

## Corretor Automático de Redações em Língua Portuguesa: um mapeamento sistemático de literatura

Luciana F. da Costa<sup>1</sup>, Elaine Harada T. de Oliveira<sup>1</sup>, Alberto N. de Castro Júnior<sup>1</sup>

<sup>1</sup>Instituto de Computação (IComp)  
Universidade Federal do Amazonas (UFAM) – Manaus, AM – Brasil

{luciana, elaine, alberto}@icompu.ufam.edu.br

**Abstract.** *Language proficiency is a skill that must be worked on throughout a person's life and, in Brazil, ENEM is an exam that evaluates this skill through written proposals. There are many challenges for teachers to work efficiently with this type of textual evaluation in the classroom and automatic assessment is one option to assist this process. In this context, a systematic literature mapping was carried out in order to get an overview of the state of the art on automatic essays evaluation for Portuguese. We have found out that there is much to be explored in the area, such as to deal with many linguistic aspects related to the competences evaluated by ENEM.*

**Resumo.** O domínio da linguagem é uma habilidade que deve ser trabalhada durante toda a vida de um indivíduo, sendo o ENEM um dos exames que avaliam essa habilidade por meio de propostas de redação no Brasil. São muitos os desafios para professores trabalharem com esse tipo de avaliação eficientemente em sala de aula, sendo o uso de correção automática uma opção para auxiliar o processo. Nesse contexto, foi realizado um mapeamento sistemático da literatura a fim de obter um panorama do estado da arte dos corretores automáticos de redações em Língua Portuguesa. Observou-se que há um grande espaço ainda a se explorar na área, como o atendimento a aspectos linguísticos relacionados às várias competências avaliadas pelo ENEM.

### 1. Introdução

O domínio e uso formal da linguagem são habilidades que devem ser trabalhadas durante toda a vida acadêmica e profissional de uma pessoa, pois muitas são as situações em que essas são requeridas, como na inserção no mercado de trabalho ou no ingresso em um curso de ensino superior. No Brasil, no decorrer dos anos escolares, os professores de Português têm o dever de auxiliar no desenvolvimento dessas habilidades, ao qual promovem diversas atividades didáticas de leitura, interpretação e produção textual.

O cenário atual brasileiro denota um professor que tende a assumir mais trabalho que sua capacidade permite, por exemplo, lecionando para muitas turmas em um mesmo período letivo. Em um perfil traçado entre os censos de 2009, 2013 e 2017 [Carvalho 2018], é detectado que a maioria dos professores no contexto do Ensino Médio leciona para 6 turmas ou mais (a porcentagem ultrapassa 70%). Isso compromete o trabalho docente, principalmente quando outras atividades necessitam de tempo exclusivo para sua realização, como planejamento de aulas, atendimento às dúvidas dos estudantes, correção de exercícios, entre outras. Com isso, atividades de produção textual com correção e retorno de periodicidade constante tornam-se praticamente inviáveis, pois dependem de um processo de correção mais profundo e demandam um *feedback* mais

assertivo, muitas vezes incompatível com a situação ao qual estão inseridos.

Dentre a gama de possibilidades que podem auxiliar a questão apresentada, uma abordagem possível que pode facilitar a atividade didática citada é o uso de correção automática de textos (*essay-grading*). Segundo Vicari (2019), *essay-grading* é uma tecnologia considerada um subproduto do Processamento de Linguagem Natural (PLN), voltada à compreensão da língua escrita e que utiliza algoritmos de análise sintática, semântica e pragmática, além de algoritmos estatísticos de PLN baseados em corpora. Esta tecnologia tem bastante expressão em relação à Língua Inglesa, tendo um vasto número de aplicações, inclusive comerciais, que representam a consistência do trabalho neste idioma. Como exemplos, pode-se citar e-Rater, criado pela Educational Testing Service (ETS) e a IntelliMetric, da Vantage Learning [Weigle 2013], que trazem resultados avançados para o processamento do idioma inglês. Para a Língua Portuguesa, o cenário é diferente. Uma pesquisa bibliográfica exploratória, realizada pelos autores anteriormente, indica que são poucos os trabalhos encontrados que tratam dessa questão.

Buscando aprofundar a visão sobre corretores automáticos para o idioma português, foi realizada uma revisão bibliográfica em formato de Mapeamento Sistemático de Literatura (MSL). Este tipo de estudo secundário tem o objetivo de identificar oportunidades de pesquisa no conjunto de trabalhos selecionados e ser um ponto de partida para estudos mais profundos na área pesquisada [Budgen et al. 2008]. Deste modo, pretende-se apresentar um panorama abrangente do estado da arte da área de corretores automáticos para a Língua Portuguesa.

Nas próximas seções, serão apresentados o objetivo e o processo metodológico do MSL, com o protocolo utilizado; dados sobre a condução, os resultados obtidos pelo estudo e as considerações finais, apontando caminhos para a pesquisa na área.

## 2. Processo Metodológico

O objetivo deste estudo foi assim definido: analisar o panorama do estado da arte em relação à área de correção automática de redações em língua portuguesa. Pretendeu-se descrever como estão os estudos de corretores automáticos no idioma português de forma geral, quais as principais estratégias utilizadas nas soluções apresentadas e os aspectos linguísticos avaliados; quais as bases de dados de referência e as métricas mais utilizadas para avaliação dessas soluções.

Para alcançar o objetivo traçado para este mapeamento sistemático, foi montado um protocolo com definições e passos a guiarem o estudo. Os principais elementos serão apresentados nas subseções a seguir.

### 2.1. Questões de Pesquisa

Busca-se respostas para as seguintes questões de pesquisa:

**Tabela 1. Questões de Pesquisa do Mapeamento Sistemático de Literatura.**

<b>Q1:</b> Quais os principais algoritmos, técnicas, métodos ou ferramentas utilizadas?
<b>Q2:</b> Quais aspectos linguísticos são avaliados?
<b>Q3:</b> Quais as bases de dados utilizadas para realizar experimentos e testes?
<b>Q4:</b> Quais as métricas utilizadas para a avaliação dos resultados?

## 2.2. Expressões de Busca e Fontes Digitais de Pesquisa

A expressão de busca foi construída a partir das seguintes palavras-chave:

- Automático(a): Automated (e suas derivações);
- Redação ou Dissertação (dissertativo): Essay(s);
- Correção – Graduação: Scoring, Grading, Evaluation (e suas derivações).

Para a elaboração da expressão de busca, foram utilizados os radicais das palavras-chaves com o intuito de oferecer a oportunidade de retorno de maiores variações dos vocábulos por meio do uso de expressões regulares. Foram construídas duas variações – levando em consideração que a pesquisa tem como foco a Língua Portuguesa e alguns das publicações encontradas anteriormente serem neste idioma. São elas:

- **Inglês:** (automat\*) AND (essay\*) AND (scor\* OR grad\* OR evaluat\*);
- **Português:** (automat\*) AND (reda\* ou dissert\*) AND (corre\* OR avalia\*).

Para esta revisão, foram utilizadas duas fontes digitais de pesquisa:

- **Scopus:** <<https://www.scopus.com/search/form.uri?display=basic>>;
- **IEEE Xplore Digital Library:** <<https://ieeexplore.ieee.org/Xplore/home.jsp>>.

A primeira foi escolhida por indexar e agregar várias fontes digitais (como a ACM, Springer e a própria IEEE); e a segunda foi escolhida para ter mais uma opção de pesquisa e retirar algum possível viés no uso em uma fonte apenas. Além disso, as duas fontes digitais possuem mecanismos de busca que aceitam expressões regulares, facilitando a abrangência do retorno da pesquisa.

## 2.3. Filtros e Critérios de Inclusão e Exclusão

Para selecionar os artigos que fizeram parte do estudo secundário, todas as publicações retornadas passaram por dois filtros de seleção, com critérios baseados nas características esperadas pelos trabalhos do estado da arte da área. O primeiro filtro se ateve à leitura apenas do Título, Resumo (*Abstract*) e Palavras-chaves, enquanto no segundo filtro as publicações selecionadas no primeiro filtro tiveram a leitura do texto integral. Nas tabelas a seguir, são apresentados os critérios de seleção utilizados nos dois filtros:

**Tabela 2. Critérios de Inclusão do 1º Filtro.**

CI-01 - Publicações que descrevam uma solução de correção automática de redações em língua portuguesa.
CI-02 - Publicações que descrevam uma revisão de literatura ou um <i>survey</i> que tenham relação com a correção automática de redações em língua portuguesa.

**Tabela 3. Critérios de Exclusão do 1º Filtro.**

CE-01 – Publicações em que as palavras-chave da expressão de busca não apareçam no título, no <i>abstract</i> ou na lista de palavras-chaves dos autores da publicação.
CE-02 – Publicações em que as palavras-chaves não tenham significado equivalente a: (i) automático(a) ou automatizado(a); (ii) texto dissertativo ou texto discursivo ou redação; ou (iii) avaliação ou correção ou pontuação.

<b>CE-03</b> – Publicações em que o objeto principal da avaliação apresentada (i) não seja textos dissertativos e, sendo textos; (ii) não seja em língua portuguesa; ou (iii) não sejam submetidos a um processo automatizado ou uma avaliação automática.
<b>CE-04</b> – Publicações que não estejam relacionadas às áreas de IA, especialmente: Processamento de Linguagem Natural, Aprendizagem de Máquina, Redes Neurais, entre outros; ou Linguística.
<b>CE-05</b> – Publicações que descrevam e/ou apresentem <i>keynote speeches</i> , livros, textos completos de dissertações e teses, editoriais, artigos de opinião, tutoriais, relatórios, cursos e similares.
<b>CE-06</b> – Publicações que não estejam escritas nos idiomas definidos para leitura nesta pesquisa (inglês e português).
<b>CE-07</b> – Publicações duplicadas.

**Tabela 4. Critérios de Inclusão do 2º Filtro.**

<b>CI-01</b> – Publicações que descrevam solução de correção automática de redações <u>explicitamente</u> em língua portuguesa.
<b>CI-02</b> – Publicações que descrevam uma revisão de literatura ou um <i>survey</i> que tenham relação com a correção automática de redações em língua portuguesa.

**Tabela 5. Critérios de Exclusão 2º Filtro.**

<b>CE-01</b> – Publicações em que é possível identificar que o idioma da solução não é o português (baseado em vocábulos, referências a outras línguas, etc).
<b>CE-02</b> – Publicações em que não é possível identificar para qual idioma a solução está baseada.
<b>CE-03</b> – Publicações em que o objeto principal da avaliação apresentada (i) não seja textos dissertativos e, sendo textos, (ii) não seja explicitamente em língua portuguesa; ou (iii) não sejam submetidos a um processo automatizado ou uma avaliação automática.
<b>CE-04</b> – Publicações não disponíveis para leitura.

Alguns critérios, especialmente de exclusão do 2º Filtro, foram criados para se adequarem à realidade dos trabalhos retornados durante estudos prévios. Identificou-se que uma grande quantidade de publicações não cita diretamente a que língua natural as soluções descritas estão voltadas. Considerando que o português tem características singulares, optou-se por selecionar apenas os trabalhos explicitamente voltados para este idioma. Também foi priorizada a recuperação de artigos publicados em periódicos e eventos científicos, tendo, assim, passado pelo crivo dos revisores de tais fontes.

### 3. Condução e Resultados

Foram executados estudos prévios em que foram identificados vários pontos de ajuste que levaram o protocolo a ser adaptado algumas vezes, a fim de atender a características da área pesquisada que foram conhecidas durante o processo. A busca final nas fontes digitais selecionadas ocorreu na última semana de maio de 2019. Todos os artigos

indexados nesta busca passaram pelo crivo dos dois filtros planejados para este estudo.

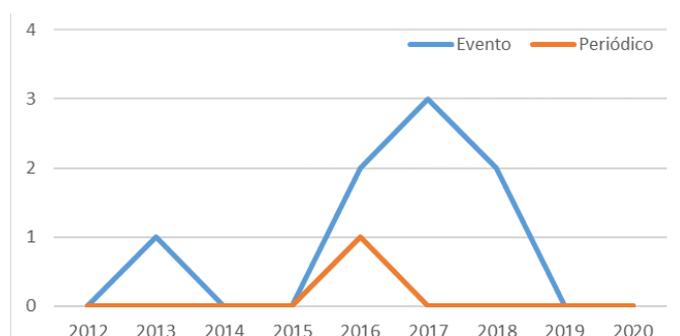
Com isso, o estudo final possui os seguintes dados:

**Tabela 6. Quantidade de Publicações Analisadas.**

SCOPUS	IEEE	TOTAL BUSCA	FILTRO 1	FILTRO 2	INSERÇÃO MANUAL	TOTAL FINAL
651	136	787	293	6	4	<b>10</b>

No gráfico a seguir, observa-se os trabalhos distribuídos no decorrer do tempo. É notável que as publicações estão fortemente presentes em eventos (como simpósios e *workshops*), enquanto temos apenas uma publicação em periódico.

**Gráfico 1. Trabalhos por Anos e Tipo de Publicação.**



A seguir, é feita uma breve apresentação dos trabalhos selecionados e as respostas para as questões de pesquisa levantadas por este estudo secundário.

### 3.1. Trabalhos selecionados

Como já apresentado anteriormente, 10 trabalhos aderiram aos critérios propostos deste MSL. Dos 10, 6 foram selecionados durante a condução do estudo, atendendo adequadamente aos critérios propostos pela pesquisa. Os 4 trabalhos restantes foram inseridos manualmente. Estes foram encontrados em pesquisa bibliográfica exploratória anterior e se adequaram ao estudo levantado, atendendo aos critérios levantados pelo mapeamento. Porém, não são indexados nas fontes digitais utilizadas pelo estudo.

Alguns trabalhos abordam o problema apenas no que tange o domínio da modalidade escrita formal do português. Pode-se citar: um analisador ortográfico-gramatical utilizando PLN e Algoritmos Genéticos [Santos, Paiva e Bittencourt 2016] e um sistema de avaliação automática utilizando PLN e Aprendizagem de Máquina [Júnior e Oliveira 2016], [Júnior, Spalenza e Oliveira 2017]. Um trabalho investiga questões relacionadas à aderência de uma redação a um tema, além de questões como estrutura argumentativa e coesão textual [Filho et al. 2018].

Para a avaliação do texto de maneira geral, temos um estudo experimental com a utilização de um classificador bayesiano [Bazelato e Amorim 2013], um estudo exploratório de um analisador multi-aspectos, adotando várias técnicas de IA [Amorim e Veloso 2017] e um estudo experimental e comparativo entre o uso de engenharia de atributos e *deep learning* [Fonseca et al. 2018].

Por fim, há trabalhos que fixam sua correção automática em relação à similaridade de um texto a outro, de referência, considerado como “padrão” ou “correto”, como [Frinhani et al. 2016], [Freitas et al. 2017] e [Figueira et al. 2013], com bases de dados

extraídas em experimentos em sala de aula. Não foram encontradas revisões de literatura ou *surveys* que atendessem aos critérios levantados para este estudo.

Na Tabela 7, estão listados os 10 trabalhos selecionados:

**Tabela 7. Trabalhos selecionados no Mapeamento Sistemático de Literatura.**

#	Referência	Título
1	[Santos, Paiva e Bittencourt 2016]	Avaliação Léxico-Sintática de Atividades Escritas em Algoritmo Genético e Processamento de Linguagem Natural: Um Experimento no ENEM
2	[Júnior e Oliveira 2016]	Proposta de um Sistema de Avaliação Automática de Redações do ENEM
3	[Júnior, Spalenza e Oliveira 2017]	Proposta de um Sistema de Avaliação Automática de Redações do ENEM Utilizando Técnicas de Aprendizagem de Máquina e Processamento de Linguagem Natural
4	[Filho et al. 2018]	An approach to evaluate adherence to the theme and the argumentative structure of essays
5	[Bazelato e Amorim 2013]	A Bayesian Classifier to Automatic Correction of Portuguese Essays
6	[Amorim e Veloso 2017]	A multi-aspect analysis of automatic essay scoring for Brazilian Portuguese
7	[Fonseca et al. 2018]	Automatically Grading Brazilian Student Essays
8	[Frinhani et al. 2016]	An automatic essay correction for an active learning environment
9	[Freitas et al. 2017]	Evaluation of an Automatic Essay Correction System Used as an Assessment Tool
10	[Figueira et al. 2013]	Module of evaluation automatic essay questions on virtual learning environment LabSQL

Nas subseções a seguir, as questões de pesquisa serão discutidas, identificando as características dos trabalhos selecionados de forma mais aprofundada.

### 3.2. Principais algoritmos, técnicas, métodos ou ferramentas utilizadas (Q1)

A primeira questão de pesquisa pretendia investigar quais as principais estratégias computacionais foram utilizadas para solucionar a correção automática de redações em português. Os tipos mais recorrentes de abordagens para a criação das soluções são: Processamento de Linguagem Natural (9 ocorrências), Aprendizagem de Máquina (6 ocorrências) e outros tipos (5 ocorrências), como o uso de Algoritmos Genéticos.

Algumas das estratégias mais utilizadas relacionadas a PLN são as ferramentas CoGrOO<sup>1</sup>, presente em 5 trabalhos (1, 2, 3, 4 e 6), e Apache OpenNLP<sup>2</sup>, utilizada em 3 trabalhos (1, 2, 3). A partir delas e das demais estratégias PLN encontradas, tem-se predominantemente o tratamento do texto e a identificação de termos.

Em relação a Aprendizagem de Máquina, foi observado o uso de classificação (aprendizagem supervisionada) como a melhor estratégia para detecção de graduação

<sup>1</sup> <http://ccsl.ime.usp.br/cogroo/>

<sup>2</sup> <https://opennlp.apache.org/>

(pontuação) de notas. Destacam-se os modelos classificadores baseados no algoritmo SVM (*Support Vector Machine*) como um dos mais utilizados (2, 3, 4). Além do mais, foi utilizada também regressão em duas abordagens – uma com um regressor SVM (4) e uma abordagem com *deep neural networks* (7). Em relação a outras estratégias utilizadas, é importante citar o uso de Algoritmos Genéticos, presente em 3 trabalhos (1, 8, 9).

É relevante ressaltar que quase todos os trabalhos apresentam arquiteturas híbridas, ou seja, utilizam mais de um tipo de abordagem para a estratégia da solução do problema, sendo apenas um (5) que implementa apenas uma técnica (classificação bayesiana).

### 3.3. Aspectos linguísticos avaliados (Q2)

A segunda questão de pesquisa é voltada aos aspectos linguísticos do português que são avaliados pelos trabalhos selecionados, sendo eles pontos mais específicos da língua ou uma avaliação feita de uma maneira mais abrangente, englobando vários aspectos. Foi observado que 7 dos 10 (1,2,3,4,5,6,7) trabalhos citam como referência para a avaliação de textos as competências descritas pelo Exame Nacional do Ensino Médio (ENEM).

As competências do ENEM são [INEP 2019]:

- Demonstrar domínio da modalidade escrita formal da língua portuguesa (C1);
- Compreender a proposta de redação e aplicar conceitos das várias áreas de conhecimento para desenvolver o tema, dentro dos limites estruturais do texto dissertativo-argumentativo em prosa (C2);
- Selecionar, relacionar, organizar e interpretar informações, fatos, opiniões e argumentos em defesa de um ponto de vista (C3);
- Demonstrar conhecimento dos mecanismos linguísticos necessários para a construção da argumentação (C4);
- Elaborar proposta de intervenção para o problema abordado, respeitando os direitos humanos (C5).

Observou-se um equilíbrio na abordagem de aspectos linguísticos entre os trabalhos: três (1, 2, 3) são voltados apenas ao domínio da norma padrão da Língua Portuguesa, que corresponde à Competência 1 (C1) do ENEM; três trabalhos (5, 6, 7) abrangem todas competências de maneira geral, incluindo a C1; e três trabalhos (8, 9, 10) efetuam a correção em comparação de sentidos (texto “correto” versus texto a ser corrigido). Este último aspecto pode ser relacionado com a aderência ao tema (compreender a proposta de redação), voltado a C2 do ENEM, porém, isto deve ser analisado mais atentamente.

Além disso, apenas um trabalho (4) aborda aspectos que estão relacionados a C2 do ENEM: estrutura argumentativa, coesão textual e diversidade léxica e conectivos (limites estruturais do texto dissertativo-argumentativo em prosa); e aderência ao tema.

### 3.4. Bases de dados para experimentos e testes (Q3)

Em relação à terceira questão de pesquisa, procurou-se identificar as bases de dados utilizadas para experimentos e testes de correções de redações em Língua Portuguesa, tendo em consideração a lacuna de bases de dados em outras línguas que não a língua inglesa, observada em [Passero et al. 2017].

Foi detectado que mais da metade dos trabalhos utilizou o Banco de redações do portal UOL<sup>3</sup> como fonte de dados – 6 das 10 publicações analisadas (1, 2, 3, 4, 5 e 6). Este disponibiliza correções gratuitas feitas por especialistas e ficam disponíveis ao público. Outro portal com as mesmas características, utilizado por um trabalho (4), foi o Banco de Redações do portal Brasil Escola<sup>4</sup>. Um estudo (7) utilizou uma base privada de redações. A forma de classificação de cada redação destas bases é por nota, seguindo a classificação de notas por competência do ENEM. As notas são dadas por especialistas em língua portuguesa. Os demais trabalhos analisados construíram a própria base a partir dos experimentos e testes das soluções propostas pelos próprios.

Em relação à quantidade de redações disponíveis, bases de acesso mais restrito apontam um número consideravelmente maior de textos que as bases de acesso público. A maior base contém 56644 redações (7), seguido de 17000 questões discursivas (10). A maior base de conteúdo público contém 4547 redações (3), enquanto os demais trabalhos contam com menos de 2000 redações cada.

### 3.5. Métricas utilizadas na avaliação dos resultados (Q4)

A última questão de pesquisa (Q4) refere-se às métricas mais utilizadas para medição dos experimentos e testes. Estas métricas estão descritas na Tabela 8:

**Tabela 8. Métricas mais utilizadas na avaliação de resultados.**

Métrica	Quantidade	Métrica	Quantidade
<i>Pearson's correlation</i>	4 (2, 4, 5, 8)	<i>accuracy</i>	4 (2, 4, 5, 10)
<i>precision</i>	4 (1, 2, 3, 4)	<i>quadratic weighted kappa</i>	2 (6, 7)
<i>recall</i>	4 (1, 2, 3, 4)	<i>mean absolute error</i>	2 (2, 3)

Pode-se observar que as métricas com maior número de utilizações são *Pearson's correlation*, *precision*, *recall* e *accuracy*. A métrica *Pearson's correlation* é principalmente usada para avaliar a relação entre os resultados da correção de uma solução e os dados de correção de um especialista ou de um texto considerado como correto. *Precision* é utilizada para indicar o quanto dos resultados apontados como corretos pela solução também são considerados corretos pelos especialistas. Por sua vez, *recall* indica o quanto, dentro dos resultados considerados corretos pelos especialistas, a solução também apontou como correto. Por fim, *accuracy* é usada para avaliar a corretude dos resultados da solução de maneira geral. Ainda podemos citar *mean absolute error*, que verifica a média das diferenças entre as correções dos especialistas e as dadas por uma solução. Uma outra métrica bastante discutida, *quadratic weighted kappa* (QWK), dada como predominante na avaliação de soluções para outras línguas [Zupanc e Bosnic 2015], é apenas utilizada em dois dos trabalhos selecionados.

## 4. Considerações Finais

Este artigo apresentou um Mapeamento Sistemático de Literatura com o objetivo de obter um panorama do estado da arte dos corretores automáticos para Língua Portuguesa. Tal estudo denota que há um grande espaço a explorar nesta área para o citado idioma. Foram selecionados 10 trabalhos que apresentam bons resultados, utilizando, principalmente, arquiteturas híbridas para a solução automática (ferramentas de PLN e modelos de

<sup>3</sup> <https://educacao.uol.com.br/bancoderedacoes/>

<sup>4</sup> <https://vestibular.brasescola.uol.com.br/banco-de-redacoes>

aprendizagem supervisionada em conjunto, por exemplo). Porém, mais da metade destes trabalhos são ainda focados em poucos aspectos da linguagem, centrados na abordagem de aspectos linguísticos relacionados a apenas duas competências do ENEM (C1 e C2) – usadas como referência de avaliação pela maioria dos trabalhos. É notável também a escassez de bases com uma expressiva quantidade de dados para experimentação e testes, sendo esta uma limitação que deve ser observada para futuras pesquisas na área.

Um caminho a ser explorado é aprofundar o estudo em corretores automáticos de redações que englobem, por exemplo, mais competências requeridas no ENEM, conseqüentemente, mais aspectos linguísticos, como, por exemplo, os relacionados à semântica. Os aspectos semânticos podem ser considerados os mais desafiadores em relação aos corretores automáticos no geral, pois, mesmo em outras línguas, ainda é um campo a ser melhor explorado [Zupanc e Bosnic 2015].

Outro caminho possível é a utilização deste tipo de solução computacional no apoio ao ensino do português. Em um cenário em que iniciativas de uso de ferramentas online crescem<sup>5</sup> e se diversificam para atender vários contextos em sala de aula, faz-se importante ampliar a investigação no tema, visto que a maior parte dos trabalhos encontrados neste estudo encontram-se com experimentos baseados apenas nas bases de dados de referência enquanto os estudos que apresentam dados de sala de aula não são necessariamente relacionados à disciplina de Língua Portuguesa.

Em relação ao estudo secundário, uma continuação em modo manual em periódicos e anais de eventos da área computacional e educacional no Brasil, como a Revista Brasileira de Informática na Educação (RBIE) e o Simpósio Brasileiro de Informática na Educação (SBIE), é sugerida, vista a possibilidade de que o idioma possa acarretar um peso maior em publicações no âmbito brasileiro.

### **Agradecimento**

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

### **Referências**

- Amorim, E. e Veloso, A. (2017). “A Multi-Aspect Analysis of Automatic Essay Scoring for Brazilian Portuguese”. In: Proceedings of the Student Research Workshop at the 15th Conference of the European Chapter of the Association for Computational Linguistics. Valencia, Espanha.
- Bazelato, B. S. e Amorim, E. C. F. (2013). “A Bayesian Classifier to Automatic Correction of Portuguese Essays.” In: Conferência Internacional sobre Informática na Educação (TISE), 2013, V. 18, p. 779–782, Porto Alegre, RS, Brasil.
- Budgen, D., Turner, M., Brereton, P., e Kitchenham, B. A. (2008). “Using mapping studies in software engineering”. In PPIG, volume 8, p. 195–204.
- Carvalho, M. R. V. (2018). “Perfil do professor da educação básica.” Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Brasília, Brasil.
- Figueira, A. S., Silva, A. G., Melo, B. M. e Lino, A. P. e Lobato, F. R. L. e Favero, E. L. (2013) “Module of evaluation automatic essay questions on virtual learning

---

<sup>5</sup> <http://www.aen.pr.gov.br/modules/noticias/article.php?storyid=108021>

- environment LabSQL”. In: 8th Iberian Conference on Information Systems and Technologies (CISTI), 2013, p. 510–514, Lisboa, Portugal.
- Filho, A. H., Prado, H. A., Ferneda, E. e Nau, J. (2018) “An approach to evaluate adherence to the theme and the argumentative structure of essays”. In: Computational Processing of the Portuguese Language. (KES) 2018, p. 788–797, Belgrado, Sérvia.
- Fonseca, E., Medeiros, I., Kamikawachi, D. e Bokan, A. (2018) “Automatically Grading Brazilian Student Essays”. In: International Conference on Knowledge Based and Intelligent Information and Engineering Systems. (PROPOR), 2018, p. 170–179, Canela, RS, Brasil.
- Freitas, S. A. A., Canedo, E. D., Frinhani, C. L., Vidotti, M. F. e Silva, M. C. (2017) “Evaluation of an Automatic Essay Correction System Used as an Assessment Tool”. In: Universal Access in Human–Computer Interaction: Human and Technological Environments. (UAHCI), 2017, p. 210–222, Vancouver, Canadá.
- Frinhani, C. L., Freitas, S. A. A., Fernandes, M. V. e Canedo, E. D. (2016) “An automatic essay correction for an active learning environment”. In: IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA), 2016, Agadir, Marrocos.
- INEP (2019) “A Redação no ENEM 2019 – Cartilha do Participante”. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Brasília, DF, Brasil.
- Júnior, C. R. C. A. e Oliveira, E. (2016) “Proposta de um sistema de avaliação automática de redações do ENEM”. Workpedia 2016 - Workshop de Pesquisa e Desenvolvimento em IA, Inteligência Coletiva e Ciência de Dados, Niterói, RJ, Brasil.
- Júnior, C. R. C. A., Spalenza, M. A. e Oliveira, E. (2017) “Proposta de um Sistema de Avaliação Automática de Redações do ENEM Utilizando Técnicas de Aprendizagem de Máquina e Processamento de Linguagem Natural”. In: Proceedings of Computer on the Beach 2017, p. 474–483, Florianópolis, SC, Brasil.
- Passero, G., Ferreira, R., Haendchen Filho, A., Dazzi, R. L. S. (2017) “Off-Topic Essay Detection: A Systematic Review”. Anais do XXVIII Simpósio Brasileiro de Informática na Educação.
- Santos, J., Paiva, R. e Bittencourt, I. I. (2016). “Avaliação Léxico-Sintática de Atividades Escritas em Algoritmo Genético e Processamento de Linguagem Natural: Um Experimento no ENEM”. In: Revista Brasileira de Informática na Educação. V. 24, N. 2, Sociedade Brasileira de Computação.
- Vicari, R. M. (2019) “Inteligência Artificial aplicada à Educação”. In: Informática na Educação: técnicas e tecnologias computacionais. Organizado por Pimentel, Mariano; Sampaio, Fábio F.; Santos, Edméa O, Sociedade Brasileira de Computação, Porto Alegre, <http://ieducacao.ceie-br.org/inteligenciaartificial/>.
- Weigle, S. C. (2013) “English as a Second Language Writing and Automated Essay Evaluation”. In: Handbook of Automated Essay Evaluation, Current Applications and New Directions, editado por Mark D. Shermis e Jill Burstein, Routledge, <https://www.routledgehandbooks.com/doi/10.4324/9780203122761.ch3>
- Zupanc, K. e Bosnic, Z. (2015). “Advances in the field of automated essay evaluation.” Informatica, vol. 39, n. 4, p. 383–395, Eslovênia.