

Plataforma de Acompanhamento Personalizado da Aprendizagem Aplicando Inteligência Artificial

Carlos S. Portela¹, Maurício R. A. Souza², Rodrigo L. Pereira³, Edson K. K. Yasojima³, Thiago D. Cordeiro⁴, Albino R. B. Junior⁴, Alan P. Silva⁴, Diego D. M. C. Matos⁴, Leonardo B. Marques⁴, Rafael F. L. Mello⁵

¹Faculdade de Sistemas de Informação – Universidade Federal do Pará (UFPA)

²Departamento de Ciência da Computação – Universidade Federal de Lavras (UFLA)

³Universidade Federal Rural da Amazônia (UFRA)

⁴Instituto de Computação / NEES – Universidade Federal de Alagoas (UFAL)

⁵Universidade Federal Rural de Pernambuco (UFRPE)

csp@ufpa.br, mauricio.ronny@ufla.br, {rodrigo.lisboa, koiti.yasojima}@ufra.edu.br, {thiago.cordeiro, albino.junior}@nees.ufal.br, alanpedro@ic.ufal.br, diego.matos@famed.ufal.br, leonardo.marques@cedu.ufal.br, rafael.mello@ufrpe.br

Abstract. *The personalized learning monitoring platform allows the digitization, correction, and pedagogical diagnosis of handwritten texts in Portuguese produced by students from public or private schools. To meet this objective, Artificial Intelligence (AI) algorithms were developed with computer vision techniques for transcribing handwritten texts; and natural language processing techniques for correcting texts in Portuguese. The proposed solution can effectively assist teachers in correcting texts, allowing students to produce texts more frequently.*

Resumo. *A plataforma de acompanhamento personalizado da aprendizagem permite a digitalização, correção e diagnóstico pedagógico de textos manuscritos, em língua portuguesa, produzidos por estudantes de escolas públicas ou privadas. A fim de atender esse objetivo, foram desenvolvidos algoritmos de Inteligência Artificial (IA) com técnicas de visão computacional, para transcrição de textos manuscritos; e técnicas de processamento de linguagem natural, para correção dos textos em língua portuguesa. A solução proposta pode auxiliar efetivamente os professores na correção de textos, permitindo que estudantes produzam textos com maior frequência.*

1. Visão Geral

De acordo com a Política Nacional de Alfabetização (PNA) [Brasil 2019], a produção de escrita consiste na habilidade de escrever palavras e produzir textos. A consolidação da produção escrita de uma criança na alfabetização acontece à medida que ela pratica a atividade cotidianamente. Estima-se que para que esses estudantes melhorem seus textos, é necessário praticar por meio da escrita de 2 ou mais redações por semana [Fernandes 2019]. No entanto, durante a pandemia da COVID-19, aumentou o déficit educacional dos estudantes da educação pública, em particular em Língua Portuguesa.

O estudo *Perda de Aprendizagem na Pandemia* estimou que, no ensino remoto, os estudantes aprenderam, em média, apenas 38% do conteúdo de Língua Portuguesa, em comparação com as aulas presenciais [Agência Senado 2021]. Diante da gravidade da problemática, destaca-se a necessidade de auxiliar tanto os gestores escolares na recuperação da aprendizagem quanto os professores na produção e correção de textos.

A Plataforma de Acompanhamento Personalizado da Aprendizagem (APA) possui várias tecnologias associadas que, quando executadas em conjunto de forma sequencial, permitem a digitalização, correção e diagnóstico pedagógico de textos manuscritos em português. Essas soluções encontram-se disponíveis em <http://plataforma-integrada.nees.ufal.br/>. A arquitetura da solução é organizada em módulos específicos que utilizam Kubernetes, uma plataforma que permite o gerenciamento de cargas de trabalho e serviços distribuídos em contêineres.

Inicialmente, foi criado layout padrão para folha de escrita, a qual utiliza QR Code para permitir a focalização e captura da câmera de um dispositivo móvel. Essa captura é intermediada por um aplicativo mobile que permite reconhecer o QR Code. As imagens capturadas são acessadas por algoritmos de Inteligência Artificial (IA). De maneira geral, o processo é dividido em duas etapas: a extração, que é composta por algoritmos de Visão Computacional; e a análise, que é realizada por algoritmos de Processamento de Linguagem Natural (PLN). Na etapa de extração, é realizada a binarização para melhoria da qualidade da imagem e eliminação de elementos da folha do manuscrito. Então, é realizada a segmentação da imagem em linhas e palavras, com o objetivo de facilitar a anotação do texto em um banco de dados. Após a anotação, ocorre a extração do manuscrito em português utilizando-se da técnica de OCR (*Optical Character Recognition*). Por fim, na etapa de análise ocorre a avaliação dos aspectos linguísticos do idioma português para realizar a correção dos manuscritos e gerar um diagnóstico que identifica plágio, textos incoerentes e/ou desvios de temas. Após o diagnóstico, é possível buscar meios de intervenção na prática educativa a partir de uma comunidade de práticas, um módulo colaborativo da Plataforma APA para elaboração de materiais para recuperação de aprendizagem. Permite vincular os itens dos materiais às habilidades da BNCC e disponibiliza funcionalidade de curadoria para aprovação desse material.

A relevância da solução está relacionada a sua aderência à Política Nacional para Recuperação das Aprendizagens na Educação Básica, instituída pelo Decreto Nº 11.079, de 23 de maio de 2022. Nesse contexto, a solução foi adotada pelas escolas do Programa Brasil na Escola, o que permitiu validar a solução a partir de um projeto piloto e de 2 ciclos de avaliação na rede pública de ensino. Além da relevância no contexto educacional, a solução pode ser expandida para empresas e startups que atuam nas áreas de: i) avaliação de provas escritas de concursos; ii) análise de textos jurídicos; e iii) contexto da saúde, onde são escritos prontuários e receitas médicas de maneira manual. Por fim, existe relação com os Grandes Desafios de Pesquisa em Sistemas de Informação no Brasil 2016-2026 à medida que a solução adota uma visão sistêmica que considera não somente a parte tecnológica, mas também a complexidade social da problemática enfrentada (recuperação da aprendizagem) e atores envolvidos (governo e escolas).

A prova de conceito foi conduzida através de um projeto piloto realizado em Dezembro de 2021 em Brasília: CEF 03 DE PLANALTINA (INEP: 53005961) e CEF 11 DE TAGUATINGA (INEP: 53003500). O projeto piloto serviu para validar tanto o layout da folha de avaliação quanto o aplicativo de digitalização dos textos manuscritos.

O escopo da avaliação de redações do Ciclo 1 contemplou os anos finais do Ensino Fundamental (6º ao 9º ano), com uma taxa de participação de 72% do total de escolas do Programa Brasil na Escola, considerando um total de 915.598 estudantes. A coleta de imagens dos textos, via aplicativo de digitalização, foi realizada em maio de 2022, com aproximadamente 240 mil digitalizações. Já o Ciclo 2 de avaliação ocorreu em julho de 2022, adicionando mais cerca de 96 mil digitalizações na base do projeto.

2. Público-Alvo

Professores da rede pública que conhecem as dificuldades de seus alunos na escrita de textos, mas estão sobrecarregados com as demandas de sala de aula. Adicionalmente, gestores escolares comprometidos com a efetiva recuperação da aprendizagem pós-pandemia. Sendo assim, necessitam de apoio e recursos para realizar essas ações.

3. Minibio dos Autores

Carlos Portela é doutor em Ciência da Computação pela UFPE. Professor da Faculdade de Sistemas de Informação da UFPA e Gerente de Qualidade na Plataforma APA.

Maurício Souza é doutor em Ciência da Computação pela UFMG. Professor do Departamento de Ciência da Computação da UFLA e Engenheiro de Qualidade na APA.

Rodrigo Lisbôa é doutor em Computação Aplicada pela UFPA. Professor da UFRA, Campus Paragominas e Vice-Coordenador do Projeto da Plataforma APA.

Edson Yasojima é doutor em Computação Aplicada pela UFPA. Professor da UFRA, Campus Capitão Poço e Desenvolvedor Full-Stack na Plataforma APA.

Thiago Cordeiro é doutor em Engenharia Elétrica pela UFCG. Professor do Instituto de Computação da UFAL e Coordenador do Projeto da Plataforma APA.

Albino Junior é mestrando do Programa de Pós-Graduação em Informática (PPGI) da UFAL. Atua como Gerente de Banco de Dados da Plataforma APA.

Alan Silva é doutor em Ciência da Computação pela UFCG. Professor do Instituto de Computação da UFAL e Diretor Geral do NEES.

Diego Matos é doutor em Ciência da Computação pela UFCG. Professor do PPGI da UFAL e Vice-Diretor Geral do NEES.

Leonardo Marques é doutor em Psicologia pela UFSCAR. Professor do Centro de Educação da UFAL e Membro do Conselho do NEES.

Rafael Mello é doutor em Ciência da Computação pela UFPE. Professor do PPGIA da UFRPE e Membro do Conselho do NEES.

4. Referências

Agência Senado (2019) “Pandemia acentua deficit educacional e exige ações do poder público”, <https://www12.senado.leg.br/noticias/infomaterias/2021/07/pandemia-acentua-deficit-educacional-e-exige-acoes-do-poder-publico>, Fevereiro.

Brasil (2019) “Política Nacional de Alfabetização (PNA)”, http://portal.mec.gov.br/images/banners/caderno_pna_final.pdf, Janeiro.

Fernandes, L. (2019). A Escrita na Prova de Redação do ENEM: Um Olhar sobre a Prática Docente. In *Simpósio Internacional de Ensino da Língua Portuguesa*.