ISSN: 2358-8829

Análise automatizada de padrões de erro em Ciências com apoio de Inteligência Artificial a partir de dados educacionais da rede pública

Laisa Daiane da Cruz Sales [1]
Leonardo Baltazar Cantanhede [2]
Severina coelho da silva Cantanhede [3]

INTRODUÇÃO

A investigação de dados educacionais em larga escala tem se consolidado como uma relevante estratégia para a compreensão dos processos de aprendizagem e das dificuldades enfrentadas pelos estudantes ao longo de sua trajetória escolar (SOARES, 2011; INEP, 2020). Na perspectiva do Ensino de Ciências, essa análise torna-se ainda mais pertinente, considerando a complexidade dos conceitos científicos e a necessidade de desenvolvimento de habilidades relacionadas à interpretação, argumentação e ao raciocínio científico (DELIZOICOV; ANGOTTI; PERNAMBUCO, 2011; DRIVER et al., 1994).

Avaliações externas aplicadas em redes públicas de ensino apresentam um conjunto significativo de informações que evidenciam o nível de desempenho dos estudantes em distintas habilidades. Contudo, tais dados são frequentemente utilizados de forma predominantemente descritiva, sem análises mais aprofundadas (SOARES, 2011) que permitam compreender os erros cometidos e os fatores que contribuem, direta ou indiretamente, para essas dificuldades. Nessa perspectiva, a utilização de ferramentas de Inteligência Artificial configura-se como uma possibilidade promissora para o tratamento e análise desses dados, viabilizando a identificação de padrões de erro, relações entre respostas e tendências de aprendizagem de forma mais sistematizada.

Graduando do Curso de Licenciatura em Pedagogia da Universidade Federal - UFMA, laisa.sales@discente.ufma.br;
2Doutor em Ciências, área de concentração Química Inorgânica, pela Universidade Federal de São Carlos– UFSCar, leonardo.cantanhede@ifma.edu.br;
3Professora orientadora: Doutora pelo Programa de Doutorado em Educação em Ciências e Matemática PPGECEM/REAMEC, severina.cantanhede@ufma.br.

ISSN: 2358-8829

Desse modo, este plano de trabalho propõe a análise automatizada de padrões de erro em Ciências, a partir de dados provenientes de avaliações aplicadas na rede municipal de ensino,

buscando não apenas identificar os erros cometidos, mas também relacioná-los às habilidades não alcançadas e compreender as possíveis causas associadas às dificuldades dos estudantes.

Objetivo Geral

Analisar padrões de erro em Ciências, com suporte de Inteligência Artificial, a partir de dados de avaliações aplicadas na rede municipal de ensino, estabelecendo relações com às habilidades não atingidas e às possíveis causas associadas as dificuldades de aprendizagem.

Objetivos Específicos

Organizar e analisar os dados provenientes das avaliações de Ciências da rede municipal de ensino público;

Identificar as habilidades não atingidas pelos estudantes com base nas matrizes de referência das avaliações;

Classificar e analisar os tipos de erro presentes nas respostas dos estudantes;

Utilizar técnicas de análise automatizada com suporte da Inteligência Artificial para identificar padrões de erro e tendências de aprendizagem;

Verificar as possíveis causas dos erros, considerando aspectos conceituais, linguísticas e interpretativas, com base em categorias analíticas previamente estabelecidas.

A presente pesquisa terá abordagem de caráter qualitativo, com suporte na análise quantitativa de dados educacionais. Serão utilizados dados provenientes de avaliações aplicadas na rede municipal pública de ensino, disponibilizados pela Secretaria Municipal de Educação, Ciência e Tecnologia do Maranhão (SEMECTI), respeitando os princípios éticos e garantindo o anonimato dos participantes.

ISSN: 2358-8829

Inicialmente, os dados serão estruturados para a análise, considerando apenas as questões relacionadas à Ciência e as respectivas habilidades avaliadas. Posteriormente, a identificação dos erros cometidos pelos estudantes será realizada (BROUSSEAU, 2008), por meio da análise das alternativas selecionadas e, quando possível, das justificativas apresentadas. A pesquisa busca, ainda, estabelecer relações entre os erros identificados e as habilidades não atingidas, bem como compreender as possíveis causas associadas às dificuldades de aprendizagem, tais como dificuldades conceituais, problemas de interpretação dos enunciados, limitações na compreensão da linguagem científica e dificuldades no raciocínio científico (MORTIMER, 2000; DRIVER et al., 1994).

Quanto à análise automatizada, serão utilizadas ferramentas computacionais de suporte, incluindo recursos de Inteligência Artificial e o software IRaMuTeQ, empregado na análise estatística de dados textuais (RATINAUD, 2009; CAMARGO; JUSTO, 2013). Para essa análise, serão utilizadas técnicas como a Classificação Hierárquica Descendente (CHD), a Análise de Similitude e a Análise Fatorial de Correspondência (AFC), com o objetivo de identificar

padrões lexicais, relações entre termos e estruturas discursivas presentes nas respostas dos estudantes.

Os resultados serão interpretados com base em categorias analíticas de erros previamente estabelecidas, de modo a permitir a articulação entre análise automatizada e interpretação qualitativa, bem como a identificação de padrões recorrentes de erro e de dificuldades de aprendizagem. O processo de operacionalização da pesquisa será desenvolvido por meio de uma sequência de etapas independentes, conforme descrito a seguir:

1ª Etapa – Organização e tratamento dos dados

Organização e sistematização dos dados provenientes das avaliações de Ciências.

2ª Etapa – Identificação das habilidades avaliadas

Análise das questões com base nas matrizes de referência, identificando as habilidades envolvidas

3ª Etapa – Classificação dos erros

Identificação e classificação dos erros cometidos pelos estudantes, com base em categorias analíticas definidas.

4ª Etapa – Análise automatizada dos dados

Aplicação de ferramentas computacionais de análise textual, com uso do software Iramuteq.

5ª Etapa – Análise dos padrões de erro

Identificação de padrões recorrentes de erro e tendências de aprendizagem.

6ª Etapa – Sistematização e interpretação dos resultados

Integração dos resultados da análise automatizada com a interpretação qualitativa.

7ª Etapa – Produção científica

Sistematização dos resultados e elaboração de relatórios parciais e final.

4. Resultados Esperados

- Identificação de padrões de erro em Ciências no nível Fundamental da Educação Básica;
- Relação entre os erros cometidos e as habilidades não atingidas pelos estudantes;
- Identificação de possíveis causas das dificuldades dos estudantes;
- Construção de base de dados estruturada para análise educacional;
- Proposta de recursos e auxílios para o desenvolvimento de intervenções pedagógicas;

Contribuições para a melhoria do ensino de Ciências na rede municipal de ensino público

5. Cronograma de Execução das Atividades

As etapas que serão desenvolvidas a partir deste cronograma de atividades, previstas para 12 meses, poderão ocorrer de forma parcialmente concomitante, de acordo com o andamento da pesquisa

.

Etapa/Atividade	1º Set	2º Out	3º Nov	4º Dez	5º Jan	6º Fev	7º Mar	8º Abr	9º Mai	10º Jun	11º Jul	12º Ago
1. Organização e tratamento dos dados – Organização e sistematização dos dados provenientes das avaliações de Ciências.	X	X										

2. Identificação das habilidades avaliadas – Análise das questões com base nas matrizes de referência, identificando as habilidades envolvidas.		X	X									
3. Classificação dos erros – Identificação e classificação dos erros cometidos pelos estudantes, com base em categorias analíticas definidas.			X	X	X							
4. Análise automatizada dos dados – Aplicação de ferramentas computacionais de análise textual - software Iramuteq e IA.					X	X	X	X				
5. Análise dos padrões de erro – Identificação de padrões recorrentes de erro e tendências de aprendizagem.								X	X	X		
6. Sistematização e interpretação dos resultados – Integração dos resultados da análise automatizada com a interpretação qualitativa.										X	X	
7. Produção científica – Sistematização dos resultados e elaboração de relatórios parciais e final.											X	X

ISSN: 2358-8829

Referências Bibliográficas

BROUSSEAU, Guy. *Teoria das situações didáticas*. Buenos Aires: Libros del Zorzal, 2008.

CAMARGO, Brigido Vizeu; JUSTO, Ana Maria. IRaMuTeQ: um software gratuito para análise de dados textuais. *Temas em Psicologia*, Ribeirão Preto, v. 21, n. 2, p. 513–518, 2013.

DELIZOICOV, Demétrio; ANGOTTI, José André; PERNAMBUCO, Marta Maria. *Ensino de Ciências: fundamentos e métodos*. São Paulo: Cortez, 2011.

DRIVER, Rosalind; ASOKO, Hilary; LEACH, John; MORTIMER, Eduardo Fleury; SCOTT, Philip. Constructing scientific knowledge in the classroom. *Educational Researcher*, v. 23, n. 7, p. 5–12, 1994.

INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Relatório SAEB*. Brasília: INEP, diferentes edições.

MORTIMER, Eduardo Fleury. *Linguagem e formação de conceitos no ensino de ciências*. Belo Horizonte: UFMG, 2000.

RATINAUD, Pierre. *IRaMuTeQ: Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires*. 2009. Disponível em: http://www.iramuteq.org. Acesso em: 19/03/2026.

CASTRO, P. A.; SOUSA ALVES, C. O.. Formação Docente e Práticas Pedagógicas Inclusivas. **E-Mosaicos**, V. 7, P. 3-25, 2019.

BAPTISTA, C. R. et al. **Inclusão e escolarização: múltiplas perspectivas**. 2 ed. Porto Alegre: Mediação, 2015.

BRASIL. Conselho Nacional da Educação. Câmera de Educação Básica. Resolução nº 2, de 11 de setembro de 2001. **Diretrizes Nacionais para Educação Especial na Educação Básica**. Diário Oficial da União, Brasília, 14 de setembro de 2001. Seção IE, p. 39-40. Disponível em: http://portal.mec.gov.br/cne/arquivos/pdf/CEB0201.pdf. Acesso em: 06 fev. 2020.