

# SisTIRA: Plataforma Web de Tutoria Inteligente para Avaliação Automática de Respostas Discursivas

Avelar Rodrigues de Sousa<sup>1</sup>, Rafael Torres Anchiêta<sup>2</sup>

<sup>1</sup>Instituto Federal de Ciência, Educação e Tecnologia do Piauí (IFPI) – Campus Picos  
CEP 64650-500 – Picos – PI – Brasil

<sup>2</sup>Instituto Federal de Ciência, Educação e Tecnologia do Maranhão (IFMA) – Campus Caxias  
CEP 65600-310 – Caxias – MA – Brasil

avelarrodrigues89@gmail.com, rafael.torres@ifma.edu.br

**Resumo.** *Este trabalho apresenta o Sistema de Tutoria Inteligente de Respostas Automáticas (SisTIRA), uma plataforma web para avaliação automática de respostas discursivas. Desenvolvido com NestJS, Prisma e integrado ao Google GenAI (modelo Gemini-1.5-flash-8b), o sistema gerencia provas, bancos de questões e usuários, atribuindo notas de 0.0 a 1.0 com feedback breve. Em testes iniciais com vinte questões, o sistema demonstrou alta consistência e pertinência nos resultados, confirmando sua aplicabilidade em avaliações síncronas. O artigo descreve a arquitetura da solução, os testes de validação e perspectivas de aprimoramento.*

## 1. Introdução

A pandemia de COVID-19 evidenciou o papel crítico dos Ambientes Virtuais de Aprendizagem (AVAs) para a continuidade educacional, levando à adoção massiva de plataformas como *Google Classroom*, *Moodle* e *Google Forms* e impactando mais de 1,6 bilhão de estudantes [Bank et al. 2021, Gomes and Pimentel 2021]. Embora eficientes para avaliações objetivas, esses sistemas ainda encontram barreiras na correção automática de respostas discursivas, que exige compreensão semântica e critérios consistentes [Almeida and Moura 2024, Oliveira et al. 2020].

Sistemas Tutores Inteligentes (STIs) oferecem *feedback* adaptativo, mas privilegiam questões objetivas, carecendo de modelos robustos para produções textuais [Nkambou et al. 2010, Murray 2003], limitando a escalabilidade e comprometendo a padronização avaliativa.

Este artigo apresenta o Sistema de Tutoria Inteligente de Respostas Automáticas (SisTIRA)<sup>1</sup>, uma plataforma que integra técnicas de Processamento de Linguagem Natural (PLN) e Aprendizado de Máquina para a avaliação automática de respostas discursivas em português. O sistema atribui notas de 0.0 a 1.0 e um *feedback* breve, favorecendo agilidade e padronização.

O restante deste artigo está organizado da seguinte forma: Seção 2 aborda trabalhos relacionados; Seção 3 descreve a arquitetura e funcionalidades do SisTIRA; Seção 4 apresenta a análise preliminar de desempenho; e Seção 5 traz as considerações finais e perspectivas futuras.

---

<sup>1</sup>Vídeo apresentando a ferramenta: [https://drive.google.com/drive/folders/10Cwh\\_txfEr8gNlMaTCba4sC9oOeBeGu0x?usp=sharing](https://drive.google.com/drive/folders/10Cwh_txfEr8gNlMaTCba4sC9oOeBeGu0x?usp=sharing)

## 2. Trabalhos Relacionados

A correção automática de respostas discursivas em AVAs pode ser dividida em dois principais eixos:

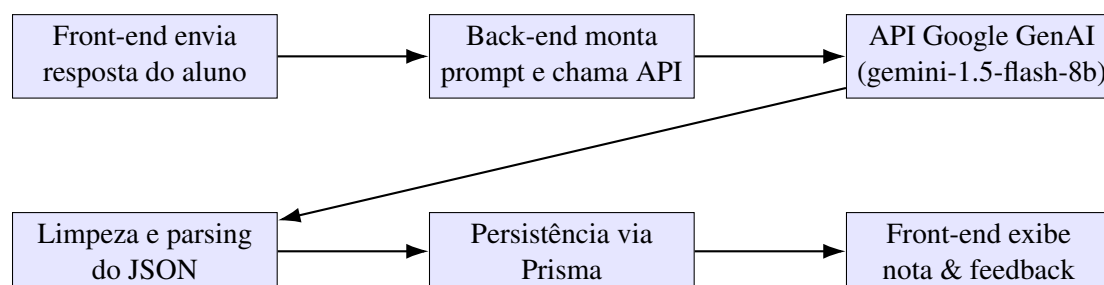
- **Extensão de AVAs por meio de plugins:** Plataformas como o *Moodle* oferecem *plugins* que atribuem notas preliminares a partir de padrões textuais (palavras/frases, expressões-alvo, erros comuns). Apesar de úteis, exigem instalação/configuração e apresentam sensibilidade ao idioma e ao domínio [Bateson nd, Moodle Docs nd].
- **Aplicações de PLN:** Baseiam-se em traços linguísticos e representações semânticas para estimar nota e *feedback*. No português, destacam-se métodos de similaridade como *SpaCy*, *Word Mover's Distance*, *Sentence-Transformers* e *embeddings* do Gemini [Almeida and Moura 2024, Oliveira et al. 2020]. Em inglês, há soluções consolidadas como *e-rater* (ETS) e *IntelliMetric* [Educational Testing Service nd, Attali and Burstein 2006, Rudner and Liang 2002].

Diante de limitações como falta de transparência, portabilidade e robustez ao idioma, este trabalho propõe o SisTIRA, um avaliador voltado ao português brasileiro que combina representações semânticas, alinhamento com gabaritos e geração de *feedback* breve.

## 3. Ferramenta: SisTIRA

O SisTIRA foi implementado como uma Interface de Programação de Aplicações (API) no estilo Transferência de Estado Representacional (RESTful) em `NestJS`, utilizando o Mapeamento Objeto-Relacional (ORM) `Prisma` para acesso ao banco de dados. A solução integra o serviço de PLN Google GenAI para realizar a correção automática de respostas discursivas.

O sistema possui módulos para gerenciar provas, cadastrar e organizar questões em bancos reutilizáveis e de cadastro de usuários, com funcionalidades de registro e autenticação.



**Figura 1. Fluxograma horizontal em duas linhas do processo de correção automática no SisTIRA, mostrando o caminho completo da resposta do aluno até a exibição do feedback.**

Como ilustrado na Figura 1, o fluxo inicia quando o front-end envia a resposta do aluno ao back-end, que gera o *prompt* e aciona a API do Google GenAI. A resposta

recebida é processada em formato JSON, armazenada no banco de dados via Prisma e devolvida ao front-end com a nota e o *feedback*, garantindo um processo de avaliação integrado e automatizado.

Por padrão, o SisTIRA atribui notas em uma escala de 0.0 a 1.0, definida em incrementos de 0.1. Essa escolha simplifica a adaptação para outros formatos, como 0 a 10 (multiplicando por 10) ou porcentagem (multiplicando por 100).

#### 4. Análise Preliminar

Foram realizadas 20 rodadas de avaliação com o modelo `gemini-1.5-flash-8b` sobre um conjunto de 20 questões (níveis básico, médio e difícil), sempre com o mesmo *prompt*. Em 19/20 itens, o desvio-padrão foi 0.00; apenas a questão #16 apresentou 0.05. O *feedback* variou pouco, geralmente entre 1–2 versões por questão, sendo mais diversificado nos casos #9 e #18.

##### Exemplos representativos:

**Básico** — Pergunta: “O que significa CPU em computação?” Resposta do aluno: “CPU significa *Central Processing Unit*, ou Unidade Central de Processamento, responsável por executar instruções e processar dados no computador.” Nota: 1.0 | Feedback: “Resposta completa e precisa, demonstrando total compreensão do conceito.”

**Médio** — Pergunta: “Qual a diferença entre HTTP e HTTPS?” Resposta do aluno: “HTTPS usa criptografia mas funciona de forma semelhante ao HTTP.” Nota: 0.7 | Feedback: “A resposta menciona a criptografia, mas não explica a diferença fundamental de segurança e integridade da comunicação.”

**Difícil** — Pergunta: “O que é o Teorema CAP em bancos de dados distribuídos?” Resposta do aluno: “O Teorema CAP diz que não dá pra ter consistência, disponibilidade e tolerância a partições ao mesmo tempo em sistemas distribuídos.” Nota: 0.9 | Feedback: “Resposta correta e concisa. A menção de que a escolha entre consistência e disponibilidade é necessária em caso de partições reforçaria ainda mais a resposta.”

##### Prompt de correção utilizado:

Questão: <texto da questão>

Respostas de referência:

- ERRADA: <texto da resposta ERRADA>
- MÉDIA: <texto da resposta MÉDIA>
- CORRETA: <texto da resposta CORRETA>

Resposta do aluno:

<texto da resposta do aluno>

Instruções:

1. Atribua uma nota entre 0.0 e 1.0 (em incrementos de 0.1).
2. Forneça um feedback breve (máx. 2 frases) justificando a nota.
3. Retorne apenas um JSON válido (sem texto extra) com as chaves "nota" e "feedback".
4. Use ponto (.) como separador decimal.

**Tabela 1. Métricas de análise preliminar por questão (20 rodadas, 1 rodada/questão).**

ID	Score Médio	Desvio-Padrão	Feedbacks Únicos
1	1.00	0.00	2
2	0.00	0.00	1
3	0.70	0.00	1
4	0.70	0.00	1
5	0.00	0.00	1
6	1.00	0.00	1
7	0.90	0.00	1
8	0.20	0.00	3
9	0.90	0.00	5
10	0.90	0.00	1
11	0.90	0.00	1
12	0.70	0.00	1
13	0.60	0.00	4
14	0.00	0.00	2
15	0.50	0.00	1
16	0.96	0.05	2
17	0.90	0.00	1
18	0.60	0.00	7
19	0.70	0.00	1
20	0.60	0.00	2

Os dados da Tabela 1 confirmam que, em 20 rodadas para 20 questões, as notas médias permaneceram estáveis (com desvio-padrão nulo em 19 itens) e os *feedbacks* exibiram baixa variação textual, corroborando a reprodutibilidade do processo de correção do SisTIRA com o `gemin-1.5-flash-8b`.

## 5. Conclusão

Este trabalho apresentou o SisTIRA, uma plataforma web para avaliação automática de respostas discursivas em português. Desenvolvido com `NestJS` e `Prisma` e integrado ao Google GenAI (modelo `Gemini-1.5-flash-8b`), o sistema cadastra provas, gerencia bancos de questões e atribui notas de 0.0 a 1.0 com *feedback* breve. Os testes indicaram alta consistência nas notas, com desvio-padrão quase nulo e *feedbacks* estáveis (Tabela 1), confirmando a confiabilidade do modelo para correção síncrona.

Como trabalhos futuros, prevê-se a validação com avaliações humanas (Coeficiente *Cohen's Kappa*), o refinamento dos *prompts*, a melhoria da interface, a ampliação das funcionalidades e a análise de desempenho sob carga real.

Em síntese, o SisTIRA se apresenta como solução viável e escalável para automatizar a correção de respostas discursivas, promovendo agilidade, padronização e justiça em AVAs.

## Referências

- Almeida, J. and Moura, R. (2024). Investigação de métodos de similaridade textual no contexto da avaliação automática de questões discursivas. In *Anais da XII Escola Regional de Computação do Ceará, Maranhão e Piauí*, pages 110–118, Parnaíba/PI. SBC.
- Attali, Y. and Burstein, J. (2006). Automated essay scoring with e-rater® v.2. In Litman, D. and Pinker, S., editors, *Proceedings of the 2nd Workshop on Innovative Use of NLP for Building Educational Applications*, pages 9–16, New York, USA. Association for Computational Linguistics.
- Bank, T. W., UNESCO, and UNICEF (2021). The state of the global education crisis: A path to recovery. Technical report, The World Bank, UNESCO, UNICEF, Washington D.C., Paris, New York.
- Bateson, G. (n.d.). Essay (auto-grade) question type. Moodle Plugins Directory. [https://moodle.org/plugins/qtype\\_essayautograde](https://moodle.org/plugins/qtype_essayautograde) (Acesso em: 15 ago. 2025).
- Educational Testing Service (n.d.). e-rater®: Automated writing evaluation engine. ETS Research. <https://www.ets.org/research/topics/erater> (Acesso em: 15 ago. 2025).
- Gomes, A. S. and Pimentel, E. (2021). Ambientes virtuais de aprendizagem para uma educação mediada por tecnologias digitais. In *Informática na Educação: ambientes de aprendizagem, objetos de aprendizagem e empreendedorismo*. Sociedade Brasileira de Computação, Porto Alegre.
- Moodle Docs (n.d.). Essay (auto-grade) question type. Moodle Documentation. [https://docs.moodle.org/en/Essay\\_\(auto-grade\)\\_question\\_type](https://docs.moodle.org/en/Essay_(auto-grade)_question_type) (Acesso em: 15 ago. 2025).
- Murray, T. (2003). An overview of intelligent tutoring system authoring tools: Updated analysis of the state of the art. In Murray, T., Blessing, S. B., and Ainsworth, S., editors, *Authoring Tools for Advanced Technology Learning Environments*, pages 491–544. Springer, Dordrecht.
- Nkambou, R., Bourdeau, J., and Psyché, V. (2010). Building intelligent tutoring systems: An overview. In Nkambou, R., Bourdeau, J., and Mizoguchi, R., editors, *Advances in Intelligent Tutoring Systems*, pages 377–406, Berlin, Heidelberg. Springer.
- Oliveira, D., Pozzebon, E., and Santos, T. (2020). Aplicação de técnicas de processamento de linguagem natural na automatização de correção de questões discursivas. In *Anais do Congresso Brasileiro de Informática na Educação*.
- Rudner, L. M. and Liang, T. (2002). Automated essay scoring using intellimetric™. Technical Report ED465287, ERIC. <https://eric.ed.gov/?id=ED465287>.