

Título: Análise Linguística de Respostas Abertas em Questionários de Professores com Apoio de IA

João Augusto Ferreira Freire – CAEd/UFJF joao.freire@fundacaocaed.org.br

César Pedrosa Soares – CAEd/UFJF cesar.soares@fundacaocaed.org.br

Fillipe Guedes Soares – CAEd/UFJF fillipe.soares@fundacaocaed.org.br

Rayssa Maluf de Souza – CAEd/UFJF rayssa.souza@fundacaocaed.org.br

Gabriela Paschoal da Silva – CAEd/UFJF gabriela.silva@fundacaocaed.org.br

Resumo

Este estudo, desenvolvido pelo Centro de Políticas Públicas e Avaliação da Educação (CAEd), aborda metodologias de Processamento de Linguagem Natural (PLN) para analisar respostas de questionários aplicados a professores, com foco nas reflexões sobre suas práticas pedagógicas. Foi ajustado um modelo BERT em português para classificação das respostas, que foi capaz de distinguir quatro níveis de reflexão docente com êxito. A distribuição das classes indica que a maior parte dos professores formula respostas que revelam algum grau de intenção pedagógica (níveis 1 e 2), mas ainda com predomínio de formulações superficiais. A frequência de n-gramas revelou a ampla utilização de material concreto como apoio e de sala de aula invertida. Esses resultados oferecem ao CAEd uma visão integrada das estratégias pedagógicas e das interpretações docentes, fornecendo subsídios para futuras iniciativas de avaliação e aprimoramento da prática docente.

Palavras-chave: Processamento de Linguagem Natural; Educação; Análise de texto; Reflexão docente.

Introdução

A expansão de avaliações educacionais que coletam respostas abertas de professores gera dados qualitativos que, quando analisados, podem revelar aspectos importantes sobre a prática docente. Contudo, o processamento manual destes dados é demorado e, muitas vezes, por conta do volume de informação, inviável (Schuh et al., 2016). Este estudo propõe um modelo de Processamento de Linguagem Natural (PLN) orientado por IA capaz de classificar o nível de reflexão pedagógica de respostas de professores coletadas a partir de questionários aplicados nacionalmente pelo Centro de Políticas Públicas e Avaliação da Educação (CAEd). Tais questionários solicitaram que o participante nomeasse (um item) e justificasse (outro item) a prática pedagógica priorizada em sala de aula para tratar conteúdos complexos, fornecendo indícios do

grau de reflexão sobre a prática docente. Nesse sentido, o modelo de classificação proposto visa filtrar respostas inválidas e medir, em larga escala, a profundidade das reflexões docentes.

Metodologia

O PLN é o campo da computação dedicado a criar métodos que permitam ao computador interpretar, compreender e até gerar linguagem natural de forma semelhante à humana. Entre suas aplicações estão a análise automatizada de sentimentos, a criação de *chatbots* e o desenvolvimento de Grandes Modelos de Linguagem (*Large Language Models – LLMs*), como *ChatGPT*, *Llama-3* e *DeepSeek* (Hadi et al., 2023).

O fluxo padrão do PLN inicia-se com o pré-processamento dos textos, etapa importante para organizar o corpus antes de análises mais detalhadas. As principais operações realizadas incluíram: normalização (conversão para minúsculas), remoção de pontuação, símbolos e *stop-words*, tokenização e lematização, que reduz palavras às suas raízes semânticas. Esses passos são importantes, pois dados de texto podem conter erros ortográficos, variações de caixa e formatação heterogênea (Jakhotiya et al., 2022).

Após o pré-processamento, procede-se à vetorização, transformando o texto em representações numéricas manipuláveis por algoritmos. Métodos tradicionais incluem *Bag-of-Words* e *Term Frequency-Inverse Document Frequency (TF-IDF)*. Já os *embeddings* distribucionais capturam relações semânticas, representando palavras em espaços vetoriais de alta dimensão (Pilehvar e Camacho-Collados, 2020). No presente estudo, foi adotado um modelo de *embedding* pré-treinado (Devlin et. al, 2019) em português, apelidado BERTimbau (Souza et. al, 2020), com o intuito de processar as respostas dos questionários de professores.

Classificação das Respostas dos Professores

O questionário aplicado aos professores solicitou que o docente indicasse e justificasse a prática pedagógica priorizada em sala de aula, em contextos de conteúdo complexo e de difícil aprendizado. Estas respostas foram utilizadas para estruturar um modelo de classificação capaz de identificar o nível de reflexão pedagógica dos professores, com base em chaves de correção (apresentadas na seção de resultados) que refletem o nível de reflexão do docente com relação às suas práticas pedagógicas.

Para tanto, após o processamento das respostas mencionado na subseção anterior, a metodologia abrangeu: (i) classificação inicial a partir de um LLM (Llama-3-70B), baseada em chaves de correção determinadas com base no nível de reflexão do respondente acerca da prática

apresentada; (ii) verificação manual de uma amostra aleatória (n = 50 casos), (iii) análise de desbalanceamento, (iv) geração de respostas sintéticas para balancear as classes, (v) *fine-tuning* de do BERTimbau, (vi) teste em um subconjunto de respostas reais, e (vii) ciclos iterativos de refinamento até atingir o melhor desempenho.

Resultados e Discussão

O modelo final apresentou um alto nível de acerto (>90%), validado qualitativamente pela equipe do CAEd. Em termos de distribuição, considerando as 62.278 respostas avaliadas de professores atuando em diferentes escolas públicas brasileiras, observou-se o seguinte panorama das respostas docentes em relação ao nível de reflexão pedagógica, em uma escala de 0 a 3:

Nível de Reflexão	Chave de correção	N	%
0	Resposta inválida	3001	4,8%
1	Justifica e apresenta relação entre a prática e os resultados esperados, mas de forma vaga	36376	58,4%
2	Justifica e apresenta argumentação entre a relação da prática pedagógica, as características dos conteúdos complexos e os resultados esperados	18420	29,6%
3	Justifica e apresenta argumentação entre a relação da prática pedagógica, as características dos conteúdos complexos e os resultados esperados e apresenta situações de impacto da prática pedagógica na aprendizagem dos estudantes	4481	7,2%

Fonte: CAEd/UFJF (2025)

O Quadro abaixo exhibe exemplos de respostas referentes a cada um destes níveis.

Quadro 1: Exemplo de respostas dos questionários para cada nível

9	Exemplo
0	Vídeos e aulas práticas. Pela didática
1	Vídeos e aulas extraclasse quando possível. Favorece o aprendizado.
2	Vídeos e atividades lúdicas para completar o entendimento da matéria. Aplicando intervenções quando necessário. Vídeos e atividades lúdicas são importantes para o processo de ensino aprendizagem, pois contribuem para o melhor desenvolvimento cognitivo, melhorando a atenção, a memorização e o aprendizado do aluno.
3	Utilizo: Aprendizagem ativa baseada em problemas: O método de atividades com problemas do mundo real, incentivando os estudantes a resolverem esses desafios em grupo. Mapas conceituais e gráficos visuais: Ferramentas visuais, como mapas mentais e esquemas são muito úteis para ajudar os estudantes a organizarem e compreenderem conceitos inter-relacionados. Tecnologias educacionais: Plataformas que oferecem simulações interativas e atividades gamificadas ajudam a contextualizar temas difíceis, especialmente em áreas como ciências exatas, onde o visual e o prático tornam o aprendizado mais acessível e motivador. Ensino interativo com feedback contínuo: uso de quiz, testes rápidos e discussões dirigidas para verificar a compreensão dos alunos em tempo real. Considero essas práticas mais adequadas para conteúdos complexos porque elas promovem uma aprendizagem mais profunda e significativa, essencial para a compreensão de temas difíceis.

Fonte: CAEd/UFJF (2025)

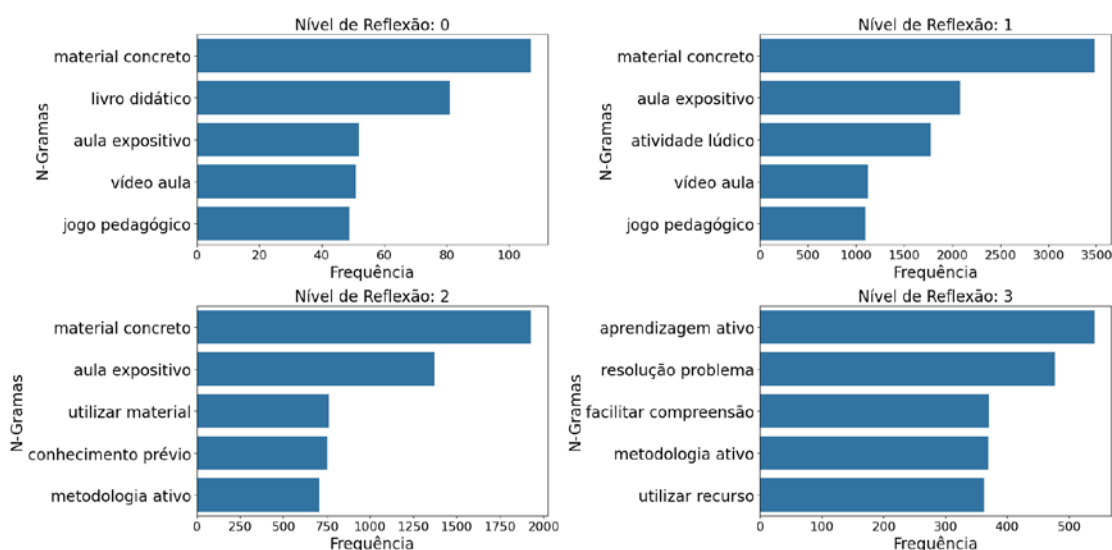
Dados os níveis de reflexão dos professores, é possível mapear quais são as práticas pedagógicas mais utilizadas num contexto de temas complexos e difícil aprendizagem, por nível. Para fazer esse mapeamento, escolhemos uma análise frequentista com “bigramas” e “trigramas”. Os n-gramas são sequências de “n” palavras que aparecem consecutivamente em um texto. Quando falamos de bigramas, estamos nos referindo a pares de palavras (n=2), e os trigramas são trios de

palavras (n=3). Essas sequências são particularmente úteis porque capturam o contexto e as relações entre as palavras, algo que os unigramas (palavras únicas) não conseguem fazer.

Vale ressaltar que nós removemos palavras e frases muito frequentes que não trazem sentido (como por exemplo “trabalhar conteúdos complexos”, que está no título da pergunta e é bastante repetida pelos respondentes. Lembrando também que, dado o processo de lematização, as palavras aparecem como sua raiz semântica, sem flexão de tempo verbal, número ou gênero.

Aqui, a ideia é observar como as práticas pedagógicas preferidas pelos professores, num contexto de conteúdos complexos e de difícil aprendizagem, varia de acordo com seu nível de reflexão acerca de suas práticas pedagógicas. Ou seja: nós calculamos bigramas e trigramas para cada nível de reflexão das práticas pedagógicas, com base no modelo descrito anteriormente. Dessa forma, conseguimos mapear as práticas preferidas por cada grupo de professores, apresentadas nos gráficos a seguir.

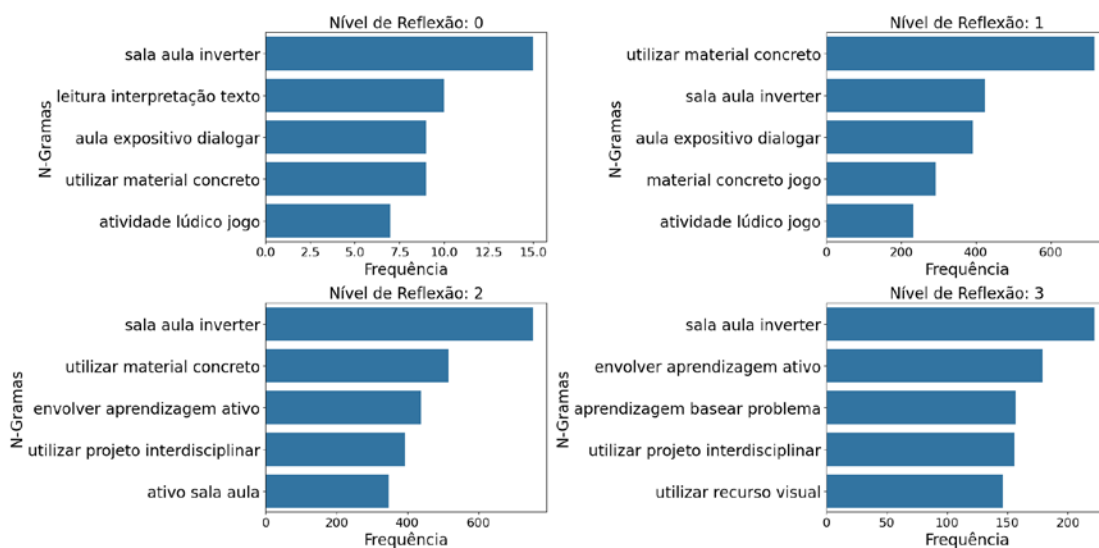
Gráfico 1 – Cinco bigramas mais frequentes na resposta ao primeiro item



Fonte: CAEd/UFJF (2025)

No caso dos bigramas, constata-se a presença expressiva do termo “material concreto”, sendo o mais citado nas respostas classificadas nos níveis 0, 1 e 2. Esse termo se refere objetos físicos e manipuláveis que ajudam os alunos a compreender conceitos abstratos por meio da experiência prática. A presença de “aula expositivo” nos três primeiros níveis reforça a ausência de práticas pedagógicas engajadoras específicas no vocabulário mapeado para esses níveis. No nível de reflexão mais elevado, os bigramas abrangem tipos de práticas ativas, com “resolução problema” sendo a prática específica mais citada.

Gráfico 2 – Cinco trigramas mais frequentes na resposta ao primeiro item



Fonte: CAEd/UFJF (2025)

Para os trigramas, “sala aula inverter” (representando sala de aula invertida) aparece como a prática mais frequente nos níveis 0, 2 e 3, com “utilizar material concreto” aparecendo também nos três primeiros níveis. Um trigrama interessante é “utilizar projeto interdisciplinar”, o 4º mais citado nos níveis 2 e 3, sendo uma prática mais dinâmica. No nível 3, “aprendizagem basear problema” conversa bem com “resolução problema”, que aparece nos bigramas, enquanto “utilizar recurso visual” representa uma prática específica para trazer a atenção dos estudantes.

Discussão e Conclusão

A distribuição das classes indica que a maior parte dos professores apresentam algum grau de intenção pedagógica (níveis 1 e 2), mas ainda com predomínio de formulações superficiais. A baixa incidência de respostas nos níveis 3 aponta para o desafio de promover práticas reflexivas mais elaboradas. O cenário traçado reforça a importância do modelo de classificação como instrumento de leitura diagnóstica, servindo de insumo para estratégias de formação docente.

A implementação do modelo de linguagem mostrou-se eficaz para interpretar as respostas abertas dos docentes, oferecendo um diagnóstico confiável sobre o nível de suas reflexões pedagógicas. O fluxo de desenvolvimento do modelo de classificação possibilitou a criação de um conjunto de dados robusto para o ajuste fino de um modelo específico para a língua portuguesa. Esse processo, no qual dados reais e sintéticos se complementam, permitiu atenuar desequilíbrios de classe e refinar sucessivamente o desempenho do modelo de classificação. A validação qualitativa conduzida pela equipe do CAEd confirmou que as classificações do modelo se alinham,

de forma consistente, aos critérios pedagógicos estabelecidos, assegurando a credibilidade dos resultados para uso prático.

A análise por n-gramas revelou que a utilização de material concreto é uma das práticas mais utilizadas pelos professores. No entanto, práticas de caráter mais engajador não aparecem com frequência, sendo identificadas apenas nos trigramas para respostas classificadas no maior nível de reflexão.

Como desdobramento, o modelo de classificação estruturado torna-se uma ferramenta estratégica para as políticas de avaliação e desenvolvimento docente. Além disso, por ser modular e baseado em tecnologias amplamente disponíveis, o processo pode ser facilmente replicado em outros questionários, etapas de ensino ou áreas do conhecimento. Assim, a experiência demonstra que soluções de PLN, quando combinadas a ciclos de validação qualitativa, não apenas ampliam a escala de análise, mas também mantêm a qualidade interpretativa indispensável para orientar a melhoria das práticas pedagógicas.

Referências

- DEVLIN, Jacob 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. 2019. p. 4171-4186. Disponível em: <https://arxiv.org/pdf/1810.04805>. Acesso em: 19 maio 2025.
- HADI, Muhammad Usman et al. **Large language models: a comprehensive survey of its applications, challenges, limitations, and future prospects**. 2023. Disponível em: https://www.researchgate.net/publication/372258530_Large_Language_Models_A_Comprehensive_Survey_of_its_Applications_Challenges_Limitations_and_Future_Prospects Acesso em: 19 maio 2025.
- JAKHOTIYA, Aachal et al. Text pre-processing techniques in natural language processing: a review. **International Research Journal of Engineering and Technology (IRJET)**, v. 9, n. 2, p. 878-880, fev. 2022. Disponível em: <https://www.irjet.net/archives/V9/i2/IRJET-V9I2150.pdf>. Acesso em: 19 maio 2025.
- PILEHVAR, Mohammad Taher; CAMACHO-COLLADOS, Jose. **Embeddings in Natural Language Processing: Theory and Advances in Vector Representation of Meaning**. [S.l.]: Synthesis Lectures on Human Language Technologies, [2023?]. Disponível em: https://josecamachocollados.com/book_embNLP_draft.pdf. Acesso em: 19 maio 2025.
- SCHUH, John H. et al. **Assessment in student affairs**. 2. ed. Hoboken: Wiley, 2016. Disponível em: <https://www.wiley.com/en-us/Assessment+in+Student+Affairs%2C+2nd+Edition-p-9781119049609> Acesso em: 19 maio 2025.
- SOUZA, F., Nogueira, R., LOTUFO, R. (2020). BERTimbau: Pretrained BERT Models for Brazilian Portuguese. In: Cerri, R., Prati, R.C. (eds) Intelligent Systems. BRACIS 2020. **Lecture Notes in Computer Science()**, vol 12319. Springer, Cham. https://doi.org/10.1007/978-3-030-61377-8_28 Acesso em: 19 maio 2025.