



MTA: Uma Ferramenta apoiada em IA para Aprimoramento da Escrita de Projetos de Pesquisa Acadêmica

Ederson Bastiani, UFRGS - edersonbastiani@gmail.com - [0000-0002-2830-6568](tel:0000-0002-2830-6568)
Antonio Miguel Faustini Zarth, IFSC, miguelzarth@gmail.com, [0009-0001-1318-6145](tel:0009-0001-1318-6145)
Eliseo Reategui, PPGIE/UFRGS, eliseoreategui@gmail.com, [0000-0002-5025-9710](tel:0000-0002-5025-9710)

Resumo: A escrita acadêmica desempenha um papel crucial no compartilhamento e na disseminação do conhecimento científico. No entanto, escrever de forma acadêmica, mais especificamente projetos científicos, apresenta desafios significativos. Além da necessidade de clareza e precisão, os textos devem seguir rigorosos padrões metodológicos e argumentativos. Desta forma, diversos trabalhos e tecnologias têm sido propostos com o intuito de facilitar e, sobretudo, melhorar a qualidade dos textos escritos. Neste contexto, este trabalho propõe a ferramenta computacional MTA, um *plug-in* apoiado por mineração de textos e aprendizagem de máquina, com o objetivo de qualificar e propor *feedbacks* formativos visando o aprimoramento da escrita de projetos científicos. A avaliação da ferramenta, confirmada com o apoio de um professor especialista, mostrou que ela pode oferecer suporte significativo aos alunos, promovendo autonomia e melhorando a coesão e consistência da escrita acadêmica, com *feedbacks* detalhados e construtivos.

Palavras-Chave: escrita acadêmica, projetos científicos, mineração de textos

Abstract: Academic writing plays a crucial role in sharing and disseminating scientific knowledge. However, writing academically, specifically scientific projects, poses significant challenges. In addition to the need for clarity and precision, texts must adhere to rigorous methodological and argumentative standards. Thus, various works and technologies have been proposed with the aim of facilitating and, above all, improving the quality of written texts. In this context, this paper proposes the computational tool MTA, a plug-in supported by text mining and machine learning, aiming to qualify and provide formative feedback to enhance the writing of scientific projects. The evaluation of the tool, confirmed with the support of a specialist teacher, showed that it can offer significant support to students, promoting autonomy and improving the cohesion and consistency of academic writing, with detailed and constructive feedback.

Keywords: academic writing, scientific projects, text mining

1. Introdução

A escrita se configura como uma habilidade fundamental em uma sociedade pautada pelo conhecimento. Ela possibilita a expressão de pensamentos, sentimentos e ideias, sendo crucial para o desenvolvimento individual e coletivo. No contexto acadêmico, a escrita científica assume um papel ainda mais relevante, servindo como ferramenta de comunicação, socialização do conhecimento e reconhecimento profissional em todos os cursos superiores (Araújo; Dieb; Costa, 2017). Para isso, textos científicos necessitam de uma linguagem clara, coerente, formal e precisa para descrever os objetivos, métodos utilizados, os resultados obtidos e as conclusões tiradas.

No entanto, a produção de textos acadêmicos de qualidade é um desafio que muitos alunos enfrentam, especialmente aqueles que não estão familiarizados com as convenções da escrita científica. Diversos fatores podem contribuir para as dificuldades enfrentadas pelos alunos na produção destes textos. A heterogeneidade cultural e da



trajetória educacional dos estudantes, por exemplo, assim como o domínio da linguagem e a familiaridade com as normas acadêmicas são alguns destes fatores (Assis, 2014b).

Produzir um texto acadêmico não é uma tarefa trivial mesmo para escritores experientes. É fundamental considerar a perspectiva do professor, que assume o papel de orientador no processo de pesquisa e escrita. Cabe ao docente auxiliar o estudante na construção de um trabalho que siga os padrões do gênero textual em questão, apresenta coerência interna e utilize metodologias adequadas para alcançar os objetivos propostos (Praça, 2015). Contudo, raramente os professores dispõem da quantidade de tempo necessária para planejar, criar materiais, revisar e fornecer *feedback* aos alunos. Como resultado, é natural que os estudantes sejam encorajados a desenvolver seus textos de maneira independente.

Com estes desafios em evidência, torna-se compreensível a necessidade de se investigar e analisar diferentes estratégias, metodologias e tecnologias emergentes que possam auxiliar a gerenciar o processo de escrita e contribuir na construção textual científica. O principal objetivo é garantir que essas estratégias e tecnologias beneficiem de maneira otimizada acadêmicos que estão imersos nesse processo – desde aqueles que estão dando seus primeiros passos na complexa jornada da escrita científica, até profissionais experientes que orientam ou que estão constantemente buscando refinar suas habilidades.

Este trabalho apresenta uma nova ferramenta computacional, apoiada por técnicas de mineração de texto e análise semântica, com o objetivo de aprimorar a redação de projetos de pesquisa ao intervir no processo de revisão textual. Diante da complexidade das normas acadêmicas e da dificuldade de se gerenciar uma ampla gama de informações durante a escrita de um projeto de pesquisa, evidencia-se esta proposta com a necessidade de aprofundar novas possibilidades que beneficiem os processos cognitivos de escrita fomentando a autonomia de seus escritores, sobretudo dos estudantes.

2. Tecnologias para Análise e Mineração de Textos

O avanço das tecnologias de inteligência artificial (IA) tem revolucionado a maneira como analisamos e extraímos informações de grandes volumes de texto. Técnicas de processamento de linguagem natural (PLN), aprendizado de máquina (ML, do inglês *Machine Learning*) e análise semântica têm sido fundamentais para a análise textual, permitindo identificar padrões, tendências e padrões ocultos em dados textuais. Neste contexto, a mineração de texto destaca-se como uma técnica de processamento de linguagem natural para descobrir e extrair padrões e informações pertinentes de textos semi estruturados e não estruturados (Feldman; Sanger, 2006). Essa abordagem permite aos pesquisadores e analistas identificar relações entre conceitos, detectar sentimentos e descobrir *insights* valiosos que podem ser utilizados em diversas áreas para distintos propósitos.

A mineração de texto tem recebido destaque crescente na literatura recente, impulsionada pelo aumento exponencial de dados textuais originados de diversas fontes, incluindo redes sociais, registros médicos, dados de seguros de saúde e agências de notícias (Allahyari et al., 2017). No contexto educacional, a mineração de texto também pode ter um impacto significativo. Ela vem se popularizando principalmente devido a ampliação do uso de ambientes virtuais de aprendizagem e o aumento da quantidade de informações armazenada relacionada aos estudantes, docentes, interações, entre outros (Ferreira-Mello et al., 2019). Como exemplos, pode ser utilizada para analisar grandes



volumes de texto de recursos educacionais, como livros didáticos, redações, artigos e projetos de pesquisa com o objetivo de identificar padrões, tendências e conhecimentos não triviais destes documentos.

Em um estudo publicado em 2011, Reategui et al. (2011) apresentaram a ferramenta Sobek, um minerador de textos para obter os conceitos mais relevantes em um texto e suas respectivas relações. Diversos trabalhos têm utilizado a ferramenta em contextos educacionais, como no auxílio à escrita de resumos (Reategui et al. 2012) ou no apoio ao processo de escrita acadêmica (Oliveira et al., 2022).

Para aprimorar a extração de conhecimento a partir da mineração de texto, é importante combinar esse processo com outras análises. Para isso, no campo do processamento de linguagem natural, principalmente considerando relações de coerência, é comum utilizar técnicas de representação de palavras, como o algoritmo TD-IDF (do inglês *Term Frequency — Inverse Document Frequency*) e o Word2Vec. Essas técnicas permitem transformar palavras em vetores numéricos (*word embeddings*), capturando as relações semânticas entre elas, possibilitando que os sistemas de processamento de linguagem natural realizem análises mais aprofundadas e façam associações e previsões mais precisas.

Desenvolvido por Mikolov et al. (2013), o Word2Vec é uma das abordagens para a criação de word embeddings que utiliza redes neurais para aprender representações a partir de grandes volumes textuais. A eficácia do word2vec reside na sua capacidade de representar relações semânticas e sintáticas complexas de maneira simples e eficiente através de *word embeddings*. O Word2Vec tem contribuindo significativamente para avanços em várias aplicações de PLN, como tradução automática e análise de sentimentos (Mikolov et al., 2013; Rong, 2014).

Esta combinação de mineração de dados com análise contextual semântica é utilizada no cerne deste projeto, cujo principal objetivo é investigar como uma ferramenta computacional, apoiada por técnicas de mineração de texto, pode qualificar a escrita de projetos de pesquisa atuando no processo de revisão textual. A próxima seção detalha tecnologias e trabalhos relacionados que também utilizam técnicas de PLN e ML para análise de projetos e textos acadêmicos.

3. Trabalhos Relacionados

Vários softwares têm sido criados e se destacam com o objetivo principal de aprimorar a qualidade do texto produzido. Feltrim (2004) apresentou o sistema SciPo, uma ferramenta integrada para apoiar a estruturação e redação de resumos e introduções em português, especialmente para teses e dissertações em Ciência da Computação. O SciPo oferece recursos como seleção de estrutura esquemática, crítica automática de textos, e sugestões para melhorar a escrita acadêmica. Como limitação, no entanto, a ferramenta não contempla aspectos como coerência, coesão e estilo, o que seria de suma importância para a escrita de projetos de pesquisa.

Outro proposta que buscou auxiliar a escrita de projetos de pesquisa, teses de doutorado e outros trabalhos acadêmicos é o Thesis Writer. Através de recursos como geração de texto, revisão e organização de ideias, o Thesis Writer se propõe a otimizar o processo de escrita por meio de materiais instrucionais, tutoriais e suporte logístico, guiando os alunos em todo processo em um ambiente online e colaborativo. Apesar de ter funcionalidades relevantes, esta ferramenta não permite assegurar que as seções do projeto formem um objeto de pesquisa consistente.

Com foco na análise da coesão e coerência textual, o trabalho López e López-López (2014) utilizou a análise semântica latente (LSA, do inglês *Latent*

Semantic Analysis) para analisar a coerência global entre diversas seções de uma tese acadêmica na área da ciência da computação e tecnologia da informação. Neste contexto específico, a ferramenta foi capaz de indicar se algum conceito era externo à área (computação), além de analisar e indicar a coerência em 3 níveis: baixa, média ou alta. Embora apresente uma abordagem moderna e promissora, esta ferramenta não foi testada com estudantes reais.

Outros sistemas e estudos, por sua vez, focaram principalmente na análise léxica como sistema tutor. O Academic Word Suggestion Machine (AWSuM) objetiva facilitar a composição de textos ao empregar um conjunto léxico que permite a proposição de palavras e expressões frequentemente utilizadas em textos acadêmicos. Adicionalmente, essa ferramenta é capaz de identificar e sugerir sinônimos para termos corriqueiros, auxiliando os usuários a evitar redundâncias e ampliar a diversidade lexical de suas produções textuais.

De forma semelhante, o ambiente TURET (Tutor for Thesis Writing), proposto por García-Gorrostieta; López-López; González-López (2017), realiza a identificação e avaliação de parágrafos argumentativos em textos de projetos de pesquisa, fornecendo um feedback formativo. No entanto, apesar de apontarem resultados promissores, tanto o TURET quanto o AWSuM não estão isentos de limitações. A precisão das sugestões fornecidas pelas ferramentas depende em grande parte da qualidade e da relevância do texto de entrada fornecido pelo usuário. Além disso, os experimentos do TURET não foram realizados diretamente com estudantes, não sendo possível assim, identificar os ganhos reais com relação à escrita acadêmica.

4. Metodologia

Para a construção do arcabouço computacional objeto deste estudo, cujo objetivo é apoiar o estudante durante o processo de escrita de seu projeto de pesquisa, foi desenvolvida a ferramenta chamada Plug-in MTA, uma extensão para o editor de textos on-line Google Docs. O resumo do fluxo de funcionamento é representado na Figura 1. A etapa 1, representada em branco, representa a escrita do texto em si e é independente do ambiente. As etapas 2, 3 e 4, apresentadas com fundo verde, são realizadas a partir da interação com a ferramenta. As tecnologias, algoritmos, fontes utilizadas para a implementação destes módulos e metodologia de avaliação serão descritas a seguir.

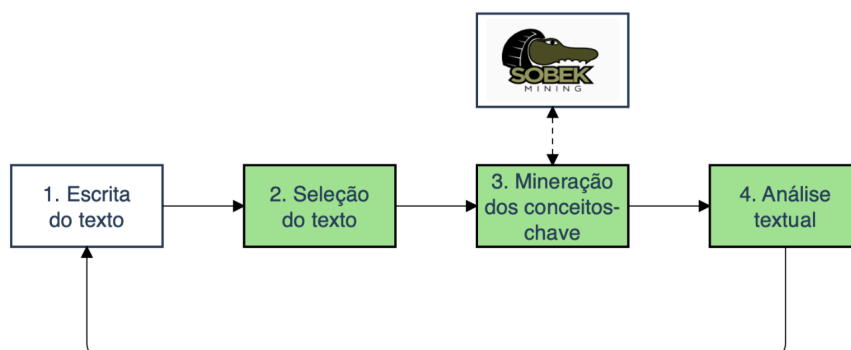


Figura 1: Fluxo de funcionamento do sistema

O plugin MTA foi desenvolvido utilizando a Google Apps Script, uma plataforma online de desenvolvimento baseada em JavaScript que permite adicionar funcionalidades em diferentes produtos do Google Workspace. O plugin foi

desenvolvido e disponibilizado por meio da plataforma *Google Workspace Marketplace*, seguindo as diretrizes do Google.

A interface do usuário, onde o usuário poderá selecionar e delimitar as partes relevantes de seu projeto de pesquisa, foi implementada utilizando as linguagens de marcação HTML5, CSS e Javascript, além dos frameworks *Bootstrap* e *Font Awesome*. Para registro e posterior avaliação dos feedbacks apresentados pelo plug-in, um módulo foi desenvolvido utilizando *Adonis JS* e *MySQL*.

Após a escrita, seleção e definição dos elementos textuais, o plugin MTA utiliza a ferramenta Sobek para a extração e apresentação dos conceitos-chave encontrados. Os elementos utilizados para extração são o título, a questão de pesquisa e o objetivo geral. A escolha destes elementos textuais se deu pelo entendimento de que, a partir da delimitação do tema, eles compõem o eixo estruturante de um projeto de pesquisa, necessitando estar relacionados entre si (Araújo; Dieb; Costa, 2017) por um mesmo conjunto de conceitos.

Para a construção do módulo especialista, cuja responsabilidade no arcabouço computacional é armazenar as regras de conhecimento que representam um projeto de pesquisa coerente, foram utilizadas 4 fontes. A primeira, um conjunto de 10 teses de doutorado do sistema Lume da Universidade Federal do Rio Grande do Sul. A segunda fonte foram projetos de pesquisa de graduação, comentados pelos respectivos professores responsáveis. Como terceira fonte foi utilizada literatura relevante acerca de projetos de pesquisa e metodologia científica, onde foram identificados padrões de escrita, explicações e recomendações sobre a elaboração de projetos. Por fim, foram adicionadas heurísticas e regras adicionais relacionadas à pesquisa acadêmica, propostas pelos autores, como última fonte deste módulo. O fluxo de desenvolvimento do módulo especialista é representado na figura 2.

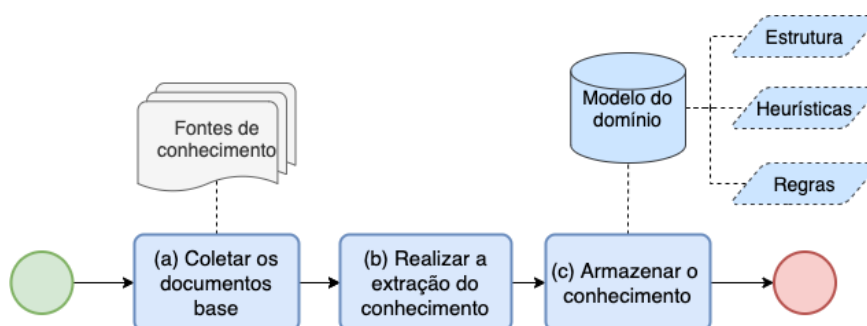


Figura 2: Desenvolvimento do módulo de especialista

Utilizando estas fontes de dados, foram formuladas métricas para avaliar como os elementos de pesquisa se relacionam dentro de um texto. Para isso, foi utilizado o Word2Vec para a construção das representações semânticas dos termos e o algoritmo TD-IDF para computação da similaridade semântica entre título e o objetivo, título e a questão norteadora, e objetivo e a questão norteadora. Os referências quantitativos de análise foram obtidos a partir da codificação das similaridades entre estas seções das teses de doutorado utilizadas. Estas relações foram calculadas a partir da similaridade dos cossenos, dada pela equação:

$$\cos(\theta) = (A \cdot B) / (||A|| * ||B||)$$

onde A e B representam os vetores de palavras a serem comparados e $\cos(\theta)$ indica a similaridade, com resultado entre 0 e 1, sendo o valor 1 indicando alta similaridade e 0 representando ortogonalidade, ou seja, os vetores são completamente distintos em

termos de significado.

A partir das heurísticas do módulo especialista, também foram implementadas funções de análise léxica e de coerência entre os elementos textuais: objetivo, questão da pesquisa, referencial e metodologia. Para o objetivo, o sistema avalia se o texto começa com verbo infinitivo e não possui outros verbos, indicando a precisão do objetivo em apenas uma linha de investigação pelo pesquisador. O módulo também avalia a presença ou não de aspectos metodológicos no objetivo e, finalmente, a presença de pronomes indefinidos, o que indicaria uma referência imprecisa ou indeterminada.

Na questão de pesquisa, o módulo especialista avalia se existe um, e somente um, termo interrogativo, se está formatada como uma pergunta e se não apresenta desnecessariamente aspectos metodológicos. Também são avaliadas as relações de coerência entre a questão de pesquisa, objetivo e título do projeto.

Na metodologia do projeto, o sistema avalia se as seguintes métricas foram abordadas e bem definidas: caracterização e abordagem da pesquisa, estratégia de pesquisa a ser utilizada, procedimentos de coleta de dados, sujeitos da pesquisa, período de realização e, por fim, coerência entre os conceitos-chave do projeto. O último elemento textual analisado é o referencial teórico. Nesta parte o sistema especialista também avalia se o texto está coerente com os conceitos-chaves encontrados anteriormente.

Para avaliação e validação do plug-in MTA em relação ao uso e cumprimento de seu objetivo, foram aplicadas diversas atividades em um conjunto de projetos de pesquisa em elaboração por estudantes de graduação e pós-graduação, analisando a interação do plug-in com os estudantes, seus textos e recomendações do sistema. A partir desta interação, a metodologia de análise dos resultados obtidos neste trabalho foi feita em dois passos:

- I. Análise de interação dos usuários a partir do log;*
- II. Concordância entre as sugestões do sistema com um professor especialista.*

Os registros de usuários considerados foram aqueles que realizaram, ao menos, um ciclo de interação completa com a ferramenta. A análise de concordância das recomendações foi feita por um professor com experiência de mais de 10 anos em docência em disciplinas relacionadas à Metodologia Científica, com o objetivo de qualificar como adequada ou não as indicações propostas pelo sistema e as alterações feitas pelos estudantes.

5. Resultados

A utilização do plug-in MTA, ferramenta computacional desenvolvida neste estudo, começa com a escrita do projeto e posterior determinação das seções no editor *Google Docs*. Nesta etapa, o estudante deverá selecionar trechos específicos do seu projeto e, em seguida, clicar na sua respectiva representação textual, sendo estas: título, questão de pesquisa, objetivo geral, referencial teórico e metodologia. A disposição dos elementos textuais a serem definidos foram escolhidas nesta sequência pois seguem, geralmente, a mesma disposição que consta nos projetos de pesquisa utilizados na metodologia.

Após a seleção e definição das seções necessárias, o estudante avança para a etapa de extração e apresentação dos conceitos-chave, que são identificados a partir dos elementos estruturais do projeto: título, a questão da pesquisa e o objetivo. A escolha desses elementos textuais reside no fato de que a definição do tema é crucial na

elaboração do projeto de pesquisa, pois influencia diretamente a identificação do problema central do estudo (Araújo; Pimenta; Costa, 2015). Por sua vez, Reategui (2020) aponta que a seleção de um título adequado é decisiva, devendo este refletir claramente o conteúdo da pesquisa. Esse cuidado estende-se à formulação do objetivo geral, que deve elucidar claramente o propósito da investigação, e à definição da principal questão de pesquisa a ser respondida.

A partir da extração dos conceitos-chave através do minerador de textos Sobek, o estudante poderá adicionar ou excluir termos encontrados. Uma vez concluída esta etapa, o estudante poderá avançar para a etapa de análise textual do projeto.

Utilizando-se das regras, heurísticas e base de conhecimento do módulo especialista, o plug-in apresenta nesta etapa avaliações e orientações textuais na forma de *feedbacks* de reflexão e *feedbacks* de reforço positivo, ilustrado na figura 3. A adoção desse modelo de interação com o aluno reflete os apontamentos de Hayes et al. (1980), que destacam a tendência dos escritores em encontrar dificuldades na identificação de falhas em suas próprias redações, em comparação com a revisão de textos de outros autores. Esta dificuldade é atribuída à familiaridade do autor com seu próprio texto, o que pode levar à falsa percepção de ausência de falhas. Esse fenômeno ilustra a importância do feedback reflexivo em colaborar para encontrar eventuais limitações e, também, do feedback positivo, buscando orientar o aluno e reforçar sua confiança no seu processo de escrita. A partir destes *feedbacks*, o estudante poderá reescrever ou ajustar texto para nova submissão das etapas de avaliação do plug-in.

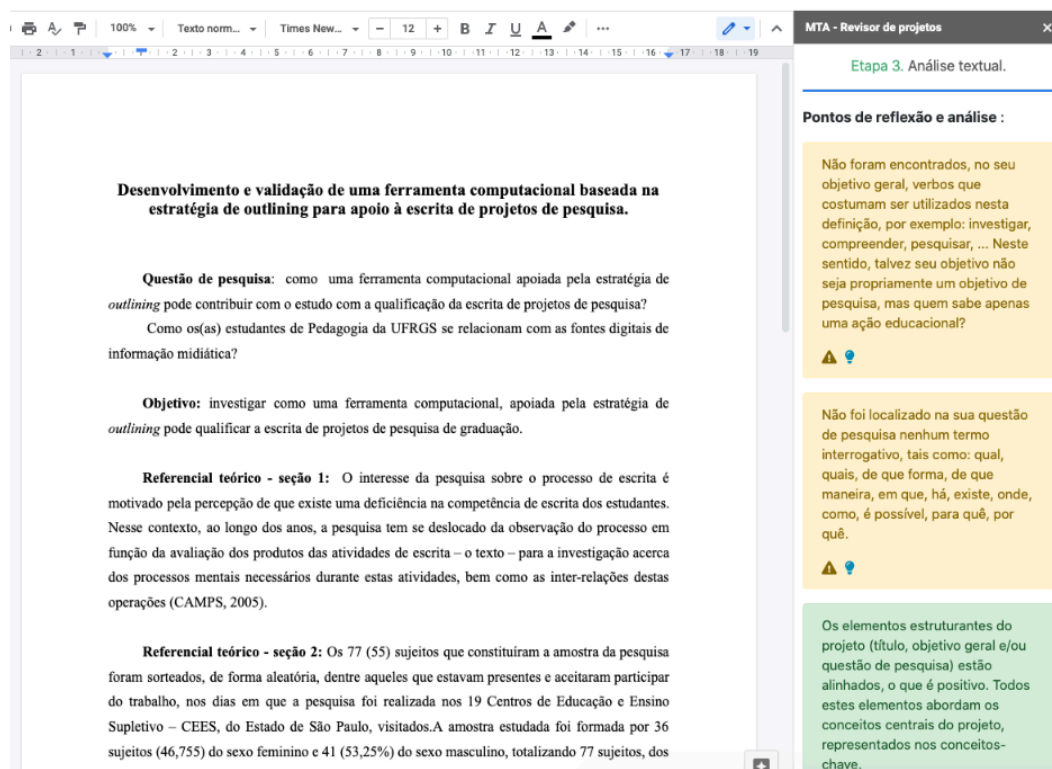


Figura 3: Apresentação dos *Feedbacks*.

A partir da análise pelo professor especialista das interações dos estudantes e respectivos *feedbacks*, foi possível encontrar as ações mais frequentes e identificar os pontos fracos e fortes do sistema proposto. A análise foi dividida em 3 fases de funcionamento: a seleção dos trechos do texto, a extração dos conceitos-chave e a análise dos elementos textuais.

A avaliação inicial dos resultados contou com a revisão dos registros de *log* de



34 estudantes que utilizaram o ambiente, os quais foram qualificados por terem completado a análise textual. A partir desses registros, procurou-se identificar evidências de adequação dos projetos, avaliando quais estudantes realizaram alterações após concluírem todas as etapas de utilização do plug-in MTA. Foram identificados 51 ciclos completos realizados por 16 estudantes, resultando em duas alterações no título, 19 alterações no objetivo geral, 16 alterações na questão de pesquisa, 9 alterações no referencial teórico e 10 modificações na redação da metodologia dos projetos. É importante ressaltar que os estudantes que receberam sugestões do plug-in, mas não utilizaram novamente a ferramenta para reavaliação do projeto, não puderam ser contabilizados devido à ausência de registro no *log*. Portanto, a avaliação quantitativa das contribuições pode ser ainda mais significativa.

Após a investigação das interações através dos logs, a adequação das alterações feitas pelos estudantes foram classificadas por um professor com mais de 10 anos de docência em disciplinas relacionadas à Metodologia de Pesquisa. A tabela 1 apresenta as similaridades e discordâncias entre a avaliação do professor e a do sistema, nos critérios definidos e explicitados na metodologia deste estudo. É possível perceber que, majoritariamente, a avaliação do sistema está em consonância com a do professor especialista, com um índice médio de 71,21% de concordância. Também é possível identificar que o sistema apresentou um número muito pequeno de discordâncias completas (4, de um total de 132 itens avaliados, ou 3,03%).

Tabela 1 — Níveis de concordância.

Elemento textual	Concordância		Concordância parcial		Discordância		Total
Objetivo	27	65,85%	13	31,71%	1	2,44%	41
Questão de pesquisa	32	76,19%	10	23,81%	0	0,00%	42
Referencial teórico	16	64,00%	6	24,00%	3	12%	25
Metodologia	19	79,17%	5	20,83%	0	0,00%	24
Total	94	71,21%	34	25,76%	4	3,03%	

A concordância total ocorreu nas situações onde professor e sistema sinalizaram problemas no texto, ou ambos sinalizaram que não havia nenhuma necessidade de alteração. A concordância parcial ocorreu nos casos em que o sistema não identificou problemas no elemento textual, mas o professor localizou alguma, de menor importância, fragilidade na elaboração do projeto. A discordância, por sua vez, ocorreu quando o sistema não identificou problemas na escrita, mas o professor registrou problemas relevantes. Não houve registros de discordância de falso negativo pelo sistema, indicando problemas onde não existiam, em comparação com a avaliação do professor.

Em relação a comparação da análise do projeto do estudante pelo plug-in e o professor, também foi possível perceber que a ferramenta identificou detalhes na escrita com critérios e nível de detalhes que seria difícil perceber, principalmente quando se tem um grande volume de textos para serem avaliados ou revisados. O trecho abaixo foi verbalizado pelo professor especialista:

“Na descrição dos procedimentos metodológicos deste projeto, eu não havia percebido que a caracterização do objetivo não tinha sido declarada. A definição de um estudo de casos múltiplos, com abordagem qualitativa já me



parecia ok. Mas a classificação do objetivo mais adiante como descritivo melhorou ainda mais a descrição dos procedimentos metodológicos, ajudou a formalizar mais a ideia. Percebo que, como professor, muitas vezes deixo passar detalhes como este porque acabo tendo que ler muitos trabalhos, uma turma tem muitas vezes 20 ou mais projetos para serem revisados. Não uso um checklist, então esses detalhes acabam passando. Já o computador tem uma lista de critérios que são buscados. Não deixa passar esse tipo de detalhe”.

6. Considerações Finais

Este trabalho apresentou uma ferramenta computacional com o propósito de auxiliar os estudantes no aprimoramento de suas habilidades de escrita acadêmica. A investigação demonstrou que o uso de uma ferramenta apoiada por técnicas de mineração de texto pode oferecer um suporte significativo aos alunos durante o processo de escrita de projetos de pesquisa. Essa abordagem não apenas facilita a revisão textual, mas também promove a autonomia e a autoeficácia dos estudantes, permitindo que eles recebam *feedbacks* detalhados e construtivos, essenciais para o desenvolvimento de uma escrita mais coesa e consistente.

A avaliação do uso da ferramenta pelos estudantes, corroborada por um professor especialista, foi bastante positiva. Observou-se um baixo nível de discordância entre as avaliações do professor e as avaliações da ferramenta. Além disso, na análise automatizada da escrita dos alunos, a ferramenta demonstrou a capacidade de identificar problemas com um nível de detalhamento e critérios que seriam difíceis de ser seguidos para um professor, considerando a carga de trabalho e o número de textos que precisam ser avaliados por ele.

A escrita tem se revelado uma tarefa desafiadora tanto para professores quanto para estudantes, exigindo não apenas conhecimento linguístico, mas também a ativação de processos cognitivos complexos. A ferramenta apresentada é proposta como um complemento à orientação docente, permitindo que os professores se atenham a aspectos mais críticos da escrita e específicos da orientação acadêmica. Como sugestão para trabalhos futuros, propõe-se a implementação de um sistema de feedback personalizado e adaptativo, utilizando algoritmos de aprendizado de máquina para fornecer recomendações detalhadas e específicas com base no desempenho individual de cada aluno ao longo do tempo.

Referências

- ARAÚJO, J.; DIEB, M.; COSTA, S. M.. O QNP e as dificuldades de construção do objeto de pesquisa: uma experiência de aprendizagem mediada sobre o gênero projeto de pesquisa. **DELTA: Documentação de Estudos em Linguística Teórica e Aplicada**, v. 33, n. 3, p. 729–757, jul. 2017.
- ARAÚJO, J.; PIMENTA, A. A.; COSTA, S.. A proposta de um quadro norteador de pesquisa como exercício de construção do objeto de estudo. **Interações (Campo Grande)**, v. 16, n. 1, p. 175–188, jan. 2015.
- ASSIS, J. A. Ações do professor e do universitário nas práticas de ensino e de aprendizagem da escrita acadêmica: o papel da avaliação e da reescrita no processo de apropriação do gênero resenha. **Eutomia**, v. 1, n. 13, p. 543–561, 2014.
- ALLAHYARI, M. et al. **Text Summarization Techniques: A Brief Survey**. Disponível em: <<https://arxiv.org/abs/1707.02268>>.



FERREIRA-MELLO, R. et al. Text mining in education. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, v. 9, n. 6, 4 ago. 2019.

FELDMAN, R.; SANGER, J. **The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data**. Cambridge: Cambridge University Press, p. 94-130, 2006.

FELTRIM, V. D.. **Uma abordagem baseada em corpus e em sistemas de crítica para a construção de ambientes Web de auxílio à escrita acadêmica em português**. 2004. Tese (Doutorado em Ciências de Computação e Matemática Computacional) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2004.

GONZÁLEZ-LÓPEZ, Samuel et al. Coverage, Opinion and Speculation: Key Features Analyzed by TURET 2.0. **Advances in Intelligent Learning Environments**, p. 27, 2017.

HAYES, J. R.; FLOWER, L. S. Identifying the Organization of Writing Processes. **Cognitive Processes in Writing**, p. 5–30, 1980.

LÓPEZ, S. G.; BETHARD, S.; LÓPEZ-LÓPEZ, A. Identifying Weak Sentences in student drafts: A tutoring system. In: Mascio, T., Gennari, R., Vitorini, P., Vicari, R., de la Prieta, F. (eds) **Methodologies and Intelligent Systems for Technology Enhanced Learning**. Advances in Intelligent Systems and Computing, vol 292. Springer, Cham. 2014.

MIKOLOV, T.; CHEN, K.; CORRADO, G.; DEAN, J. **Efficient estimation of word representations in vector space**. In: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS, 1., 2013, Scottsdale, Arizona. Proceedings [...]. Ed. BENGIO, Y.; LECUN, Y. [S.l.: s.n.], 2013.

OLIVEIRA, S. ; REATEGUI, E.; BARCELLOS, P. S. C. C. ; BIGOLIN, M.; CARNIATO, M. Improving academic writing with a method for text revision supported by text mining. **International Journal of Emerging Technologies in Learning (iJET)**, v. 17, n. 21, p. 150–163, 2022.

PRAÇA, F. S. G. Metodologia da pesquisa científica: Organização estrutural e os desafios para redigir o trabalho de conclusão. **Revista Eletrônica Diálogos Acadêmicos**, [S. l.], v. 08, n. 1, p. 72–87, 2015.

REATEGUI, E. et al. Sobek: A text mining tool for educational applications. In: **Proceedings of the International Conference on Data Science (ICDATA)**. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), 2011. p. 1.

REATEGUI, E.; KLEMMANN, M.; FINCO, Mateus David. **Using a text mining tool to support text summarization**. In: 2012 IEEE 12th International Conference on Advanced Learning Technologies. IEEE, 2012. p. 607-609.

REATEGUI, Eliseo. **Escrita de uma Dissertação/Tese em Informática na Educação**. In: JAQUES, Patrícia Augustin; PIMENTEL, Mariano; SIQUEIRA, Sean; BITTENCOURT, Ig. (Org.). Metodologia de pesquisa científica em informática na educação: concepção de pesquisa. Porto Alegre: SBC, 2020.

RONG, Xin. **word2vec parameter learning explained**. arXiv preprint arXiv:1411.2738, 2014.