1)</td> <td>(0.27)</td> <td>(0.63)</td> </tr> <tr> <td colspan="5">=====</td> </tr> <tr> <td>mean</td> <td>84.8</td> <td>4.9351</td> <td>91.3629</td> <td>77.0064</td> </tr> </tbody> </table> Time taken to build model (full training data) : 0.14 seconds					Attribute	Cluster	0	1	2		Full Data					(131.0)	(0.1)	(0.27)	(0.63)	=====					mean	84.8	4.9351	91.3629	77.0064
Attribute	Cluster	0	1	2																									
	Full Data																												
	(131.0)	(0.1)	(0.27)	(0.63)																									
=====																													
mean	84.8	4.9351	91.3629	77.0064																									

=== Model and evaluation on training set ===

**Clustered Instances**

<b>0</b>	<b>13 ( 10%)</b>
<b>1</b>	<b>43 ( 33%)</b>
<b>2</b>	<b>75 ( 57%)</b>

## ANEXO F

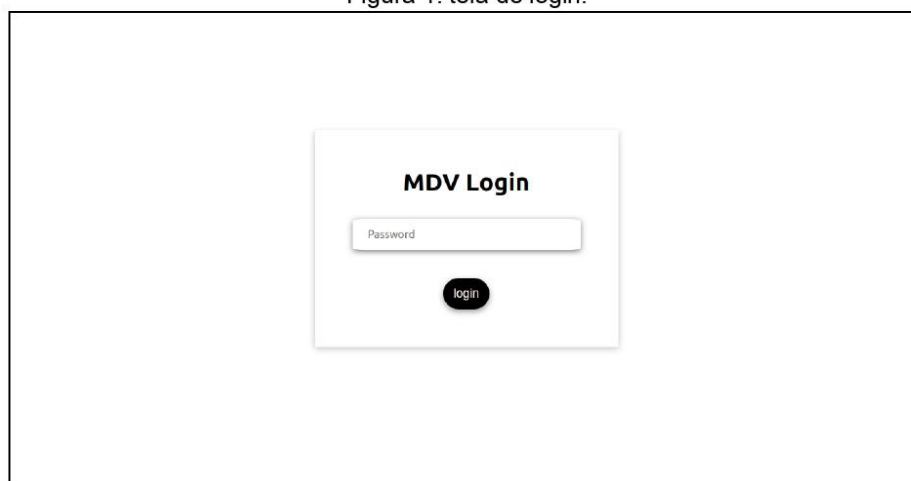
### Manual Usuário - Moodle Data Visualization

O presente guia pode ser utilizado para auxílio do usuário(a) quanto a utilização da ferramenta Moodle Data Visualization (MDV)<sup>1</sup>. A ferramenta foi desenvolvida com o objetivo de servir de apoio a professores, tutores e técnicos que venham a atuar em um curso que use a plataforma Moodle em suas disciplinas.

#### Tela de Login

Ao entrar no site da ferramenta, é exibida inicialmente a tela de login que protege o acesso à área principal. Para acessar o sistema o usuário deve digitar a senha correta e será redirecionado para a ferramenta. Quando digitada a senha incorreta, é exibida a mensagem de erro contendo a mensagem "senha incorreta" na tela e o usuário poderá tentar novamente.

Figura 1: tela de login.



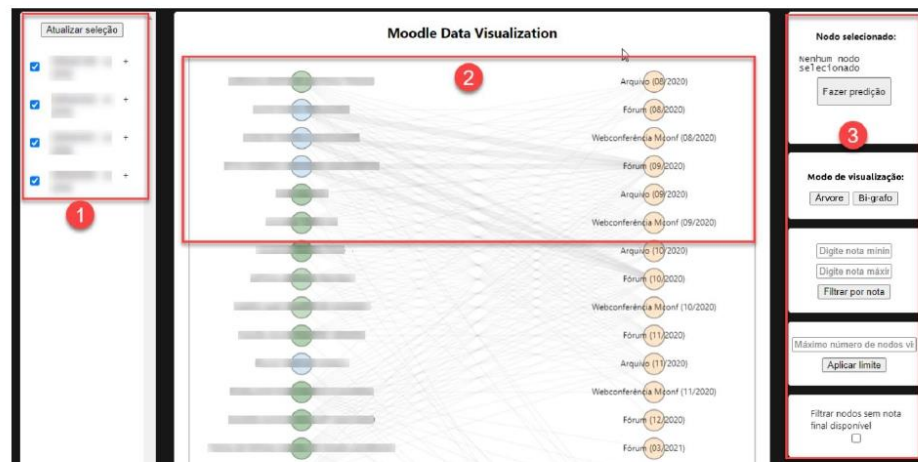
#### Visão Geral do MDV

Uma vez na área principal, a interface com a tela principal é exibida ao usuário (Figura 2). Abaixo são listadas as principais funcionalidades da ferramenta:

---

<sup>1</sup> <http://mdv.inf.ufrgs.br/>

Figura 2: tela geral MDV.



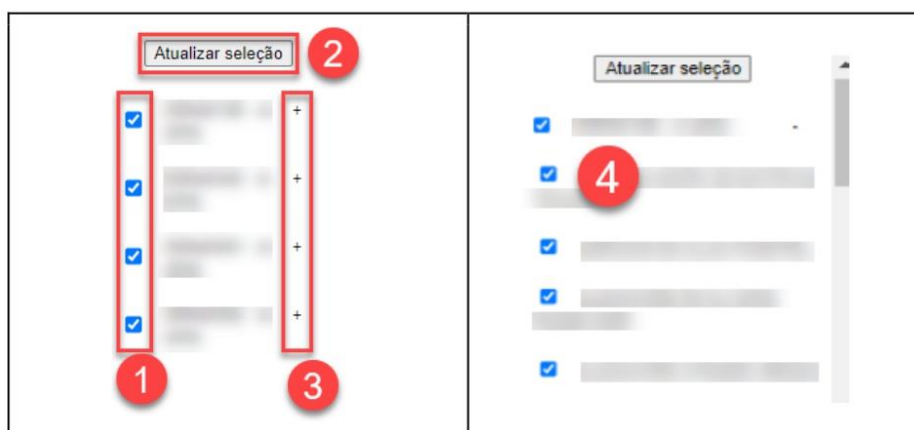
- 1) **Área de seleção:** Nesta área podemos indicar quais são os dados de interesse para a análise. Podemos escolher entre analisar uma turma, todas as turmas ou somente um subconjunto de turmas. Também podemos indicar se queremos analisar todos os alunos de uma turma ou alguns deles.
- 2) **Área de visualização e análise:** Nesta área podemos visualizar os gráficos de análise, os quais representam os indivíduos (alunos, tutores e professores) e os recursos que acessaram na(s) turma(s) escolhida(s).
- 3) **Área de ajustes, ações e parametrização:** aqui podemos parametrizar, escolher como desejamos visualizar os dados, podendo indicar filtros e seleções específicas. Também podemos realizar predições usando o módulo de IA.

A seguir, os elementos disponíveis nessas áreas serão detalhados.

### 1) ÁREA DE SELEÇÃO

Como já mencionado, aqui podemos indicar os dados de interesse para a análise. Para tanto, a área apresenta dois componentes: Seleção de itens de análise, tais como turma (1) e respectivos alunos (3), além do botão de atualização (2).

Figura 3: detalhe da área de seleção.

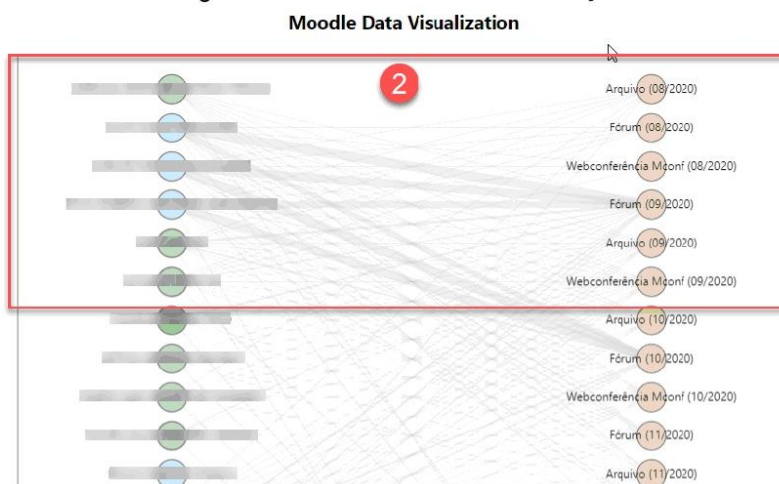


- **Seleção de turmas (1):** aqui podemos selecionar a(s) turma(s) que desejamos visualizar.
- **Atualizar seleção:** ao clicar no botão (2), são apresentados na tela os dados relativos às turmas ou alunos selecionados.
- **Escolha de alunos a serem analisados:** na mesma lista, se desejado, podemos selecionar os alunos, de cada turma, que desejamos incluir ou excluir da análise. Isso pode ser feito clicando-se no sinal de mais (+) ao lado do nome de cada turma (3). Nele, a lista será expandida (4) e podemos desmarcar ou marcar os alunos que desejamos incluir na análise. Depois de mudar a seleção, é necessário clicar no botão (2) “Atualizar seleção”.

### 2) ÁREA DE VISUALIZAÇÃO E ANÁLISE

Nesta área, como já citado anteriormente, podemos visualizar os gráficos de análise, os quais representam os recursos e os indivíduos (alunos, tutores e professores).

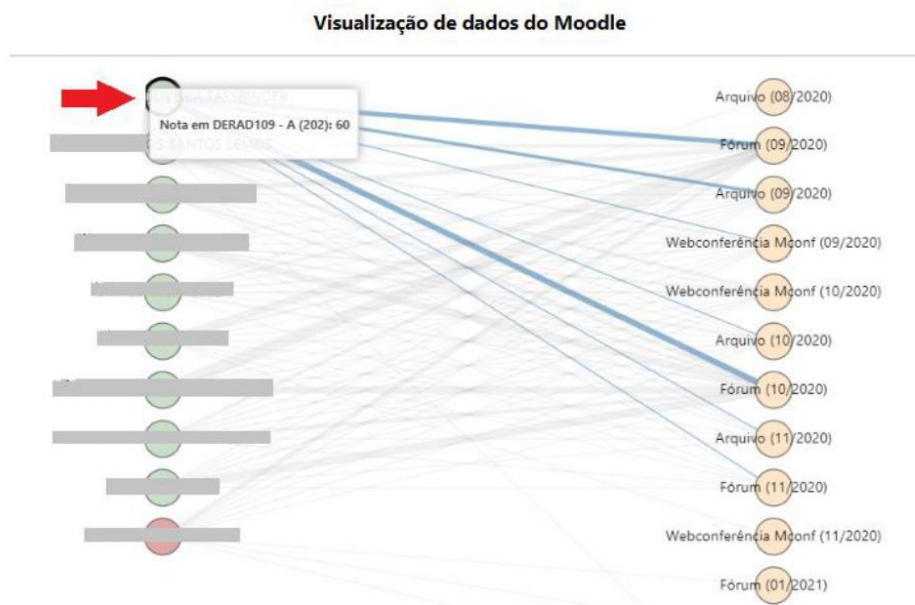
Figura 4: detalhe da área de visualização.



Na coluna à esquerda podemos visualizar os participantes do curso e à direita os recursos acessados. As linhas (ligações) entre esses nodos representam os acessos ou interações.

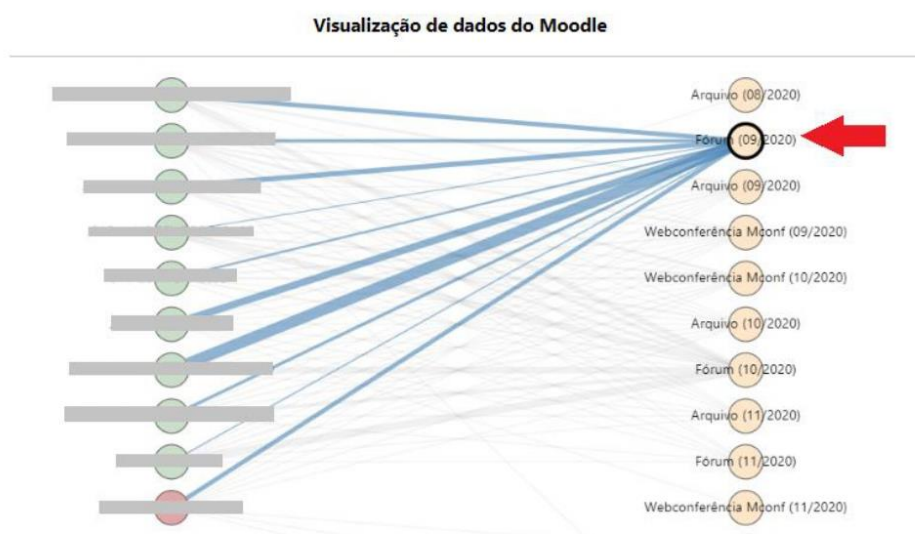
Detalhes de interação: no grafo central podemos escolher um indivíduo de maneira a destacar as interações desse indivíduo com os respectivos componentes e recursos do curso. Ao passar o cursor sobre o círculo (nodo) correspondente ao aluno, a nota do aluno é disponibilizada para visualização.

Figura 5: seleção de indivíduos



Detalhes dos recursos: os recursos da disciplina são representados por círculos na cor bege, localizados à direita da área de visualização. A partir do nodo, ou seja, ao clicar sobre ele, podemos visualizar qual ou quais participantes interagiram com aquele recurso.

Figura 6: seleção dos recursos.



As arestas (linhas que conectam indivíduos e recursos) possuem uma espessura que pode ser maior ou menor de acordo com a frequência de acesso (quanto mais acessos, mais larga é a linha).

### 3) ÁREA DE AJUSTES, AÇÕES E PARAMETRIZAÇÃO

A área destinada aos ajustes, ações e parametrização, como já apontado anteriormente, possibilita ao usuário da ferramenta parametrizar, escolher como desejamos visualizar os dados, podendo indicar filtros e seleções específicas, bem como, realizar a predição em relação a aprovação ou não do aluno.

A função **predição ou previsão** (Figura 7-a), possibilita ao usuário por meio dos dados gerados pelos alunos (e respectivas notas), predizer dois possíveis resultados: ser aprovado ou não. Essa funcionalidade serve como uma auxílio ao(s) professores, tutores e gestores para que possam identificar alunos com desempenho abaixo do



esperado e a partir dessa informação detectar aluno(s) com propensão ao abandono do curso.

Na função **modo de visualização** (Figura 7-b), podemos seleccionar entre duas formas: a opção bi-grafo ou a opção árvore.

Há ainda uma opção para **filtrar por nota mínima e máxima** (Figura 7-c), limitar a quantidade de nodos na tela (Figura 7-d).

Por fim podemos seleccionar por meio da função **filtrar nodos sem nota final disponível** (Figura 7-e) os participantes do curso que possuem nota ou remover os nodos sem nota final. Essa funcionalidade pode ser utilizada, por exemplo, quando a intenção é seleccionar somente os alunos, excluindo professores, tutores e técnicos.

Figura 7: detalhe da área de ajustes, ações e parametrização.

**Nodo selecionado:**  
Nenhum nodo selecionado  
Fazer predição **a**

**Modo de visualização:**  
Árvore Bi-grafo **b**

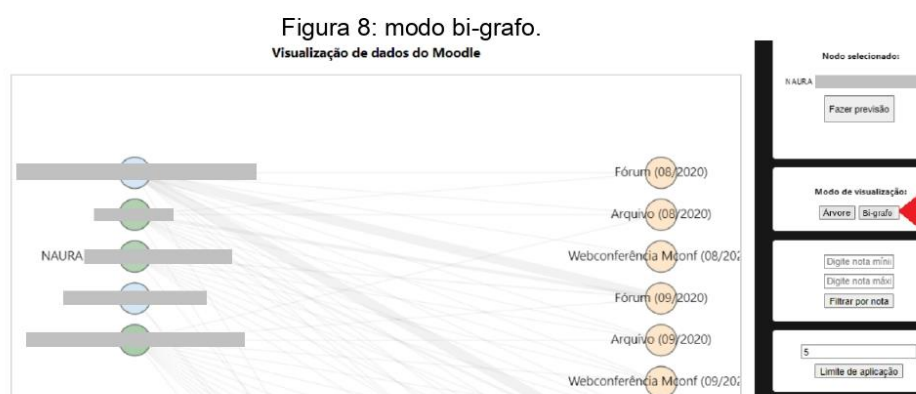
Digite nota mínim  
Digite nota máxi  
Filtrar por nota **c**

Máximo número de nodos vi  
Aplicar limite **d**

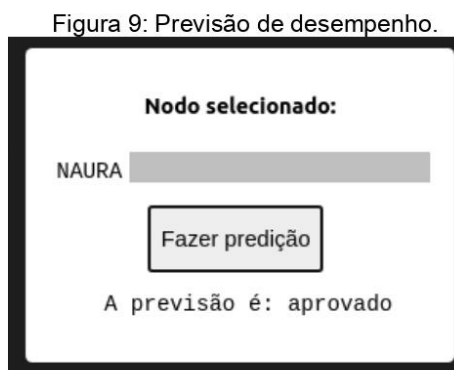
Filtrar nodos sem nota final disponível  
 **e**

**Visualização em modo Bi-grafo**

Quando selecionamos a visualização em modo bi-grafo, é permitido ao usuário visualizar duas colunas contendo nodos (Figura 8). Na coluna à esquerda, nodos relativos aos participantes do curso, ou seja, alunos, técnicos, tutores e professores. A coluna à direita apresenta os diferentes recursos da disciplina selecionada. As linhas ou arestas representam as relações entre os nodos das duas colunas. A visualização permite *zoom in* e *zoom out*, arrastar, deslocar e selecionar nodos ou arrastar a figura inteira.



Ao selecionar um nodo, o nome deste irá aparecer no topo da coluna da direita, logo abaixo de “Nodo selecionado:”. Ao clicar em fazer previsão, a ferramenta irá prever o desempenho acadêmico do aluno (aprovar ou reprovar). Na Figura 9 podemos verificar um exemplo desta funcionalidade.



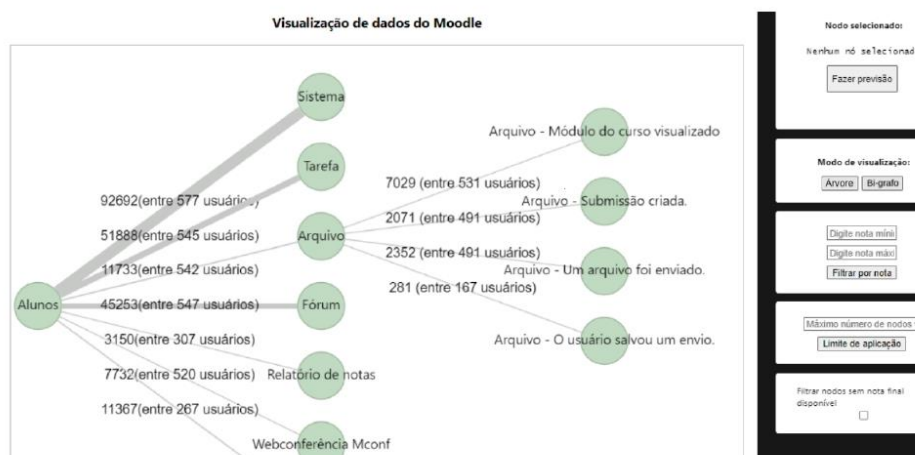
Quando selecionamos a função árvore, a tela central atualiza para um segundo modo de visualização.

## Visualização em modo Árvore

Essa visualização evidencia métricas de interação com os componentes do Moodle realizada pelo grupo de alunos selecionados. O nodo mais à esquerda representa o grupo de alunos. As arestas podem se conectar aos componentes do Moodle, e são apresentadas legendadas com o número total de interações com o componente, bem como, o número de alunos que realizaram essas interações ( informação entre parêntesis).

Ao clicar com o botão esquerdo do mouse em um nodo de componente do Moodle, este irá se expandir e exibir as ações realizadas no componente selecionado. De maneira análoga à citada anteriormente, as arestas são legendadas com o número de interações e o número de usuários que realizaram essas interações.

Figura 10: modo árvore.



Buscamos por meio deste manual facilitar e tornar a utilização da ferramenta MDV mais dinâmica. Nos colocamos a disposição dos usuários para resolução de dúvidas por meio dos contatos abaixo:

Leandro Krug Wives: [leandro.wives@ufrgs.br](mailto:leandro.wives@ufrgs.br)

Igor Kühn: [kuhnigor@gmail.com](mailto:kuhnigor@gmail.com)

Eduardo Paim:

## APÊNDICE 1

### Questionário TAM aplicado com docentes

- 1- A utilização desta ferramenta no meu trabalho permite-me realizar as tarefas mais rapidamente do que se as executasse em outros produtos semelhantes.
- <sup>30</sup> ( ) Discordo Completamente  
( ) Discordo Muito  
( ) Discordo um Pouco  
( ) Não Concordo nem Discordo  
( ) Concordo um Pouco  
( ) Concordo Muito  
( ) Concordo Completamente
- 2- A utilização desta ferramenta no meu trabalho aumenta o meu benefício, interesses ou vantagens.
- 3- Utilizar esta ferramenta torna o meu trabalho mais fácil.
- 4- Considero esta ferramenta útil no meu trabalho.
- 5- Foi fácil para mim operar esta ferramenta.
- 6- Acho fácil que esta ferramenta faça o que eu preciso.
- 7- A minha interação com esta ferramenta foi clara e compreensível.
- 8- Acho que esta ferramenta é suficientemente flexível e interativa.
- 9- Com as minhas habilidades, foi fácil utilizar esta ferramenta.
- 10- Considero que a ferramenta possui elementos visuais agradáveis e intuitivos.

---

<sup>30</sup> Cada questão poderia ser respondida em uma escala Likert de sete pontos que variando de *discordo completamente* até *concordo completamente*.

- 11- Considero adequada a forma de seleção de turmas e alunos disponível na **área de seleção**.
- 12- O modo como podemos visualizar recursos, alunos e suas relações e iterações na **área de visualização e análise** é claro e compreensível.
- 13- Em relação a **área de ajustes, ações e parametrização**, a função predição apresenta utilidade no meu trabalho como docente.
- 14- A funcionalidade **modo de visualização** apresenta duas opções para visualização das informações. As duas formas de apresentação se mostram úteis para a realização da análise das turmas e dos alunos.
- 15- O modo como selecionamos as notas na função **filtrar por nota** se mostra adequado.
- 16- O uso da função **aplicar limite**, com a opção de limitar o número de nodos apresentados na tela, torna mais fácil a visualização e compreensão das informações durante as análises.
- 17- Considero a função **filtrar nodos sem nota final disponível** fundamental durante o uso da ferramenta.
- 18- A utilização desta ferramenta me ajuda a compreender como os estudantes utilizam o Moodle.
- 19- A utilização desta ferramenta ajudou a identificar como os estudantes interagem com diferentes recursos no Moodle.

20- Acredito que com o uso desta ferramenta é possível identificar alunos com propensão a evasão.

21- Em relação à ferramenta, gostaria de sugerir as seguintes melhorias (opcional):

22- Em relação ao manual do usuário, gostaria de sugerir as seguintes melhorias (opcional):