

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
CURSO DE CIÊNCIA DA COMPUTAÇÃO

EDUARDO CHAVES PAIM

**Ferramenta para visualização de interações
de estudantes em recursos do ambiente
Moodle**

Monografia apresentada como requisito parcial
para a obtenção do grau de Bacharel em Ciência
da Computação

Orientador: Prof. Dr. Leandro Krug Wives
Coorientador: Ms. Igor Kuhn

Porto Alegre
2022

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos André Bulhões Mendes

Vice-Reitora: Prof^ª. Patricia Pranke

Pró-Reitora de Graduação: Prof. Cíntia Boll

Diretora do Instituto de Informática: Prof^ª. Carla Maria Dal Sasso Freitas

Coordenador do Curso de Ciência de Computação: Prof. Rodrigo Machado

Bibliotecária-chefe do Instituto de Informática: Beatriz Regina Bastos Haro

AGRADECIMENTOS

Agradeço à todos os familiares, amigos, professores e demais conhecidos que me acompanharam nesses últimos anos.

RESUMO

A recente pandemia de COVID-19 resultou na utilização de ambientes virtuais de aprendizagem, tal como a plataforma Moodle, em várias instituições de ensino para substituir as aulas presenciais e respeitar as restrições de isolamento social. Essa mudança gerou questionamentos a respeito da qualidade do ensino remoto e do engajamento dos estudantes com esses ambientes virtuais. Este trabalho apresenta uma ferramenta que processa dados de *logs* fornecidos pela plataforma Moodle e os apresenta em uma visualização interativa que permite identificar os comportamentos dos usuários com os diferentes recursos do Moodle e suas possíveis relações com o desempenho acadêmico. Para validar a ferramenta foi aplicado um questionário para usuários com experiência docente, composto por questões sobre a apresentação e usabilidade da ferramenta. As respostas obtidas indicam que a ferramenta cumpre a proposta sendo útil na identificação de comportamentos dos estudantes e sua interação com o ambiente virtual de ensino.

Palavras-chave: Ciência da Computação. UFRGS. Visualização de Dados. Machine Learning. Moodle. Ambiente Virtual de Ensino e Aprendizagem.

A Moodle tool for mining educational data

ABSTRACT

The recent COVID-19 pandemic resulted in an even wider adoption of virtual learning environments, such as the Moodle platform, by many educational institutions required to replace in-person classes and abide to social isolation guidelines. These changes raised concerns about the effectiveness of online learning and the commitment from students. This work introduces a web tool that processes logs from the Moodle platform and presents it in an interactive graph-based visualization that helps to identify how users interact with the different online resources available and how these interactions interfere with their academic performance. To validate this tool a questionnaire with presentation and usability questions was given to users with previous teaching experience. Answers show that the tool delivers the proposed value and is useful to identify user behaviour in virtual learning environments.

Keywords: Computer Science, UFRGS, Data Visualization, Machine Learning, Moodle, Virtual Teaching and Learning Environment.

LISTA DE FIGURAS

Figura 2.1	Log de atividade do Curso	14
Figura 2.2	Notas dos alunos ao longo do semestre	15
Figura 2.3	Exemplo de grafo bipartido	16
Figura 2.4	Exemplo de árvore	16
Figura 3.1	Arquitetura final do sistema	22
Figura 3.2	Diagrama do modelo de predição	23
Figura 3.3	Fluxo de criação de sessões	25
Figura 3.4	Fluxo de transformação dos dados	26
Figura 3.5	Modo de visualização Árvore	27
Figura 3.6	Modo de visualização Bigrafo	29
Figura 4.1	Tela de login	30
Figura 4.2	Tela geral MDV	31
Figura 4.3	Detalhe da área de seleção	32
Figura 4.4	Detalhe da área de visualização	33
Figura 4.5	Seleção de indivíduos	33
Figura 4.6	Seleção de recursos	34
Figura 4.7	Detalhe da área de ajustes, ações e parametrização	35
Figura 4.8	Modo bi-grafo	35
Figura 4.9	Previsão de desempenho	36
Figura 4.10	Modo árvore	37
Figura 5.1	Resultados do questionário para a questão 19	39
Figura 5.2	Resultados do questionário para a questão 10	39
Figura 5.3	Resultados do questionário para a questão 9	39
Figura 6.1	Interface do usuário	42
Figura 6.2	Trajetória do Aluno	45

LISTA DE TABELAS

Tabela 3.1 Acurácia nos dados de validação para diferentes valores de número de episódios e learning rate.....	24
--	----

LISTA DE ABREVIATURAS E SIGLAS

AVA	Ambiente Virtual de Aprendizagem
API	Application Programming Interface
CSV	Comma-separated values
FAQ	Frequently Asked Questions
GNU GPL	GNU General Public License
HTML	HyperText Markup Language
HTTP	Hypertext Transfer Protocol
IA	Inteligência Artificial
INF	Instituto de Informática da Universidade Federal do Rio Grande do Sul
MDV	Moodle Data Visualization
ML	Machine Learning
UFRGS	Universidade Federal do Rio Grande do Sul
CSV	Comma-Separated Values
XLSX	Microsoft Excel Spreadsheet file

SUMÁRIO

1 INTRODUÇÃO	10
2 CONCEITOS E TECNOLOGIAS	13
2.1 Conceitos	13
2.1.1 Moodle	13
2.1.2 Visualização em grafos	15
2.1.3 <i>Deep Learning e Graph Neural Networks</i>	17
2.2 Tecnologias	18
2.2.1 G6 Graph Visualization Engine	19
2.2.2 PyTorch, NetworkX e deepSNAP	19
2.2.3 Flask e Gunicorn	20
2.2.4 Containers	21
3 IMPLEMENTAÇÃO E ARQUITETURA	22
3.1 Arquitetura	22
3.2 Implementação	23
3.2.1 Módulo de IA	23
3.2.2 Autorização de usuários	24
3.2.3 Ferramenta de visualização	26
4 GUIA DE USO	30
4.1 Login	30
4.2 Visão Geral do MDV	30
4.2.1 Área De Seleção	31
4.2.2 Área de visualização e análise	32
4.2.3 Área de ajustes, ações e parametrização	34
4.3 Visualização em bi-grafo	35
4.4 Visualização em árvore	36
5 AVALIAÇÃO E VALIDAÇÃO	38
5.1 Resultados	38
6 TRABALHOS RELACIONADOS	41
6.1 Um <i>plugin</i> do tipo report para a identificação do risco de evasão na educação superior a distância que usa técnicas de visualização de dados	41
6.2 Uma ferramenta de mineração de dados educacionais para o Moodle	42
6.3 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	43
6.4 Ferramenta para visualização de trajetórias de aprendizagem de alunos em cursos da plataforma Moodle	43
7 CONCLUSÕES	46
REFERÊNCIAS	48
APÊNDICE A — FORMULÁRIO TAM	50

1 INTRODUÇÃO

Nas últimas décadas, o avanço da tecnologia e a adoção em massa de computadores, *smartphones* e similares por parte do público em geral possibilitou o surgimento de ferramentas de ensino virtuais, chamadas de Ambientes Virtuais de Aprendizado (AVAs). AVAs podem ser definidos como mídias que utilizam o ciberespaço para veicular conteúdos e permitir interação entre os atores do processo educativo (PEREIRA; SCHMITT; DIAS, 2007). Por permitirem uma melhor organização das tarefas docentes e fornecerem uma comunicação assíncrona ou instantânea entre pessoas em diferentes localidades, essas ferramentas vêm sendo amplamente adotadas pelas instituições de ensino.

A recente pandemia de COVID-19 intensificou o uso dessas tecnologias em várias instituições de ensino, à medida que restrições de circulação resultaram em uma adoção emergencial de aulas *online*. Essa mudança trouxe à tona questionamentos com relação a qualidade do ensino remoto, às mudanças em relação à metodologia tradicional e às possíveis dificuldades e problemas resultantes dessa tecnologia. Por conta disso, vários estudos têm sido feitos em instituições ao redor do mundo para entender como a mudança afetou o processo de aprendizagem, em especial em países que seguiam o modelo tradicional de ensino (BAO, 2020; CHIRIKOV et al., 2020; OLATUNDE-AIYEDUN; OGUNODE; AIYEDUN, 2021).

Muito embora AVAs não tenham como objetivo substituir o papel do professor no processo de ensino, essas ferramentas possibilitam automatizar várias tarefas que até então eram feitas pelo docente. Disponibilização de conteúdos, abertura e fechamento de tarefas, comunicados e correções são exemplos do que pode ser automatizado pelos AVAs. Milligan (1999) elenca dez funcionalidades que um AVA deve fornecer: acesso virtual a materiais didáticos; acompanhamento do progresso do estudante; controle de tempo; possibilidade de criar avaliações; ferramentas de comunicação síncrona e assíncrona; espaço para salvar materiais; possibilidade de atualizar recursos; ambientes de ajuda (como FAQs); informação de progresso para o estudante; e um guia de usuário.

Uma ferramenta que cumpre esses dez requisitos e que é amplamente adotada pela comunidade acadêmica é a plataforma Moodle (HQ, 2022a), que disponibiliza um software *open source* com uma licença GNU GPL. O Moodle é um ambiente completo, facilmente estendido ou modificado de acordo com as preferências dos administradores e que é utilizado pela Universidade Federal do Rio Grande do Sul (UFRGS) como ferramenta padrão de aprendizagem virtual. O professor responsável por determinada disciplina pode

criar uma página para agregar os alunos, tutores e outros envolvidos para centralizar a disponibilização de conteúdos, tarefas, avaliações e demais interações à critério do docente.

A grande variedade de funcionalidades e opções diferentes para melhorar a experiência de aprendizagem da plataforma Moodle acaba por introduzir uma desvantagem para os usuários. As opções são tão variadas que os alunos priorizam certos recursos em relação a outros, ignoram ou até desconhecem funcionalidades que poderiam ser úteis para o processo de aprendizagem. É possível que as preferências dos estudantes ao utilizar um AVA variem por diversos fatores, como localidade, idade e curso escolhido. Analisar as interações dos alunos da UFRGS no Moodle é uma forma de trazer luz a essas questões.

Uma das funcionalidades incluídas na plataforma Moodle é o *log* de todas as interações que os usuários fazem com o sistema, que ficam salvos em um banco de dados de uso interno da plataforma e que podem ser extraídos por usuários com permissões administrativas sobre o recurso, como o professor que criou a página da disciplina. Esses *logs* são extraídos em formatos tabulares, como CSV e XLSX. No entanto, essas informações não são disponibilizadas em formato intuitivo e é uma tarefa não-trivial extrair qualquer conclusão desses dados sem passar por algum pré-processamento para filtrar as informações desejadas.

Já que esse obstáculo introduz uma barreira técnica para que usuários não acostumados com processamento e visualização de dados utilizem essas informações, seria interessante que existisse um recurso para tornar essa tarefa factível e intuitiva. Um recurso que possibilitaria o professor visualizar de que forma as funcionalidades da plataforma são utilizadas pelos estudantes e como essas preferências afetam a aprendizagem do aluno.

Portanto, este trabalho visa a criação de uma ferramenta que processe esses *logs* e os mostre em uma visualização interativa para que o usuário possa realizar sua análise sem um conhecimento técnico de manipulação e processamento de dados. A ferramenta deve permitir a exibição de informações de forma que o usuário consiga identificar componentes do ambiente do Moodle, como fóruns, trabalhos e avaliações, e a sua frequência de acesso por parte dos estudantes. Adicionalmente, a ferramenta deve permitir identificar padrões de usabilidade diferentes entre os estudantes e as suas possíveis relações com o progresso e desempenho no processo de aprendizagem.

Considerando as necessidades de interação, filtragem e visualização de um grande volume de dados representando interações de usuários do AVA entre si ou com componentes do Moodle, é necessário apresentar uma visualização que evidencie essas interações

tanto quanto evidencia as próprias entidades envolvidas no processo. Segundo (HERMAN; MELANCON; MARSHALL, 2000), um conjunto de dados estruturados que apresenta uma relação natural entre seus elementos é um candidato aplicável para um modelo de visualização baseado em grafos. Como este trabalho visa evidenciar a relação entre dois conjuntos de entidades, que são o conjunto de alunos pertencentes a uma disciplina e o conjunto de componentes do sistema disponíveis aos estudantes nessa disciplina, a ferramenta implementa uma visualização baseada em grafos.

Também foi incluída na ferramenta a possibilidade de utilizar um modelo de *deep learning* para prever o desempenho acadêmico de um aluno baseado nas suas interações com o Moodle. O modelo foi treinado com dados de turmas já encerradas e é disponibilizado na interface como uma funcionalidade adicional.

Com isso, a ferramenta deve 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 se consiga fazer uma análise de como os discentes utilizam a plataforma. A ferramenta se propõe a auxiliar na análise e descrição dos comportamentos dos estudantes no Moodle, o que ajudaria em análises futuras envolvendo a identificação das características de estudantes no contexto brasileiro se comparado com as pesquisas feitas em outras regiões do mundo.

O restante do trabalho é dividido da seguinte forma: Seção 2 descreve conceitos relevantes ao trabalho e tecnologias utilizadas. Seção 3 descreve a arquitetura final do projeto. Seção 4 apresenta a usabilidade da ferramenta. Seção 5 mostra os resultados do processo avaliativo com usuários teste. Seção 6 discute e compara a ferramenta com projetos relacionados e seção 7 apresenta comentários finais do trabalho.

2 CONCEITOS E TECNOLOGIAS

Para um melhor entendimento sobre este trabalho, alguns conceitos e tecnologias precisam ser conhecidos. Este capítulo traz uma explicação sobre estes assuntos.

2.1 Conceitos

Esta seção apresenta a terminologia do Moodle que descreve os recursos da plataforma. Também apresenta conceitos variados relacionados às funcionalidades da ferramenta de visualização.

2.1.1 Moodle

O *design* e as funcionalidades da plataforma Moodle foram desenvolvidos baseados em uma abordagem da pedagogia conhecida como socioconstrutivismo, que define que o processo de desenvolvimento e aprendizagem humano é construído principalmente a partir de interações com outros indivíduos (HQ, 2022b). Por esse motivo, a maioria dos componentes em uma página de curso do Moodle envolvem 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 utilizados pelos cursos do Moodle da UFRGS:

- **Página:** É a página inicial de uma disciplina, que usualmente contém uma seção de avisos junto com links relevantes sobre o funcionamento das aulas, como plano de ensino, cronogramas e links para FAQs. Pode ser dividido em módulos que representam divisões no conteúdo ou no cronograma do semestre, por exemplo;
- **Arquivo:** Esse componente é o que permite o docente fazer *upload* de arquivos que ficarão disponíveis na página para acesso dos estudantes, representados por um link. É uma das formas mais comuns de disponibilizar conteúdos na plataforma;
- **Fórum:** É um fórum criado para a disciplina, que concentra dúvidas dos estudantes sobre o conteúdo ou andamento das aulas. Tanto professores e tutores quanto alunos podem responder aos tópicos criados;

- **Tarefa:** É uma seção especial do curso que inclui um espaço para descrever a o que está sendo pedido, um espaço para adicionar arquivos ou escrever um texto, uma espaço para notas e também uma seção de comentários sobre o envio que serve como uma forma de *feedback* para o aluno;
- **Pasta:** Um agrupador de componentes para melhorar a organização da página do curso;
- **Questionário:** Uma maneira de criar tarefas pelo Moodle. Permite a elaboração de questões objetivas e dissertativas e a possibilidade de utilizar uma correção automática.
- **Relatório de Notas:** Parte que concentra as notas do estudante nas tarefas e avaliações ao longo do curso.
- **Sistema:** Componente que compreende várias funcionalidades de menor relevância para os atores de 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, toda vez que um usuário visualiza a página do curso essa ação é logada pelo sistema.
- **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). Esse software é disponibilizado como um serviço pelo Centro de Processamento de dados (CPD) da UFRGS e é integrado ao Moodle através de um *plugin* que estende as funcionalidades base da plataforma.

Um usuário com permissão de administrador sobre a página do curso pode baixar os logs pelo caminho *Administration > Course administration > Reports > Logs*. Um exemplo de log é dado na figura 2.1.

Figura 2.1: 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
```

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. Em caso de não ter afetado um

usuário, a coluna é preenchida com '-'.

- **Contexto do Evento:** O curso onde o evento ocorreu.
- **Componente:** O 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 veio 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 de cada aluno na disciplina. Ela serve como parâmetro para a identificação de trajetórias que ocasionam em um baixo desempenho acadêmico. A Figura 2.2 mostra um exemplo da planilha de notas em um curso no Moodle exportada como 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.

Figura 2.2: 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

Os dados disponíveis para a realização desse trabalho foram os logs e as notas de quatro turmas das disciplinas *DERAD502 - Cooperativismo E Economia Solidária*, *DERAD501 - Segurança Alimentar E Nutricional*, *DERAD202 - Questão Ambiental E Legislação* e *DERAD109 - Métodos De Pesquisa Científica*. As disciplinas foram ministradas entre setembro de 2020 e fevereiro de 2021. Ao total, as disciplinas envolvem 528 alunos, excluindo os professores, tutores e palestrantes.

2.1.2 Visualização em grafos

Tendo em vista que a motivação deste trabalho é analisar as relações do usuário com os variados recursos do Moodle, uma visualização em grafos permite representar

de forma explícita essas interações por meio de arestas. O uso de grafos em análise de AVAs aparece em (KUHN, 2019), onde é apresentado um estudo envolvendo grafos na representação de trajetórias de aprendizagem em ambientes virtuais. Além disso, é possível utilizar recursos gráficos da visualização em grafos como cores, espessura das arestas e expansão de nodos para incluir informações adicionais que enriquecem a análise do usuário da ferramenta. Portanto, este trabalho apresenta as informações modeladas em formato de grafo.

Mais especificamente, duas estruturas de grafos são utilizadas: grafos bipartidos, aqui chamados de bi-grafos, 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. Exemplos dos dois tipos são dados nas figuras 2.3 e 2.4.

Figura 2.3: Exemplo de grafo bipartido

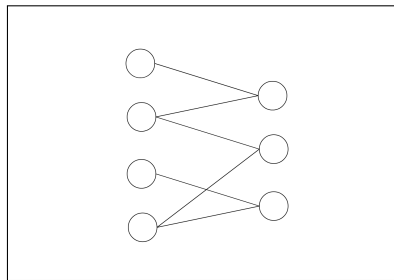
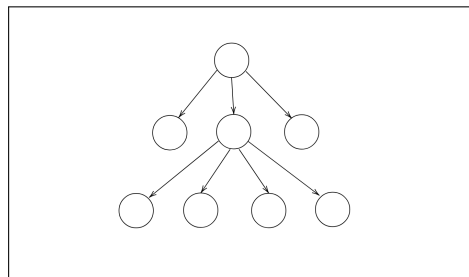


Figura 2.4: Exemplo de árvore



No trabalho de Bennett et al. (2007), é apresentado um modelo para os fatores envolvidos na interpretação de visualização em grafos. Em 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. Com isso em vista, é enunciado alguns critérios que formam uma boa visualização (SUN; 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 esses critérios foram usados para formar uma visualização que seja intuitiva para o usuário. Por ser uma visualização interativa, a manipulação do grafo permite diminuir ou aumentar a complexidade das informações mostradas, facilitando a percepção dos elementos à medida que a visualização é alterada.

2.1.3 Deep Learning e Graph Neural Networks

Dentre os vários problemas que a área de Inteligência Artificial explora, um dos mais conhecidos é a busca por um algoritmo eficiente para achar padrões em um conjunto de dados de entrada e classificá-los de acordo com a necessidade. *Feature Learning*, ou aprendizado de representação, é um conjunto de técnicas que busca resolver esse problema por meio do aprendizado automático das representações necessárias para a classificação.

Deep Learning, ou aprendizado profundo, é uma técnica de aprendizado por representação que utiliza várias camadas de módulos não-lineares para unir múltiplos níveis de abstração em sequência e, assim, aprender padrões complexos. Essa técnica se tornou famosa nas últimas décadas por produzir bons resultados em problemas de predição tradicionais, como processamento de linguagem natural e reconhecimento de fala (LECUN; BENGIO; HINTON, 2015).

A utilização de grafos como dados de entrada para um algoritmo de classificação de *deep learning* vem sendo explorada nos últimos anos. Por meio da adaptação dos

dados de entrada, um algoritmo de *graph neural network* pode realizar predição de nodos, predição de arestas ou classificação de grafos (HU et al., 2019). Neste trabalho essa técnica é utilizada para prever o desempenho acadêmico de alunos baseados nas suas interações com o Moodle representadas por um grafo.

Para o treinamento, o baixo volume de dados disponíveis requereu a utilização de técnicas para evitar *overfitting*. *Overfitting* é um problema que ocorre quando o modelo de predição captura durante a aprendizagem não só os padrões dos dados relevantes à tarefa mas também as flutuações randômicas nos exemplos, *outliers* e características específicas do dado que não são generalizáveis para um volume de treinamento maior.

Existem algumas técnicas para contornar esse problema que podem envolver criar dados novos a partir dos conhecidos para aumentar o volume de dados ou alterar a estrutura do modelo para dificultar que esses problemas ocorram. A técnica escolhida foi o *Dropout* (HINTON et al., 2012), que consiste em alterar a estrutura da árvore ao longo do treinamento para adicionar barulho e induzir o modelo a descartar características específicas de cada *datapoint*.

Por fim, o treinamento também deve levar em conta os hiperparâmetros do modelo, que são os parâmetros que influenciam na velocidade do processo de aprendizagem. Para o *graph neural network* treinado nesse trabalho, esses parâmetros são o *learning rate*, o número de episódios de treinamento, o tamanho da camada oculta de treinamento e o número de camadas. Para achar valores eficientes para esses parâmetros foi utilizada a técnica de *grid search*, que consiste em uma automatização do treinamento utilizando um conjunto de combinações de valores para os hiperparâmetros. Ao fim da busca, a combinação que obteve a melhor acurácia é escolhida.

2.2 Tecnologias

A ferramenta apresentada neste projeto é construída como uma aplicação Web. Como é acessível através de um navegador, não é necessário baixar e instalar o projeto. Abaixo são listadas as tecnologias necessárias para disponibilizar a ferramenta aos usuários.

2.2.1 G6 Graph Visualization Engine

Devido à popularidade e adoção massiva de tecnologias Web ao redor do mundo, a diversidade de projetos pode levar a problemas de compatibilidade e dificuldades para os desenvolvedores que têm que lidar com as especificações de cada tecnologia. Por esse motivo, a organização *Web Hypertext Application Technology Working Group (WHATWG)* define padrões para as várias tecnologias envolvidas no desenvolvimento web, como transferência de arquivos, Web Sockets, notificações, armazenamento e elementos visuais (WHATWG, 2022). Uma dessas especificações é o HTML, que inclui não só a *markup language* HTML mas também as bibliotecas que estendem a linguagem, como a Canvas API.

Atualmente, os navegadores mais populares suportam duas APIs que permitem adicionar elementos gráficos nas páginas da Web sem a necessidade de plugins: Canvas API e WebGL. Canvas API é uma ferramenta antiga que suporta animações somente em 2D e é recomendada para gráficos e animações simples. WebGL é uma API mais recente, que suporta gráficos em 2D e 3D e é utilizada em vários projetos envolvendo simulações e processamento de imagens, por permitir suporte a GPUs.

Tanto Canvas quanto WebGL não incluem estruturas complexas como grafos, objetos em 3D e simulações de física. Por esse motivo, vários projetos estendem as funcionalidades tradicionais dessas APIs com elementos de alto nível. Um desses projetos é o G6 Graph Visualization Engine, que fornece componentes para a visualização e manipulação de grafos (ANTV, 2022). Por estender bibliotecas que já são suportadas pelos navegadores mais populares, esse projeto também é suportado na maioria das versões dos navegadores Google Chrome, Mozilla Firefox e Microsoft Edge.

Esse trabalho utiliza o G6 para fornecer a interface gráfica de manipulação dos grafos de interações no Moodle. Os demais elementos da página foram escritos com as bibliotecas padrão do Javascript e não requerem plugins adicionais.

2.2.2 PyTorch, NetworkX e deepSNAP

O modelo de predição utilizado no trabalho foi treinado na linguagem Python. Uma das ferramentas mais conhecidas de aprendizado de máquina é o framework PyTorch, que fornece várias ferramentas para auxiliar no treinamento de modelos e na disponibilização desses para a utilização em aplicações reais (PYTORCH, 2022). Para mo-

delar os dados do trabalho em grafos foi utilizada a biblioteca NetworkX (NETWORKX, 2022), que inclui várias ferramentas para a modelagem e edição de grafos, bem como vários dos algoritmos tradicionais da área.

Foi também utilizada uma biblioteca que estende as funcionalidades do PyTorch para facilitar o desenvolvimento de *graph neural networks*, a deepSNAP (SNAP, 2022). A deepSNAP permite a integração de grafos definidos com a NetworkX com os modelos de predição definidos com o PyTorch. Para esse trabalho, ela é utilizada para inserir os grafos que representam as interações dos alunos com os componentes do Moodle no modelo de predição de desempenho acadêmico. Cada grafo é um *datapoint*, a entrada do modelo de predição.

2.2.3 Flask e Gunicorn

Para tornar o trabalho disponível para os usuários é necessário fornecer a ferramenta por um servidor. Dentre as diversas opções de framework e linguagem existentes atualmente, foi escolhido o web framework Flask, para a linguagem Python. Flask é um framework mínimo, o que significa que não inclui funcionalidades comuns como validação de formulário e abstrações para comunicação com banco de dados. Essa característica é relevante para esse trabalho, já que a aplicação não requer a maior parte das funcionalidades de um framework completo.

Visto que a aplicação envolve dados sensíveis dos usuários, como nome completo, nota final na disciplina, email e endereço de IP, é necessário proteger a aplicação de acessos não autorizados. Para isso, foi utilizado um paradigma de sessões para fornecer acesso ao conteúdo protegido do site, que são liberados por meio de uma senha. O processo de login e logout e o gerenciamento de sessões foi feito usando as extensões Flask-login e Flask-WTF.

Outro motivo que levou à escolha do Flask é a necessidade de utilizar a linguagem Python. Já que esse trabalho inclui um módulo de aprendizado de máquina e uma das bibliotecas utilizadas, a deepSNAP, é implementada somente em Python, utilizar um framework escrito na mesma linguagem simplifica a arquitetura do sistema.

Já que o Flask é um framework mínimo, é necessário instalar mais um componente que implemente a lógica do servidor no modelo WSGI. Web Server Gateway Interface (WSGI) é uma especificação que descreve como um servidor se comunica com aplicações web ou com demais aplicações que respeitem o mesmo modelo (WSGI, 2022). Foi

escolhido o Gunicorn, que implementa essa especificação.

2.2.4 Containers

A ferramenta ficou acessível a partir de uma máquina do instituto de informática da UFRGS disponível por acesso com SSH. Para transferir a aplicação para a máquina e evitar os vários problemas de compatibilidade que podem ocorrer ao executar um projeto em um ambiente diferente do que foi desenvolvido, foi utilizado um serviço de virtualização, o Docker Engine (INC, 2022).

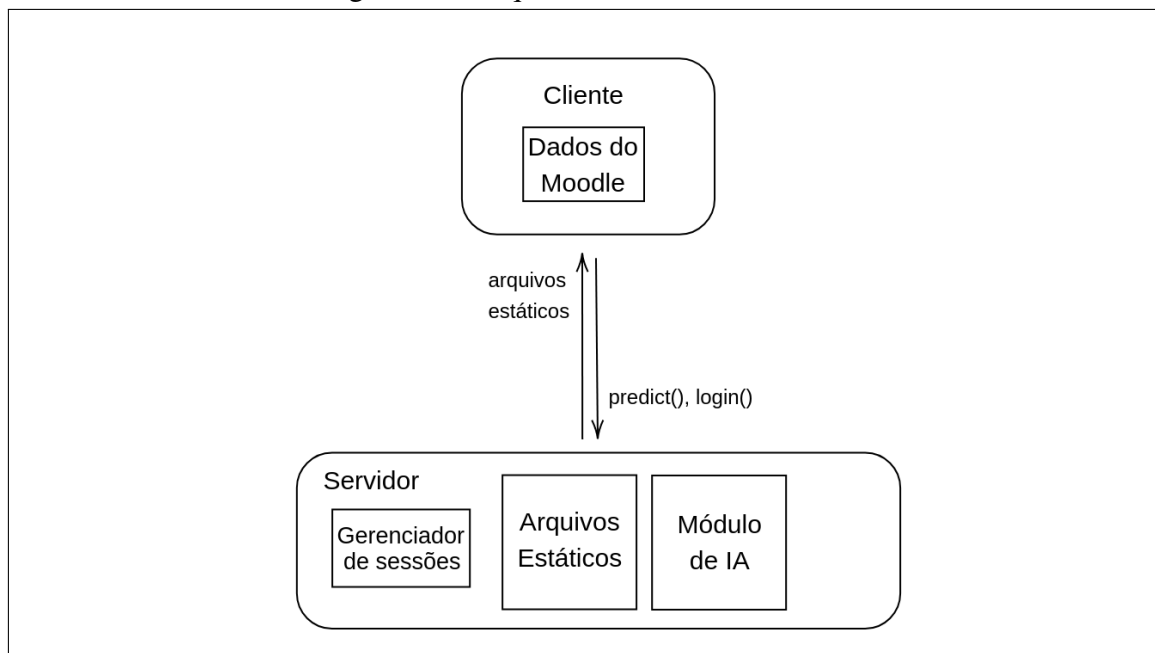
Docker Engine é um serviço que fornece virtualização a nível de sistema operacional para executar software em módulos chamados containers. Cada container possui seus próprios arquivos de configuração, caminhos, softwares instalados, usuários e demais funcionalidades de um sistema operacional. Neste trabalho, também foi utilizado a plataforma DockerHub, desenvolvida pela mesma organização, para compartilhar o projeto entre máquinas diferentes.

3 IMPLEMENTAÇÃO E ARQUITETURA

Este capítulo apresenta a arquitetura final e uma descrição do processo de implementação. A ferramenta foi disponibilizada em um subdomínio do Instituto de Informática da UFRGS através de um servidor HTTP aberto na porta 80. A figura 3.1 mostra os elementos principais do sistema.

3.1 Arquitetura

Figura 3.1: Arquitetura final do sistema



O servidor contém os arquivos estáticos e os dados brutos que serão enviados ao cliente após passar pela etapa de autorização. O módulo de sessão administra as sessões dos usuários e os processos de *login* e *logout*. O módulo de IA contém o modelo de predição de desempenho acadêmico previamente treinado com os dados disponíveis.

O cliente consegue utilizar a ferramenta após receber os arquivos estáticos correspondentes à interface gráfica e os dados necessários armazenados no servidor. Caso precise fazer uma predição de desempenho acadêmico, é feita uma requisição ao servidor contendo os dados do aluno. O servidor então executa o modelo de IA com os dados recebidos como entrada e retorna ao cliente a predição.

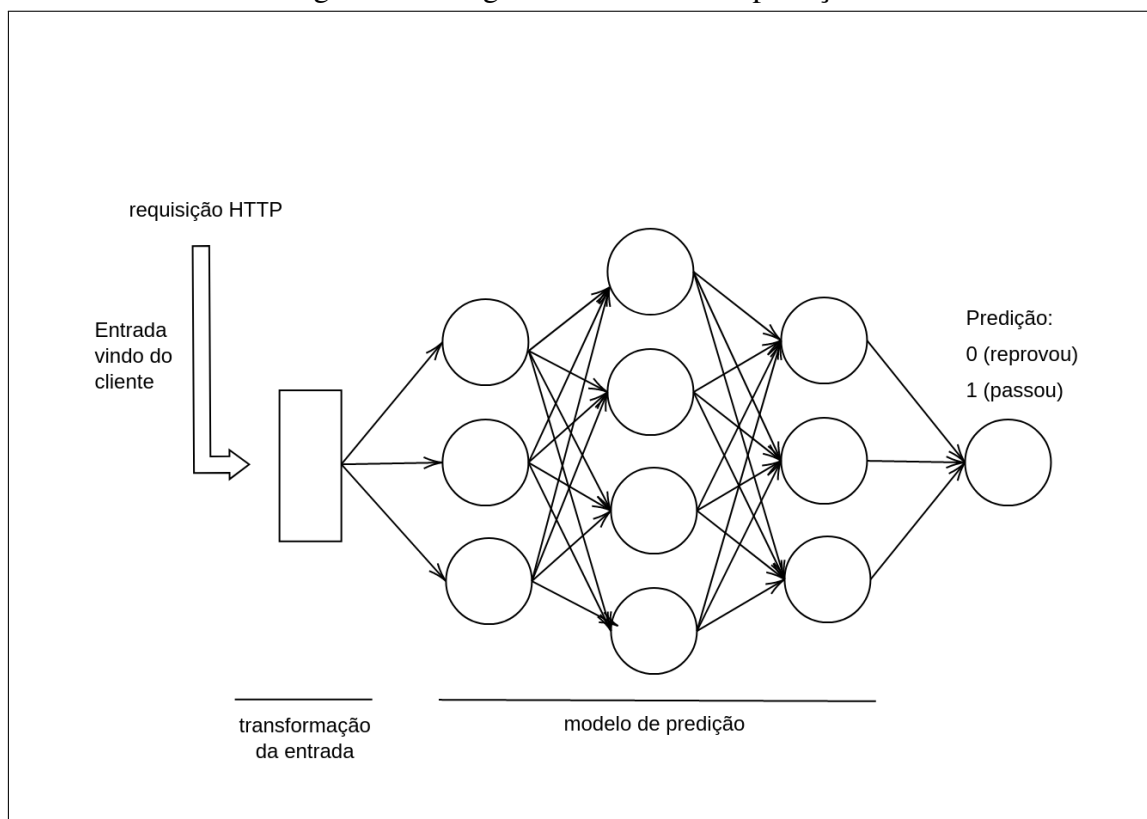
3.2 Implementação

3.2.1 Módulo de IA

Conforme comentado na seção 2.1, este módulo utiliza uma *graph neural network* para fazer previsões em cima de um grafo que representa as interações de um aluno com o Moodle. Ao contrário das demais funcionalidades da ferramenta, que funcionam sem requisições adicionais para o servidor, o modelo de previsão não é transferido para o cliente junto com os arquivos estáticos.

Esse modelo foi treinado com os logs de interação e com as notas finais dos alunos contidos nos dados disponíveis. Como o objetivo é prever 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 para a previsão ser "0", que representa não aprovado, ou "1" que representa aprovado. A figura 3.2 mostra um diagrama do processo de previsão.

Figura 3.2: Diagrama do modelo de previsão



O treinamento do modelo foi feito com os dados das quatro turmas disponíveis enunciados na seção 2.1. Ao todo, seriam 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. A tabela 3.1 mostra os resultados do grid search para os parâmetros: número de episódios e learning rate. Os demais parâmetros que não aparecem na tabela obtiveram os valores ótimos: 48 para número de nodos nas camadas ocultas e 3 para número de camadas.

Tabela 3.1: Acurácia nos dados de validação para diferentes valores de número de episódios e learning rate.

	$e_p = 50$	$e_p = 100$	$e_p = 200$	$e_p = 400$
$l_r = 0.0001$	0.571	0.571	1.000	1.000
$l_r = 0.001$	0.857	0.714	1.000	0.857
$l_r = 0.01$	0.857	1.000	1.000	1.000

Os resultados do grid search mostram valores elevados de acurácia para um learning rate superior a 0.0001 e número de episódios superior a 50. Para evitar o problema de overfitting foi selecionado o modelo com menor número de nodos que obteve um bom desempenho. Portanto, os parâmetros selecionados foram: 12 nodos na camada oculta, 100 episódios, 3 camadas e learning rate de 0.001. A acurácia nos dados de teste para esse modelo foi de 0.875.

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 cliente.

3.2.2 Autorização de usuários

Como os logs e as notas dos alunos contêm dados sensíveis, o acesso à visualização foi protegido por autenticação por senha usando sessões. Já que é necessário apenas

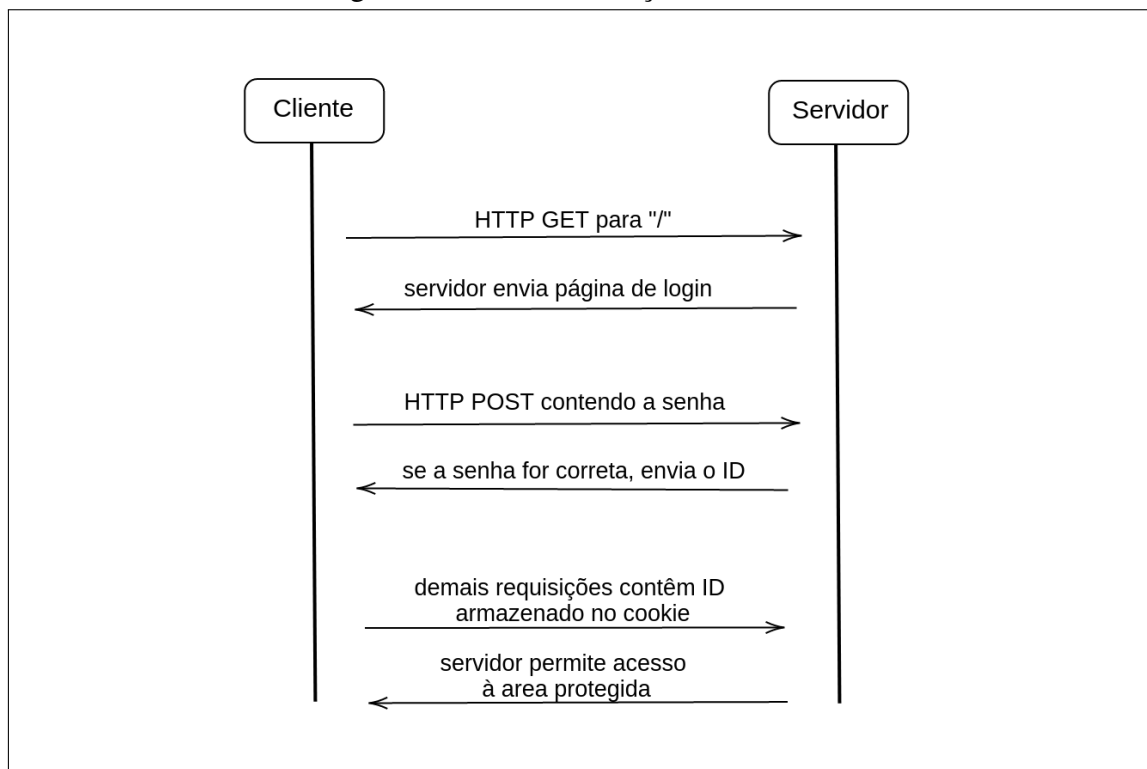
a proteção dos dados do servidor e não identificação dos clientes, a aplicação não necessita de múltiplos usuários. Portanto foi criado um usuário global que serve para criar as múltiplas sessões no servidor.

Um usuário que se conecta ao servidor recebe em resposta a página de login, que contém um único campo para inserir a chave de acesso à ferramenta. Uma vez que o usuário digita a senha e envia o formulário, o servidor valida a senha e, caso esteja errada, envia uma mensagem informando o usuário para que ele faça outra tentativa.

Caso a senha esteja correta, o servidor cria um objeto de sessão que contém um ID único para o estado do servidor. Esse ID é enviado de volta para o cliente. Então, o navegador armazena esse ID como um cookie http-only, que é enviado em todas as próximas requisições feitas pelo cliente até o período de expiração do cookie ou até o usuário requisitar um logout.

Quando o servidor recebe uma requisição em uma área protegida, ele checa os cookies enviados pelo cliente para validar o ID contido no cookie. Caso seja válido e exista um objeto de sessão correspondente, o acesso é autorizado e o conteúdo requisitado é enviado de volta para o cliente. A figura 3.3 ilustra o processo de autorização de usuários.

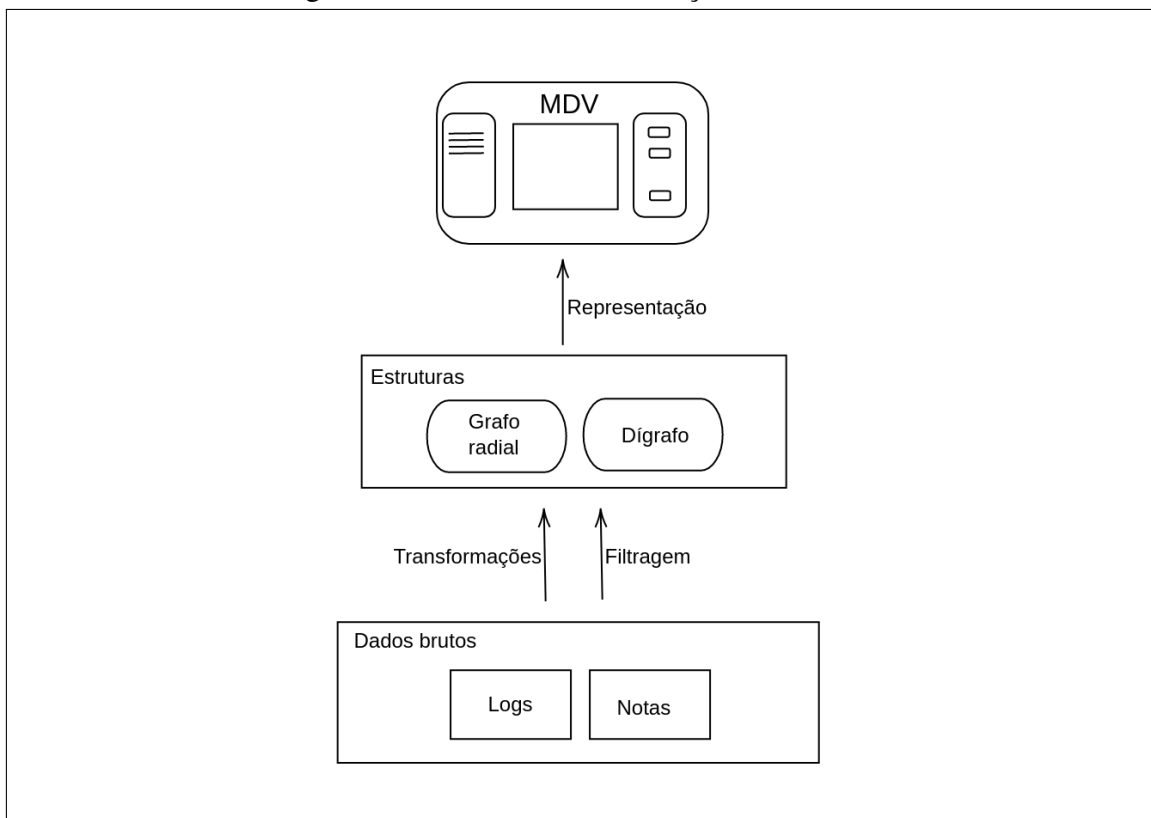
Figura 3.3: Fluxo de criação de sessões



3.2.3 Ferramenta de visualização

Uma vez que o cliente tenha acesso à ferramenta, o servidor envia os arquivos referentes à visualização e os dados de logs e notas. Conforme o usuário utiliza a ferramenta e manipula a interface, transformações vão sendo feitas nas estruturas de visualização para mostrar as informações requisitadas pelo usuário. Um diagrama desse processo está na figura 3.4.

Figura 3.4: Fluxo de transformação dos dados



De início, os dados são processados para a transformação nos grafos de visualização. O processamento envolve agrupar logs por estudante para que se consiga analisar as interações de um indivíduo, agrupar métricas gerais sobre a utilização dos componentes do sistema, descartar informações que não são utilizadas na visualização, como IP de origem; e formatar os dados para que fiquem legíveis ao usuário. Nessa etapa também é feita a filtragem de informações conforme as restrições definidas pelo usuário ao interagir com a interface.

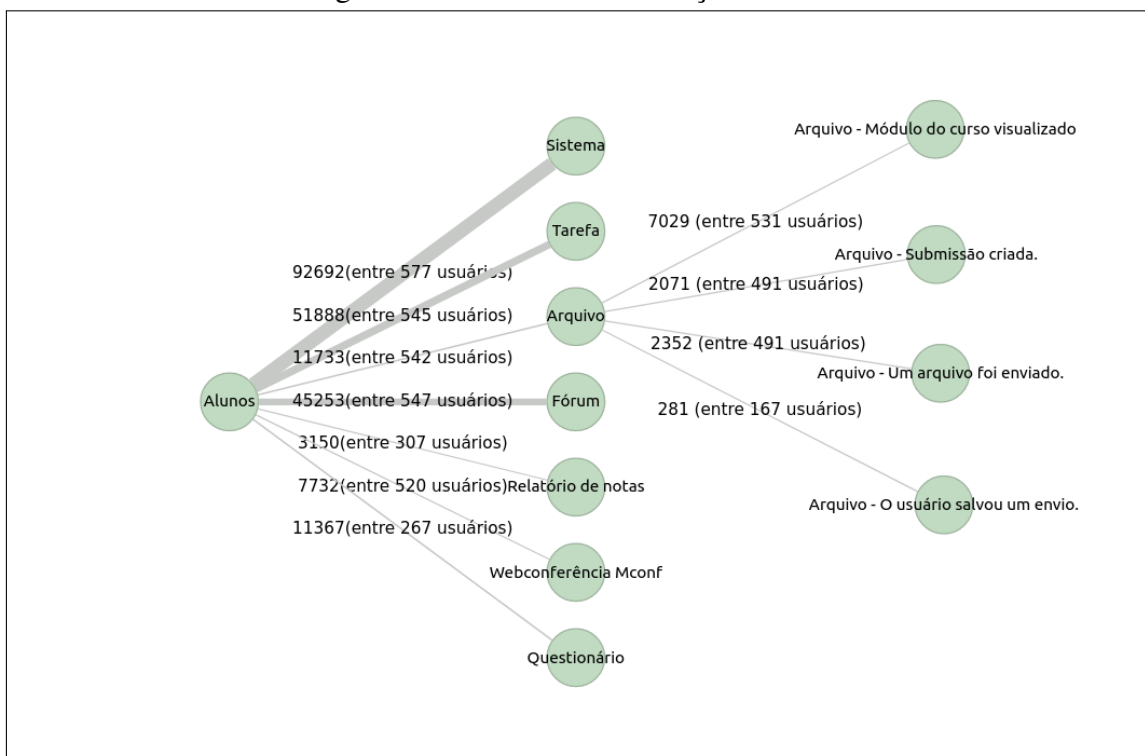
Em seguida, os grafos resultantes são enviados para a graph visualization engine. Nesta etapa os elementos que compõe a interface são definidos, como cor dos nodos, legendas, informações adicionais, métricas visíveis e a distribuição dos elementos na tela.

A engine então transforma essas configurações em elementos visuais que são exibidos ao usuário.

Por fim, a visualização é inserida no código HTML da página e mostrada ao usuário. O usuário pode então interagir com funcionalidades comuns como *zoom in* e *zoom out*, selecionar nodos e mostrar legendas. Caso ele queira informações mais específicas, os filtros permitem selecionar por modo de visualização e filtrar por nota final, nome do aluno e tipo de nodo. Quando os filtros são aplicados, o processamento dos dados é realizado novamente para extrair as informações requisitadas e a interface é atualizada com o resultado do processamento.

São criados dois modos de visualização: árvore e bigrafo. Cada um possui sua própria estrutura criada a partir dos dados brutos, mas apenas um é mostrado por vez ao usuário. O modo árvore mostra informações agregadas sobre o conjunto de alunos selecionados. Esse modo, mostrada na figura 3.5, evidencia quais componentes do Moodle recebem mais acessos pelos alunos.

Figura 3.5: Modo de visualização Árvore



Na imagem, o primeiro nodo representa uma junção das interações do grupo de estudante, e as suas arestas indicam o número total de interações seguido pelo número de usuários que fizeram essas interações. Podemos ver que, além do sistema, os componentes do Moodle que os alunos dessas disciplinas mais utilizaram foram as Tarefas, com 51.888 interações realizadas por 545 usuários, e o Fórum, com 45253 interações realizadas por

547 usuários.

Além dos componentes principais, é possível analisar também quais ações dentro de cada componente foram feitas pelos usuários. No exemplo da imagem, a maior parte das interações com o componente de Arquivos foi a visualização do recurso, seguido pelo envio de um arquivo.

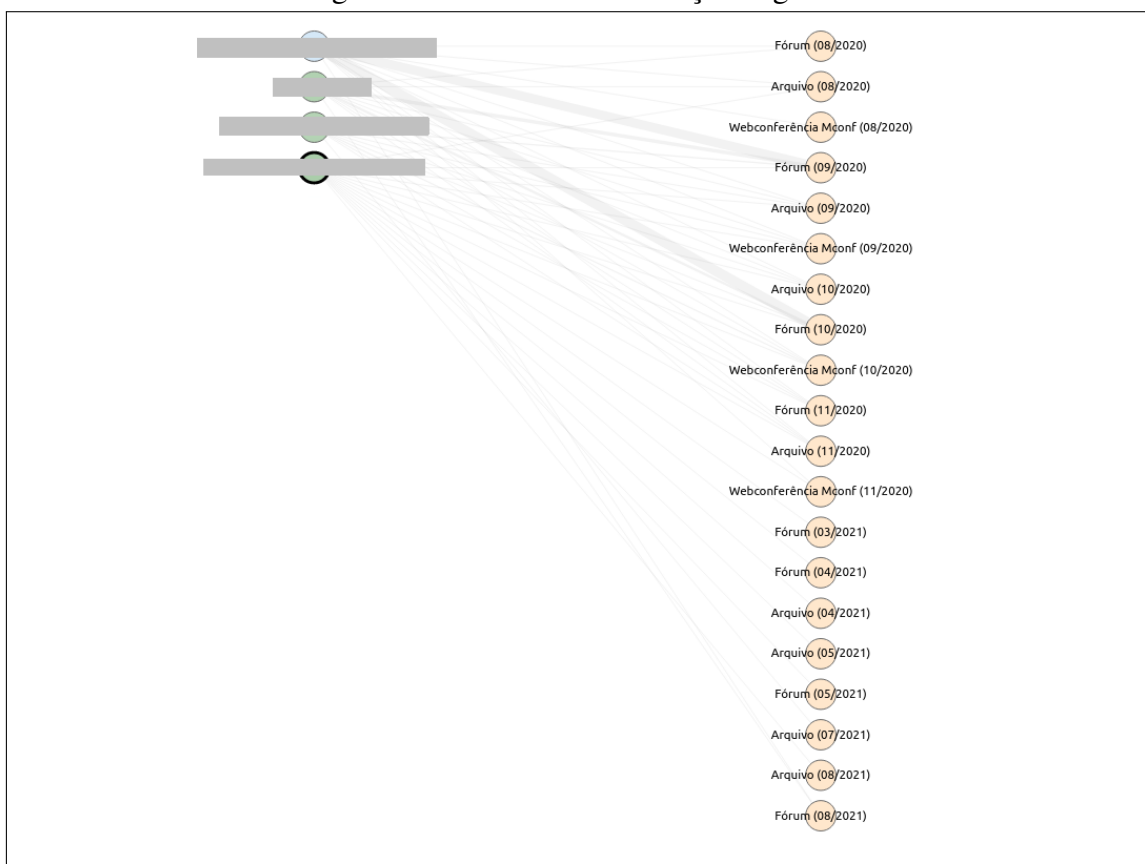
O segundo modo de visualização, bigrafo, é mostrado na figura 3.6. Esse modo evidencia as interações de indivíduos com os componentes do Moodle. A coluna da esquerda representam os usuários selecionados do sistema. Nodos azuis indicam que o indivíduo não é um aluno e não possui nota final, como é o caso de professores e tutores. Nodos verdes indicam alunos que aprovaram na disciplina. Nodos vermelhos indicam alunos que reprovaram. As arestas de cada nodo indicam a quantidade de interações realizadas pelos usuários com o componente.

A coluna da esquerda indica os componentes do sistema separados pelo mês em que foram acessados. Essa coluna evidencia o nível de engajamento do usuário ao longo do semestre. Para permitir essa análise, foi necessário reduzir o número de componentes que aparecem na tela. Nessa imagem, são mostrados os componentes Fórum, Arquivo, Webconferência e Questionário ao longo dos meses.

Os filtros alteram a estrutura do grafo em que a visualização é baseada. Remover os nodos que não são estudantes significa remover os nodos de usuário que não possuem nota final. Filtrar os nodos por nota remove os nodos que não pertencem ao intervalo especificado. Limitar o número de nodos na visualização estabelece um número máximo de nodos de usuário que podem aparecer na tela, para a visualização não ficar poluída. Selecionar um usuário específico mostra apenas o nodo desse usuário e os componentes que ele interagiu.

É também no modo de visualização bigrafo que ocorre a interação com o módulo de IA do servidor. Quando um nodo é selecionado, mostrado pelo contorno escuro como indicado no terceiro nodo de usuário da figura 3.6, é possível utilizar um botão para enviar uma requisição que retorna a predição sobre o desempenho acadêmico do usuário.

Figura 3.6: Modo de visualização Bigrafo



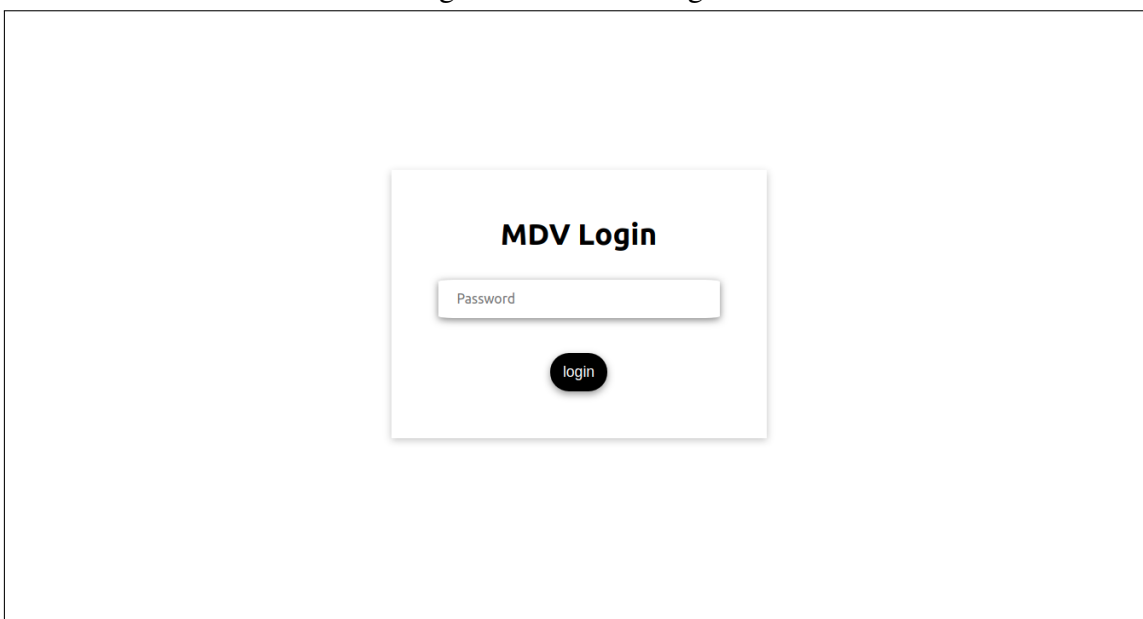
4 GUIA DE USO

Esse guia pode ser utilizado para auxílio do usuário(a) quanto a utilização da ferramenta Moodle Data Visualization (MDV). 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.

4.1 Login

Ao entrar no site da ferramenta, aparece uma tela de login que protege o acesso à área principal. Basta digitar a senha correta que o usuário será redirecionado para a ferramenta. Caso erre, aparecerá uma mensagem de erro com "senha incorreta" na tela e o usuário poderá tentar novamente.

Figura 4.1: Tela de login

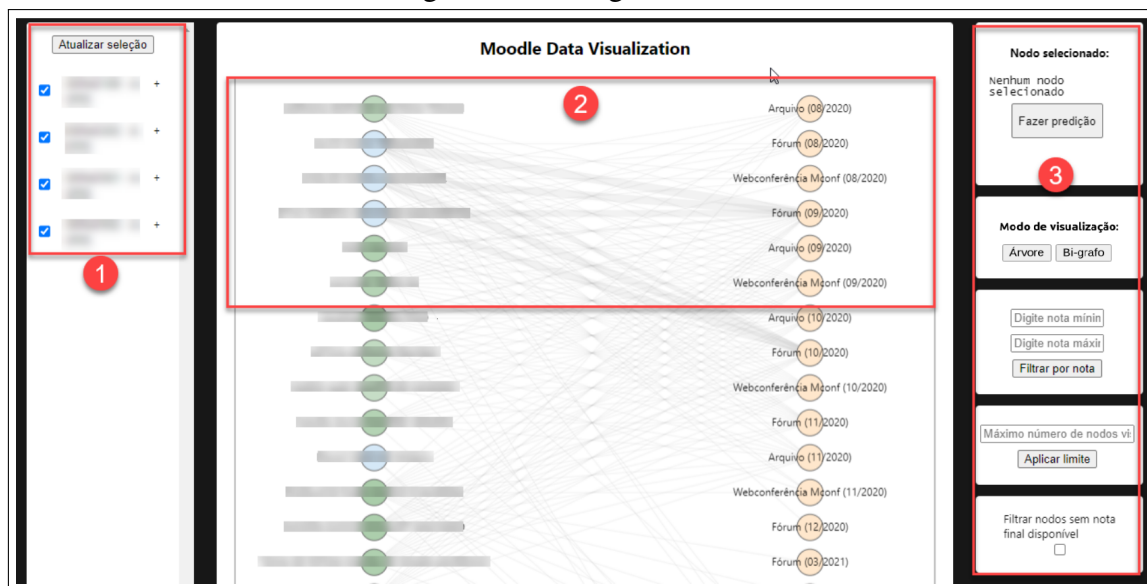


4.2 Visão Geral do MDV

Uma vez na área principal, o usuário irá se deparar com a interface mostrada na figura 4.2. Abaixo são listadas as principais funcionalidades da ferramenta:

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

Figura 4.2: Tela geral MDV



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 previsões usando o módulo de IA.

A seguir, os elementos disponíveis nessas áreas serão detalhados.

4.2.1 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).

1. Seleção de turmas (1): aqui podemos selecionar a(s) turma(s) que desejamos visualizar.
2. Atualizar seleção: ao clicar no botão (2), são apresentados na tela os dados relativos às turmas ou alunos selecionados.
3. Escolha de alunos a serem analisados: na mesma lista, se desejado, podemos sele-

Figura 4.3: Detalhe da área de seleção



cionar 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”.

4.2.2 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).

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 pode 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.

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 nele, podemos visualizar qual ou quais participantes interagiram com aquele recurso.

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,

Figura 4.4: Detalhe da área de visualização

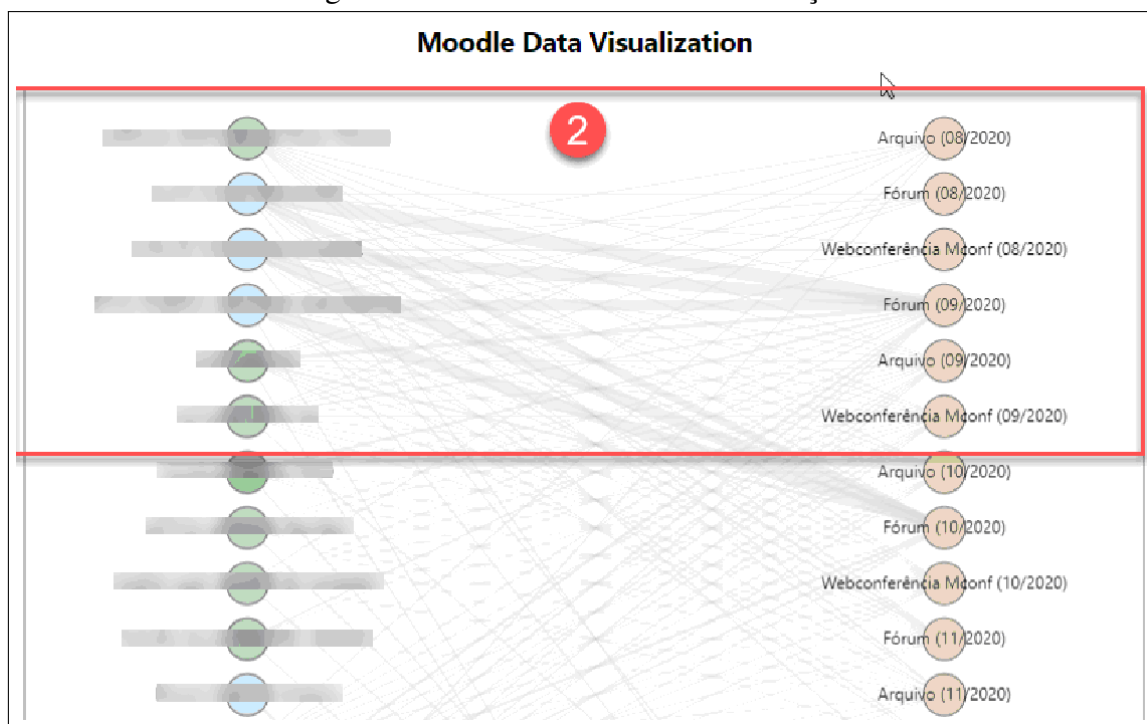
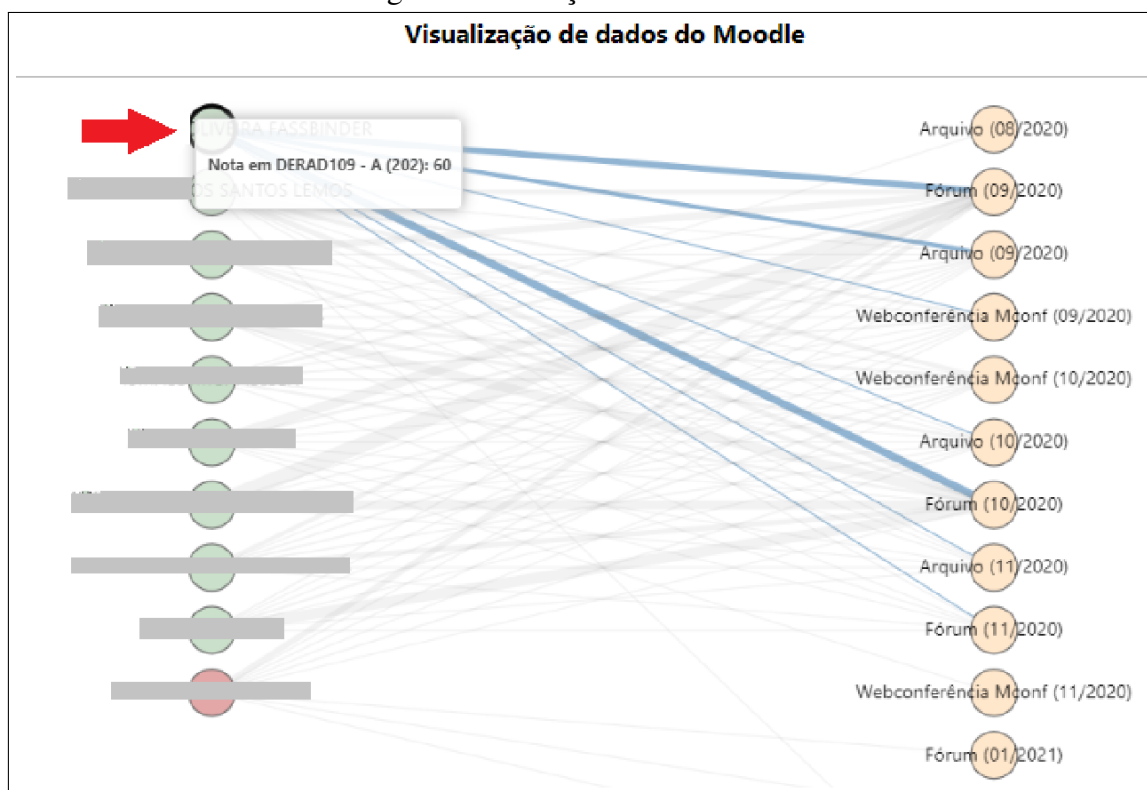
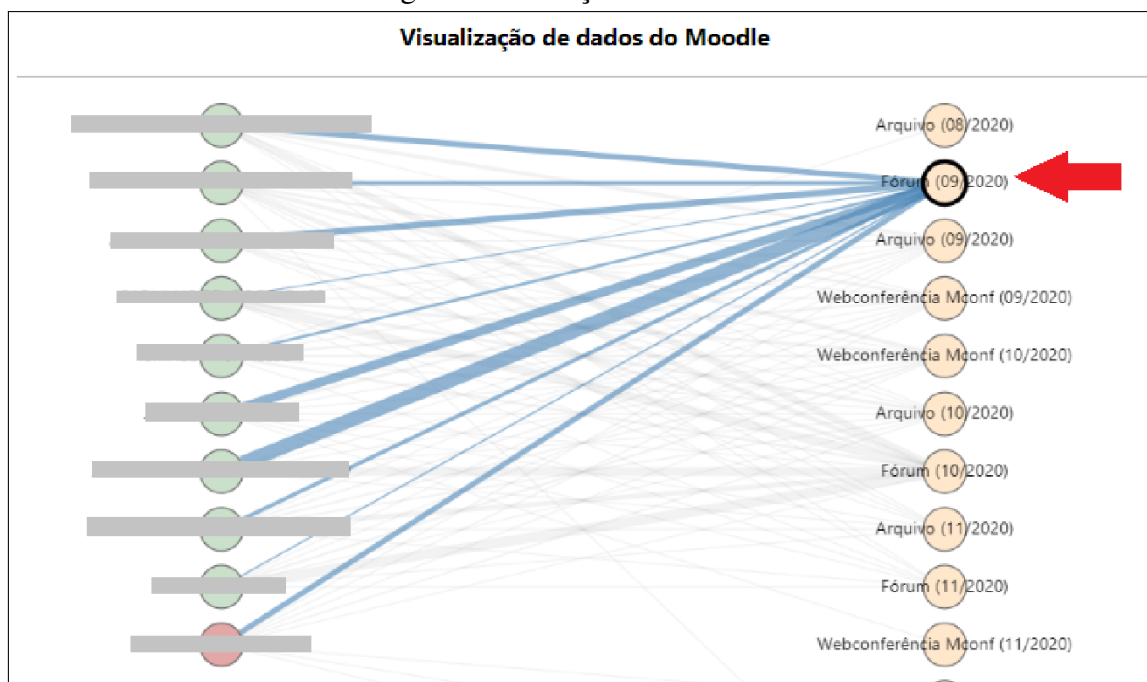


Figura 4.5: Seleção de indivíduos



mais larga é a linha).

Figura 4.6: Seleção de recursos



4.2.3 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 4.7-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 identificar aluno(s) com propensão ao abandono do curso.

Na função modo de visualização (Figura 4.7-b), podemos selecionar entre duas formas: a opção bi-grafo ou a opção árvore.

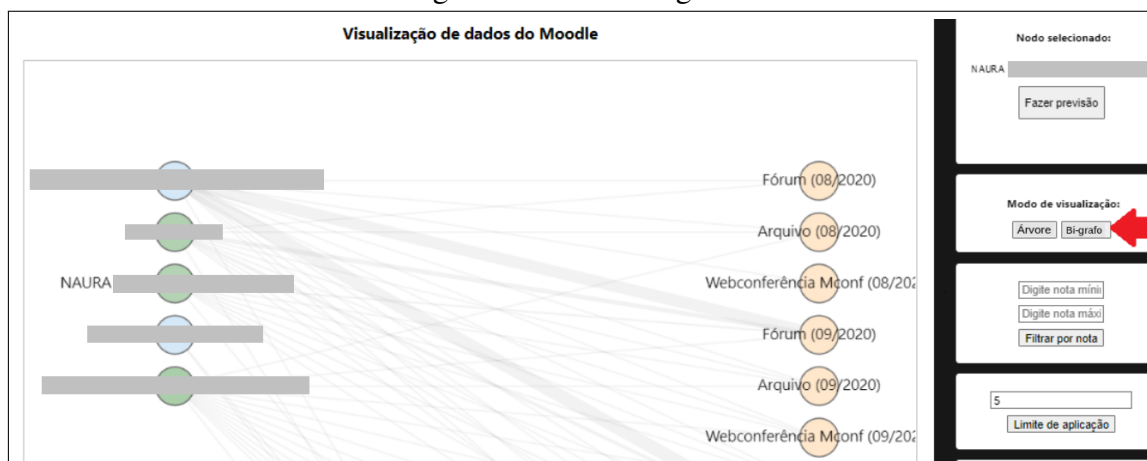
Há ainda uma opção para filtrar por nota mínima e máxima (Figura 4.7-c), limitar a quantidade de nodos na tela (Figura 4.7-d) e remover os nodos sem nota final, como é o caso de alunos que não aprovaram ou que abandonaram, por exemplo (Figura 4.7-e).

Figura 4.7: Detalhe da área de ajustes, ações e parametrização

4.3 Visualização em bi-grafo

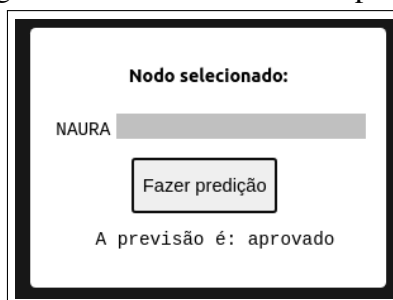
No caso da visualização em modo bi-grafo, esta permite ao usuário visualizar duas colunas contendo nodos (Figura 4.8). Na coluna à esquerda, nodos relativos aos participantes do curso, ou seja, alunos, técnicos, tutores e professores. A coluna à direita apresenta os diferentes recursos da disciplina selecionada. As linhas ou arestas representam as relações entre os nodos das duas colunas. A visualização permite zoom in e zoom out, arrastar e seleccionar nodos ou arrastar a figura inteira.

Figura 4.8: Modo bi-grafo



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). A figura 4.9 mostra um exemplo dessa funcionalidade.

Figura 4.9: Previsão de desempenho.



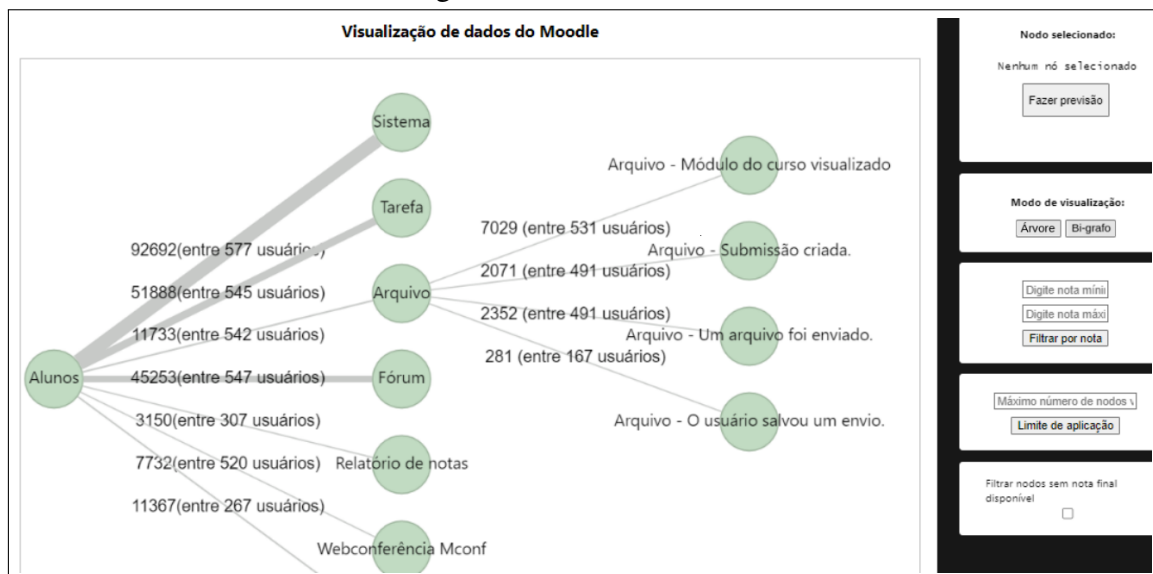
Quando selecionada a função árvore, a tela central atualiza para o segundo modo de visualização.

4.4 Visualização em árvore

Essa visualização evidencia métricas de interação com os componentes do Moodle feita pelo grupo de alunos selecionado. O nodo mais à esquerda representa o grupo de alunos. Suas arestas vão até os componentes do Moodle e são legendadas com o número total de interações com o componente mais o número de alunos que fizeram essas interações, em parêntesis.

Ao clicar com o botão esquerdo do mouse em um nodo de componente do Moodle, este irá se expandir e mostrar as ações feitas naquele componente. 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 4.10: Modo árvore



5 AVALIAÇÃO E VALIDAÇÃO

Como forma de avaliação da MDV, foi aplicado uma ferramenta de consulta em forma de questionário com os professores 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), proposto por Fred Davis em (DAVIS, 1989).

A Teoria da Ação Raciocinada trata-se de uma teoria geral socio-psicológica e comportamental utilizada para explicar um conjunto de comportamentos humanos como o voto, exercícios e o uso de preservativo. 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 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 essas respostas.

O questionário consiste em 22 questões com perguntas variadas sobre aspectos visuais e de utilização da ferramenta. Cada questão pode ser respondida em uma escala com sete opções que variam de *discordo completamente* até *concordo completamente*. Também há questões de feedback aberto para sugestões de melhoria. As perguntas do questionário estão disponíveis no apêndice.

5.1 Resultados

O questionário foi disponibilizado junto com o guia de usuário e teve um total de 6 respondentes. Os respondentes fazem parte do corpo docente do Bacharelado em Desenvolvimento Rural (PLAGEDER/ UFRGS) ou doutorandos do Programa de Pós-Graduação em Informática na Educação (PPGIE/UFRGS). A seguir, é mostrado os principais resultados e comentários fornecidos pelos usuários.

A questão 19 aborda um dos principais objetivos do trabalho, que é fornecer uma ferramenta para visualizar as formas que o Moodle é utilizado pelos usuários. Nesse quesito, as respostas indicaram que a ferramenta cumpriu com o que foi proposto. Por outro lado, a questão 10 apresentou resultados diversos e indicou uma possível dificuldade de utilização da ferramenta por parte dos usuários.

A questão 9 também indicou outro possível ponto a ser melhorado que é a possi-

bilidade da ferramenta ser mais flexível, visto que um terço dos respondentes teve uma resposta neutra à pergunta. O espaço disponível para feedback textual ajudou a definir os possíveis problemas.

Figura 5.1: Resultados do questionário para a questão 19.

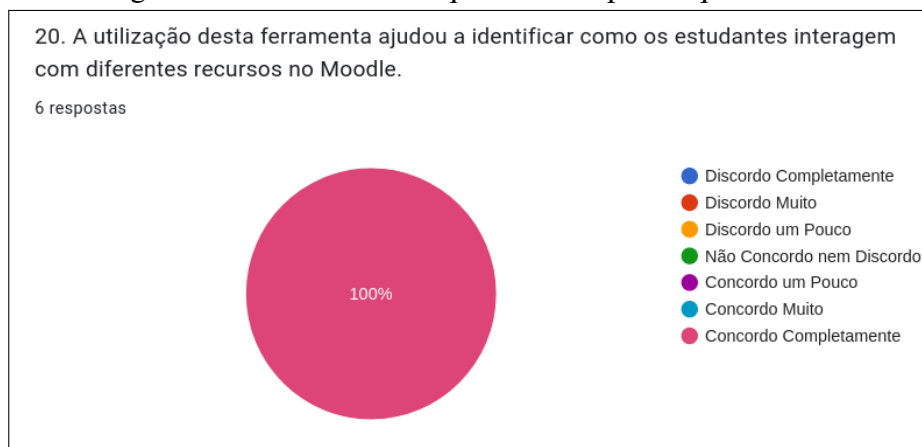


Figura 5.2: Resultados do questionário para a questão 10.

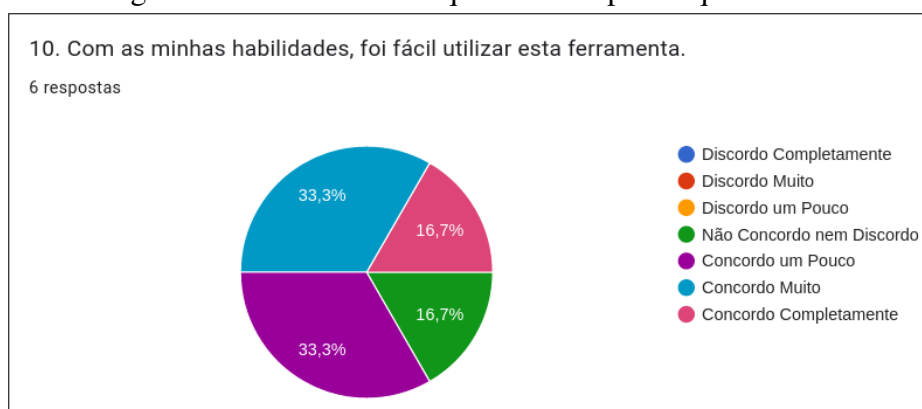
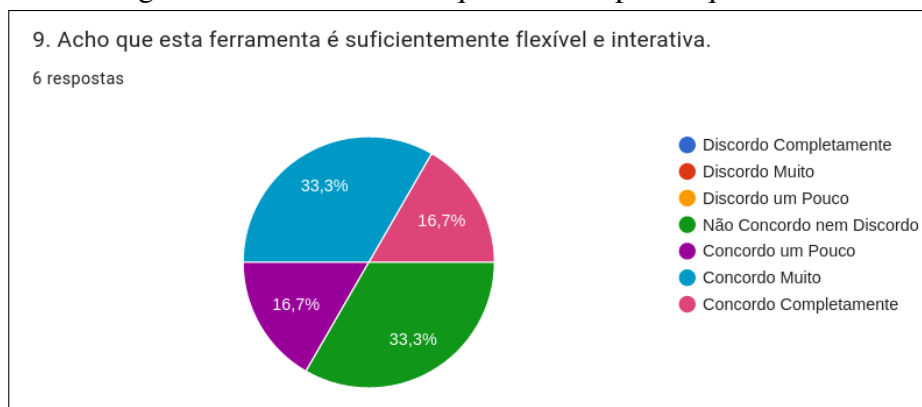


Figura 5.3: Resultados do questionário para a questão 9.



Abaixo são apresentados alguns comentários dos usuários.

"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 paratização, com objetivo de cada função."

Esse comentário sugeriu uma melhoria que seria uma explicação contida na própria ferramenta, e não só no manual de usuário, sobre a funcionalidade dos filtros de visualização. A mudança poderia ser realizada com a utilização de *Tooltips* tal como foi feito com os nodos da visualização. Dois outros usuários também sugeriram melhorias similares.

"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."

Esse comentário expressa uma dificuldade que é não estar familiarizado com visualizações em grafos. Isto é compreensível, tendo em vista que a maior parte das visualizações de dados disponíveis para o público em geral consiste em histogramas ou gráficos de linha. Uma explicação mais detalhada sobre o que significam nodos e arestas na visualização poderia ser uma maneira de melhorar esse aspecto.

"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."

Dois respondentes também comentaram sobre a possibilidade do manual de usuário conter explicações em vídeo, para que o usuário acompanhe um exemplo de utilização da ferramenta. Como tanto a ferramenta MDV quanto o manual de usuário estão disponíveis via internet e não em maneira impressa, a integração de animações junto à explicação é coerente e poderia ser útil para novos usuários.

6 TRABALHOS RELACIONADOS

Nessa seção são comentados alguns trabalhos que apresentam uma visualização de informações de usabilidade da plataforma Moodle. Serão apresentados uma breve descrição de cada trabalho, bem como as características que o diferenciam da ferramenta apresentada aqui.

6.1 Um *plugin* do tipo report para a identificação do risco de evasão na educação superior a distância que usa técnicas de visualização de dados

Em (BEZERRA, 2019), é apresentado um plugin para o Moodle que expõe visualizações a respeito do desempenho acadêmico dos alunos de uma disciplina, bem como a frequência de acesso aos recursos do Moodle pelos estudantes e uma lista indicando os alunos em possível risco de reprovação. A interface do plugin é mostrada na figura 6.1.

A frequência dos acessos ao moodle é evidenciada por um gráfico *período X número de acessos* que mostra a variação nos acessos dos estudantes ao longo do andamento da disciplina. Logo abaixo, uma visualização em *donut* mostra a distribuição de desempenho dos alunos nas atividades, mostrando em vermelho alunos que não entregaram as tarefas.

O trabalho também conta com um gráfico em barra que mostra o número de interações feita no fórum da disciplina pelo grupo de alunos. Nesse gráfico, existe a opção interativa de expandir a visualização para ver quais alunos pertencem ao grupo de altas, médias e baixas interações.

Adicionalmente, o *plugin* fornece a possibilidade de filtrar os dados para conseguir relatórios mais específicos. Existe o filtro de desempenho, que separa o grupo de alunos entre alunos que não entregaram a atividade, abaixo da média, próximo da média ou acima da média. Há também o filtro de selecionar por data inicial e final.

O plugin apresenta uma maneira extremamente prática de visualizar o desempenho acadêmico dos alunos de uma disciplina, visto que a integração dos dados de logs e notas é feita pela aplicação e não requer ações por parte do usuário. É também fácil de identificar o quanto cada aluno interage com o moodle, com o auxílio dos filtros e das listas de usuários pertencentes a cada grupo.

Comparado com o trabalho apresentado aqui, o plugin apresenta como pontos fortes a praticidade e o foco adicional em desempenho não só com a nota final da disciplina

mas também das atividades realizadas ao longo do semestre. Enquanto que o trabalho apresentado aqui foca em uma ferramenta de análise mais profunda dos diversos componentes do Moodle acessados pelos alunos, bem como possibilidades mais variadas de filtragem.

Outra diferença é que a praticidade do plugin introduz uma limitação, que é a necessidade de instalar o plugin no servidor do Moodle para ter acesso à ferramenta. No trabalho apresentado aqui basta possuir os arquivos de logs e notas de uma disciplina.

Figura 6.1: Interface do usuário



Fonte: Brito, Medeiros e Bezerra (2019)

6.2 Uma ferramenta de mineração de dados educacionais para o Moodle

Em (FEIL, 2022) é apresentado uma ferramenta de análise de dados do Moodle utilizando algoritmos de machine learning. A análise é feita pelo download e processamento dos dados de logs e a visualização dos resultados por meio de gráficos.

O trabalho utiliza três algoritmos para fazer a análise: correlação de Spearman, *k-means* e Árvore de decisão. Esses algoritmos são utilizados para encontrar correlações entre as colunas dos dados analisados que sejam interessantes para a análise, por exemplo a possível relação de notas abaixo da média nos questionários ao longo do semestre com

a aprovação ou reprovação do aluno na disciplina.

O trabalho requer uma configuração do Moodle para que possa se conectar à API da plataforma e baixar os dados solicitados. Como essa conexão por API não é habilitada por padrão, é necessário criar o serviço e adicionar o usuário que queira fazer a análise na lista de permissão da API. Já que essas configurações são específicas aos administradores do servidor do Moodle sendo utilizado, não é algo que usuários como professores e tutores consigam fazer sem auxílio externo.

Comparado com esse trabalho, a ferramenta de mineração possui como semelhanças o processamento dos *logs* para extrair informações de análise sobre o desempenho acadêmico e o uso de algoritmos de machine learning para auxiliar nesse processo. No entanto, o trabalho apresentado aqui não foca na análise e nem na automatização da análise dos dados, apenas na criação de ferramentas que possibilitem a análise por parte do usuário.

6.3 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

A proposta de tese de doutorado (KUHN, 2019) busca utilizar representações em grafos de trajetórias de alunos para analisar a aprendizagem no modelo Ensino à Distância (EaD). Técnicas de mineração de dados e visualização são empregadas em cima de dados do Moodle para estudar as lacunas que levam à evasão acadêmica no ensino à distância.

A ferramenta apresentada aqui auxilia nessa análise, pois permite a visualização e manipulação dos grafos para expor informações relevantes a respeito do problema de evasão acadêmica. Embora não seja implementada uma trajetória de aprendizagem, a ferramenta complementa o estudo trazendo representações alternativas de dados do Moodle e de interações com os componentes do sistema.

6.4 Ferramenta para visualização de trajetórias de aprendizagem de alunos em cursos da plataforma Moodle

Em (GUIMARÃES, 2021), é apresentada uma ferramenta de processamento, anonimização e visualização dos dados do Moodle que foca na trajetória de um aluno ao utilizar a plataforma. A ferramenta processa os dados de *logs* e a planilha de notas de

uma disciplina do Moodle e agrega as informações de cada aluno de forma que seja possível representar em um grafo o caminho que um estudante faz ao interagir com a página da disciplina.

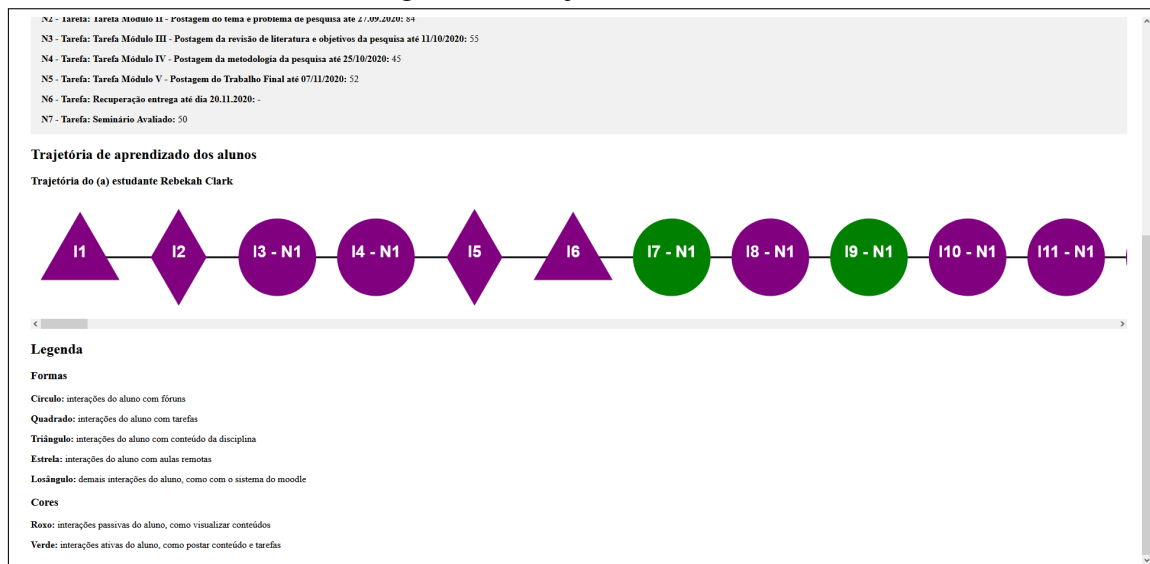
O processamento e a agregação dos dados são muito similares aos apresentados nesse trabalho. Baseado nos arquivos de logs e notas, métricas como número de interações com o Moodle são calculadas e inseridas em dicionários JSON que agregam as informações de um estudante. Após isso, essas informações são exibidas em uma visualização de trajetórias em grafos mostrando o caminho do estudante. Um exemplo desse tipo de visualização está na figura 6.2.

Essa visualização contém nodos com formatos diferentes indicando componentes distintos do Moodle. O grafo em formato de caminho mostra a sequência de componentes acessados ordenados pela data e horário de acesso. As cores diferentes indicam interações passivas, como visualizar páginas, e ativas, como postar comentários. Acima do grafo aparece as notas do estudante ao longo do semestre e também a nota final.

O trabalho apresentado aqui e a ferramenta de (GUIMARÃES, 2021) apresentam muitas semelhanças. O processamento de dados, a visualização em formato de grafos e a organização temporal de informações são ideias que foram adotadas em ambos trabalhos. Uma das limitações citadas nessa ferramenta é a dificuldade de se visualizar componentes que foram acessados mais de uma vez pelo estudante, e também a falta da opção de visualizar múltiplos alunos ao mesmo tempo para fins de comparação. A ferramenta também necessita de linha de comando para a utilização, o que afasta usuários não familiarizados com essa tecnologia.

Este trabalho visa solucionar os problemas citados e expandir as possibilidades de análise permitindo ao usuário interagir com o grafo e filtrar as informações exibidas. As diferentes estruturas de visualização visam atender às dificuldades identificadas pela validação do trabalho. Também foi facilitada a usabilidade por disponibilizar a ferramenta em um servidor web, que retira a necessidade de configurar a ferramenta para a utilização.

Figura 6.2: Trajetória do Aluno



Fonte: Guimarães (2021)

7 CONCLUSÕES

Neste trabalho é apresentada uma ferramenta de visualização de dados com o objetivo de facilitar a análise da utilização dos discentes de recursos da plataforma Moodle, utilizando logs de usuário fornecidos em formato textual. A ferramenta cumpre o objetivo processando esses dados e exibindo-os em uma visualização interativa.

Com a ferramenta disponível por meio de um navegador web, o docente consegue analisar os dados fornecidos e identificar padrões de comportamento que levam estudantes a aprovar ou reprovar nas disciplinas. Também é possível identificar recursos que estão sendo mais ou menos utilizados pelos alunos.

A avaliação apresentou resultados positivos quanto ao objetivo principal do trabalho e também indicou possíveis melhorias a serem feitas para facilitar a utilização da ferramenta. As principais melhorias se dão na integração das explicações contidas no manual de usuário junto com a ferramenta. Isso poderia ser realizado tanto com um link na ferramenta que levaria até um manual sendo exibido em uma nova guia no navegador ou com explicações na própria ferramenta que poderiam ser exibidas e ocultadas por interações com a visualização.

Outra limitação importante é a falta de um mecanismo funcional para que os usuários forneçam os próprios dados de *logs* para que a ferramenta mostre as interações realizadas pelos usuários dos dados fornecidos. Isso seria útil para conseguir informações de utilização de uma turma ou disciplina específica de interesse do usuário.

Como trabalho futuro, a implementação desse mecanismo de fornecimento de dados pelo usuário possibilitaria que a ferramenta atendesse melhor os usuários interessados em suas próprias disciplinas. Junto à isso, poderia ser implementado um mecanismo de sessão no navegador para que o usuário consiga salvar sua sessão atual e retornar posteriormente sem perder os dados e as configurações. Os dados inseridos pelo usuário também poderiam ser utilizados para ampliar o conjunto de treinamento do modelo de predição.

Também se mostra interessante ampliar as visualizações fornecidas para que outras informações fiquem em evidência. Neste trabalho foi apresentada a visualização em árvore, que mostra o quanto cada recurso é acessado pelos usuários, e bi-grafo, que mostra como essa utilização varia ao longo dos meses. Uma nova visualização poderia explorar trajetórias dos usuários, isto é, a sequência de recursos acessada pelo usuário ao realizar as tarefas comuns da disciplina, como responder questionários e enviar exercícios. Também, tendo em vista as dificuldades reportadas por alguns usuários em manipular grafos,

poderiam ser feitos estudos adicionais sobre possíveis representações em estruturas tradicionais, como tabelas ou gráficos de linha.

REFERÊNCIAS

- ANTV. **What is G6**. 2022. <<https://g6.antv.vision/en/docs/manual/introduction>>. [Online; acessado Agosto-2022].
- BAO, W. Covid-19 and online teaching in higher education: A case study of peking university. **Human behavior and emerging technologies**, Wiley Online Library, v. 2, n. 2, p. 113–115, 2020.
- BENNETT, C. et al. The aesthetics of graph visualization. In: **CAe**. [S.l.: s.n.], 2007. p. 57–64.
- BEZERRA, M. T. d. S. e. E. P. Brito e. Um plugin do tipo report para a identificação do risco de evasão na educação superior a distância que usa técnicas de visualização de dados. **Thesis (Masters) - Universidade Federal da Paraíba**, 2019.
- BRITO, M.; MEDEIROS, F.; BEZERRA, E. A report-type plugin to indicate dropout risk in the virtual learning environment moodle. In: **2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)**. [S.l.: s.n.], 2019. v. 2161-377X, p. 127–128.
- CHIRIKOV, I. et al. Online education platforms scale college stem instruction with equivalent learning outcomes at lower cost. **Science advances**, American Association for the Advancement of Science, v. 6, n. 15, p. eaay5324, 2020.
- DAVIS, F. D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. **MIS quarterly**, JSTOR, p. 319–340, 1989.
- FEIL, B. C. Uma ferramenta de mineração de dados educacionais para o moodle. **Thesis - Universidade Federal do Rio Grande do Sul**, 2022.
- GUIMARÃES, J. F. Ferramenta para a visualização de trajetórias de aprendizagem de alunos em cursos da plataforma moodle. **Thesis - Universidade Federal do Rio Grande do Sul**, 2021.
- HERMAN, I.; MELANCON, G.; MARSHALL, M. Graph visualization and navigation in information visualization: A survey. **IEEE Transactions on Visualization and Computer Graphics**, v. 6, n. 1, p. 24–43, 2000.
- HINTON, G. E. et al. **Improving neural networks by preventing co-adaptation of feature detectors**. arXiv, 2012. Available from Internet: <<https://arxiv.org/abs/1207.0580>>.
- HQ, M. **About Moodle**. 2022. <https://docs.moodle.org/400/en/About_Moodle>. [Online; acessado Agosto-2022].
- HQ, M. **Philosophy**. 2022. <<https://docs.moodle.org/400/en/Philosophy>>. [Online; acessado Agosto-2022].
- HU, W. et al. **Strategies for Pre-training Graph Neural Networks**. arXiv, 2019. Available from Internet: <<https://arxiv.org/abs/1905.12265>>.

INC, D. **Use containers to Build, Share and Run your applications**. 2022. <<https://www.docker.com/resources/what-container/>>. [Online; acessado Agosto-2022].

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

KUHN, I. 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. **Proposta de Tese de Doutorado - Universidade Federal do Rio Grande do Sul**, 2019.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, p. 436–44, 05 2015.

MILLIGAN, C. Delivering staff and professional development using virtual learning environments. **JTAP**, 1999.

NETWORKX. **Software for complex networks**. 2022. <<https://networkx.org/>>. [Online; acessado Agosto-2022].

OLATUNDE-AIYEDUN, T. G.; OGUNODE, N. J.; AIYEDUN, C. E. Effectiveness of virtual learning during covid-19 lockdown in nigerian universities. **Olatunde-Aiyedun, TG, Ogunode, NJ & Eyiolorunse-Aiyedun, CT (2021). Assessment of virtual learning during covid-19 lockdown in Nigerian public universities. Academicia Globe: Inderscience Research**, v. 2, n. 5, p. 159–175, 2021.

PEREIRA, A. T. C.; SCHMITT, V.; DIAS, M. Ambientes virtuais de aprendizagem. **AVA-Ambientes Virtuais de Aprendizagem em Diferentes Contextos. Rio de Janeiro: Editora Ciência Moderna Ltda**, p. 4–22, 2007.

PYTORCH. **End-to-end Machine Learning Framework**. 2022. <<https://pytorch.org/features/>>. [Online; acessado Agosto-2022].

SNAP. **DeepSNAP Documentation**. 2022. <<https://snap.stanford.edu/deepsnap/>>. [Online; acessado Agosto-2022].

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.

WHATWG. **The WHATWG**. 2022. <<https://whatwg.org/faq>>. [Online; acessado Agosto-2022].

WSGI. **What is WSGI?** 2022. <<https://wsgi.readthedocs.io/en/latest/what.html>>. [Online; acessado Agosto-2022].

APÊNDICE A — FORMULÁRIO TAM

1. A utilização desta ferramenta no meu trabalho permite-me realizar as tarefas mais rapidamente do que se as executasse em outros produtos semelhantes.
 - Escala com sete opções de *discordo completamente* até *concordo completamente*.
2. A utilização desta ferramenta no meu trabalho aumenta o meu benefício, interesses ou vantagens.
 - Escala com sete opções de *discordo completamente* até *concordo completamente*.
3. Utilizar esta ferramenta torna o meu trabalho mais fácil.
 - Escala com sete opções de *discordo completamente* até *concordo completamente*.
4. Considero esta ferramenta útil no meu trabalho.
 - Escala com sete opções de *discordo completamente* até *concordo completamente*.
5. Foi fácil para mim operar esta ferramenta.
 - Escala com sete opções de *discordo completamente* até *concordo completamente*.
6. Acho fácil que esta ferramenta faça o que eu preciso.
 - Escala com sete opções de *discordo completamente* até *concordo completamente*.
7. A minha interação com esta ferramenta foi clara e compreensível.
 - Escala com sete opções de *discordo completamente* até *concordo completamente*.
8. Acho que esta ferramenta é suficientemente flexível e interativa.
 - Escala com sete opções de *discordo completamente* até *concordo completamente*.
9. Com as minhas habilidades, foi fácil utilizar esta ferramenta.

- Escala com sete opções de *discordo completamente* até *concordo completamente*.
10. Considero que a ferramenta possui elementos visuais agradáveis e intuitivos.
- Escala com sete opções de *discordo completamente* até *concordo completamente*.
11. Considero adequada a forma de seleção de turmas e alunos disponível na área de seleção.
- Escala com sete opções de *discordo completamente* até *concordo completamente*.
12. O modo como podemos visualizar recursos, alunos e suas relações e iterações na área de visualização e análise é claro e compreensível.
- Escala com sete opções de *discordo completamente* até *concordo completamente*.
13. Em relação a área de ajustes, ações e parametrização, a função predição apresenta utilidade no meu trabalho como docente.
- Escala com sete opções de *discordo completamente* até *concordo completamente*.
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.
- Escala com sete opções de *discordo completamente* até *concordo completamente*.
15. O modo como selecionamos as notas na função filtrar por nota se mostra adequado.
- Escala com sete opções de *discordo completamente* até *concordo completamente*.
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.
- Escala com sete opções de *discordo completamente* até *concordo completamente*.

17. Considero a função filtrar nodos sem nota final disponível fundamental durante o uso da ferramenta.
 - Escala com sete opções de *discordo completamente* até *concordo completamente*.
18. A utilização desta ferramenta me ajuda a compreender como os estudantes utilizam o Moodle.
 - Escala com sete opções de *discordo completamente* até *concordo completamente*.
19. A utilização desta ferramenta ajudou a identificar como os estudantes interagem com diferentes recursos no Moodle.
 - Escala com sete opções de *discordo completamente* até *concordo completamente*.
20. Acredito que com o uso desta ferramenta é possível identificar alunos com propensão a evasão.
 - Escala com sete opções de *discordo completamente* até *concordo completamente*.
21. Em relação à ferramenta, gostaria de sugerir as seguintes melhorias (opcional):
 - Campo de texto aberto.
22. Em relação ao manual do usuário, gostaria de sugerir as seguintes melhorias (opcional):
 - Campo de texto aberto.