
Identificação de deficiências em textos educacionais com a aplicação de Processamento de Linguagem Natural e Aprendizado de Máquina

Identification of disabilities in educational texts with the application of Natural Language Processing and Machine Learning

Identificación de deficiencias en textos educativos con la aplicación de Procesamiento del Lenguaje Natural y Aprendizaje Automático

RESUMO

A correção de textos educacionais como redações e questões discursivas é uma tarefa importante, além disso, diversas escolas têm exigido a intensificação da atividade da escrita para a evolução do discente. Entretanto, o esforço despendido para a correção pode aumentar a carga de trabalho do professor ou até mesmo gerar custos adicionais, bem como um longo tempo de correção para instituições como o MEC (Ministério da Educação), que é responsável pela aplicação do ENEM (Exame Nacional do Ensino Médio). Em 2019 foi anunciado pelo MEC a tendência do ENEM se tornar digital, trazendo novas possibilidades para a análise e avaliação das redações elaboradas pelos estudantes. Neste contexto, algumas técnicas de inteligência artificial para análise de textos educacionais tem se revelado úteis no processo de avaliação automática da linguagem escrita. Assim, o objetivo desta pesquisa é analisar textos empregando para tanto as técnicas de Processamento de Linguagem Natural e Aprendizado de Máquina para identificar deficiências em textos educacionais. Esta pesquisa experimental consistiu na classificação de 695 redações elaboradas em língua portuguesa em 20 temas. Os resultados demonstraram que as técnicas empregadas possibilitaram a identificação de redações cujo conteúdo foge à temática proposta na prova, dentre outras informações importantes para que o docente possa identificar falhas na escrita da redação, tais como a coesão textual ou texto insuficiente. Os resultados esperados com a aplicação da solução desenvolvida neste experimento buscam otimizar o trabalho do professor, bem como reduzir o tempo e o custo do processo de avaliação de textos educativos.

PALAVRAS-CHAVE: Desenvolvimento educacional. Gestão do Conhecimento. Redação. Inteligência artificial. Tecnologia.

ABSTRACT

The correction of educational texts such as essays and discursive questions is an important task, in addition, several schools have demanded the intensification of the activity of writing for the evolution of the student. However, the effort spent on correction can increase the workload of the teacher or even generate additional costs and a long correction time for institutions such as the MEC (Ministry of Education), which is responsible for the application of ENEM (National Examination for Education Medium). In 2019, MEC announced the trend of ENEM to become digital, bringing new possibilities for evaluating and analyzing the essays prepared by students. In this context, some artificial intelligence techniques for analyzing educational texts have proven to be useful in the process of automatic assessment of written language. Thus, the objective of this research is to analyze texts using the techniques of Natural Language Processing and Machine Learning to identify

deficiencies in educational texts. This experimental research consisted of the classification of 695 essays prepared in Portuguese in 20 themes. The results showed that the techniques employed made it possible to identify newsrooms whose content differs from the theme proposed in the test, among other important information so that the teacher can identify flaws in the writing of the essay, such as textual cohesion or insufficient text. The expected results with the application of the solution developed in this experiment seek to optimize the work of the teacher, reducing the time and cost of the process of evaluating educational texts.

KEYWORDS: Educational development. Knowledge management. Essay. Artificial intelligence. Technology.

RESUMEN

La corrección de textos educativos como ensayos y preguntas discursivas es una tarea importante, especialmente porque varias escuelas han exigido la intensificación de la actividad de la escritura para la evolución del estudiante. El esfuerzo dedicado a la corrección puede aumentar la carga de trabajo del maestro o incluso generar costos adicionales, además del largo tiempo de corrección para instituciones como el MEC (Ministerio de Educación), que es responsable de aplicar el ENEM (Examen Nacional del escuela secundaria). En 2019, MEC anunció la tendencia de ENEM a convertirse en digital, brindando nuevas posibilidades para evaluar y analizar las redacciones preparadas por los estudiantes. En este contexto, algunas técnicas de Inteligencia Artificial para el análisis automático de textos educativos han demostrado ser útiles en el proceso de evaluación automática del lenguaje escrito. El objetivo de esta investigación es analizar textos utilizando las técnicas de Procesamiento del lenguaje natural y Aprendizaje automático para identificar deficiencias en los textos educativos. Esta investigación experimental consistió en la clasificación de 695 ensayos preparados en portugués en 20 temas. Los resultados mostraron que las técnicas empleadas permitieron identificar salas de redacción cuyo contenido difiere del tema propuesto en la prueba, entre otra información importante para que el maestro pueda identificar fallas en la redacción del ensayo, como la cohesión textual o texto insuficiente. Los resultados esperados con la aplicación de la solución desarrollada en este experimento buscan optimizar el trabajo del profesor, reduciendo el tiempo y el costo del proceso de evaluación de textos educativos.

PALAVRAS-CLAVE: Desarrollo educativo. Gestión del conocimiento. Redacción. Inteligencia artificial. Tecnología.

1 INTRODUÇÃO

Um dos principais desafios do professor de Língua Portuguesa é conseguir que seus alunos aprendam a escrever bem. Para isso um dos principais meios utilizados em sala de aula é a prática da redação. Entretanto, a tarefa de correção de textos e redações é árdua. No Brasil, o Exame Nacional do Ensino Médio (ENEM) possui a avaliação da redação, que compõe uma das cinco dimensões de avaliação consideradas nesta prova, sendo este o único item discursivo do referido exame. A maior preocupação dos candidatos é que lhes seja atribuída nota zero. Para que não seja atribuída a nota zero, a redação não poderá conter as seguintes características (BRASIL, 2019): 1) fuga total ao tema; 2) não obediência à estrutura dissertativo-argumentativa; 3) extensão total de até 7 linhas; 4) cópia integral de texto(s) da Prova de Redação e/ou do Caderno de Questões; 5) impropérios, desenhos e outras formas propositais de anulação, em qualquer parte da folha de redação; números ou sinais gráficos fora do texto e sem função clara; 6) parte deliberadamente desconectada do tema proposto; 7) assinatura, nome, apelido, codinome ou rubrica fora do local devidamente designado para a assinatura do participante; 8) texto predominante ou integralmente em língua estrangeira e, por fim; 9) folha de redação em branco, mesmo que haja texto escrito na folha de rascunho. Considerando-se tais problemas foi identificado que é possível minimizar o tempo e esforços gastos em correções empregando-se técnicas de Inteligência Artificial (IA), como é o caso da mineração de textos. Sua aplicação pode permitir a identificação de características no texto que já apontariam a identificação de problemas existentes na redação.

A aplicação das técnicas de IA num contexto de avaliações em papel teria sua eficácia diminuída. Porém, em julho de 2019 foi anunciado pelo MEC (Ministério da Educação) que a avaliação do ENEM deixará de ser aplicada na versão em papel até 2026. A implantação do Enem Digital será progressiva, com início em 2020 e previsão de consolidação em 2026. Em 2018 o INEP (Instituto Nacional de Estudos e Pesquisas) corrigiu cerca de 4.122.423 provas de redação, segundo indicado por Rocha e Moreno (2019). O uso de técnicas inteligentes no âmbito da Inteligência Artificial, baseadas no Processamento de Linguagem Natural (PNL) e Aprendizado de Máquina (AM) constitui-se numa alternativa interessante para reduzir o esforço e, principalmente, o tempo para identificação das principais causas

que acarretam o baixo desempenho dos estudantes, seja em redações, seja em questões discursivas.

Por outro lado, a avaliação automática do conteúdo de redações é um tema abordado por diferentes autores como Newman *et al.* (2010), Shermis *et al.* (2010) e Vilallon e Calvo (2009). Em especial, sobre o desenvolvimento de técnicas de mineração e análise de texto em português se destacam os trabalhos de Bazelato e Amorim (2010), Epstein e Reategui (2015), Nobre e Pellegrino (2010) e, mais recentemente, Candido e Webber (2018). Face ao contexto apresentado, a presente pesquisa considerou a seguinte questão-problema: Como analisar deficiências de conteúdo em textos educacionais empregando técnicas de Processamento de Linguagem Natural (PNL) e Aprendizado de Máquina (AM)?

Visando responder este questionamento foi realizado um processo de descoberta do conhecimento em uma base de redações, ou seja, extraiu-se dados disponíveis no repositório do Portal UOL, o qual possui um conjunto de 695 redações com notas já atribuídas a redações do ENEM. Os experimentos realizados tiveram como finalidade aplicar o Aprendizado de Máquina e, posteriormente, gerar conhecimento útil aos docentes. Num primeiro momento, os experimentos demonstrados neste artigo voltaram seu foco nos motivos que levam à atribuição de nota zero na redação do ENEM. Assim, inicialmente foi tratada a questão de fuga ao tema proposto para a redação. Desta forma, esta pesquisa tem como objetivo analisar textos empregando técnicas de PNL e AM para identificar deficiências de conteúdo em textos educacionais. O conhecimento adquirido decorrente da análise e experimentos efetuados nos textos educacionais se utilizará do processo KDD (*Knowledge Discovery in Databases* – descoberta de conhecimento em bases de dados) em sua metodologia.

2 REFERENCIAL TEÓRICO

2.1 Gestão do Conhecimento e Descoberta de Conhecimento em Bases de Dados (KDD)

O aumento na quantidade de pesquisas abordando o tema Gestão do Conhecimento (GC) tem sido observado desde a década passada por autores como Brewer e Brewer (2010) e Serenko *et al.* (2010). O trabalho de Bhojaraju (2019) traz contribuições sobre a gestão do

conhecimento (GC), definições, necessidades, ativos e o desafio em iniciar suas práticas em qualquer organização. O autor mostra que os conhecimentos são um importante ativo estratégico na atual Sociedade do Conhecimento. Já Gaspar *et al.* (2016) afirmam que ferramentas de tecnologia da informação e comunicação viabilizam os processos de gestão do conhecimento nas organizações. Nonaka e Takeuchi (1997) observam a ‘hierarquia’ na qual o conhecimento é gerado a partir dos dados interpretados, tornando-se informação que, por sua vez, a partir das análises dos resultados coletados é transformada em conhecimento.

Segundo Moraes e Ambrósio (2007), descobrir conhecimento significa identificar, receber informações relevantes e poder processá-las e agregá-las ao conhecimento prévio de seu usuário, mudando assim o estado de seu conhecimento atual, a fim de que determinada situação ou problema possa ser resolvido. Neste sentido, observa-se que o processo de descoberta de conhecimento está fortemente relacionado à forma pela qual informações e conhecimentos são processados. Neste contexto surge o processo *Knowledge Discovery in Databases* (KDD) - descoberta de conhecimento em bases de dados, uma proposta de Fayyad *et al.* (1995) para obter respostas que não podem ser detectadas quando se aplica métodos tradicionais na análise de dados para posterior tomada de decisão. Isto porque, em sua grande maioria, os métodos tradicionais são capazes de verificar apenas as relações explícitas por ventura existentes nos bancos de dados. Brachman *et al.* (1996) descrevem sobre a evolução decorrente da descoberta de conhecimento em bancos de dados (KDD) e técnicas de mineração de dados. Os processos KDD são usados para verificação de uma hipótese de um usuário ou para a descoberta de novos modelos e relações entre informações e conhecimentos já existentes. O processo de geração do conhecimento pode ser visualizado na Figura 1, na qual considera-se o *framework* do processo KDD.

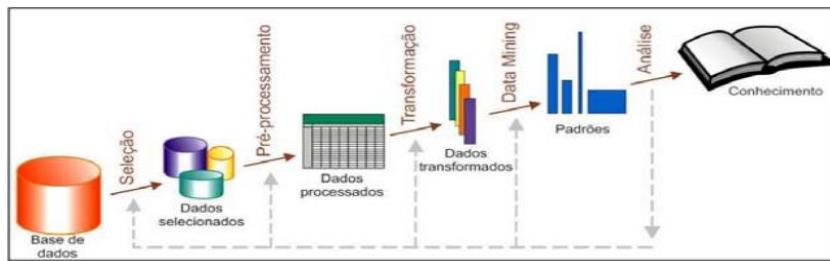


Figura 1 - Framework do processo KDD
 Fonte: Fayyad, Piatetsky-Shapiro e Smyth (1995).

Conte e Souza (2019), em sua pesquisa “Mineração de dados educacionais para avaliar os fatores que influenciam no desempenho de candidatos do ENEM”, mostram que usando de maneira correta os passos do processo de Descoberta de Conhecimento em Base de Dados (KDD) é possível analisar e caracterizar alunos e escolas de acordo com o desempenho obtido, tornando assim possível o conhecimento do perfil dos alunos e escolas que são submetidos a um determinado exame.

2.2 Processamento de Linguagem Natural (PNL) e Aprendizado de Máquina (AM) aplicados à avaliação de textos educacionais

Shermis *et al.* (2010) relatam em sua pesquisa que desde 1966 já haviam estudos para análise da escrita e acesso à tecnologia para fornecer *feedback* rápido aos professores. No entanto, a tecnologia para tais aplicações experimentou grande evolução a partir da década de 1990, quando os microcomputadores permitiram a geração de textos eletrônicos. Um estudo realizado por Araújo (2011), discute os movimentos da evolução na educação no último século e a incorporação de Tecnologias de Informação e Comunicação para entender se as tecnologias podem ajudar a promover maior qualidade e êxito na educação. Já Mungioli (2008), busca entender as transformações ocorridas com a chegada do computador em sala de aula, neste estudo, foi possível entender que a nova geração lida com a linguagem e o movimento de uma forma diferente dos adultos e percebeu uma adaptação muito fácil ao ambiente de aprendizagem com o uso da tecnologia.

As técnicas de PNL e outros mecanismos computacionais são cada vez mais pesquisados em prol de auxiliar professores no processo de correção e identificação de problemas no aprendizado. Cândido e Webber (2018) entendem que é um desafio tratar a coesão de um texto de forma automática, não obstante, em sua pesquisa os autores descrevem as possibilidades de se tratar com assertividade a coerência e coesão de redações com uso de ferramentas de PNL. O estudo realizado por eles utiliza os elementos linguísticos e técnicas computacionais para realizar a avaliação de uma redação. Os experimentos por eles realizados compararam a análise executada por um software e as avaliações feitas por dois especialistas humanos. Foram encontrados resultados convergentes em 70% dos casos analisados no experimento. Considera-se que tais resultados iniciais são promissores para o desenvolvimento de solução para avaliação automática de redações, abrindo-se então novas possibilidades de pesquisa.

Outros experimentos sobre análise automática de coesão textual em redações foram realizados por Nobre e Pellegrino (2010). Em seus estudos, identificou-se de forma automática problemas de coesão em 90% dos textos argumentativos e dissertativos analisados. Os resultados da solução automatizada aplicada no experimento foram compatíveis às notas atribuídas em correções feitas por avaliadores humanos. Os autores afirmam ainda que uma correção realizada por um programa de computador não sofre interferências externas como fadiga e alteração de humor, permitindo assim avaliar e analisar sempre de forma equânime. Entretanto, percebe-se a necessidade de revisão das expressões regulares visando detectar problemas não identificados pelo computador. Assim, o processo automatizado diminui a carga de trabalho do avaliador humano e se mostra uma ferramenta para apoio ao processo de correção executado por um avaliador humano.

O aprendizado de máquina (AM) é um campo da Inteligência artificial (IA) que permite a aplicação do PNL por meio de aprendizado baseado em experiências anteriores. Desta forma, os algoritmos empregados no aprendizado de máquina extraem o conhecimento em bases de dados reais e tomam decisões fundamentadas em classificações existentes. Este artigo aplica esta técnica num banco de redações já corrigidas e classificadas por avaliadores humanos.

2.3 Mineração de Dados aplicada à avaliação de textos educacionais

Segundo Epstein e Reategui (2015), a compreensão textual é um importante aspecto da educação do aluno. Se não for apropriadamente aprimorada no decorrer da formação educacional poderá trazer problemas futuros para o indivíduo tanto na área pessoal, quanto profissional. Os autores argumentam ainda que diferentes estudos demonstram que é possível auxiliar no processo de compreensão textual por meio da sumarização de texto e visualização de conceitos chaves. Estes mecanismos auxiliam o leitor a compreender mais facilmente o texto e realizar uma análise crítica daquilo que foi lido.

Experimentos realizados por Macedo; et.al. (2011), utilizaram técnicas de mineração de textos utilizando o software Sobeck em fóruns de ensino a distância. O sistema faz uma análise estatística de conceitos utilizados em textos escritos por alunos e traz como resultado os grafos com os conceitos principais na dissertação. O uso dessa ferramenta proporcionou redução no tempo dedicado pelo docente para a leitura de todo fórum, pois a rede de conceitos pode proporcionar indicadores valiosos, como: relevância das postagens e pertinência no que foi escrito, disponibilizando assim um maior tempo ao professor para direcionar auxílio aos discentes que registraram poucas contribuições nos indicadores citados.

Considerando-se o processo de mineração de dados (*data mining*), é feita a exploração, análise e extração de informações de um determinado volume de dados. Para Epstein e Reategui (2015, p. 1), “o minerador identifica os conceitos mais relevantes de um texto, as relações entre esses conceitos e apresenta os resultados em uma visualização de grafo”. Assim, esta pesquisa empregou o processo de mineração de dados para diagnosticar a proporção de redações com problemas tais como: pontuação, coerência e dispersão do tema proposto.

3 MATERIAIS E MÉTODOS

Para atingir o objetivo proposto nesta pesquisa quantitativa experimental, foram realizados experimentos computacionais utilizando-se como ferramenta principal a linguagem Python versão 3.7. O computador utilizado para processar as informações foi um

Acer com processador AMD A12 Quad-Core, memória RAM de 8GB e HD de 1TB. Os experimentos realizados empregaram as seguintes bibliotecas do Python: NLTK para normalização do texto, Numpy para álgebra linear e operações com matrizes, pandas para gerenciamento de dados, Scikit-learn para extração de atributos e algoritmos de aprendizado de máquina e Matplotlib para visualização dos dados.

A proposta do experimento voltou-se à aplicação de técnicas de Processamento de Linguagem Natural (PNL) e Aprendizado de Máquina (AM) utilizando algoritmos para mensurar a medida de similaridade entre as redações analisadas. Mais especificamente, empregou-se a remoção de *stop words*, *tokenização e stemming* para a normalização dos textos (MARTIN; JURAFSKY, 2009). A seguir foi analisada a frequência dos termos encontrados empregando-se a Matriz TD IDF, conforme abordagem proposta por Salton e Yang (1973), que visa medir o grau de importância de uma palavra em relação a um conjunto de documentos. Para tanto, utilizou-se a frequência (TF) da palavra e sua relevância no conjunto de textos (IDF). Isto foi feito ao se extrair a palavra do texto normalizado e gerar uma matriz de sentenças ‘x palavras’, indicando-se o valor de TF-IDF de cada palavra para cada sentença. A similaridade das sentenças então foi calculada por meio do cosseno da distância entre os vetores TF-IDF dos pares de sentenças. O cálculo dessas distâncias foi processado por meio do algoritmo *KMeans* (LOYDD, 1982), resultando em 20 agrupamentos (*clusters*) encontrados.

Para o delineamento da condução dos experimentos da pesquisa foi empregado o *framework* do processo KDD proposto por Fayyad *et al.* (1995). Assim, a proposta do experimento é obter respostas que não se podem detectar quando se aplica métodos tradicionais na análise de dados para posterior tomada de decisão, pois, em sua grande maioria, os métodos tradicionais são capazes de verificar apenas as relações explícitas nos bancos de dados.

O quadro abaixo demonstra a aplicação de cada etapa do experimento realizado.

Quadro 1 - Descrição dos métodos empregados no experimento realizado

Etapa	Descrição				
-------	-----------	--	--	--	--

Aquisição e seleção de dados	Base com 695 redações; 20 temas; 20 descrições e todas as notas de cada redação.
Pré-processamento	Remoção das <i>StopWords</i> ; <i>tokenização</i> de palavras; <i>tokenização</i> de sentenças e stemming
Formatação e transformação dos dados	Construção da Matriz TFIDF(verificar frequência de termos por documento); <i>Kmeans</i> (Classificação dos textos); Escalonamento Multidimensional.
Mineração de dados e padrões visualizados	Histogramas; Mapas de Calor; <i>Cluster Map</i> ; Matriz MDS PO; Mapas de temas
Interpretação e conhecimento	Análise dos Resultados e Conclusão

A pesquisa quantitativa utilizou uma amostra de 695 redações com 20 temas diferentes. Os dados disponibilizados para análise no experimento executado foram: número da redação; texto integral da redação escrita pelo aluno; título do aluno; texto motivador; tema da redação; nota da avaliação de cada competência (total de 5) e, por fim; nota total.

Após a aquisição dos dados foram empregadas as técnicas de remoção de *stop words*, *tokenização* e *stemming* para a preparação e normalização dos textos. Após esta etapa inicial foi realizada a construção de matrizes para verificação de frequências de termos e classificação destes com a finalidade de construir os mapas e gráficos que serão expostos nos resultados deste experimento.

4 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Após aplicação das técnicas mencionadas anteriormente foi possível averiguar que o experimento atingiu seu objetivo, ou seja, foi possível gerar informações úteis ao avaliador da redação aplicando-se técnicas de IA TFIDF, Escalonamento Multidimensional e *KMeans* demonstradas nos tópicos a seguir.

4.1 Análise a partir de Histogramas

No primeiro histograma (figura 2a) é possível visualizar o histórico das notas aplicadas, no qual foi identificado as notas médias aplicadas entre 251 e 448 em sua maioria. O segundo histograma (figura 2b) é possível visualizar a quantidade de palavras por redação, sendo que a maioria das redações teve entre 80 e 277 palavras. Na figura 2c é demonstrada a quantidade de sentenças na redação, que teve uma média de dez sentenças por redação. Entretanto, também foi identificado que há redações com menos de 2 sentenças, o que aponta para textos com quantidade de linhas insuficientes para avaliação da redação. O último histograma (figura 2d) apresenta um cruzamento entre as sentenças e palavras por sentenças, a análise média é de 35 a 43 palavras por sentença. Porém, foram encontradas sentenças com até 300 palavras, o que indica o uso inadequado de pontuação.

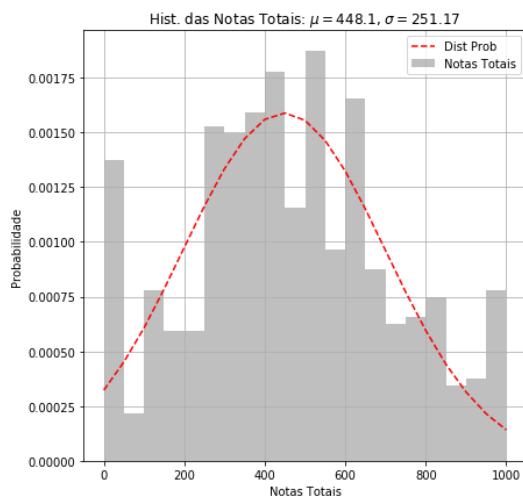


Figura 2a - Histórico de notas por redação

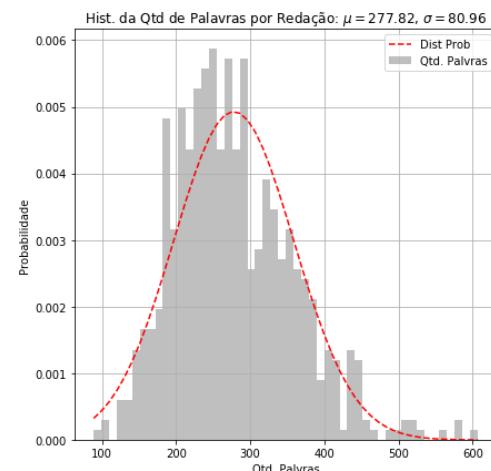


Figura 2b - Quantidade de palavras por redação

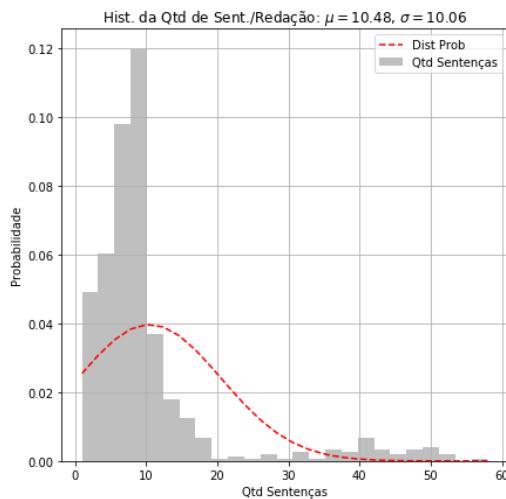


Figura 2c - Quantidade de sentenças por redação

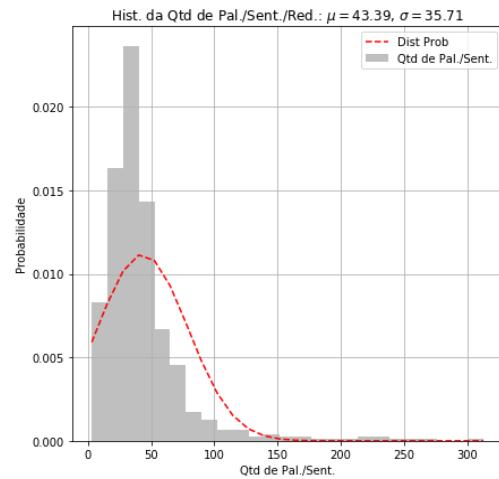


Figura 2d - Quantidade de palavras por sentença

Algumas análises podem ser levantadas a partir destes histogramas, critérios diferenciados apontam para a possibilidade de estabelecimento de correlação entre a nota atribuída numa redação e a quantidade de sentenças. Tal achado poderia auxiliar o avaliador a responder às seguintes perguntas: Há um padrão entre a quantidade de sentenças e a nota aplicada pelo avaliador? Uma sentença muito longa indica a falta de pontuação? Neste caso, a técnica de Aprendizado de Máquina pode ser utilizada com o objetivo de indicar ao avaliador em quais redações há a possibilidade de erros de pontuação e, consequentemente, problemas de coesão textual. Na aplicação deste experimento é possível avaliar hábitos de escritas, como por exemplo: regras de pontuação, repetição de palavras e sentenças e coesão textual. Tais achados estão em consonância com os estudos de Bazelato e Amorim (2010), Epstein e Reategui (2015) e Nobre e Pellegrino (2010).

4.2 Detecção de padrões por meio do emprego de técnicas inteligentes

A aplicação de técnicas de Mineração de Dados e Padrões Visualizados proporcionou dimensionar informações que serão de grande relevância no processo de correção e avaliação de redações. Observando-se os modelos de Mapa MDS (Figuras 3a e 3b), relacionando o tema proposto para a redação à abordagem usada pelo aluno, foi possível observar ocorrência de fuga do tema. A figura 3a no agrupamento na cor amarela, um exemplo de resultado no qual os alunos conseguiram escrever sua redação consoante

ao tema proposto. É possível chegar a esta conclusão pois as 20 elipses encontram-se dentro do mesmo *cluster* (agrupamento), sendo que cada uma delas representa uma redação analisada.

Já na Figura 3b é demonstrada a ocorrência de fuga em relação ao tema proposto na redação, uma vez que se observa a variação de cores no *cluster* (agrupamento). Assim, foi encontrada a presença de quatro *clusters*, o que significa que nos textos escritos pelos alunos não há similaridade com o tema proposto para a redação. Outra observação identificada é o tamanho das elipses geradas no experimento, o que demonstra diferentes notas atribuídas pelo avaliador. Estas notas atribuídas podem ser utilizadas para efeito de aprendizado de máquina para futura atribuição automática de notas pelo sistema.

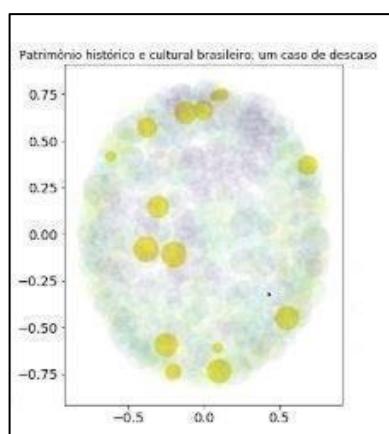


Figura 3a - Mapa MDS sem ocorrência de fuga de tema

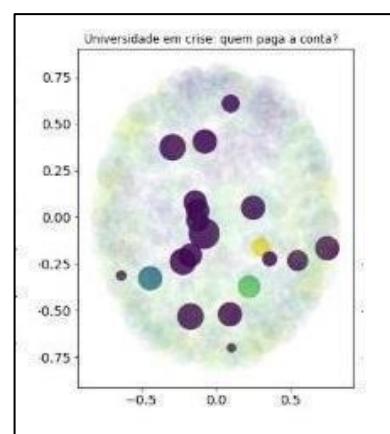


Figura 3b - Mapa MDS com ocorrência de fuga de tema

Outro experimento realizado identificou o índice de similaridade entre as redações. Nas Figuras 4a, 4b e 4c foi demonstrado o processo de *clusterização* para identificar o nível de similaridade nos 20 temas analisados nas redações consideradas no experimento.

Impeachment: a presidente deve perder o mandato?	
15	1.0

Figura 4a - Nível de Similaridade – 20 Temas

Política X Ciência: a "pílula do câncer"	
9	1.0

Figura 4b - Nível de Similaridade - 20 Temas

0 sucesso vem da escola ou do esforço individual?	
2	0.578947
16	0.157895
12	0.157895
13	0.052632
3	0.052632
Name:	cluster, dtype: float64

Figura 4c - Nível de Similaridade - 20 Temas

Nas Figuras 4a e 4b pôde-se notar que 100% dos alunos abordaram o tema que foi sugerido para a redação e se agruparam em apenas um *cluster* (respectivamente *clusters* 15 e 9). Entretanto, já na figura 4c houve dispersão do tema indicado para a redação, o que resultou em uma distribuição entre cinco *clusters* diferentes. Assim, 57,89% das redações se agruparam no *cluster* 2; 15,78% agrupadas nos *clusters* 16 e 12 e 5,26% agrupadas nos *clusters* 13 e 3. Em seu trabalho, Candido e Webber (2018) também desenvolvem técnicas de mineração e análise de texto para aplicação em texto, verificando os agrupamentos ocorridos.

Na Figura 5 são expostos os resultados alcançados após a realização dos experimentos indicados anteriormente. Por meio da figura foi possível identificar quais temáticas tiveram maior fuga ao tema proposto para a redação e, consequentemente, maior dificuldade de escrita demonstrada pelos alunos.

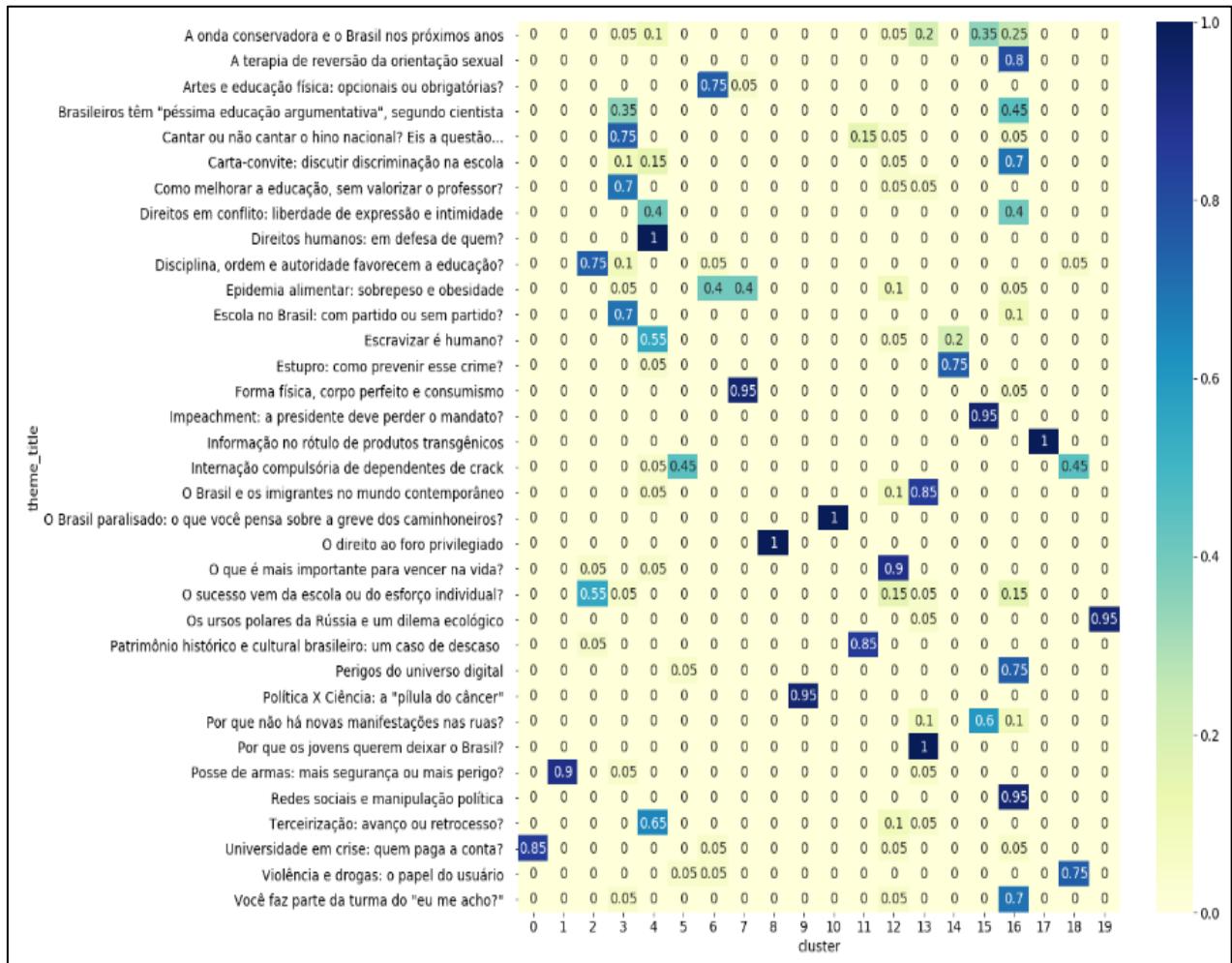


Figura 5 - Resultados dos experimentos com as 20 temáticas

Ao observar os resultados do experimento da Figura 5, foi identificado que apenas cinco temáticas não resultaram em fuga de tema proposto, sendo elas: ‘Direitos Humanos: em defesa de quem?’; ‘Informação no rótulo de produtos transgênicos’; ‘O Brasil paralisado: o que você pensa sobre a greve dos caminhoneiros’; ‘O direito ao foro privilegiado’ e, por fim; ‘Por que os jovens querem deixar o Brasil?’.

Levando-se em consideração os experimentos realizados nesta pesquisa foi possível identificar várias possibilidades para avaliação das técnicas de escrita e/ou argumentação dos alunos. Partindo-se deste princípio, entende-se que é possível mensurar o número de

redações que estão aderentes à proposta de tema informada, o que pode trazer importante conhecimento ao avaliador ou docente em relação à evolução dos alunos. Assim, é possível vislumbrar as primeiras prerrogativas de benfeitorias ao trabalho de docentes e avaliadores no processo de correção de textos educacionais.

5 CONCLUSÕES

A aplicação das técnicas de Processamento de Linguagem Natural (PNL), em especial mineração de textos, permitiu identificar padrões no conteúdo de cada redação analisada, bem como encontrar relações de similaridade entre as redações sobre um determinado tema. Os resultados encontrados no experimento realizado demonstraram que existem temas de redação com elevada coesão nos textos produzidos, ou seja, todas as redações analisadas encontram-se num mesmo *cluster* (agrupamento).

No entanto, também foi possível verificar que há outros temas que apresentaram maior dispersão, o que demonstra que os alunos podem não ter a mesma compreensão sobre o tema proposto para a redação, uma vez que seus textos apresentaram conteúdos que fugiram ao tema indicado na prova. Desta forma, foi possível cumprir o objetivo desta pesquisa ao identificar deficiências de conteúdo em textos educacionais. A partir da aplicação das técnicas de Processamento de Linguagem Natural (PNL) e Aprendizado de Máquina (AM) foi possível ainda mensurar as redações aderentes ao tema originalmente proposto. Assim, dos 20 temas nos quais foram aplicados os experimentos, apenas cinco não sofreram fuga ao tema indicado, enquanto nos demais 15 temas verificou-se fuga parcial ou total ao tema proposto.

Em complemento, outros experimentos realizados neste trabalho apontaram a quantidade e o tamanho das sentenças nas redações analisadas. Como principais resultados foi possível identificar redações com menos de duas sentenças, além de encontrar também sentenças com até 300 palavras. A primeira ocorrência aponta para texto insuficiente, enquanto a segunda ocorrência aponta para possíveis erros de pontuação e coesão textual na construção do texto. Desta forma, o processo aplicado neste trabalho possibilita identificar e analisar deficiências de conteúdo em textos educacionais de forma automatizada, o que está em consonância com trabalhos prévios empregados na fundamentação teórica desta pesquisa.

As principais contribuições deste estudo buscam permitir ao avaliador, professor ou empresas que aplicam processos seletivos avaliar as redações com menor esforço, otimizando assim o trabalho e reduzindo tempo e custo do processo de avaliação de textos educativos. Esta colaboração pode ser primordial na aplicação do ENEM digital, proporcionando ao avaliador auxílio na identificação das falhas de escrita, minimizando interferências como fadiga e alteração de humor do avaliador, sintomas estes que podem afetar a correção de um texto dissertativo. O cumprimento do objetivo proposto denota ainda contribuição sob a perspectiva acadêmica, ao servir de base para estudos que, uma vez alinhados aos conhecimentos dos profissionais de ensino, possam gerar novas abordagem que possibilitem capacitar os alunos a redigirem textos coesos com o tema proposto.

Enfim, conclui-se que técnicas de IA podem ser de grande valia para apoiar os educadores, poupando esforço e tempo, e conferindo ainda maior eficácia no processo. A pesquisa experimental executada apresenta algumas limitações, dentre as quais destaca-se que ainda não foi possível identificar todas as possibilidades que acarretam na nota zero, de acordo com as regras do Exame Nacional do Ensino Médio.

Como sugestões para estudos futuros indica-se a possibilidade de aplicação de outras técnicas para aprimorar os resultados das informações e conhecimentos oriundos da automação proposta nos experimentos realizados neste estudo. Uma alternativa é a identificação de cópia de textos motivadores e desrespeito aos direitos humanos na dissertação; dentre outros motivos que não são aceitos e desclassificam as redações produzidas pelos alunos. Outra indicação com potencial para aprimoramento dos resultados é a utilização de uma base de dados de redações maior, inclusive com maior diversidade de temáticas diferenciadas. Desta forma, o processo de aprendizado de máquina poderia trazer maior inferência na acurácia dos experimentos realizados em estudos futuros. Outra perspectiva vislumbrada para trabalhos futuros é de empregar estas técnicas construindo bases de dados com redações em colaboração com docentes, podendo assim observar como as técnicas empregadas poderiam apoiar o trabalho de análise das redações e também o aprimoramento das técnicas de IA com o apoio de especialistas.

REFERÊNCIAS

ARAÚJO, U. A quarta revolução educacional: a mudança de tempos, espaços e relações na escola a partir do uso de tecnologias e da inclusão social. **ETD - Educação Temática Digital**, v. 12, n. , p. 31-48, 2010.

BAZELATO, B. S.; AMORIM, E. C. F. A bayesian classifier to automatic correction of portuguese essays. In: Congresso Internacional de Informática Educativa (TISE), XVIII. **Anais...** Porto Alegre: Centro de Computação e Comunicação para a contrução do Conhecimento, 2013, p. 1-13.

BRACHMAN, R. J. Mining business databases. **Communications of the ACM**, Nov. 1996, p. 42.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **A redação no Enem 2019: cartilha do participante**. Brasília: INEP, 2019.

BREWER, P. D.; BREWER, K. L. Knowledge management, human resource management, and higher education: a theoretical model. **Journal of Education for Business**, v. 85, n. 6, p. 330-335, 2010.

CÂNDIDO, T.; Webber, C. Avaliação da coesão textual: desafios para automatizar a correção de redações. **RENOTE - Revista Novas Tecnologias na Educação**, v. 16, n. 1, p. 103-112, 2018.

EPSTEIN, D.; REATEGUI, E. B. Uso de mineração de textos no apoio à compreensão textual. **RENOTE: Revista Novas Tecnologias na Educação**, v. 13, n. 1, p. 1-10, 2015.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P.; UTHURUSAMY, R. **Advances in knowledge discovery & data mining**. Menlo Park: American Association for Artificial Intelligence, 1995.

GASPAR, M. A.; SANTOS, S. A.; DONAIRE, D.; KUNIYOSHI, M. S.; MAGALHÃES, F. L. F. de. Gestão do conhecimento em empresas atuantes na indústria de software no Brasil: um estudo das práticas e ferramentas utilizadas. **Informação & Sociedade: Estudos**, v.26, n.1, p. 151-166, jan/abr 2016.

GUNJAL, B. Knowledge management: why do we need it for corporates. **Malaysian Journal of Library & Information Science**, p. 1-14, Apr 2019.

MACEDO, A. L.; BEHAR, P. A.; AZEVEDO, B. F. T. (2014). Acompanhamento da interação e produção textual coletiva por meio de mineração de textos. **ETD - Educação Temática Digital**, .v 16, n. 1, p. 67-83, 2016.

MARTIN, J H.; JURAFSKY, D. **Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition**. Upper Saddle River: Pearson/Prentice Hall, 2009.

MORAIS, E. A. M.; AMBRÓSIO, A.P.L.(2007). **Mineração de textos**. Relatório técnico. Goiânia: Instituto de Informática - UFGO, 2007.

MUNGIOLI, M. C. P. Jogando com o narrador: estratégias narrativas na produção de textos em ambientes escolares informatizados. **ETD - Educação Temática Digital**, v. 10, n. 1, p. 24-48, 2009.

NEWMAN, D.; LAU, J. H.; GRIESER, K.; BALDWIN, T. Automatic evaluation of topic coherence. In: Annual Conference of the North American Chapter of the Association for Computational Linguistics, 2010. **Anais...** Association for Computational Linguistics.

NOBRE, J. C. S.; PELLEGRINO, S. R. M. ANAC: um analisador automático de coesão textual em redação. In: Brazilian Symposium on Computers in Education - SBIE), 2010. **Anais...** SBC, 2010, p. 1-12.

NONAKA, I; TAKEUCHI, H. **Criação de conhecimento na empresa**. Rio de Janeiro: Elsevier, 1997.

MEC. **Enem terá aplicação digital em fase piloto em 2020 e deixará de ter versão em papel em 2026**. Brasília: MEC, 03 jul. 2019.

NSTC - National Science and Technology Council (Committee on Technology). **Preparing for the Future of Artificial Intelligence**. Washington DC: NSTC, 2016.

ROCHA, G.; MORENO, A. C. **Enem 2018: número de redações nota mil volta a crescer, e cai o número de notas zero.** Rio de Janeiro: Portal G1, 18 jan. 2019.

SALTON, G.; YANG, C. S. On the specification of term values in automatic indexing. **Journal of Documentation**, v. 29, n. 4, p. 351-372, 1973.

SERENKO, A.; BONTIS, N.; BOOKER, L. D.; SADEDDIN, K. W. A scientometric analysis of knowledge management and intellectual capital academic literature (1994-2008). **Journal of Knowledge Management**, v. 14, n. 1, p. 3-23, 2010.

SHERMIS, M. D.; BURSTEIN, J.; HIGGINS, D.; ZECHNER, K. Automated essay scoring: Writing assessment and instruction. **International Encyclopedia of Education**, v. 4, n. 1, p. 20-26, 2010.

VILLALON, J.; CALVO, R. A.. Concept extraction from student essays, towards concept map mining. In: IEEE International Conference on Advanced Learning Technologies, 9th, 2009. **Proceedings...** Jul. 2009, p. 221-225.