

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

HENRIQUE ALVES BILO

**Ferramenta para análise de sentimentos em  
postagens no Moodle**

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

Orientador: Prof. Dr. Leandro Krug Wives  
Coorientadora: Dr<sup>a</sup>. Brenda Salenave Santana

Porto Alegre  
2023

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos André Bulhões Mendes

Vice-Reitora: Prof<sup>a</sup>. Patricia Pranke

Pró-Reitora de Graduação: Prof. Cíntia Boll

Diretora do Instituto de Informática: Prof<sup>a</sup>. Carla Maria Dal Sasso Freitas

Coordenador do Curso de Ciência de Computação: Prof. Marcelo Walter

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

## **AGRADECIMENTOS**

Gostaria de agradecer inicialmente ao meu orientador Leandro Wives e minha coorientadora Brenda Salenave por todo auxílio, apoio e paciência durante a construção desse trabalho. Gostaria também de agradecer minha namorada Fernanda Azevedo por me incentivar a fazer o meu melhor. Por fim, gostaria de agradecer meus familiares Ricardo Bilo, Viviane Alves e Theresinha Alves pelo grande incentivo, força e apoio.

## RESUMO

Este trabalho tem como objetivo melhorar a relação entre professores e alunos por meio da compreensão dos sentimentos dos alunos por parte do professor. Entender os sentimentos dos alunos é crucial para que os professores possam adaptar sua metodologia de ensino, identificar problemas e tomar medidas para melhorar a aprendizagem dos alunos. Diante disso, este trabalho propõe o desenvolvimento de uma ferramenta para coletar e analisar dados das postagens dos alunos no ambiente virtual de aprendizagem Moodle. Para tanto, são aplicadas técnicas de processamento de linguagem natural para classificar os diferentes tipos de sentimentos expressos pelos alunos em suas postagens, fóruns, mensagens diretas ao professor e mensagens via bate-papo. Os resultados mostraram que a ferramenta de monitoramento de sentimentos dos alunos foi eficaz ao possibilitar aos professores compreenderem as emoções dos estudantes. Através da instalação do software nas máquinas dos professores, acompanhada de um arquivo PDF detalhado com instruções de uso, a ferramenta permitiu identificar os sentimentos dos alunos. Os resultados foram obtidos a partir da participação de três professores no experimento. Ao compreender melhor os sentimentos dos alunos, os professores podem adaptar sua metodologia de ensino e fornecer uma experiência educacional mais personalizada e satisfatória, potencialmente levando a um melhor desempenho dos alunos e uma maior satisfação com o processo de aprendizagem.

**Palavras-chave:** Análise de sentimentos. Moodle. Análise de dados.

## **Tool for sentiment analysis in student posts on Moodle**

### **ABSTRACT**

This work aims to improve the relationship between teachers and students through the understanding of students' emotions by the teacher. Understanding students' emotions is crucial for teachers to adapt their teaching methodology, identify issues, and take measures to improve student learning. In this regard, this work proposes the development of a tool to collect and analyze data from students' posts in the Moodle virtual learning environment. To achieve this, natural language processing techniques are applied to classify the different types of emotions expressed by students in their posts, forums, direct messages to the teacher, and chat messages. The results showed that the student sentiment monitoring tool was effective in enabling teachers to understand students' emotions. By installing the software on teachers' machines, accompanied by a detailed PDF file with instructions for use, the tool allowed for the identification of students' emotions. The results were obtained from the participation of three teachers in the experiment. By better understanding students' emotions, teachers can adapt their teaching methodology and provide a more personalized and satisfactory educational experience, potentially leading to improved student performance and higher satisfaction with the learning process.

**Keywords:** Sentiment analysis, Moodle, Data analysis.

## LISTA DE FIGURAS

Figura 2.1	Extração de emoções a partir do texto.....	22
Figura 3.1	Pipeline do projeto.....	26
Figura 3.2	Fluxo de dados da análise.....	29
Figura 3.3	Gráfico de Polaridade.....	32
Figura 3.4	Gráfico de Métricas Gerais.....	33
Figura 3.5	Gráfico de Relatório Geral dos Alunos.....	34
Figura 4.1	Perfil dos usuários.....	36
Figura 4.2	Respostas da facilidade de instalação e objetivo do <i>software</i> .....	37
Figura 4.3	Respostas da facilidade de uso e velocidade de processamento.....	38
Figura 4.4	Respostas da utilidade das métricas extraídas pela ferramenta proposta.....	39
Figura 4.5	Respostas da probabilidade de se utilizar a ferramenta.....	39
Figura 5.1	Tela do usuário sistema GWPulse.....	42
Figura 5.2	fluxo do sistema GWPulse.....	42
Figura 5.3	Resultado comparação SentiStrength com outros algoritmos.....	47
Figura A.1	Alerta ao tentar executar o instalador.....	54
Figura A.2	Tela apresentada ao usuário para realizar login.....	55
Figura A.3	Tela de seleção de disciplina.....	55
Figura A.4	Tela de carregamento.....	55

## LISTA DE TABELAS

Tabela 2.1 Tabela de funções utilizadas API Moodle .....	17
Tabela 2.2 Bibliotecas do Python utilizadas .....	17
Tabela 5.1 Funcionalidades entre os trabalhos estudados e o do autor. ....	48

## LISTA DE ABREVIATURAS E SIGLAS

API	<i>Application Programming Interface</i>
AVA	Ambiente Virtual de Aprendizagem
CNN	<i>Convolutional Neural Networks</i>
CSV	<i>Comma-Separated Values</i>
EaD	Ensino a Distância
HTML	<i>HyperText Markup Language</i>
LeIA	Léxico para Inferência Adaptada
NRC	National Research Council Canada
PLN	Processamento de Linguagem Natural
SUS	<i>System Usability Scale</i>
VADER	<i>Valence Aware Dictionary and sEntiment Reasoner</i>

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>10</b>
<b>2 CONCEITOS E TECNOLOGIAS</b> .....	<b>13</b>
<b>2.1 Conceitos</b> .....	<b>13</b>
2.1.1 Processamento de Linguagem Natural.....	13
2.1.1.1 Pré-processamento de dados textuais.....	13
2.1.2 Mineração de texto.....	14
2.1.3 Análise de sentimentos .....	14
2.1.3.1 Polaridade .....	15
2.1.3.2 Sentimentos.....	15
2.1.3.3 Emoções.....	15
<b>2.2 Tecnologias</b> .....	<b>16</b>
2.2.1 Moodle .....	16
2.2.2 Linguagem de programação Python .....	16
2.2.3 LeIA .....	18
2.2.4 NRCLex .....	19
2.2.5 EmoRoBERTa.....	20
2.2.6 Google Perspective .....	21
2.2.7 Yake.....	23
2.2.8 Dash .....	24
<b>3 DESENVOLVIMENTO</b> .....	<b>25</b>
<b>3.1 Login do Usuário e seleção de disciplinas</b> .....	<b>25</b>
<b>3.2 Coleta de dados</b> .....	<b>26</b>
<b>3.3 Pré-processamento</b> .....	<b>28</b>
<b>3.4 Análise de sentimentos</b> .....	<b>28</b>
<b>3.5 Visualização de dados</b> .....	<b>31</b>
3.5.1 Criação e exibição dos gráficos.....	32
<b>4 AVALIAÇÃO E VALIDAÇÃO</b> .....	<b>35</b>
<b>4.1 Questionário</b> .....	<b>35</b>
<b>4.2 Resultados</b> .....	<b>36</b>
<b>5 TRABALHOS RELACIONADOS</b> .....	<b>41</b>
<b>5.1 GWpulse: Supporting Learner Modeling and Group Awareness in Online Forum with Sentiment Analysis</b> .....	<b>41</b>
<b>5.2 Mineração de Dados Educacionais e Análise de Sentimentos em Ambien- tes Virtuais de Aprendizagem: um Mapeamento Sistemático</b> .....	<b>43</b>
<b>5.3 An ontology based text approach for culture aware emotion mining: A Moodle plugin</b> .....	<b>44</b>
<b>5.4 Sentiment Strength Detection in Short Informal Text</b> .....	<b>46</b>
<b>5.5 Análise comparativa</b> .....	<b>48</b>
<b>6 CONCLUSÃO</b> .....	<b>49</b>
<b>6.1 Limitações</b> .....	<b>49</b>
<b>6.2 Trabalhos futuros</b> .....	<b>50</b>
<b>REFERÊNCIAS</b> .....	<b>51</b>
<b>APÊNDICE A — GUIA DE USO</b> .....	<b>54</b>
<b>APÊNDICE B — QUESTIONÁRIO</b> .....	<b>57</b>
<b>B.1 Pesquisa de perfil</b> .....	<b>57</b>
<b>B.2 Feedback</b> .....	<b>57</b>
<b>B.3 Avaliação utilizando o padrão SUS</b> .....	<b>59</b>

## 1 INTRODUÇÃO

Compreender como os alunos estão se sentindo é uma tarefa cada vez mais desafiadora para professores e educadores, especialmente em cursos EaD ou em aulas híbridas, exigindo que eles mantenham uma atenção constante para entender as necessidades e sentimentos dos alunos. É importante que os professores se empenhem em compreender as emoções de seus alunos, pois isso pode ajudar a melhorar o processo de ensino e aprendizagem, bem como o relacionamento entre alunos e professores (LOPES, 2011). Nas modalidades de ensino a distância, o contato do professor com o aluno é limitado e muitas vezes se dá apenas por meio de tecnologias (como, por exemplo, e-mails, conversas via chat da plataforma de ensino utilizada e fóruns). Isso torna ainda mais difícil perceber as emoções do aluno e identificar possíveis problemas que estejam dificultando o aprendizado. À medida que a tecnologia se desenvolve e a sociedade muda, os alunos enfrentam novos desafios e pressões, e estão cada vez mais em risco devido a fatores diversos, como problemas na vida pessoal ou até financeiros. Todos esses fatores podem afetar significativamente seus sentimentos e emoções. Segundo Muir-Herzig (2004):

...O termo "em risco" alude ao fato de que um aluno tem um problema, mas as escolas não têm certeza exata da origem do problema. Os fatores de predição de risco são fatores socioeconômicos, fatores de saúde, familiares e escolares. As escolas geralmente usam os seguintes critérios para identificar alunos em risco: notas baixas, baixo GPA e/ou alto índice de absenteísmo de um aluno. O problema ao usar identificadores é que nenhum aluno em risco é igual. Como a escola se prepara para ensinar esses alunos é a chave para o sucesso do aluno em risco na escola.

Muitas vezes os professores não conseguem visualizar de maneira clara o que cada aluno está sentindo e isso pode acabar omitindo informações, como por exemplo, se o aluno está pensando em desistir, se ele está triste ou feliz. Dar uma boa aula está muito longe de ser apenas transmitir uma ideia com clareza; se o indivíduo estiver com algum problema, possivelmente não irá conseguir absorver o conteúdo necessário. A pandemia enfatizou ainda mais esse problema de comunicação entre alunos e professores, pois ambos tiveram que se adaptar rapidamente a uma nova realidade de ensino remoto, o que tornou a identificação e a resolução dos problemas ainda mais desafiadoras. De acordo com Palmeira e Sousa (2021):

Os professores sofreram uma imensa pressão nesse período, já que, de maneira instantânea, tiveram que se adaptar às ferramentas virtuais, preparar as atividades para que os alunos se mantivessem interessados e, concomitantemente, precisaram estar disponíveis para pais e alunos, esclarecendo eventuais

dúvidas acerca do novo sistema que estava sendo implementado, o sistema de educação a distância.

Saber como os alunos estão se sentindo é importante em diversas situações, pois pode ajudar a desenvolver estratégias de ensino mais eficazes, bem como melhorar a motivação e o desempenho escolar (OSTI; BRENELLI, 2013). Além disso, a detecção precoce de problemas emocionais pode ajudar a prevenir a evasão escolar e promover a saúde mental dos alunos (BÓBÓ et al., 2019). No entanto, a falta de tempo é um grande desafio para os professores, que muitas vezes não conseguem dedicar a atenção necessária a cada aluno. Essa lacuna é particularmente preocupante, pois pode levar à omissão de problemas e sentimentos importantes dos alunos. Diante dessa realidade, identificou-se que seria de grande utilidade uma ferramenta que auxiliasse os professores a terem uma visão geral da turma, podendo facilitar a compreensão de forma prática e eficiente dos sentimentos dos alunos. Para que isso seja possível, é necessária uma ferramenta auxiliar, como o Moodle, que é um sistema utilizado por diversas instituições para criar e administrar cursos, fornecendo ferramentas para os alunos para que possam interagir, colaborar entre si e também comunicar-se com os professores. Assim como mencionado por Legoinha, Pais e Fernandes (2006):

O conceito fundamental do Moodle consiste numa página, onde professores disponibilizam recursos e desenvolvem atividades com e para os alunos. Uma eventual metáfora para a página Moodle poderia ser a sala de aula ubíqua. A cada utilizador registado está associado um perfil e uma fotografia podendo comunicar com qualquer outro, reforçando a componente social desta plataforma.

Com o objetivo de auxiliar na observação e compreensão dos sentimentos dos alunos a partir de interações textuais na plataforma Moodle, este trabalho propõe o desenvolvimento de uma ferramenta de análise de sentimentos. A finalidade dessa ferramenta é fornecer suporte aos professores responsáveis por uma turma na plataforma, permitindo-lhes compreender os sentimentos expressos pelos alunos em suas interações nesta. Utilizando a ferramenta, o professor terá acesso as postagens em fóruns, chats e mensagens privadas enviadas para o mesmo. Vale destacar que a ferramenta pode ser utilizada por outras pessoas além dos professores, como alunos, por exemplo. No entanto, no caso de um aluno, ele terá acesso somente às mensagens postadas em fóruns e chats, e não terá acesso às mensagens enviadas diretamente ao professor. Para os experimentos realizados neste trabalho, foi imposta uma limitação de acesso às mensagens coletadas pela ferramenta apenas a professores, que utilizaram a ferramenta para monitorar apenas as

mensagens ocorridas nas disciplinas ministradas por eles. Para atingir esse objetivo, foi criada uma ferramenta que faz uso da API (acrônimo de *Application Programming Interface*) disponibilizada pelo Moodle às universidades, cuja função neste trabalho é coletar as mensagens trocadas em disciplinas de interesse no Moodle. A API fornecida pelo Moodle às universidades permite que desenvolvedores acessem e manipulem dados armazenados no sistema.

A partir da API do Moodle, a coleta de mensagens é realizada e seguimos para a próxima etapa; a análise de sentimentos. Para realizar essa análise, foram utilizadas 4 API's diferentes:

- LeIA (Léxico para Inferência Adaptada);
- NRCLex (National Research Council Canada);
- EmoRoBERTa (Emotions + BERT's language);
- Google Perspective.

Por fim, após estimar os sentimentos, são exibidos três gráficos interativos de forma que o usuário (professor, no caso) possa selecionar o aluno, o período e até mesmo um sentimento que deseje obter detalhes. Com isso, é possível ter uma visão complementar da turma por completo através de análises visuais.

O restante do trabalho encontra-se estruturado da seguinte forma. No Capítulo 2 são descritos os conceitos e tecnologias utilizados no desenvolvimento deste projeto. Além disso, são descritos os sentimentos utilizados para a etapa de análise de sentimentos do trabalho. No Capítulo 3 é descrito o desenvolvimento geral da ferramenta proposta neste trabalho, incluindo detalhes de como os dados foram coletados, analisados e ainda os resultados visuais das análises de sentimentos extraídos. O Capítulo 4 descreve experimentos realizados para validação da ferramenta proposta através de *feedbacks* de professores. Além disso, apresenta uma análise e interpretação dos resultados da análise de sentimentos observados pelos professores. No Capítulo 5 são apresentados e discutidos trabalhos relacionados ao tópico endereçado neste trabalho. Por fim, o Capítulo 6 apresenta a conclusão 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 esses assuntos.

### 2.1 Conceitos

Nesta Seção são descritos os conceitos relacionados com o trabalho e mencionados no restante do documento.

#### 2.1.1 Processamento de Linguagem Natural

De acordo com a definição de Vieira e Lopes (2010), processamento de Linguagem Natural (PLN) é uma área de Ciência da Computação que estuda o desenvolvimento de programas de computador que analisam, reconhecem e/ou geram textos em linguagens humanas, ou linguagens naturais. Em outras palavras, PLN é a capacidade de um computador entender o significado e a intenção das palavras usadas em um texto. Isso inclui tarefas como tradução automática (SERRANO, 2022), análise de sentimentos (PENCZ-KOSKI; PENTEADO, 2019) e geração de texto (EMMERT-STREIB et al., 2020). Neste trabalho, focamos na análise de sentimentos, com o intuito de analisar os textos coletados do Moodle (explicado em maiores detalhes na Seção 2.2.1).

##### 2.1.1.1 Pré-processamento de dados textuais

O pré-processamento de texto é um processo de transformação do texto bruto em uma forma mais adequada para a atividade desejada. O objetivo é preparar o texto de forma que fique mais fácil para as máquinas analisarem. Assim como diz Andrade (2019):

É de extrema importância utilizar técnicas de processamento de linguagem natural para pré-processamento dos textos. O texto bruto se apresenta de forma esparsa e desestruturada, que não permite a execução de classificadores de maneira eficiente.

Uma das técnicas de pré-processamento é a quebra do texto em pequenas partes, chamadas de *tokens*, que resultam da segmentação do texto em elementos individuais,

como palavras, números, pontuações e outros caracteres. Essa quebra em *tokens* permite a identificação de padrões gramaticais e de relacionamento entre as palavras, que são utilizados para entender melhor o texto e extrair informações relevantes. O pré-processamento envolve uma série de atividades que buscam garantir que o texto esteja em um formato adequado para a análise, como a eliminação de palavras que não acrescentam sentido e a correção de erros ortográficos (SANTANA et al., 2023). Este conjunto de palavras eliminadas é denominado de *stopwords*, que são palavras comuns na língua, podendo ser preposições e conjunções, as quais são frequentemente usadas em textos, porém não agregam significado à análise. Ao final da etapa de pré-processamento, o texto está preparado para ser submetido a diversas abordagens de PLN, como a análise de sentimentos.

### **2.1.2 Mineração de texto**

A mineração de texto, também conhecida como mineração de dados textuais, é uma técnica de análise de informações que consiste em extrair conhecimento significativo a partir de textos, como e-mails, documentos, publicações em redes sociais, comentários em blogs, entre outros (MORAIS; AMBRÓSIO, 2007). Para realizar essa análise, a mineração de texto utiliza algoritmos de PLN que possibilitam interpretar e identificar padrões, tendências, sentimentos, opiniões e outras informações relevantes presentes nos textos. Dessa forma, as informações obtidas podem ser usadas para diversas finalidades, incluindo a análise de sentimentos, que é o objetivo central deste trabalho.

### **2.1.3 Análise de sentimentos**

A análise de sentimentos é uma área de estudo que se concentra em identificar e extrair informações subjetivas dos dados textuais, como sentimentos, opiniões e emoções (MEJOVA, 2009). Esses elementos subjetivos são comumente expressos por meio de palavras, frases ou sentenças isoladas em um texto. Por essa razão, a análise de sentimentos geralmente se concentra em unidades linguísticas menores, como termos específicos, expressões ou construções gramaticais, que transmitem um sentimento positivo, negativo ou neutro em relação a um determinado tópico ou objeto. O objetivo deste trabalho é realizar a análise de sentimentos. Nesta Seção, serão apresentadas os significados dos termos de polaridade, sentimentos e emoções. Como sentimentos e emoções muitas vezes

se referem à mesma ideia, neste trabalho usaremos esses termos de forma intercambiável.

### *2.1.3.1 Polaridade*

A polaridade é um aspecto importante na análise de sentimentos, pois representa o grau de positividade e negatividade de um texto (BENEVENUTO; RIBEIRO; ARAÚJO, 2015). No entanto, a polaridade nem sempre é claramente positiva ou negativa. Em alguns casos, um texto pode ter uma polaridade fraca, ou seja, não ser nem claramente positiva, nem claramente negativa. Assim, no contexto de análises textuais, a polaridade é geralmente utilizada para classificar textos em positivos, negativos ou também neutros, permitindo que se observe se uma frase ou um texto tem um sentido bom ou ruim. Como exemplos é possível citar as seguintes frases: “Como você está bonita hoje” é positivo, enquanto “Hoje é um péssimo dia” é negativo, e a frase “Hoje é 21 de Outubro” é considerada neutra.

### *2.1.3.2 Sentimentos*

Sentimentos e emoções são termos frequentemente usados de forma intercambiável, mas existem diferenças importantes entre eles. Segundo Hansen (2005), os sentimentos são experiências emocionais subjetivas que uma pessoa pode sentir, e que podem ser expressas através de palavras para outras pessoas que também já passaram por situações semelhantes e, portanto, podem compreender melhor a complexidade dessas emoções. Os sentimentos são influenciados pelas experiências pessoais e pela interpretação individual de uma situação. Os sentimentos geralmente surgem após a emoção e podem ter uma duração mais longa do que a resposta emocional inicial. Eles são, portanto, um importante aspecto da regulação emocional e ajudam a fornecer uma estrutura para a compreensão e comunicação das experiências emocionais. Damasio e Dolan (1999) destacam que, através da identificação e expressão dos sentimentos, os indivíduos podem melhor compreender e gerenciar suas emoções e, assim, tomar decisões mais adaptativas e saudáveis.

### *2.1.3.3 Emoções*

De acordo com Hansen (2005) as emoções são vistas como mecanismos instintivos, rápidos e subconscientes que controlam as respostas individuais a diversas situações. Essas respostas são geradas automaticamente pelo sistema nervoso autônomo e podem incluir mudanças fisiológicas, como expressões faciais, aumento da frequência cardíaca e

respiração acelerada. Embora as emoções sejam consideradas primitivas, elas são fundamentais para a adaptação do indivíduo ao ambiente e podem ser influenciadas por processos cognitivos conscientes e experiências anteriores. Neste trabalho as emoções analisadas são referenciadas nas Seções 2.2.4, 2.2.5 e 2.2.6.

## **2.2 Tecnologias**

Aqui são descritas as tecnologias utilizadas no desenvolvimento do trabalho.

### **2.2.1 Moodle**

O Moodle é uma plataforma de ensino online e um sistema de gerenciamento de aprendizagem para criar e gerenciar cursos eletrônicos (VASCONCELOS; JESUS; SANTOS, 2020). Professores utilizam esta plataforma para gerenciar as disciplinas que ministram, permitindo-lhes compartilhar conteúdos em diferentes formatos, esclarecer dúvidas e avaliar os alunos inscritos nas disciplinas de maneira eficiente e prática. A plataforma Moodle também oferece outros recursos, como fóruns, bate-papo e mensagens privadas para ajudar a construir um aprendizado mais eficiente.

A API fornecida pelo Moodle é uma ferramenta que possibilita ler e escrever dados neste ambiente virtual. Para ser disponibilizada no servidor onde o Moodle foi instalado, o uso de ferramentas externas, como é o caso da API, deve ser habilitado <sup>1</sup>. As funções utilizadas para este trabalho estão na Tabela 2.1.

### **2.2.2 Linguagem de programação Python**

A linguagem de programação Python foi criada por Guido van Rossum em 1991 (ROSSUM; JR, 1995) e chamou atenção por ser uma linguagem de fácil interpretabilidade e de uso. Esta linguagem é uma das principais escolhas para atuar na área de dados devido ao seu grande número de bibliotecas, sendo possível realizar a instalação de uma maneira rápida e eficiente, sem complicações.

Para o trabalho em questão, a linguagem Python foi essencial para obter um bom desempenho, visto que são utilizadas diversas bibliotecas para a manipulação dos dados

---

<sup>1</sup><<https://success.vitalsource.com/hc/en-us/articles/360055861733-Moodle-API-Integration-Roster-Exchange-Setup>>

Tabela 2.1: Tabela de funções utilizadas API Moodle

Função	Entrada	Saída
core_user_get_users_by_field	id_usuario	Informações sobre o usuário
core_enrol_get_users_courses	id_usuario	Cursos que o usuário participa
mod_chat_get_chats_by_courses	id_curso	Chats contidos no curso
mod_chat_get_sessions	id_chat	Sessões contidas no chat
mod_chat_get_session_messages	id_chat, datas_início, data_fim	Mensagens das sessões
core_enrol_get_enrolled_users	id_curso	Usuários de um curso
core_message_get_messages	id_usuario_destino, id_usuario_origem, tipo_mensagem, mensagens_lidas	Mensagens de um usuário para outro
mod_forum_get_forums_by_courses	ids_cursos	Fóruns do curso
mod_forum_get_forum_discussions	id_fórum	Discussões do fórum
mod_forum_get_discussion_posts	id_discussão	Postagens em determinado fórum

Fonte: Tabela elaborada pelo autor a partir da documentação da API do Moodle <sup>2</sup>

coletados dos alunos e professores. Além disso, é importante destacar que o código-fonte do programa contém um arquivo denominado “*requirements*” que especifica as versões das bibliotecas Python utilizadas, as quais são descritas na Tabela 2.2 juntamente com suas respectivas funcionalidades dentro do programa.

Tabela 2.2: Bibliotecas do Python utilizadas

Biblioteca	Funcionalidade no programa
requests	Chamada de API's
csv	Leitura/Escrita de arquivos CSV
urllib	Trabalha com URL's para criação do token do Moodle
os	Funcionalidades do sistema operacional
pwininput	Monitora inputs (esconde senha do usuário)
threading	Controle de leitura e escrita da thread
concurrent.futures	Criação e controle de threads
json	Serializa objetos para o formato json
unidecode	Retira acentos das mensagens
time	Controle de tempo
collections	Preenche e manipula estrutura de dados
googletrans	Traduz as mensagens
pandas	Manipulação de DataFrames
datetime	Conversões de datas
bs4 (BeautifulSoup)	Remoção de elementos HTML das mensagens
numpy	Manipulação de arrays

Fonte: Tabela elaborada pelo autor

### 2.2.3 LeIA

Segundo Elbagir e Yang (2019) VADER (*Valence Aware Dictionary and sEntiment Reasoner*) é uma ferramenta de análise de sentimentos baseada em dicionário e regras, especialmente ajustada aos sentimentos expressos nas mídias sociais. É uma ferramenta totalmente gratuita de código aberto <sup>3</sup>.

A ferramenta LeIA é adaptada para a língua portuguesa (ALMEIDA, 2018), e utiliza o VADER como base para ajudar a monitorar e compreender a opinião de um público sobre um determinado assunto. Ela utiliza técnicas para classificar os textos como positivos, negativos ou neutros. De acordo com Amin et al. (2019) as técnicas utilizadas são as seguintes:

- *Preprocessing*: O texto de entrada é dividido em palavras individuais (*tokenização*). Inicialmente é feita uma divisão de palavras, *emojis* e sinais de pontuação. É importante ressaltar que alguns sinais de pontuação afetam a valência das palavras (pontuação que decide se a palavra é positiva, negativa ou neutra); por exemplo uma frase que tem muitos sinais de exclamação pode ser interpretada, dependendo do contexto, como algo intensamente negativo ou positivo.
- *Boosting*: Após a *tokenização*, todos *tokens* são verificados para fins de aumento de valência. Se uma palavra de reforço como “extremamente”, “muito”, “ótimo” for encontrada, a valência da palavra será aumentada. Em seguida, todas palavras escritas em letras maiúsculas são avaliadas levando em consideração o contexto em que estão inseridas, e sua polaridade pode variar entre positiva, neutra ou negativa. Além disso, a valência da palavra também pode ser aumentada ou diminuída em função do contexto.
- *Valence Calculation*: Nesta etapa final, a valência de uma frase é medida em uma escala que varia entre -4 a +4. Este valor é calculado e é posteriormente normalizado para variar entre -1 a +1. Desta forma, cada sentença é atribuída com sua respectiva polaridade. Com isso, a polaridade positiva é classificada como  $\geq 0.05$ , a negativa  $\leq -0.05$  e a neutra fica entre  $-0.05$  e  $0.05$ .

---

<sup>3</sup><<https://github.com/cjhutto/vaderSentiment>>

#### 2.2.4 NRCLex

NRCLex é uma biblioteca de processamento de linguagem natural, desenvolvida pelo NRC (MOHAMMAD; TURNEY, 2010). É utilizada para análise de emoções, classificação de texto e extração de informações. Para fazer essas análises, os modelos são treinados com grandes quantidades de mensagens etiquetadas com emoções específicas. Dessa forma, a biblioteca aprende a reconhecer os padrões linguísticos associados a cada emoção, através de abordagens de aprendizado de máquina. Essa biblioteca realiza a identificação dos seguintes sentimentos: raiva, desgosto, medo, alegria, surpresa, ansiedade e confiança. Neste trabalho, fazemos uso dessa biblioteca para buscar identificar duas emoções, pois outra biblioteca referenciada na seção 2.2.5 já cobre as outras. Portanto, com a aplicação do NRCLex buscou-se identificar as seguintes emoções:

- Ansiedade: Comentários com expectativas, desejo intenso e impaciente.
- Confiança: Crença de que algo não falhará, de que é bem-feito ou forte o suficiente para cumprir sua função.

É importante ressaltar que, como todos os sistemas baseados em dicionários, a precisão da API NRCLex pode ser afetada por fatores como o uso de palavras de duplo sentido ou sarcasmo. Além disso, é importante considerar que o contexto em que uma palavra é utilizada exerce grande influência em sua interpretação e significado. No entanto, abordagens léxicas não consideram o contexto, o que limita sua eficácia nesse tipo de análise e pode fazer com que a API traga resultados errôneos.

Entrando um pouco mais a fundo sobre o funcionamento dessa API, podemos relatar que o processo geral é dividido em três etapas principais: pré-processamento, mapeamento para emoções e cálculo dos resultados. Na primeira etapa, o texto é limpo e pré-processado para remover pontuação e palavras irrelevantes como *stopwords*. Na sequência, o texto é dividido em *tokens* (palavras) onde cada um será analisado nas próximas etapas. No mapeamento para emoções, cada *token* é comparado com as palavras presentes no dicionário NRCLex, que é uma base de dados de palavras classificadas quanto a sua emoção, como as emoções citadas acima, “Ansiedade” e “Confiança”. Quando uma mensagem é analisada, cada palavra ou frase é comparada com as entradas na base de dados. Se uma palavra ou frase é encontrada, sua classificação de emoção é associada à mensagem. Por fim, a pontuação geral da mensagem é calculada com base na soma das classificações de emoções de cada palavra ou frase encontrada na base de dados.

### 2.2.5 EmoRoBERTa

A biblioteca EmoRoBERTa (KAMATH et al., 2022) é uma biblioteca de processamento de linguagem natural a qual utiliza o modelo RoBERTa, pré-treinado com vastas quantidades de texto. Essa biblioteca realiza a análise das seguintes emoções (DEMSZKY et al., 2020):

- Admiração: Encontrar algo impressionante ou digno de respeito.
- Diversão: Encontrar algo engraçado ou ser entretido.
- Raiva: Forte sentimento de desagrado ou antagonismo.
- Aborrecimento: Raiva leve, irritação.
- Aprovação: Ter ou expressar uma opinião favorável.
- Cautela: Exibindo precaução para evitar dano, transtorno ou perigo.
- Confusão: Falta de compreensão, incerteza.
- Curiosidade: Forte desejo de saber ou aprender algo.
- Desejo: Forte sentimento de querer algo ou desejando que algo aconteça.
- Desapontamento: Tristeza ou desagrado causado pela não realização de esperanças, ou expectativas.
- Desaprovação: Ter ou expressar uma opinião desfavorável.
- Desgosto: Repulsa ou forte desaprovação despertada por algo desagradável, ou ofensivo.
- Constrangimento: Autoconsciência, vergonha ou desajeitamento.
- Empolgação: Sentimento de grande entusiasmo e vontade.
- Medo: Estar com medo ou preocupado.
- Gratidão: Sentimento de agradecimento e apreciação.
- Sofrimento: Intensa tristeza, especialmente causada pela morte de alguém.
- Alegria: Um sentimento de prazer e felicidade.
- Amor: Forte emoção positiva de estima e afeição.
- Nervosismo: Apreensão, preocupação, ansiedade.
- Otimismo: Esperança e confiança no futuro ou no sucesso de algo.
- Orgulho: Prazer ou satisfação devido às suas próprias realizações, ou às realizações daqueles com quem se está intimamente associado.
- Realização: Tornar-se consciente de algo.

- Alívio: Tranquilidade e relaxamento após a libertação da ansiedade ou sofrimento.
- Remorso: Sentimento de arrependimento ou culpa.
- Tristeza: Dor emocional, tristeza.
- Surpresa: Sentir-se espantado, surpreso com algo inesperado.
- Neutro: Comentário imparcial ou que não foi possível detectar nenhuma outra emoção na mensagem

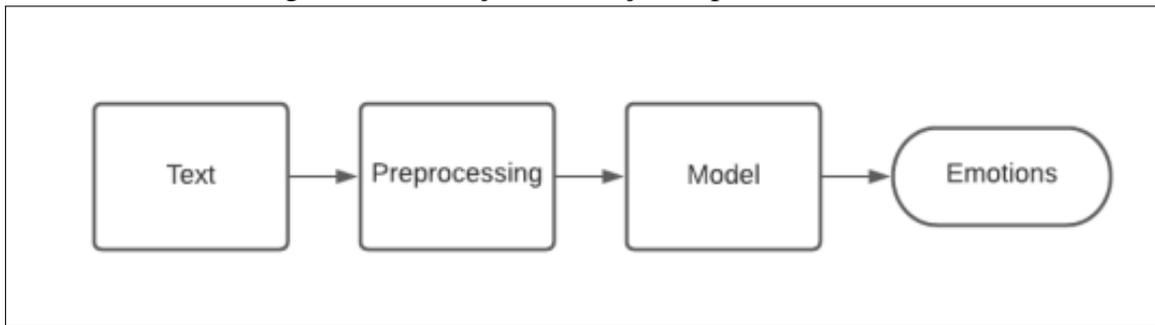
Embora essa API cubra uma ampla gama de emoções, ela não é perfeita. Como todos os sistemas baseados em aprendizado de máquina, a precisão da API EmoRoBERTa pode ser afetada por fatores como a qualidade dos dados de treinamento e a complexidade do modelo escolhido.

A metodologia utilizada para identificar as emoções na API EmoRoBERTa é um pouco mais elaborada em comparação a outras APIs, exigindo um processo mais extenso. Na Figura 2.1 é possível ver o resumo do processo. Após coletar o texto, o mesmo é passado por um pré-processamento, assim como as demais APIs. Essa etapa inclui os seguintes processos: remoção de *emojis*, transformação do texto em minúsculo, remoção de colchetes, remoção de links de sites, remoção de pontuações, remoção de palavras contendo números, remoção de caracteres especiais, corrigir erros ortográficos comuns, remoção espaços extras. Por fim, o texto estará pronto para ser analisado. Além disso, o texto é dividido em *tokens*. Com o texto pré-processado, a análise é iniciada utilizando o modelo RoBERTa (LIU et al., 2019). Este modelo foi desenvolvido para superar algumas limitações do modelo BERT da Google (DEVLIN et al., 2019). Embora ele tenha como base o modelo BERT, que é um modelo pré-treinado com um grande volume de dados, o RoBERTa passou por um processo de treinamento mais intensivo com um volume de dados maior ainda, o que resultou em um modelo mais robusto para classificação de texto e análise de sentimentos. Além disso, o RoBERTa utiliza o conjunto de dados GoEmotions (DEMSZKY et al., 2020) o qual oferece um conjunto de 27 emoções e a neutralidade listadas anteriormente nesta seção.

### 2.2.6 Google Perspective

A Perspective API (LEES et al., 2022) é uma ferramenta de processamento de linguagem natural fornecida pela Google, a qual identifica comentários abusivos utilizando

Figura 2.1: Extração de emoções a partir do texto



Fonte: Kamath et al. (2022)

aprendizado de máquina<sup>4</sup>. Essa ferramenta permite avaliar 3 diferentes tipos de categorias: atributos de produção, atributos experimentais e atributos do *New York Times*. Neste trabalho, recorreremos aos atributos de produção, que foram treinados através do treinamento em grandes quantidades de comentários de diversas fontes, e que apresentam seis características distintas:

- Toxicidade: Um comentário rude, desrespeitoso ou irracional que provavelmente fará as pessoas abandonarem uma discussão.
- Toxicidade severa: Um comentário muito odioso, agressivo, desrespeitoso ou de outra forma muito provável de fazer um usuário deixar uma discussão, ou desistir de compartilhar sua perspectiva. Esse atributo é muito menos sensível a formas mais leves de toxicidade, como comentários que incluem usos positivos de palavrões.
- Ataque de identidade: Comentários negativos ou odiosos direcionados a alguém por causa de sua identidade.
- Insulto: Comentário insultuoso, inflamatório ou negativo em relação a uma pessoa, ou a um grupo de pessoas.
- Profanidade: Palavrões ou outra linguagem obscena.
- Ameaça: Descreve a intenção de infligir dor, lesão ou violência contra um indivíduo, ou grupo.

Os atributos do *New York Times* foram treinados apenas em comentários dessa fonte específica. Eles apresentam características específicas, tais como:

- Ataque ao autor: Ataque ao autor de um artigo ou postagem.
- Ataque ao comentarista: Ataque ao colega comentarista.
- Incoerente: Difícil de entender, sem sentido.

<sup>4</sup>Ver <<https://perspectiveapi.com/how-it-works/>>

- Inflamatório: Com a intenção de provocar ou inflamar.
- Provável de rejeitar: Medida geral da probabilidade de o comentário ser rejeitado de acordo com a moderação do *New York Times*.
- Obsceno: Linguagem obscena ou vulgar, como palavrões.
- Spam: Conteúdo comercial irrelevante e não solicitado.
- Não substancial: Comentários triviais ou curtos.

Por outro lado, os atributos experimentais ainda estão em fase de teste e, por isso, não foram tão avaliados quanto os de produção. Eles possuem 2 novas características, tais como:

- Conteúdo sexual: Contém referências a atos sexuais, partes do corpo ou outro conteúdo obsceno.
- Flerte: Cantadas, elogios à aparência, insinuações sexuais sutis.

Em relação ao processo de treinamento, a Google Perspective utiliza um modelo já mencionado antes que é o BERT (DEVLIN et al., 2019). O conjunto de dados utilizado pela Google Perspective possui várias fontes baseadas na web, como, por exemplo, fóruns (*Wikipedia*) e o jornal *The New York Times* (STEVENS; RASUL; OH, 2022). Após o pré-processamento, são utilizadas Redes Neurais Convolucionais (CNNs) individuais para o idioma escolhido. Essas redes neurais são capazes de analisar sequências de palavras em um texto e detectar padrões que possam indicar a presença de características ou tipos de linguagem (LI et al., 2022). Para o caso em questão, a rede é capaz de identificar as emoções citadas acima. Neste trabalho, os textos foram traduzidos para a língua inglesa, já que, apesar de terem sido feitos testes com mensagens na língua portuguesa, os resultados obtidos para o idioma inglês foram significativamente melhores devido à disponibilidade de um conjunto de dados maior. Cabe ressaltar, contudo, que a API possui suporte para diversos outros idiomas, incluindo o português.

### 2.2.7 Yake

Yake é uma biblioteca que permite extrair palavras-chave do texto (CAMPOS et al., 2020). É uma técnica de processamento de linguagem natural que se concentra em identificar palavras ou frases relevantes e importantes em um texto. Para isso, o Yake utiliza uma abordagem baseada em linguística estatística, analisando a frequência e a dis-

tribuição das palavras no texto para determinar sua importância. Neste trabalho, essa ferramenta foi utilizada para coletar as palavras-chave de mensagens muito grandes utilizadas na API EmoROBERTa, as quais o número de caracteres ultrapassou 2000. Caso passasse disso, a API gerava um erro. Portanto, uma verificação de quantidade de caracteres foi adicionada para evitar esse problema.

O processo empregado para a identificação de palavras-chave utilizado pelo Yake é dividido em 4 partes: pré-processamento, geração de candidatos a palavras-chave, cálculo da importância das palavras-chave e seleção das palavras-chave. Primeiramente é feita a limpeza do texto e a identificação de palavras mais utilizadas. Para isso, caso a palavra seja uma *stopword*, ela é removida do texto. Em sequência é feito a separação dos *tokens*. Para a próxima etapa, a ferramenta Yake gera candidatos a palavras-chave com *n*-gramas, que são sequências de *n* palavras consecutivas em um texto. Esses *n*-gramas são avaliados para ver se pertencem ao mesmo trecho de texto e, assim, formam uma palavra-chave candidata. Para realizar o cálculo da importância das palavras-chave, o Yake usa uma série de métricas para calcular a importância relativa de cada candidato a palavra-chave, incluindo a frequência da palavra, a diversidade de contextos em que a palavra aparece, a centralidade da palavra no texto e outros fatores. Por fim, as palavras-chave mais importantes de acordo com as métricas calculadas são selecionadas e retornadas como resultado.

### 2.2.8 Dash

Dash é um framework de aplicações web desenvolvido em Python, projetado para criar elementos interativos de análise de dados e visualizações. Ele é construído baseado em tecnologias como Flask<sup>5</sup>, React<sup>6</sup> e Plotly.js<sup>7</sup>, o que permite com que sejam criados gráficos, tabelas, entre outros elementos interativos. Para esse trabalho, todos os elementos gráficos foram feitos utilizando o framework Dash<sup>8</sup>.

---

<sup>5</sup><<https://flask.palletsprojects.com/en/2.2.x/>>

<sup>6</sup><<https://reactjs.org>>

<sup>7</sup><<https://plotly.com>>

<sup>8</sup><<https://www.dash.org>>

### 3 DESENVOLVIMENTO

Este capítulo traz uma descrição do processo de elaboração da ferramenta proposta neste trabalho assim como os algoritmos implementados. Os resultados da aplicação da ferramenta em disciplinas reais são apresentados no capítulo 4.

Neste trabalho, propomos uma ferramenta para analisar os sentimentos dos alunos. Para a implementação dessa proposta, foi desenvolvido um *script* escrito na linguagem Python<sup>1</sup>. Primeiramente é solicitado ao usuário seu número de matrícula UFRGS e sua senha, conforme cadastrados no Moodle, e com isso é gerado um *token* para realizar a conexão com a plataforma. Esse *token* é uma política de segurança do Moodle que garante a autenticação do usuário para acessar a plataforma, garantindo que somente usuários autorizados possam acessar os dados. Na sequência, é verificado e exibido em tela quais disciplinas o usuário faz parte. Com isso, o usuário escolhe a disciplina para a qual deseja realizar as análises e então as ferramentas responsáveis por coletar as mensagens e analisar os sentimentos são acionadas. Ao final, é aberta uma conexão com um servidor local no qual o resultado é exibido visualmente em gráficos interativos, onde o usuário pode filtrar diferentes datas, alunos e sentimentos. Dentre esses gráficos interativos, existe um relatório geral que exhibe a média de todos os sentimentos detectados para um determinado aluno. O pipeline de funcionamento completo da ferramenta proposta é sumarizado na Figura 3.1.

#### 3.1 Login do Usuário e seleção de disciplinas

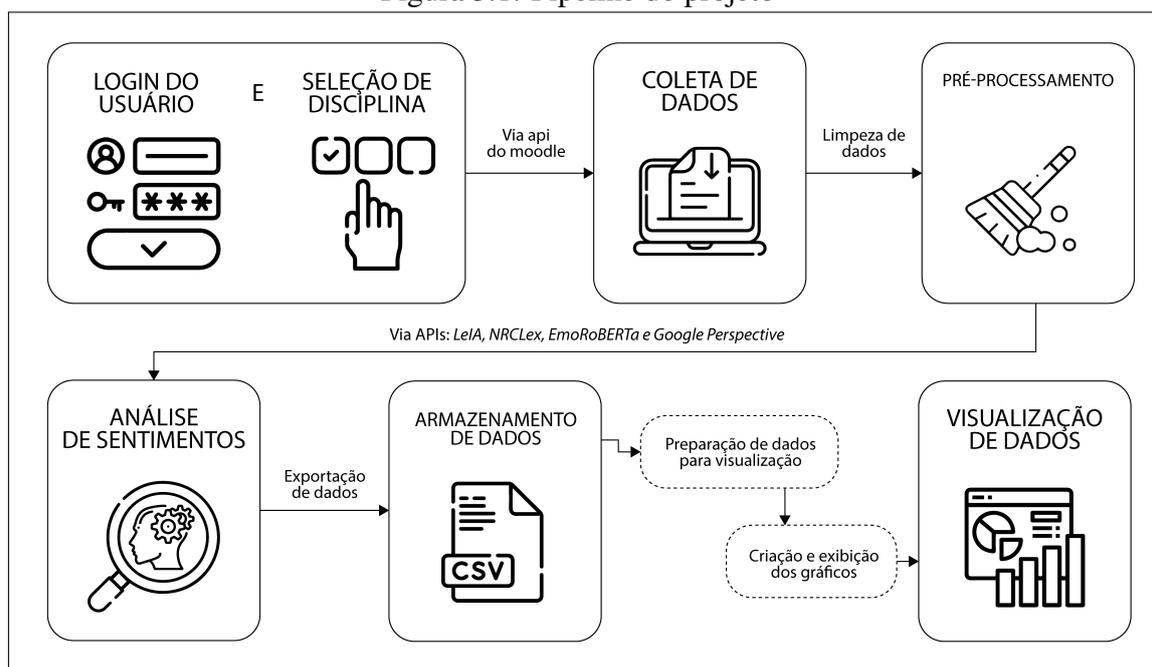
Para realizar o login, solicita-se usuário e senha para a criação de um *token* que garante acesso aos dados do Moodle. Caso o usuário ou a senha informados estejam incorretos, o programa não consegue gerar o *token* e então é solicitado novamente o login. Após a geração desse *token*, é utilizada a função *core\_user\_get\_users\_by\_field*<sup>2</sup>, que envia a matrícula informada como parâmetro para pegar o identificador do usuário. Esse identificador é um número gerado pelo próprio Moodle, onde cada usuário possui um identificador único. Após obter o identificador, a função *core\_enrol\_get\_users\_courses* é acionada e utilizada este identificador para verificar quais disciplinas aquele usuário está

---

<sup>1</sup>Disponível em: <<https://github.com/HenriqueBilo/Analises-Textuais-Moodle>>

<sup>2</sup>Consultar a Tabela 2.1 para uma descrição técnica dessa e das demais operações realizadas pelas funções aplicadas

Figura 3.1: Pipeline do projeto



Fonte: Elaborado pelo autor

participando. Como resultado, coletamos o identificador do curso e seu nome.

Após a conclusão dos processos mencionados, a etapa inicial de autenticação e identificação das disciplinas que o usuário está vinculado é concluída. Em seguida, as disciplinas vinculadas são exibidas para que o usuário selecione uma, a ser analisada.

### 3.2 Coleta de dados

Após o usuário selecionar a disciplina que deseja analisar, é iniciada a etapa de coleta de dados desta disciplina. Os dados coletados no Moodle possuem sua origem em *fóruns*, mensagens diretas ao professor da disciplina e *chats* abertos (bate-papo) à turma. É importante destacar que, para garantir a anonimização dos dados coletados neste trabalho, nenhuma informação como cartão UFRGS ou nome do usuário foi coletada. Em vez disso, foi utilizado o identificador do usuário obtido via API do Moodle, a fim de preservar a identidade de todos os envolvidos.

A primeira coleta feita é para pegar os dados dos chats, é primeiramente utilizada a função `mod_chat_get_chats_by_courses`<sup>3</sup> para obter os chats do curso, é necessário passar o identificador do curso e, como resultado, são obtidos os seguintes dados: identificador do chat, nome do curso e a Seção. Na sequência, é utilizada a função

<sup>3</sup>Consultar a Tabela 2.1 para uma descrição técnica dessa e das demais operações realizadas pelas funções aplicadas

`mod_chat_get_sessions`. Para essa função é passado o parâmetro identificador do chat, tendo como retorno todas as sessões daquele chat. Para cada chat, é possível que haja mais de 1 sessão. Os dados importantes obtidos nessa função são: número de usuários na sessão, horário inicial da sessão, horário final da sessão. Por fim, é utilizada a função `mod_chat_get_session_messages` que coleta as mensagens da sessão referenciadas, passando os parâmetros obtidos nas funções anteriores: identificador do chat, horário inicial e final da sessão. Todas mensagens obtidas são armazenadas, juntamente com a data correspondente, em um arquivo CSV de chats.

A próxima coleta é para obter os dados das mensagens que os alunos mandam diretamente para o professor, ou seja, pelo chat privado. Inicialmente é chamada a função `core_enrol_get_enrolled_users`, é passado o identificador do curso como parâmetro para obter a sua função na disciplina. Em outras palavras, é necessário saber se o usuário é aluno ou tem algum cargo de ensino para que o foco do trabalho seja apenas em alunos. Na sequência, é chamada a função `core_message_get_messages`, onde é necessário passar o usuário que está recebendo a mensagem (nesse caso, é o usuário do professor), o usuário que vai enviar (todos alunos na disciplina), um parâmetro informando que desejamos obter as conversas e, por último, um parâmetro informando que desejamos ver todas as mensagens já lidas. Como resultado, obtemos as mensagens e a data de cada mensagem. Esses dados serão armazenados em um arquivo CSV de mensagens diretas.

Por fim, a última etapa de coleta é responsável por pegar os dados das mensagens postadas em fóruns criados pelos professores. A primeira função a ser chamada é a `mod_forum_get_forums_by_courses`, a qual possui um único parâmetro obrigatório, o identificador do curso. Como resultado, obtemos uma lista de fóruns criados no curso informado. Na sequência, é chamada a função `mod_forum_get_forum_discussions` e, para cada fórum contido na disciplina escolhida, são obtidas as discussões. Ou seja, cada fórum possui tópicos de discussões com diferentes assuntos abordados. Como parâmetro, é passado o identificador do fórum e, como resultado, é obtido o identificador da discussão e as mensagens da mesma. Com o identificador das discussões, é chamada a função `mod_forum_get_discussion_posts` para pegar as mensagens e a data de criação das mesmas. Contudo, o formato das mensagens vem com alguns elementos da linguagem de marcação HTML5. Por conta disso, é utilizada a biblioteca BeautifulSoup, listada na tabela 2.2, para remover esses elementos e então obter a mensagem de forma limpa. Por fim, um arquivo CSV é criado armazenando os seguintes dados: identificador do post, assunto do post, mensagem, autor e a data da postagem.

Com todas essas etapas realizadas, temos arquivos CSV para cada origem de mensagem diferente e, com isso, o programa está pronto para realizar as análises de sentimentos e emoções propostas.

### **3.3 Pré-processamento**

Antes de seguir para a etapa de análise, é necessário realizar uma limpeza dos dados. Nesta limpeza é feita a eliminação de acentos das mensagens e tradução para inglês com o auxílio da ferramenta `Googletrans`. Com exceção da API `LeIA`, todas outras foram utilizadas no idioma inglês. Foram realizados testes com mensagens tanto em português quanto em inglês, e os resultados obtidos com as mensagens em inglês foram significativamente mais precisos. Isso ocorre porque muitos dos serviços de APIs são mais compatíveis ou possuem um conjunto de dados de treinamento mais abrangente para o idioma inglês, sendo assim mais preparados para esse idioma. Além disso, algumas APIs podem aceitar apenas mensagens codificadas em inglês, o que pode impactar na qualidade dos resultados obtidos para outros idiomas, como o português.

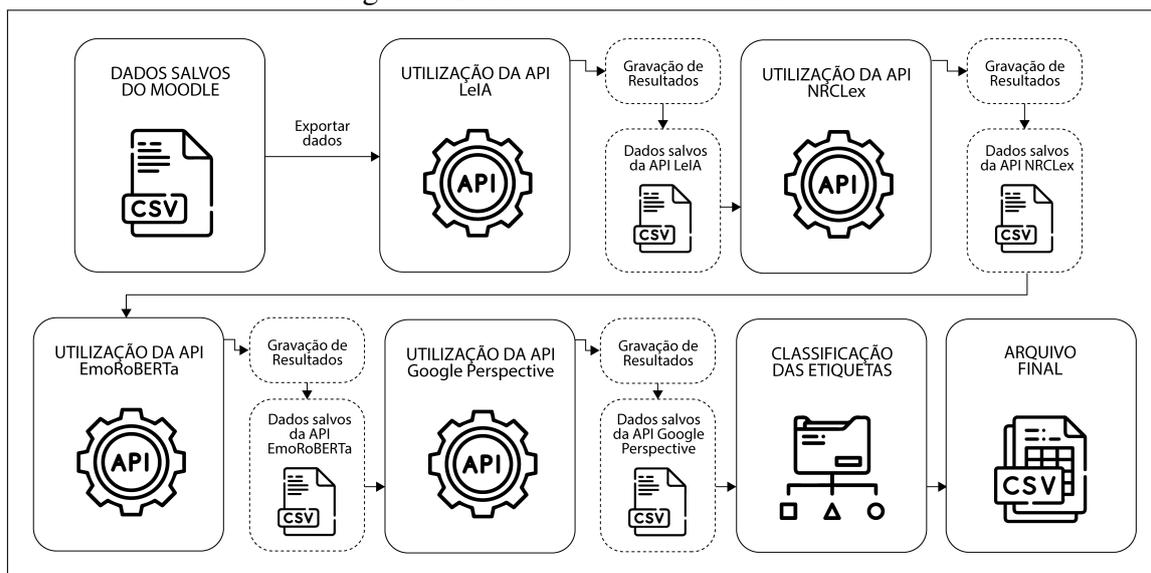
### **3.4 Análise de sentimentos**

Nesta seção, são apresentados os procedimentos realizados na análise dos dados coletados no Moodle. O objetivo da análise proposta é identificar a polaridade, os sentimentos e as emoções presentes nos textos coletados. Para isso, são utilizadas APIs específicas para análise de sentimentos e processamento de linguagem natural, as quais recebem dados coletados do Moodle e retornam os resultados correspondentes. São detalhadas as etapas necessárias para a integração das APIs, bem como as estratégias adotadas para a escolha das ferramentas e técnicas utilizadas na análise. O fluxo de desenvolvimento seguido e detalhado nesta Seção é ilustrado na Figura 3.2.

Para conduzir a análise, serão carregados os arquivos que contêm dados extraídos na etapa anterior advindos de chats, mensagens diretas e postagens a fim de obter as informações necessárias para realizar a análise de sentimentos e emoções.

Inicialmente, é efetuada a análise de sentimento com a ajuda da API `LeIA`, cujos detalhes são descritos na Seção 2.2.3. Para cada mensagem registrada, a API é consultada, retornando um valor de polaridade. O valor da polaridade identificado pelo uso da

Figura 3.2: Fluxo de dados da análise



Fonte: Elaborado pelo autor

API LeIA é então armazenado junto a respectiva mensagem, para utilização posterior na geração de gráficos (discutido em maiores detalhes na Seção 3.5).

Na etapa seguinte, é realizada a análise de emoções com o auxílio da API NRCLex. As informações extraídas nessa etapa para cada mensagem são armazenadas juntamente com as informações de polaridade anteriormente armazenadas. Após a leitura, o processo de análise será iniciado. Cada mensagem será avaliada por meio de uma *thread* separada, devido ao tempo de resposta de API, que pode ser longo dependendo do número de mensagens. A biblioteca `ThreadPoolExecutor` é utilizada para gerenciar essas *threads*, conforme descrito na tabela 2.2. Antes de submeter a mensagem a ser analisada para a API, as *threads* realizam o pré-processamento referenciado na seção 3.3. Em seguida, a mensagem é enviada e o resultado são as emoções descritas na seção 2.2.4 e seus respectivos valores. Assim, ao final dessa etapa são armazenados arquivos contendo a polaridade, e ainda o nome e valor das emoções coletadas pela NRCLex.

A API EmoRoBERTa funciona de forma semelhante à NRCLex. O programa então configura a API para sua correta execução. Inicialmente, é atribuído um arquivo de pré-treinamento, que contém informações importantes de configuração, incluindo a especificação das emoções a serem analisadas com o uso dessa API (ver seção 2.2.5) e o modelo RoBERTa que é utilizado para detectar essas emoções.

Com os parâmetros corretamente definidos, é hora de instanciar a chamada para a API EmoRoBERTa. Isso é feito utilizando a função pipeline da mesma, a qual recebe parâmetros como o modelo e a configuração desejada. Para realizar as chamadas para

cada mensagem, foi utilizada a mesma estratégia de *threads* utilizada na API NRCLex, visando otimizar o tempo de espera. Também foi utilizado o pré-processamento descrito na Seção 3.3. No entanto, para mensagens com mais de 2000 caracteres, foi necessário a utilização da ferramenta Yake, descrita no Item 2.2.7, para extrair palavras-chave relevantes e juntá-las em uma nova mensagem. Essa nova mensagem provavelmente não fará sentido para um humano, pois terá palavras soltas e não conectadas, porém a API EmoRoBERTa mesmo assim conseguirá fazer a análise e detectar os sentimentos aproximados. Finalmente, a mensagem é enviada para análise e então os sentimentos e seus respectivos valores são armazenados juntamente com as informações anteriormente coletadas.

Na próxima etapa, a API Google Perspective é usada para analisar emoções mais agressivas, como comentários maldosos, ameaças e toxicidade (ver Seção 2.2.6 para maiores detalhes). Para cada mensagem a ser analisada, *threads* são criadas para seu processamento juntamente com os tratamentos vistos na seção de pré-processamento 3.3. Em seguida, a ferramenta é acionada e os sentimentos e seus valores são armazenados.

Por último, as mensagens são classificadas em categorias ou etiquetas definidas pelo autor. Essas etiquetas fornecem uma visão consolidada de diferentes métricas com significados aproximados umas das outras e são organizadas com base em critérios específicos listados abaixo. É importante ressaltar que o valor de 0.7 foi escolhido com o objetivo de equilibrar a visibilidade das etiquetas. Foram testados outros valores, como 0.5, que deixavam as etiquetas muito visíveis, e 0.8, que as tornavam pouco frequentes.

- **Elogio** : Ocorre caso uma das seguintes métricas seja maior que 0.7: admiração, aprovação, gratidão ou amor
- **Interesse** : Ocorre caso uma das seguintes métricas seja maior que 0.7: curiosidade, desejo, empolgação, alegria, otimismo, realização ou confiança
- **Agressão** : Ocorre caso uma das seguintes métricas seja maior que 0.7: ameaça, insulto, profanidade ou ataque de identidade
- **Reclamação** : Ocorre caso uma das seguintes métricas seja maior que 0.7: aborrecimento, raiva, toxicidade, toxicidade severa
- **Insatisfação** : Ocorre caso uma das seguintes métricas seja maior que 0.7: desapontamento, confusão, desaprovação, desgosto, constrangimento, sofrimento ou tristeza
- **Preocupação** : Ocorre caso uma das seguintes métricas seja maior que 0.7: ansiedade, medo, nervosismo, alívio, remorso ou cautela

Caso nenhuma dessas etiquetas seja identificada na mensagem, isto é, os limiares definidos não forem atingidos, os valores das métricas não são registrados, a mensagem é então considerada sendo neutra. Ao final de todo o processo de análise de sentimentos, armazena-se a classificação (etiqueta) atribuída a cada mensagem junto as demais informações extraídas em uma tabela.

### 3.5 Visualização de dados

Para começar a criação dos gráficos, são lidas as informações extraídas e armazenadas anteriormente, que são:

- idUsuario: Identificador do usuário.
- mensagem: A mensagem que o usuário digitou.
- data: A data que a mensagem foi inserida.
- polaridade: O valor da polaridade da mensagem, analisada pela API LeIA.
- NRC\_EMOTIONS: O nome e valor das emoções analisadas pela API NRCLex, no formato nome:valor\*nome:valor para facilitar a leitura.
- EMO\_ROBERTA: O nome e valor das emoções analisadas pela API EmoRoBERTa, no formato nome:valor\*nome:valor para facilitar a leitura.
- GooglePerspectiveMetrics: O nome e valor das emoções analisadas pela API Google Perspective, no formato nome:valor\*nome:valor para facilitar a leitura.
- classificacao: A etiqueta atribuída a mensagem.

Na sequência, com o intuito de preparar os dados já processados para serem visualizados como resultado da ferramenta proposta, tarefas de ajustes e pré-processamento são realizadas. As tarefas necessárias para isso são descritas a seguir:

- Formatação da data para o formato dd/mm/aaaa.
- Conversão do valor das polaridades para numérico.
- Tradução do nome das emoções obtidas pelas APIs NRCLex, EmoRoBERTa e Google Perspective de inglês para português.
- Criação de colunas adicionais na tabela de informações para indicar se a mensagem possui determinada emoção. Como exemplo podemos utilizar a emoção “tristeza”: para identificar a presença dessa emoção, é criada uma coluna denominada “TEM\_TRISTEZA”. Para cada mensagem é verificado se a mesma possui esse sen-

timento e isso é sinalizado de forma binária na coluna.

Foi criado também uma tabela de referência contendo o nome de todas as emoções traduzidas para português, de forma que possa ser utilizado para popular o filtro de métricas que o usuário escolhe.

Para a criação de um relatório geral da turma analisada, foi criada uma tabela auxiliar contendo além das informações até então extraídas, a média de todas as emoções de cada usuário identificado na turma. Foi necessário acessar cada emoção contida no *dataframe* juntamente com o usuário; após isso, é feita uma média de todas as emoções de um determinado usuário.

### 3.5.1 Criação e exibição dos gráficos

No gráfico apresentado na Figura 3.3 podemos ver a métrica Polaridade sendo analisada. Essa métrica significa o quão positivo (barra verde) ou negativo (barra vermelha) foi o texto analisado. É possível notar que, para esse gráfico, possuímos 2 tipos de filtros: (1) **filtro por aluno**, onde é possível selecionar um aluno específico para ver o andamento do mesmo; e (2) **filtro por período**, que permite selecionar o período de data desejado para a análise. No lado direito do gráfico, conforme mostrado na figura, há um campo responsável por exibir as mensagens que geraram tais resultados. Para isso, basta clicar na barra desejada e o campo de mensagem será preenchido. É possível que exista mais de uma mensagem postada no mesmo dia, caso isso aconteça, ambas mensagens são exibidas.

Figura 3.3: Gráfico de Polaridade

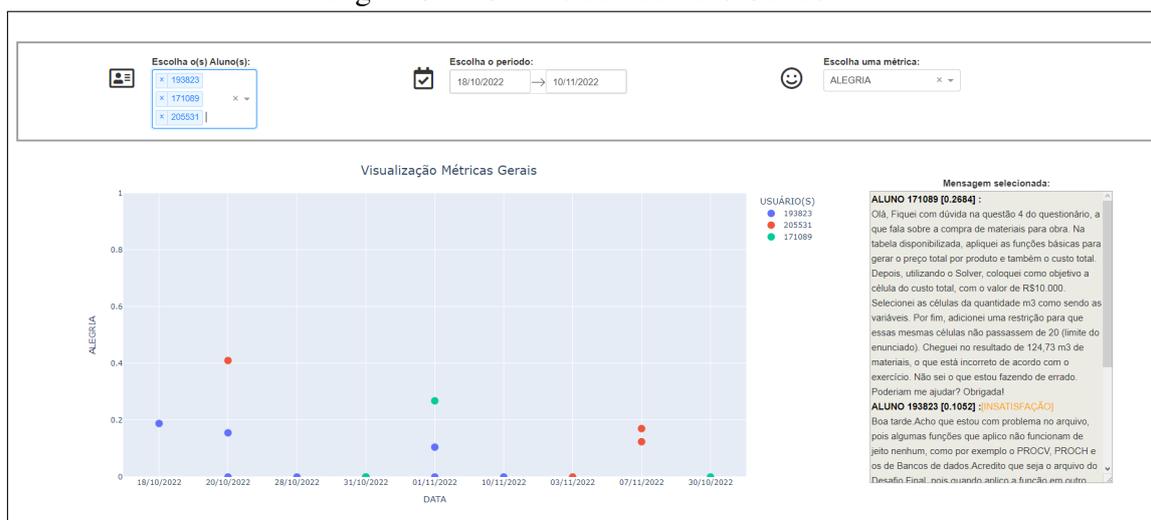


Fonte: Elaborado pelo autor

Já no gráfico apresentado na Figura 3.4, estão presentes outros três filtros: (1) **Filtro de aluno**, que ao contrário do gráfico anterior, permite filtrar mais de um aluno, de forma que seja possível comparar dois ou mais alunos. A cada aluno é atribuída uma cor diferente no gráfico para permitir a diferenciação entre alunos (legenda ao lado do gráfico). (2) **Filtro de período**, similar ao gráfico anterior, realizando o filtro das datas. (3) **Filtro de métrica**, no qual é possível escolher, qual métrica se deseja analisar de determinado(s) aluno(s). Por exemplo, é possível selecionar a métrica “Medo” caso queira ver se o aluno ficou com medo em determinado momento. É importante ressaltar que o termo “métrica” foi utilizado pois a ideia é medir o valor das emoções em questão. Uma observação que pode ser feita a respeito desse filtro, é que as métricas só vão aparecer se tiver pelo menos uma mensagem que foi detectada uma determinada métrica. Exemplo: Se não tiver nenhuma mensagem classificada com a métrica “Medo”, a mesma não vai aparecer como opção para filtrar. Isso foi implementado com o objetivo de simplificar a navegabilidade do usuário, de forma que o usuário só consiga visualizar as métricas que possuem algum resultado.

O entendimento do gráfico é semelhante ao do anterior, basta clicar nos pontos e ele irá apresentar as mensagens correspondentes. Caso o usuário queira omitir algum aluno previamente selecionado, é possível clicar no círculo colorido correspondente ao lado direito do gráfico, isso fará com que os pontos daquela respectiva cor sejam omitidos momentaneamente.

Figura 3.4: Gráfico de Métricas Gerais



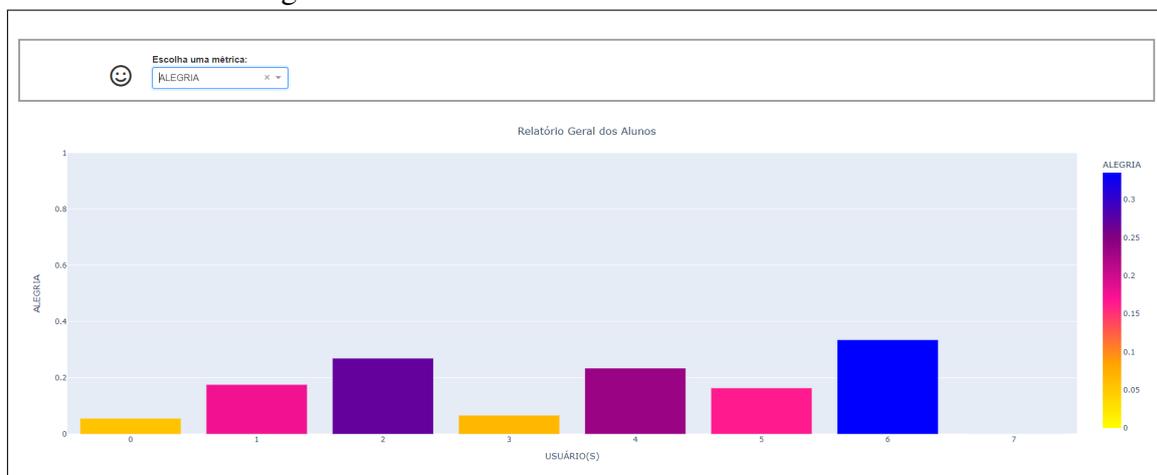
Fonte: Elaborado pelo autor

O terceiro e último gráfico, ilustrado na Figura 3.5, permite visualizar a média de todos os alunos para a emoção selecionada. Neste gráfico cada barra representa um aluno.

A barra disposta no lado direito apresenta através da variação de cores, a intensidade identificada. Sendo a barra azul a mais alta e a amarela a mais baixa. É importante ressaltar que esse gráfico segue a mesma lógica do gráfico anterior, apenas as métricas identificadas para pelo menos uma mensagem são exibidas.

Nesse gráfico também é possível clicar em cada barra. Ao contrário dos gráficos anteriores, esse não possui uma área que mostra as mensagens ao lado. Ao invés disso, os gráficos acima serão preenchidos com o usuário e a emoção que está sendo filtrada no momento.

Figura 3.5: Gráfico de Relatório Geral dos Alunos



Fonte: Elaborado pelo autor

## 4 AVALIAÇÃO E VALIDAÇÃO

Para validar a ferramenta nesse trabalho, foi proposto um experimento com foco na utilização da ferramenta por parte dos professores. O experimento consistiu na instalação da ferramenta nos computadores de cada professor para que eles pudessem realizar análises nas disciplinas desejadas. Um guia de uso foi fornecido oferecendo instruções detalhadas sobre a instalação e utilização da ferramenta (as orientações completas fornecidas para o experimento podem ser encontradas no Anexo A). Para validar esse experimento, foi proposto um questionário na Seção 4.1 para que os professores participantes deste experimento conseguissem dar *feedbacks*.

### 4.1 Questionário

A coleta de opiniões de diferentes professores por meio do questionário foi fundamental para avaliarmos se o *software* atendeu efetivamente os objetivos e expectativas.

O questionário está disponível para visualização no apêndice B e é separado em 3 seções:

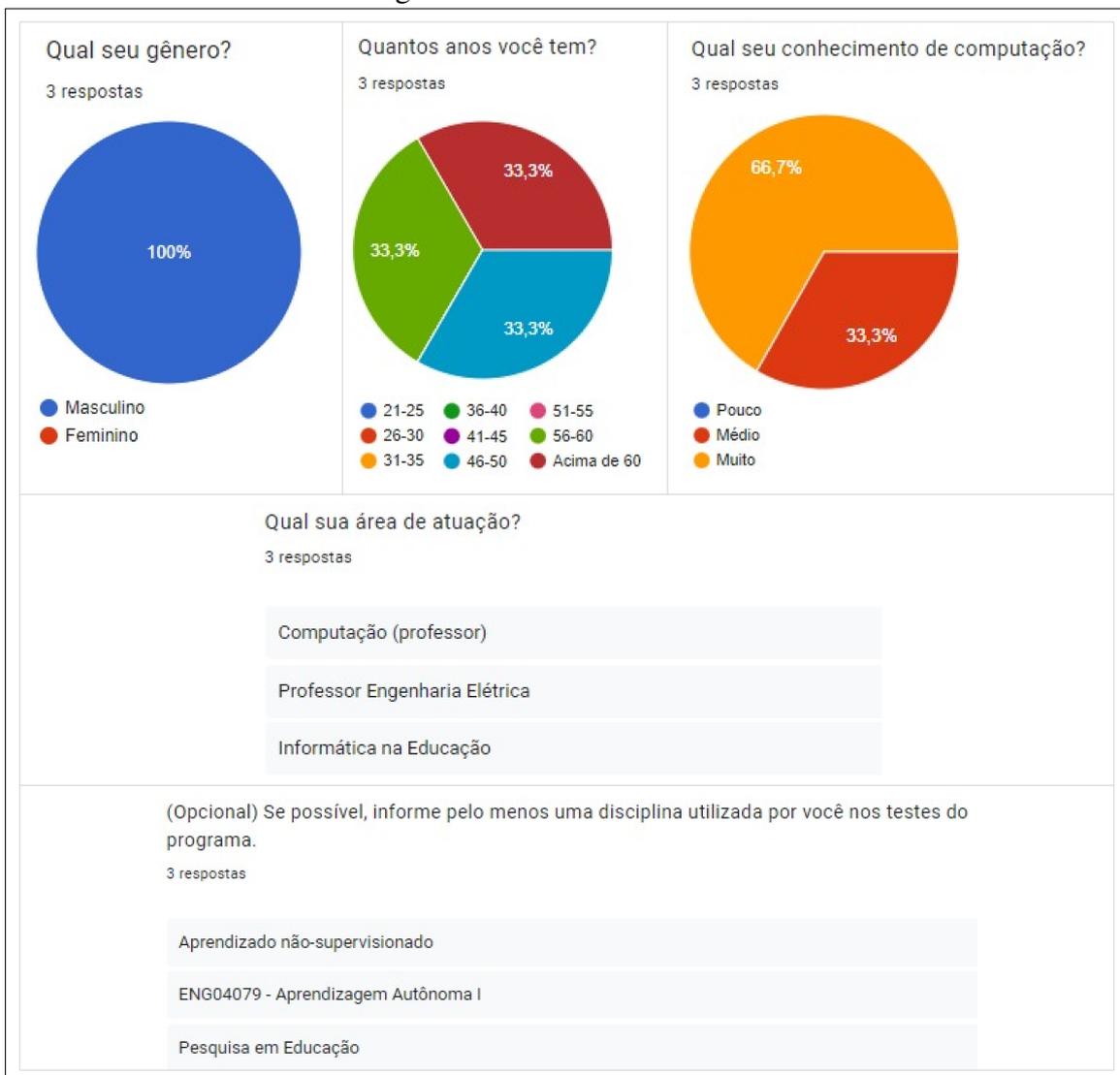
1. Pesquisa de perfil: Nessa parte o objetivo era conhecer o usuário que estava utilizando a ferramenta proposta. Assim, foram questionados feitas perguntas relacionadas a gênero, idade, área de atuação e grau de conhecimento em computação.
2. *Feedback*: Nessa seção do questionário foram abordadas questões referentes a facilidade de instalação do *software*, se o objetivo do sistema foi alcançado, a facilidade de uso, probabilidade de utilizar o sistema e também foi reservado um espaço para uma opinião mais detalhada para coletar possíveis melhorias ou problemas encontrados.
3. Avaliação utilizando o padrão SUS (BROOKE, 2013): O padrão SUS é um padrão que avalia a usabilidade e a satisfação do usuário ao utilizar o *software*. Ele consiste em um conjunto de perguntas que avaliam aspectos diferentes da usabilidade de um sistema, como a facilidade de uso, a clareza das informações e a satisfação geral com o sistema. Cada pergunta é avaliada em uma escala likert de 1 a 5, onde 1 representa “discordo totalmente” e 5 representa “concordo totalmente”. Após o preenchimento das perguntas, as pontuações são somadas e transformadas em uma pontuação geral de usabilidade, variando de 0 a 100.

## 4.2 Resultados

O questionário em questão foi enviado para 5 professores. Porém, somente 3 responderam a pesquisa.

A Figura 4.1 apresenta uma sumarização das respostas fornecidas pelos avaliadores as perguntas demográficas feitas para traçar um perfil. A partir desta figura podemos notar que a maioria dos participantes possui um nível adequado de conhecimento em computação (i.e., entre médio e alto), a faixa etária varia de 46 até uma idade superior a 60 e que ministram disciplinas em diferentes áreas. Todos os participantes se identificam como sendo do gênero masculino.

Figura 4.1: Perfil dos usuários

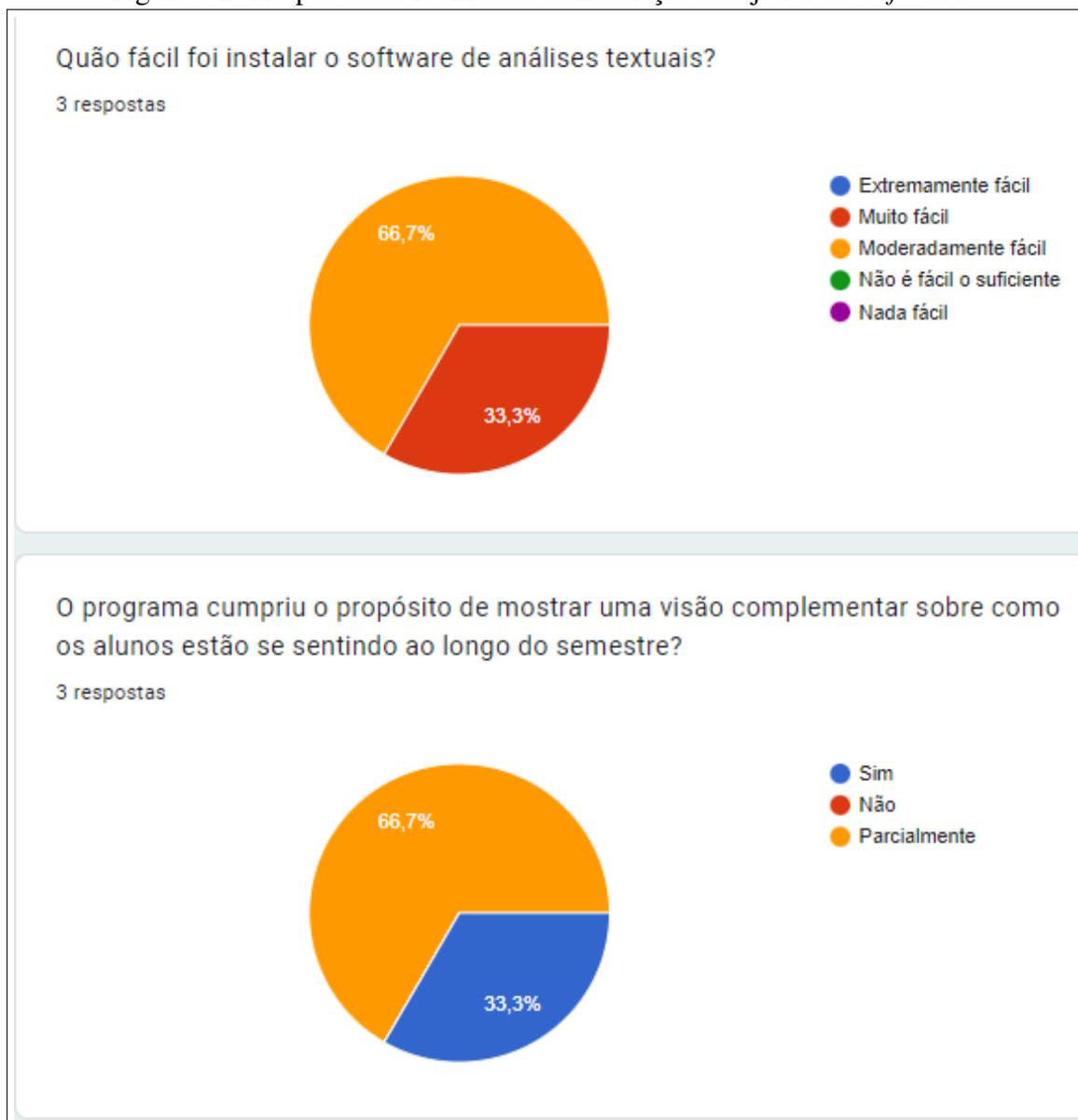


Fonte: Elaborado pelo autor

Na Seção de *feedback* é possível notar que, no geral, as pessoas acharam o *soft-*

ware moderadamente fácil de instalar, conforme ilustrado na Figura 4.2. Elas também acharam que o sistema cumpriu seu objetivo parcialmente (eventualmente por que nem todas as frases pareceram ser 100% precisas), conforme ilustrado no 2º gráfico da imagem.

Figura 4.2: Respostas da facilidade de instalação e objetivo do *software*



Fonte: Elaborado pelo autor

A Figura 4.3 ilustra as respostas dos usuários sobre a facilidade de uso e a velocidade de processamento da ferramenta. Todos os usuários concordaram que é bem simples utilizar a ferramenta. Em relação a velocidade de processamento, um usuário relatou que o tempo de aquisição dos dados pode desencorajar o uso, enquanto os outros consideraram a velocidade adequada.

Figura 4.3: Respostas da facilidade de uso e velocidade de processamento



Fonte: Elaborado pelo autor

A Figura 4.4 apresenta um gráfico com os resultados obtidos através da escala likert, onde as opções de resposta variam de 1 a 5. A opção 1 representa “Discordo completamente”, enquanto a opção 5 significa “Concordo completamente”. Todos os participante selecionaram a opção 4, que representa “Concordo”. Com isso, é possível notar que os professores que participaram da pesquisa concordam que as métricas avaliadas são úteis para compreender o que o aluno está sentindo. Como *feedback*, os professores relataram o seguinte, sendo cada resposta de um professor diferente:

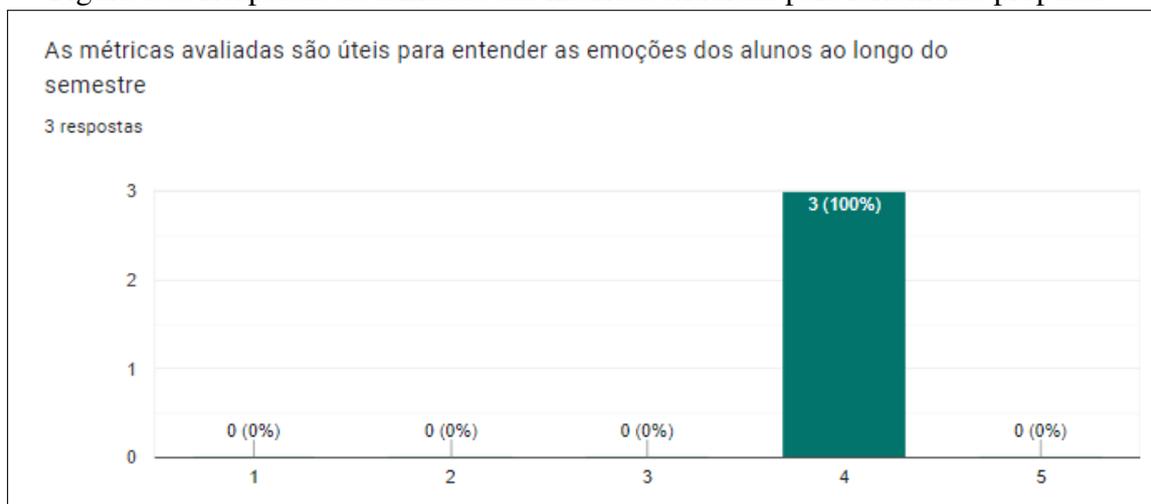
Foi interessante ver as métricas, permitem que a gente tenha uma visão geral sobre o "ânimo da turma", mas também de alunos específicos. Mesmo percebendo que essas classificações não são 100% precisas, me pareceram na maior parte das vezes muito pertinentes. Mas acabei focando em algumas das métricas, que pareciam me ajudar a compreender melhor a satisfação e eventuais problemas com a turma, por exemplo: Alegria, Ansiedade, Confusão.

Ajudam, mas falta alguma maneira de combinar ou agregar melhor os indica-

dores em um único gráfico ao invés de só poder olhar um por um.

As métricas são boas, mas os resultados necessitariam uma aderência maior com a minha percepção sobre o estado afetivo do estudante relacionado à mensagem.

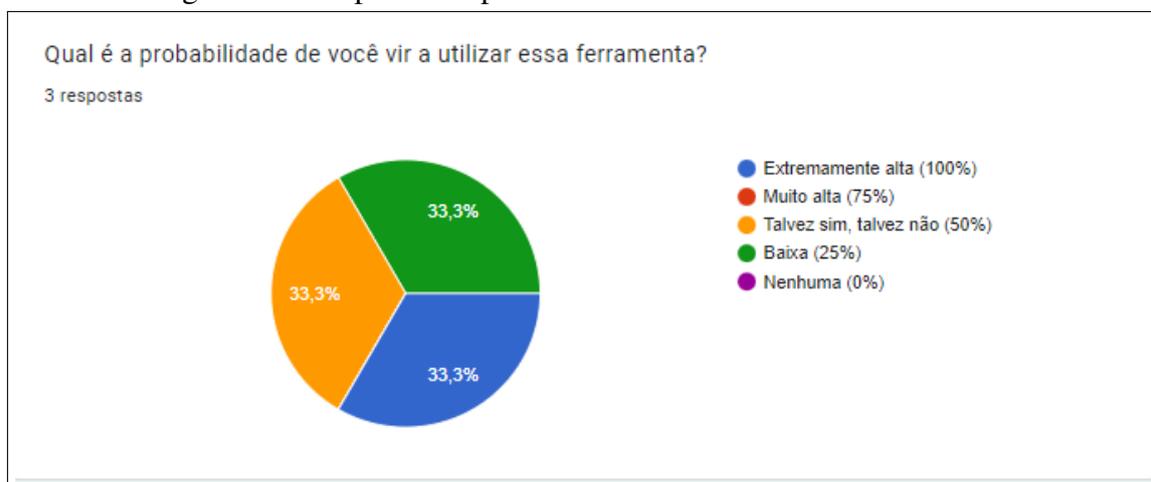
Figura 4.4: Respostas da utilidade das métricas extraídas pela ferramenta proposta



Fonte: Elaborado pelo autor

A Figura 4.5 apresenta as respostas dos avaliadores sobre a probabilidade de utilizar a ferramenta no futuro como auxílio em suas disciplinas. Houve divergências nas respostas: um avaliador relatou que há 100% de chance de utilizar a ferramenta, outro avaliador informou que a probabilidade é de 50%, enquanto o terceiro avaliador informou que a probabilidade é bem baixa (25%).

Figura 4.5: Respostas da probabilidade de se utilizar a ferramenta



Fonte: Elaborado pelo autor

Após reservar um espaço para relatar problemas encontrados, os avaliadores relataram uma demora para inicializar o *software*, isso ocorreu pois a ferramenta verifica a

necessidade de instalação de bibliotecas necessárias a sua execução. Durante os testes, em alguns casos, a ferramenta não encontrou determinado pacote previamente instalado, o que exigiu a instalação do mesmo e, conseqüentemente, ocasionou um tempo de espera para a correta execução da ferramenta. Seguem alguns relatos encontrados pelos participantes:

O software demorou um pouco, após eu ter escolhido a disciplina e ele ter baixado os dados. Creio que estava instalando um componente (o servidor web).

Tenho realmente interesse em usar a ferramenta, tanto do ponto de vista pessoal (como docente, para acompanhar o andamento de minhas turmas), quanto do ponto de vista de pesquisador em área relacionada à área do trabalho. Por questões de privacidade, talvez o nome dos alunos tenha sido removido da interface. Mas no dia a dia do professor, a análise de suas turmas/seus alunos sem poder ver seus nomes parece um problema que precisa ser solucionado.

Por fim, na seção de avaliação utilizando o padrão SUS, foi calculado a pontuação do questionário para cada usuário de acordo com a seguinte regra especificada por (BROOKE, 2013):

- Para as perguntas ímpares (1,3,5,7,9), subtrair 1 da resposta do usuário;
- Para as perguntas pares (2,4,6,8,10), subtrair 5 (5-X) da resposta do usuário;
- Somar os valores de cada usuário (individualmente) e multiplicar por 2,5;
- Somar a resposta de cada usuário e dividir por 3 (número de participantes);
- O resultado final vai variar de 0 a 100.

De acordo com os cálculos apresentados, foram encontrados os seguintes resultados para cada avaliador:

- Avaliador 1: 57,5 pontos;
- Avaliador 2: 57,5 pontos;
- Avaliador 3: 90 pontos.

A média final da resposta então fica em 68,3, bem pouco acima da média 68 estabelecida pela literatura. É possível verificar que dois avaliadores, apesar de terem respondido de forma diferente, chegaram ao mesmo resultado e, que o último avaliador ficou bem acima da média. Dos resultados, é possível concluir que o sistema tem potencial, mas há espaço para melhorias.

## 5 TRABALHOS RELACIONADOS

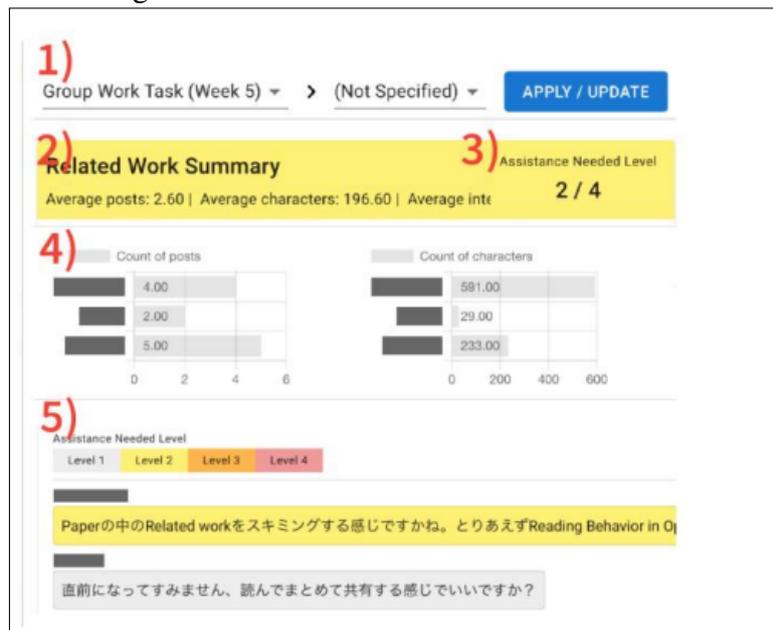
Este trabalho é focado na análise dos sentimentos e emoções (via mineração de texto) de alunos em disciplinas ministradas no Moodle. A mineração de texto tem sido objeto de muitas pesquisas na área de processamento de linguagem natural e inteligência artificial. Com isso, outros trabalhos foram estudados e se relacionam a este, trazendo visões semelhantes além de pontos interessantes a serem somados em futuros projetos.

### 5.1 GWpulse: Supporting Learner Modeling and Group Awareness in Online Forum with Sentiment Analysis

O sistema GWpulse, proposto por NAKAMIZO et al. (2022) tem como função armazenar e visualizar os resultados de análise de registros de conversas em fóruns no Moodle. O objetivo principal do trabalho é ajudar professores e alunos a entender o status das atividades em grupo e apoiar a tomada de decisão com informações que possam melhorar o desempenho dos trabalhos. A análise de sentimento é implementada para fornecer um indicador de quanta intervenção do professor é necessária.

Na Figura 5.1, é possível visualizar a interface do usuário, onde é realizada a seleção do fórum e/ou discussão (1). Na área (2), são exibidas as informações de média da contagem de postagens, contagem de caracteres e intervalo entre postagens do grupo. Além disso, é possível observar o nível de necessidade de assistência (3), que é calculado com base na média de negatividade obtida pela análise de sentimentos de cada postagem. Esse nível é dividido em quatro categorias com limites de [0,00, 0,20], (0,20, 0,40], (0,40, 0,55), (0,55, 1,00], respectivamente. Para cada usuário, a interface apresenta gráficos de contagem de postagens, contagem de caracteres e intervalo de postagens (4). Por fim, é exibido o nível de postagem ANL (Nível de Necessidade de Assistência) para cada discussão (5). A análise de sentimento é realizada por meio da biblioteca Python especializada em classificação positiva/negativa de sentenças japonesas, denominada “asari”.

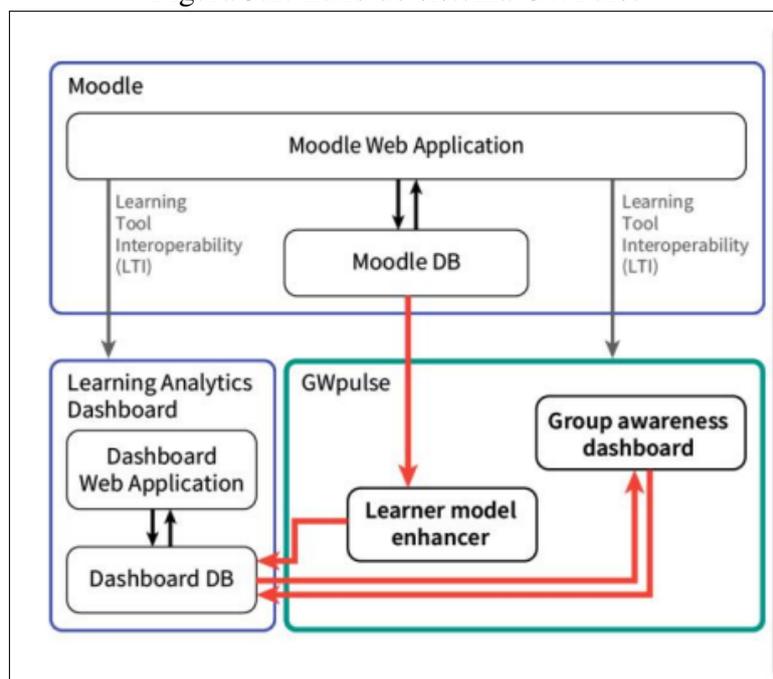
Figura 5.1: Tela do usuário sistema GWPulse



Fonte: (NAKAMIZO et al., 2022)

Na Figura 5.2 é exibido o fluxo de dados que o sistema utiliza, utilizando o Moodle e as ferramentas de análise do LEAF, que é uma plataforma que permite a coleta, análise e visualização de dados de aprendizagem para ajudar a entender como os alunos estão progredindo.

Figura 5.2: fluxo do sistema GWPulse



Fonte: (NAKAMIZO et al., 2022)

Com isso, o projeto é separado em 3 fases:

- Fase de avaliação: Ao examinar o painel de conscientização do grupo, os estudantes têm a oportunidade de avaliar suas próprias atividades, bem como as de seus colegas, de maneira objetiva e visual. Em resumo, ao analisar os dados apresentados pelo GWpulse durante a fase de avaliação, é esperado que cada participante faça avaliações fundamentadas em evidências.
- Fase de Reflexão: Usando o GWpulse, os alunos podem revisar suas próprias atividades, bem como as atividades de outros alunos ou grupos, quando estão refletindo sobre sua participação nas atividades em grupo, especialmente após receberem feedback de colegas ou do professor.
- Fase de Formação: Os professores podem usar indicadores agregados calculados a partir do GWpulse que descrevem as características dos alunos no processo de aprendizagem em grupo como parâmetros ao criar grupos para a próxima atividade em grupo.

Este trabalho segue a mesma linha deste trabalho de conclusão proposto, que é fornecer ao professor uma visão complementar de seus alunos. Porém, analisa apenas os sentimentos positivos e negativos, deixando de lado métricas que poderiam ser fundamentais como alegria, tristeza, entre outras. Além disso, o trabalho é focado em fóruns e não abrange outras formas que o aluno possa usar para se comunicar; como mensagens diretas ao professor ou até chats da turma.

## **5.2 Mineração de Dados Educacionais e Análise de Sentimentos em Ambientes Virtuais de Aprendizagem: um Mapeamento Sistemático**

Proposto por Vivian et al. (2022), este trabalho realizou um mapeamento sistemático de pesquisas sobre análise de sentimentos em Ambientes Virtuais de Aprendizagem (AVA). A ideia é fornecer uma visão geral das técnicas, métodos, algoritmos, bibliotecas e ferramentas utilizadas em análises de sentimentos e emoções nesse contexto. Com isso, o trabalho propôs três questões para realizar a pesquisa:

- QP1 Quais são as técnicas, métodos, algoritmos, bibliotecas e ferramentas de mineração de dados educacionais para análise de base de dados referentes a sentimentos e emoções do estudante no processo de aprendizagem em AVA?
- QP2 Com qual objetivo foi empregada a análise de sentimentos em AVA?
- QP3 Quais tipos de emoções e sentimentos têm sido considerados nos estudos?

Assim, foram obtidos dados como: título, ano de publicação, fonte, periódico/evento, autores, resumo e recursos discutidos. Além disso, todos os artigos foram categorizados com base na classificação de Zelkowitz e Wallace (1998) no qual foram considerados: estudos de caso, projetos experimentais, lições aprendidas e simulações.

Com isso, foram obtidos 20 estudos relevantes para responder ambas as questões propostas. Para a QP1, os algoritmos mais utilizados são o Support Vector Machine (SVM), Naive Bayes e o K-Nearest Neighbours (KNN). Quanto às bibliotecas, as mais utilizadas são NLTK, CaTools do R e NumPy, empacotadas com Vader e Tensorflow. Quanto às ferramentas, os resultados apontaram para as seguintes opções: a linguagem de programação R, o Weka e o AntConc. No que diz respeito à QP2, os motivos mais comuns para realizar a análise de sentimentos em AVA foram: (1) investigar as opiniões dos alunos sobre os cursos, professores e materiais didáticos, (2) aprimorar o processo de ensino e (3) fornecer uma experiência de aprendizagem personalizada. Por fim, em relação à QP3, os autores informam que os sentimentos/emoções mais frequentemente considerados nos estudos foram: (1) análise de polaridade, (2) emoções como raiva, antecipação, nojo, medo, alegria, tristeza, surpresa e confiança e (3) estados afetivos como frustração, desmotivação, ansiedade e tédio.

Embora seja um estudo geral e não uma aplicação prática específica, fica evidente que o objetivo desse trabalho é identificar os sentimentos e emoções mais comumente abordados em estudos sobre o tema em questão. Com isso, é possível notar que o diferencial do trabalho de conclusão proposto, seria uma abordagem de mais métricas que o usual, sendo totalizadas em 36 métricas enquanto os outros trabalhos na área se limitam a menos emoções.

### **5.3 An ontology based text approach for culture aware emotion mining: A Moodle plugin**

Proposto em Labidi, Ouamani e Saoud (2021), o principal objetivo do trabalho é apresentar uma abordagem para a mineração de emoções em ambientes de aprendizagem, utilizando teorias e modelos de emoção, algoritmos de análise de aprendizado existentes e uma ontologia da emoção. Além disso, foi desenvolvido um *plugin* para o ambiente de aprendizagem Moodle, que foi usado para avaliar e validar a abordagem proposta.

No trabalho é dito que processos cognitivos humanos, como percepção, concentração, aprendizagem, memorização, raciocínio e resolução de problemas, são todos afetados

pela emoção. As emoções deste trabalho foram classificadas como primárias, secundárias e terciárias. Uma emoção é chamada de primária quando ocorre após uma reação inconsciente que é quase instantânea e sem qualquer avaliação. Uma emoção é considerada secundária quando é gerada após avaliação cognitiva consciente de um evento ou estímulo. Uma emoção é chamada de terciária se for uma combinação de uma emoção primária e uma emoção secundária.

O *plugin* criado tem como objetivo capturar entradas de texto do aluno no Moodle em chats e fóruns e, através disso, informar ao usuário o tipo de emoção e a intensidade da mesma. Para este trabalho foi utilizado a API NRC Emoticon Lexicon word. O algoritmo foi desenvolvido em PHP e JavaScript. O *plugin* foi experimentado com vinte alunos do ensino fundamental no 6º ano. Foi pedido aos alunos que fizessem um conjunto de exercícios de conjugação onde seria feita a análise de emoções individualmente.

Como resultado da análise do *plugin*, o sistema possui reações para cada emoção analisada, essas emoções precisam ter valor maior que 0,7. As emoções juntamente com a reação do sistema estão listadas abaixo:

- Sentimento negativo  $> 0,7$ : O sistema atribui um líder de classe.
- Raiva  $> 0,7$ : O sistema exibe a mensagem “Ei! Vá com calma! Você pode escrever na caixa acima suas reclamações...”. Em seguida, é exibido uma caixa para que os alunos possam escrever o motivo de suas reclamações.
- Desgosto  $> 0,7$ : O sistema exibe comentários como “Eu sei que você está ficando cansando... É melhor você fazer uma pausa. Estarei esperando por você, não se preocupe...”. Em seguida, o sistema exibe um vídeo engraçado na tentativa de fazer o aluno descontraír.
- Felicidade  $> 0,7$ : O sistema exibe comentários positivos para manter o aluno motivado.

Ao comparar o presente trabalho com o trabalho do *plugin*, é evidente que ambos usam o Moodle para coletar dados textuais de fóruns e chats. No entanto, o diferencial reside na capacidade de coletar dados de outras fontes, como as mensagens diretas ao professor, além de uma gama muito mais ampla de emoções. Enquanto o *plugin* utilizado anteriormente tinha reações limitadas a apenas quatro emoções/sentimentos, a ferramenta atual é capaz de capturar 36 diferentes emoções. Além disso, ela fornece gráficos que permitem o acompanhamento do desenvolvimento emocional dos alunos ao longo do semestre inteiro, uma funcionalidade que não está disponível no *plugin*, que apenas mede

as emoções momentâneas.

#### 5.4 Sentiment Strength Detection in Short Informal Text

O estudo proposto por (THELWALL et al., 2010) apresentou o algoritmo SentiStrength, que tem como objetivo extrair a intensidade do sentimento de textos informais em inglês. Esse algoritmo pode ser usado para analisar comentários em redes sociais, fóruns, salas de bate-papo e outras mídias de comunicação eletrônica nas quais a comunicação baseada em texto costuma ignorar as regras de gramática e ortografia.

O SentiStrength foi avaliado em comentários do site da rede social MySpace, que é conhecido por seus membros jovens, sua orientação musical e seus padrões de comunicação informais. O algoritmo usa um dicionário de palavras sentimentais com medidas de força associadas e é otimizado por meio de um algoritmo de treinamento que ajusta os pesos das palavras sentimentais para maximizar a precisão das classificações de sentimentos.

Em relação ao dicionário de palavras utilizado pelo algoritmo, ele possui 298 termos positivos e 465 negativos, cada um classificado com uma força de sentimento de 2 a 5. O algoritmo inclui procedimentos para ajustar as forças de sentimento usando um conjunto de dados de treinamento. Além disso, o algoritmo inclui um procedimento de correção ortográfica para palavras escritas incorretamente. Esse procedimento identifica as grafias corretas de palavras que foram escritas incorretamente com letras repetidas. Por exemplo, “helllllooo” seria identificado como “hello” por esse algoritmo. O algoritmo é capaz de excluir as letras repetidas acima de duas vezes, como em “helllo” para “hello”, e também é capaz de excluir letras repetidas duas vezes para letras raramente duplicadas em inglês, como em “niice” para “nice”.

Cada frase foi analisada separadamente em relação às emoções mais positivas e mais negativas identificadas nela. Depois, para cada comentário, foi selecionada a emoção mais positiva e a emoção mais negativa das frases que o compunham. As frases foram separadas com base em quebras de linha nos comentários ou pontuação, excluindo emoticons.

Para validar os resultados obtidos pelo algoritmo SentiStrength, o autor realizou comparações com outros algoritmos a fim de comparar qual algoritmo é melhor. O resultado é mostrado na figura 5.3

Os resultados do estudo sugerem que o algoritmo SentiStrength pode ser útil para

Figura 5.3: Resultado comparação SentiStrength com outros algoritmos

<b>Algorithm</b>	<b>Optimal features</b>	<b>Accuracy</b>	<b>Accuracy +/- 1 class</b>	<b>Corr.</b>	<b>Mean % absolute error</b>
<b>SentiStrength</b> (standard configuration, 30 runs)	-	60.6%	96.9%	.599	22.0%
Simple logistic regression	700	<b>58.5%</b>	96.1%	<b>.557</b>	<b>23.2%</b>
SVM (SMO)	800	<b>57.6%</b>	<b>95.4%</b>	<b>.538</b>	<b>24.4%</b>
J48 classification tree	700	<b>55.2%</b>	<b>95.9%</b>	.548	24.7%
JRip rule-based classifier	700	<b>54.3%</b>	96.4%	<b>.476</b>	<b>28.2%</b>
SVM regression (SMO)	100	<b>54.1%</b>	97.3%	<b>.469</b>	<b>28.2%</b>
AdaBoost	100	<b>53.3%</b>	<b>97.5%</b>	<b>.464</b>	<b>28.5%</b>
Decision table	200	<b>53.3%</b>	96.7%	<b>.431</b>	<b>28.2%</b>
Multilayer Perceptron	100	<b>50.0%</b>	94.1%	<b>.422</b>	<b>30.2%</b>
Naïve Bayes	100	<b>49.1%</b>	<b>91.4%</b>	<b>.567</b>	<b>27.5%</b>
Baseline	-	<b>47.3%</b>	<b>94.0%</b>	-	<b>31.2%</b>
Random	-	<b>19.8%</b>	<b>56.9%</b>	<b>.016</b>	<b>82.5%</b>

Fonte: (THELWALL et al., 2010)

ajudar empresas e organizações a monitorar a opinião pública e avaliar a satisfação do cliente, bem como detectar comportamentos inadequados ou ameaçadores em fóruns, redes sociais e outras mídias de comunicação eletrônica.

Ao compararmos o presente trabalho com o SentiStrength, podemos notar que ambos têm como objetivo realizar a análise de polaridade de frases, ou seja, identificar se elas são positivas ou negativas. No entanto, o SentiStrength utiliza seus próprios métodos, como dicionários que contêm a intensidade de cada palavra, para realizar essa verificação. Já o presente trabalho utiliza uma ferramenta externa chamada LeIA para realizar a análise de polaridade.

Entretanto, o presente trabalho vai além da análise de polaridade e inclui a análise de outros sentimentos, que não se limitam apenas a positivo ou negativo. Além disso, apresenta de forma visual para o usuário final os valores de cada sentimento, que podem ser filtrados de acordo com a necessidade do usuário. Em outras palavras, o presente trabalho oferece uma análise de sentimentos mais completa e acessível, permitindo uma compreensão mais aprofundada das emoções expressas em um texto.

## 5.5 Análise comparativa

Foi criada a Tabela 5.1 para proporcionar uma comparação mais clara entre os trabalhos, exibindo as características de cada um deles. Vale ressaltar que o trabalho da seção 5.2 não foi incluído na tabela pois trata-se de uma pesquisa, não tendo implementações de funcionalidades.

Tabela 5.1: Funcionalidades entre os trabalhos estudados e o do autor.

<b>Funcionalidade</b>	Nakamizo et al.	Labidi, Ouamani e Saoud	THELWALL et al.	Do autor
Leitura de dados de fóruns no Moodle	X	X		X
Leitura de dados de conversas privadas no Moodle				X
Leitura de dados de bate-papos no Moodle		X		X
Leitura de dados em outros locais (MySpace)			X	
Análise de polaridade	X	X	X	X
Análise de sentimentos/emoções		X		X
Ampla variedade de sentimentos/emoções				X
Plug-in integrado ao Moodle		X		
Gráficos interativos	X			X
Filtro por data				X
Filtro por aluno				X

Fonte: Tabela elaborada pelo autor

## 6 CONCLUSÃO

Neste trabalho foi apresentada e descrita uma ferramenta de análise de dados desenvolvida com o objetivo ajudar os professores a compreenderem os diferentes sentimentos pelos quais seus alunos estão possivelmente passando no momento que estão cursando determinada disciplina, utilizando dados obtidos na plataforma Moodle. A ferramenta foi implementada através de um *script* em Python que coleta as mensagens do Moodle e utiliza API's para realizar a análise de dados. Após a análise, a ferramenta gera gráficos que são armazenados em um servidor local, permitindo que o usuário visualize os resultados.

Os gráficos gerados exibem as interações dos alunos em uma linha do tempo, representada em gráficos. Nessa linha do tempo, é possível clicar em pontos que representam os valores de cada sentimento analisado, para ver a mensagem que gerou aquele resultado. Além disso, há um gráfico de barras que mostra o relatório geral da turma, apresentando a média de cada sentimento para cada aluno. Selecionando uma das barras, o usuário pode verificar as mensagens que levaram àquela média. Com o acesso a essa ferramenta, os professores podem ter uma visão mais abrangente dos sentimentos dos alunos ao longo do semestre. Isso pode vir a permitir um acompanhamento mais amplo e assim auxiliar a evitar frustrações e reprovações futuras.

### 6.1 Limitações

Há certas limitações na ferramenta proposta, como a necessidade de gerar um *token* do Moodle cada vez que o usuário realiza o *login*. Isso poderia ser aprimorado utilizando o protocolo OAuth 2.0 (HARDT, 2012), no qual os usuários podem autenticar-se no Moodle e obter um *token* de acesso de longa duração (por exemplo, uma vida útil de várias semanas). Isso faria com que o usuário pudesse utilizar as funções do Moodle sem precisar fazer login novamente. O mesmo ocorre para a ferramenta Google Perspective, ela precisa de um *token* para funcionar corretamente. Pela dificuldade de automatização da geração da chave, os testes foram feitos utilizando a chave do autor.

Uma outra limitação da ferramenta seria a remoção dos acentos na etapa de pré-processamento descrito na Seção 3.3, o que poderia causar ou não uma mudança no sentido da palavra. Embora a tradução para o inglês feita depois da remoção dos acentos avalie o contexto em que a palavra foi utilizada, seria interessante analisar isso em um trabalho futuro para verificar se de fato ocorre essa mudança de sentido.

## 6.2 Trabalhos futuros

Considerando as inúmeras possibilidades oferecidas pela inteligência artificial, seria interessante para trabalhos futuros explorar o potencial do chatGPT (WANG et al., 2023) para análise de sentimentos. Neste trabalho foram exploradas as APIs LeIA, NR-CLex, EmoRoBERTa e Google Perspective. Porém, a capacidade do chatGPT de processar grandes volumes de dados de forma rápida e eficiente, torna-o interessante para analisar sentimentos em um grande volume de mensagens. Portanto, para trabalhos futuros, comparar essas diferentes ferramentas no contexto de análise de sentimentos seria uma possibilidade a ser explorada.

Outro possível trabalho futuro seria focar no desenvolvimento de uma interface mais amigável para o usuário, com o objetivo de melhorar a experiência de uso do sistema. Atualmente, o usuário é mantido em uma tela preta até que selecione a disciplina desejada, o que pode ser incômodo e resultar em uma pontuação de usabilidade abaixo do esperado. Ao investir em uma interface mais intuitiva, com *feedback* visual e guias claras, é possível facilitar a navegação do usuário e tornar o sistema mais acessível e fácil de usar. Dessa forma, espera-se que a pontuação de usabilidade, como relatada na Seção 4.2, possa ser significativamente aprimorada, contribuindo para uma experiência mais positiva e eficiente para os usuários do sistema.

Além dos possíveis trabalhos mencionados anteriormente, seria extremamente valioso realizar uma avaliação minuciosa das métricas analisadas, que foram obtidas de fontes já existentes, como as API's analisadas. Isso possibilitaria uma compreensão mais aprofundada da eficácia e relevância dessas métricas no contexto específico do estudo. A avaliação poderia incluir uma revisão crítica da literatura existente sobre as métricas, bem como a coleta e análise de dados para validar sua aplicabilidade e robustez.

## REFERÊNCIAS

- ALMEIDA, R. J. A. **LeIA - Léxico para Inferência Adaptada**. [S.l.]: GitHub, 2018. <<https://github.com/rafjaa/LeIA>>.
- AMIN, A. et al. Bengali vader: A sentiment analysis approach using modified vader. In: **2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)**. [S.l.: s.n.], 2019. p. 1–6.
- ANDRADE, R. R. Utilização de técnicas de aprendizado de máquina para detecção de fake news. **Ciência da Computação-Tubarão**, 2019.
- BENEVENUTO, F.; RIBEIRO, F.; ARAÚJO, M. Métodos para análise de sentimentos em mídias sociais. **Sociedade Brasileira de Computação**, 2015.
- BÓBÓ, M. et al. Identificação do perfil emocional do aluno através de análise de sentimento: Combatendo a evasão escolar. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. [S.l.: s.n.], 2019. v. 30, n. 1, p. 1431.
- BROOKE, J. Sus: a retrospective. **Journal of usability studies**, Usability Professionals' Association Bloomington, IL, v. 8, n. 2, p. 29–40, 2013.
- CAMPOS, R. et al. Yake! keyword extraction from single documents using multiple local features. **Information Sciences**, Elsevier, v. 509, p. 257–289, 2020.
- DAMASIO, A.; DOLAN, R. J. The feeling of what happens. **Nature**, [London: Macmillan Journals], 1869-, v. 401, n. 6756, p. 847–847, 1999.
- DEMSZKY, D. et al. Goemotions: A dataset of fine-grained emotions. **arXiv preprint arXiv:2005.00547**, 2020.
- DEVLIN, J. et al. BERT: Pre-training of deep bidirectional transformers for language understanding. In: **Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)**. Minneapolis, Minnesota: Association for Computational Linguistics, 2019. p. 4171–4186. Available from Internet: <<https://aclanthology.org/N19-1423>>.
- ELBAGIR, S.; YANG, J. Twitter sentiment analysis using natural language toolkit and vader sentiment. In: **Proceedings of the international multiconference of engineers and computer scientists**. [S.l.: s.n.], 2019. v. 122, p. 16.
- EMMERT-STREIB, F. et al. An introductory review of deep learning for prediction models with big data. **Frontiers in Artificial Intelligence**, Frontiers Media SA, v. 3, p. 4, 2020.
- HANSEN, F. Distinguishing between feelings and emotions in understanding communication effects. **Journal of Business Research**, v. 58, n. 10, p. 1426–1436, 2005. ISSN 0148-2963. La Londe Seminar 2003. Available from Internet: <<https://www.sciencedirect.com/science/article/pii/S0148296304001080>>.

- HARDT, D. **The OAuth 2.0 authorization framework**. [S.l.], 2012.
- KAMATH, R. et al. An enhanced context-based emotion detection model using roberta. In: IEEE. **2022 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)**. [S.l.], 2022. p. 1–6.
- LABIDI, A.; OUAMANI, F.; SAOUD, N. B. B. An ontology based text approach for culture aware emotion mining: A moodle plugin. In: IEEE. **2021 International Conference of Women in Data Science at Taif University (WiDSTaif)**. [S.l.], 2021. p. 1–6.
- LEES, A. et al. A new generation of perspective api: Efficient multilingual character-level transformers. **Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining**, 2022.
- LEGOINHA, P.; PAIS, J.; FERNANDES, J. O moodle e as comunidades virtuais de aprendizagem. In: VII CONGRESSO NACIONAL DE GEOLOGIA 2006. SOCIEDADE GEOLÓGICA DE PORTUGAL. [S.l.], 2006.
- LI, Z. et al. A survey of convolutional neural networks: Analysis, applications, and prospects. **IEEE Transactions on Neural Networks and Learning Systems**, v. 33, n. 12, p. 6999–7019, 2022.
- LIU, Y. et al. **RoBERTa: A Robustly Optimized BERT Pretraining Approach**. arXiv, 2019. Available from Internet: <<https://arxiv.org/abs/1907.11692>>.
- LOPES, R. d. C. S. A relação professor aluno e o processo ensino aprendizagem. **Obtido a**, v. 9, n. 1, p. 1–28, 2011.
- MEJOVA, Y. Sentiment analysis: An overview. **University of Iowa, Computer Science Department**, 2009.
- MOHAMMAD, S.; TURNEY, P. Emotions evoked by common words and phrases: Using Mechanical Turk to create an emotion lexicon. In: **Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text**. Los Angeles, CA: Association for Computational Linguistics, 2010. p. 26–34. Available from Internet: <<https://aclanthology.org/W10-0204>>.
- MORAIS, E. A. M.; AMBRÓSIO, A. P. L. Mineração de textos. **Relatório Técnico–Instituto de Informática (UFG)**, 2007.
- MUIR-HERZIG, R. G. Technology and its impact in the classroom. **Computers & Education**, Elsevier, v. 42, n. 2, p. 111–131, 2004.
- NAKAMIZO, Y. et al. Gwpulse: Supporting learner modeling and group awareness in online forum with sentiment analysis. 2022.
- OSTI, A.; BRENELLI, R. P. Sentimentos de quem fracassa na escola: análise das representações de alunos com dificuldades de aprendizagem. **Psico-USF**, SciELO Brasil, v. 18, p. 417–426, 2013.
- PALMEIRA, L. V. S.; SOUSA, A. G. O. de. O papel das emoções no processo de ensino-aprendizagem. **Educação do Ceará em Tempos de Pandemia**, p. 248, 2021.

PENCZKOSKI, R.; PENTEADO, R. J. **Comparação de ferramentas de processamento de linguagem natural para análise de sentimento em português: um estudo de caso em avaliações online de hotéis**. Dissertation (B.S. thesis) — Universidade Tecnológica Federal do Paraná, 2019.

ROSSUM, G. V.; JR, F. L. D. **Python tutorial**. [S.l.]: Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands, 1995.

SANTANA, B. et al. A survey on narrative extraction from textual data. **Artificial Intelligence Review**, Springer, 2023. Available from Internet: <<https://doi.org/10.1007/s10462-022-10338-7>>.

SERRANO, P. H. S. M. Esquemas de comunicação no processamento de linguagem natural. **Leitura: Teoria & Prática**, v. 40, n. 86, p. 41–54, 2022.

STEVENS, H.; RASUL, M. E.; OH, Y. J. Emotions and incivility in vaccine mandate discourse: Natural language processing insights. **JMIR Infodemiology**, v. 2, n. 2, p. e37635, Sep 2022. ISSN 2564-1891. Available from Internet: <<http://www.ncbi.nlm.nih.gov/pubmed/36188420>>.

THELWALL, M. et al. Sentiment strength detection in short informal text. **Journal of the American society for information science and technology**, Wiley Online Library, v. 61, n. 12, p. 2544–2558, 2010.

VASCONCELOS, C. R. D.; JESUS, A. L. P. de; SANTOS, C. de M. Ambiente virtual de aprendizagem (ava) na educação a distância (ead): um estudo sobre o moodle. **Brazilian Journal of Development**, v. 6, n. 3, p. 15545–15557, 2020.

VIEIRA, R.; LOPES, L. Processamento de linguagem natural e o tratamento computacional de linguagens científicas. **Em corpora**, p. 183, 2010.

VIVIAN, R. L. et al. Mineração de dados educacionais e análise de sentimentos em ambientes virtuais de aprendizagem: um mapeamento sistemático educational data mining and sentiment analysis in virtual learning environments: a systematic mapping. **EAD em Foco. Rio de Janeiro, RJ. Vol. 12, n. 2 (2022), e1786, p. 1-15**, 2022.

WANG, F.-Y. et al. What does chatgpt say: the dao from algorithmic intelligence to linguistic intelligence. **IEEE/CAA Journal of Automatica Sinica**, IEEE, v. 10, n. 3, p. 575–579, 2023.

## APÊNDICE A — GUIA DE USO

Para conseguir executar a ferramenta<sup>1</sup> é necessário baixar o instalador do programa. Nele já estarão todas as bibliotecas utilizadas instaladas. Após baixar e instalar o mesmo, é possível seguir os passos abaixo:

- Por ser um arquivo .exe, pode ser que o Windows impeça a execução do programa por questões de segurança. Pode ocorrer de uma tela como a representada na figura A.1 aparecer.

Figura A.1: Alerta ao tentar executar o instalador



Fonte: Elaborado pelo autor

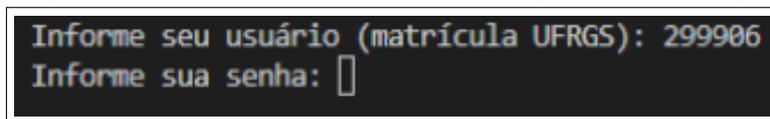
Caso isso aconteça, basta clicar em “Mais informações” e, em seguida, clicar em “Executar assim mesmo”. Lembrando que a execução do programa não oferece nenhum risco ao computador ou às credenciais do usuário.

- Após efetuar a instalação, você terá uma pasta chamada “AnalisesTextuais” no diretório escolhido. Para executar o programa, entre na pasta “AnalisesTextuais”, no local onde foi instalado, procure o arquivo “iniciarPrograma.exe” e execute-o.
- O programa irá solicitar a matrícula UFRGS e sua senha para gerar o token necessário para se comunicar com a API do MOODLE (suas credenciais não serão

<sup>1</sup>Disponível em: <[https://drive.google.com/file/d/1gKXWj-aRiP\\_qgW0YEfhaUNL-6uOZM3vb/view?usp=share\\_link](https://drive.google.com/file/d/1gKXWj-aRiP_qgW0YEfhaUNL-6uOZM3vb/view?usp=share_link)>

armazenadas). Pode ser que inicialmente demore um pouco para aparecer pois o programa carrega uma série de bibliotecas consigo. A figura A.2 ilustra a tela de login.

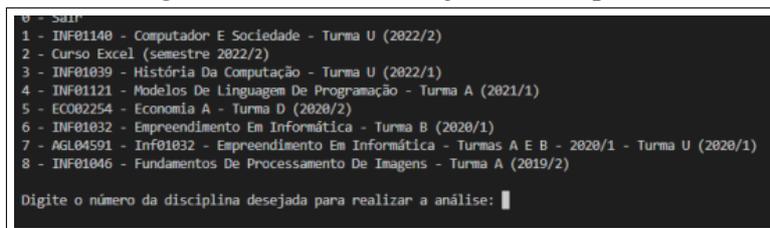
Figura A.2: Tela apresentada ao usuário para realizar login



Fonte: Elaborado pelo autor

- Após a geração do token por parte do software, será gerada uma conexão com o MOODLE, e então serão listadas todas as disciplinas das quais você faz parte. Selecione a disciplina que você deseja analisar inserindo o número correspondente da mesma. A figura A.3 ilustra a tela de seleção de disciplina.

Figura A.3: Tela de seleção de disciplina



Fonte: Elaborado pelo autor

- Após digitar o número correspondente a disciplina, o programa irá fazer as análises propostas. Barras de carregamento serão exibidas para mostrar ao usuário o andamento do processo. A figura A.4 mostra a tela exibida ao usuário para informar que está sendo feita a análise.

Figura A.4: Tela de carregamento



Fonte: Elaborado pelo autor

- Após alguns segundos, será aberta uma aba no seu navegador contendo os resultados.
- Caso você queira alterar a disciplina analisada, basta pressionar CTRL+C na tela preta que está rodando o programa. Após pressionar, o programa irá solicitar suas credenciais e irá listar as disciplinas novamente.
- Para encerrar a execução do programa, basta fechar a tela preta ou selecionar a opção 0 na lista de disciplinas. Também será necessário fechar a aba no navegador

caso alguma análise já tenha sido feita.

## APÊNDICE B — QUESTIONÁRIO

### B.1 Pesquisa de perfil

1. Qual seu gênero?

- Masculino
- Feminino

2. Quantos anos você tem?

- 21-25
- 26-30
- 31-35
- 36-40
- 41-45
- 46-50
- 51-55
- 56-60
- Acima de 60

3. Qual sua área de atuação?

- Campo para digitação

4. Qual seu conhecimento em computação?

- Pouco
- Médio
- Muito

5. Se possível, informe pelo menos uma disciplina utilizada por você nos testes do programa.

- Campo para digitação

### B.2 *Feedback*

1. Quão fácil foi instalar o software de análises textuais?

- Extremamente fácil
  - Muito fácil
  - Moderadamente fácil
  - Não é fácil o suficiente
  - Nada fácil
2. O programa cumpriu o propósito de mostrar uma visão complementar sobre como os alunos estão se sentindo ao longo do semestre?
- Sim
  - Não
  - Parcialmente
3. Quão fácil é o uso da interface do software?
- Extremamente fácil
  - Muito fácil
  - Moderadamente fácil
  - Não é fácil o suficiente
  - Nada fácil
4. Qual é a sua opinião sobre a velocidade de processamento da ferramenta?
- Positiva
  - Neutra
  - Negativa
5. As métricas avaliadas são úteis para entender as emoções dos alunos ao longo do semestre
- Escala de 1 a 5 com os seguintes extremos:
    - Discordo completamente
    - Concordo completamente
6. Complemente com alguma opinião referente a pergunta anterior se necessário
- Campo para digitação
7. Qual é a probabilidade de você vir a utilizar essa ferramenta?
- Extremamente alta (100%)

- Muito alta (75%)
  - Talvez sim, talvez não (50%)
  - Baixa (25%)
  - Nenhuma (0 %)
8. Espaço reservado para deixar sua opinião (melhorias/problemas encontrados ao utilizar o programa)
- Campo para digitação

### **B.3 Avaliação utilizando o padrão SUS**

1. Eu acho que gostaria de usar esse sistema com frequência.
  - Escala de 1 a 5 com os seguintes extremos:
    - Discordo completamente
    - Concordo completamente
2. Eu acho o sistema desnecessariamente complexo.
  - Escala de 1 a 5 com os seguintes extremos:
    - Discordo completamente
    - Concordo completamente
3. Eu achei o sistema fácil de usar.
  - Escala de 1 a 5 com os seguintes extremos:
    - Discordo completamente
    - Concordo completamente
4. Eu acho que precisaria de ajuda de uma pessoa com conhecimentos técnicos para usar o sistema.
  - Escala de 1 a 5 com os seguintes extremos:
    - Discordo completamente
    - Concordo completamente
5. Eu acho que as várias funções do sistema estão muito bem integradas.
  - Escala de 1 a 5 com os seguintes extremos:

- Discordo completamente
  - Concordo completamente
6. Eu acho que o sistema apresenta muita inconsistência.
- Escala de 1 a 5 com os seguintes extremos:
    - Discordo completamente
    - Concordo completamente
7. Eu imagino que as pessoas aprenderão como usar esse sistema rapidamente.
- Escala de 1 a 5 com os seguintes extremos:
    - Discordo completamente
    - Concordo completamente
8. Eu achei o sistema atrapalhado de usar.
- Escala de 1 a 5 com os seguintes extremos:
    - Discordo completamente
    - Concordo completamente
9. Eu me senti confiante ao usar o sistema.
- Escala de 1 a 5 com os seguintes extremos:
    - Discordo completamente
    - Concordo completamente
10. Eu precisei aprender várias coisas novas antes de conseguir usar o sistema.
- Escala de 1 a 5 com os seguintes extremos:
    - Discordo completamente
    - Concordo completamente