# Uma abordagem computacional de análise de opinião para identificação de preconceito em redações

Sebastião Rogerio da S. Neto<sup>1</sup>, Anderson Pinheiro<sup>2</sup>, Evandro de B. Costa<sup>1</sup>, Rafael Ferreira<sup>2</sup>, Máverick A. D. Ferreira<sup>2</sup>

<sup>1</sup>Instituto de Computação – Universidade Federal de Alagoas (UFAL) Caixa Postal s/n – 57072-900 – Maceió – AL – Brasil

<sup>2</sup>DEINFO – Universidade Federal Rural de Pernambuco (UFRPE) Caixa Postal s/n – 52171-900 – Recife – PE – Brasil.

Abstract. The large adoption of the ENEM grade as an instrument to evaluate students in higher education, the number of people performing this test increased significantly. Among the evaluation methods, this test includes an essay. In this exam is required to respect human rights. In this context, this work proposes the analysis of different algorithms for the identification of hate speech in essays. It adopts a hybrid approach combining machine learning techniques and lexical dictionaries. Several experiments were conducted and in the best case the proposal reached a f-measure of 82.10%. Finally, a qualitative study was carried out where essays evaluators indicated that the adoption of the proposed tool would aid in the correction of essays.

Resumo. Com a grande adoção da nota ENEM como instrumento para avaliar estudantes no ingresso do ensino superior, o número de pessoas que passou a fazer essa prova aumentou significativamente. Dentre as etapas da prova existe uma redação que avalia, entre outras coisas, se o estudante fez o seu texto respeitando os direitos humanos. Diante disto, este trabalho propõe a análise de diferentes algoritmos para identificação de frases preconceituosas em redações. Para isso, foi utilizada uma abordagem híbrida combinando técnicas de aprendizagem de máquina e dicionários léxicos. Foram realizados diversos experimentos, no melhor caso a proposta atingiu uma f-measure de 82,10%. Por fim, foi realizado um estudo qualitativo onde avaliadores de redação indicaram que a adoção da ferramenta proposta auxiliaria na correção de redações.

## 1. Introdução

No Brasil o Exame Nacional do Ensino Médio (ENEM) é uma avaliação do sistema de ensino implantada em nível nacional no país, sendo uma ferramenta de auxílio ao Ministério da Educação (MEC) para elaborar políticas de melhoria do ensino escolar. O ENEM permitiu a unificação das provas de vestibular e o processo de seleção e ingresso nas universidades federais brasileiras a partir do Sistema de Seleção Unificada (SISU). É importante destacar que a cada ano o quantitativo de inscritos no ENEM tem crescido, ultrapassando a escala de milhões[INEP 2017].

DOI: 10.5753/cbie.sbie.2017.1187

Uma das etapas do exame é a produção de uma redação, do tipo dissertativaargumentativa, sobre um tema de ordem social, científica, cultural ou política. Para
tanto, o MEC desenvolveu métricas para avaliação dessas redações que estão dispostas no
Guia do Participante do ENEM [Guia 2012], nele são apresentados cinco competências:
Competência 1 - demonstrar domínio da modalidade escrita formal da Língua Portuguesa; dissertativo-argumentativo em prosa; Competência 2 - compreender a proposta da
redação e aplicar conceitos das várias áreas de conhecimento para desenvolver o tema,
dentro dos limites estruturais do texto; Competência 3 - selecionar, relacionar, organizar
e interpretar informações, fatos, opiniões e argumentos em defesa de um ponto de vista;
Competência 4 - demonstrar conhecimento dos mecanismos linguísticos necessários para
construção da argumentação; Competência 5 - elaborar proposta de intervenção para o
problema abordado, respeitando os direitos humanos.

Dentro das competências citadas na correção de redações do ENEM a quinta busca analisar, entre outros aspectos, se o autor está respeitando os direitos humanos. Em outras palavras, o avaliador deve identificar se algum comentário preconceituoso foi escrito na redação.

A identificação automática de preconceito é uma linha de pesquisa que foi pouco explorada ao longo dos anos. Contudo, com o desenvolvimento das redes sociais surgiram trabalhos que tratam do tema utilizando diferentes abordagens, como: utilização de técnicas de classificação de texto e criação de dicionários léxicos relacionado ao domínio [Gitari et al. 2015].

Tendo em vista a pouca quantidade de trabalhos na área de identificação de preconceito e a grande quantidade de redações produzidas nas provas do ENEM, o presente trabalho propõe a análise de diferentes algoritmos para identificação de frases com indício de preconceito em redações em português.

Para a avaliação foi criado um banco de dados de postagens de Twitter classificadas como preconceituosas ou não. Diferentes configurações de características e algoritmos de aprendizagem de máquina foram testados, alcançando uma taxa de acerto de 82,10%. Além disso, foi realizado uma análise qualitativa para identificação de sentenças com indícios de preconceito em redações que seguem o formato do ENEM.

## 2. Conceitualização

Nesta seção são apresentados os conceitos relacionados ao trabalho, dividindo-se em mineração de texto, técnicas de classificação de texto, mineração de opinião e identificação de preconceito.

#### 2.1. Mineração de Texto

A mineração de texto é um ramo da mineração de dados, pode ser definida como um conjunto de técnicas e processos para descoberta de conhecimento inovador a partir de dados textuais [Rezende et al. 2011].

Segundo [Nassif 2012] o processo de mineração de textos pode ser dividido em várias etapas, dentre elas há a extração das características relevantes e utilização de algoritmos de aprendizagem de máquina, onde são descobertos padrões válidos e úteis nos textos.

## 2.2. Técnicas de Classificação

Nesta seção são apresentados resumidamente os classificadores adotados para realização dos experimentos.

# 2.2.1. Redes Bayseanas

Esse classificador é uma técnica probabilística baseada no Teorema de Bayes [Aggarwal and Zhai 2012]. Ele calcula a probabilidade que uma amostra desconhecida pertença a cada uma das classes possíveis, predizendo a classe mais provavél. Para isto, o classificador baseado em rede baysesiana calcula uma distribuição geradora para cada classe do problema através da análise das relações entre as características envolvidas e as classes de cada instância.

#### 2.2.2. RBF

Rede neural artifical é uma técnica de aprendizagem de máquina que simula o funcionamento de um sistema nervoso [Aggarwal and Zhai 2012]. Para isso, conta com a presença de neurônios artificiais interligados entre si por meio de sinapses (na computação pesos). Cada neurônio recebe entradas e, associados a estas, pesos que representam a força do sinal sináptico. A partir das entradas e de seus respectivos pesos, um somatório ponderado é realizado no núcleo do neurônio e com base em um limiar de ativação é verificado se a entrada será ou não propagada para neurônios das camadas adjacentes a camada atual.

## 2.2.3. SVM

A máquina de vetor de suporte (SVM) é um classificador baseado na teoria de aprendizado estatístico de Vapnik [Vapnik 2013]. Para efetuar classificações/reconhecimento de padrões o SVM constrói hiperplanos em um espaço multidimensional objetivando separar casos de diferentes classes. Cada hiperplano é considerado como uma separação ótima que separa os vetores das classes sem erro e com distância máxima para com os vetores mais próximos.

#### 2.2.4. KNN

O K-Nearest Neighbors técnica de classificação de padrões que consiste em atribuir uma classe a um elemento desconhecido usando a classe da maioria de seus vizinhos mais próximos, segundo uma determinada distância (no espaço de atributos). O algoritmo KNN é baseado em analogia, um objeto é classificado pelo voto da maioria de seus vizinhos [Tan 2005].

## 2.2.5. Árvore de Decisão

É uma técnica de aprendizado de máquina que utiliza uma estrutura de árvore para avaliar os atributos de uma entrada e retorna uma predição baseada nos valores desses atributos.

A árvore é estruturada através de vários nós, onde cada nó corresponde a um teste do valor de uma característica do dado de entrada. Os nós da árvore são ligados por ramos, os quais identificam os possíveis valores do teste realizado em cada nó. Por fim, cada nó da folha da árvore representa um valor de retorno [Aggarwal and Zhai 2012].

## 2.3. Mineração de Opinião e Identificação de Preconceito

A tarefa de analisar opiniões, sentimentos, avaliações, atitudes e emoções das pessoas em relação a entidades como produtos, serviços, organizações, indivíduos, problemas é conhecida por mineração de opinião ou análise de sentimento [Liu 2012]. Para a extração de opinião normalmente subjetivos em textos, é feito a sumarização das opiniões realizada através das classificações das opiniões com base em categorias (Polaridade): positiva, negativa e neutra [Nascimento 2014].

Afim de identificar opiniões são utilizadas abordagens baseadas em aprendizagem de máquina ou técnicas que utilizam dicionários lexicos [Martin and Jurafsky 2000]. Nele cada termo lexico é identificado como um lexema, que consiste de uma forma ortográfica e fonológica com uma forma de representação de significado. Um exemplo de dicionário léxico é o WordNet Affect BR que consiste em uma base com palavras de emoções na língua portuguesa [Pasqualotti 2008]. Nessa base são apresentados grupos de palavras relacionadas a diferentes emoções.

A identificação de preconceito pode ser considerada uma sub-área da mineração de opinião [Bermingham et al. 2009]. Nela as técnicas são aplicadas para identificar textos positivos (sem preconceito, normalmente chamado de texto neutro) ou negativo (texto com comentários preconceituosos. Existem dicionários léxicos específicos para termos preconceituosos [Gitari et al. 2015, Tulkens et al. 2015].

## 3. Trabalhos Relacionados

Foram identificados poucos trabalhos na área de identificação automática de textos preconceituosos.

Um dos primeiros trabalhos encontrados na literatura propõe a utilização das palavras do texto como entrada para um classificador identificar o comentário racista [Greevy and Smeaton 2004]. Num trabalho trabalho mais recente o classificador utiliza diferentes características do texto como: tópicos, sentimento e marcações semânticas [Agarwal and Sureka 2016].

Outros trabalhos utilizam dicionários léxicos para realizar a análise, ambos tratam principalmente preconceito relacionados a raça, nacionalidade e religião [Gitari et al. 2015, Tulkens et al. 2015]. Gitari et al. propõe a extensão de dicionários léxicos como WordNet e SentiWordNet para a criação de um novo focado na identicfãoica de palavras relacionadas a preconceito. Enquanto em [Tulkens et al. 2015] é proposto um dicionário de termos em Alemão. Por fim, foi encontrado um trabalho que combina técnicas de análise de redes sociais e sentimento para identificar comentários racistas [Bermingham et al. 2009].

Como não foi encontrado trabalhos que tratassem diretamente da identificação de preconceito em textos educacionais, alguns trabalhos considerados relevantes utilizam análise de sentimento.

Em [Coutinho et al. 2016] apresentou uma análise de sentimento em mensagens de texto, para isso foram capturadas mensagens de um chat de um Ambiente Virtual de Aprendizagem - AVA, foi determinado a polaridade das mensagens como positiva, negativa e neutra e classificadas com o algorítmo Naive Bayes que utiliza probabilidades para determinar qual categoria melhor se enquadra a um determinado texto de entrada.

[Dosciatti et al. 2013] apresentaram uma abordagem utilizando Máquinas de Vetores de Suporte (SVM) na identificação de emoções em textos escritos em Português do Brasil. A base de dados utilizada no experimento foi composta por notícias extraídas de um jornal online. Os textos foram previamente rotulados e submetidos a um classificador SVM com configuração multiclasse, obtendo uma taxa de acerto de 61%.

Afim de otimizar a mediação pedagógica do professor no acompanhamento das produções textuais, [Panceri and de Menezes 2015] propõe a construção do núcleo de processamento de textos do SMA Alpes, através da combinação de técnicas de recuperação de informação, mineração de textos, análise semântica latente e clusterização. Para verificar a aderência da proposta foi aplicado o SMA Alpes a um debate de teses realizado em 2013, apresentando resultados satisfatórios.

A proposta deste trabalho é criar uma abordagem híbrida que utiliza técnicas de aprendizagem de máquina e dicionários léxicos para idenciticação de textos com indicação de preconceito em redações. Desta forma, este trabalho se diferencia dos encontrados na literatura.

# 4. Metodologia

O presente trabalho foi dividido em cinco etapas: coleta da base de dados, préprocessamento, identificação de termos preconceituosos, classificação e resultados.

#### 4.1. Coleta da Base de Dados

Esta etapa diz respeito a criação da base de dados utilizadas neste trabalho. A base de dados foi retirada através de uma filtragem manual do Twitter, a base contêm 500 tweets, ela foi dividida em, 200 tweets com sentenças com preconceituosas e 300 tweets com sentenças neutras. Essa rotulação foi criada por dois analisadores, quando houve divergência entre eles um terceiro analisador foi consultado.

#### 4.2. Pré-Processamento

Nesta etapa a ferramenta divide as sentenças para serem avaliadas. Cada sentença passa por um pré-processamento onde serão eliminadas palavras com pouco valor para o texto, essas palavras são definidas como *Stopwords*, palavras não relevantes na análise de textos e incluem, normalmente, preposições, pronomes, artigos, advérbios, e outras classes de palavras [Berry and Castellanos 2008].

Além disso, todas as palavras foram normalizadas, sendo retirado os acentos, cedilhas e as palavras maiúsculas foram convertidas para minúsculas. Por fim, também foi realizado o processo de *stemming*, onde as palavras são reduzidas ao seu radical [Berry and Castellanos 2008].

## 4.3. Identificação de Termos Preconceituosos

Uma das principais contribuições do trabalho proposto foi a identificação dos termos preconceituosos para o idoma português brasileiro. Essa etapa dividiu-se em dois passos: (i) criação de um dicionário léxico para o português; (ii) análise das palavras mais utilizadas no banco de dados proposto.

Para a primeira etapa o trabalho se baseou nos dicionários propostos em [Pasqualotti 2008] e [Tulkens et al. 2015]. Como já foi dito, o primeiro é um dicionário geral para análise de sentimento em português. Dele foram retirados os grupos de palavras que podem indicar comentários preconceituosos.

O segundo trabalho apresenta um recurso léxico para identificação de comentários racistas em alemão. Ele foca principalmente em palavras relacionados a raça, nacionalidade e religião. Neste caso, foi realizado uma tradução e adequação dos termos em alemão para o português. É importante frisar que nesse dicionário as palavras são categorizadas como termos racistas e termos neutros.

Após a criação dos dicionários realizou-se um estudo sobre as palavras mais frequentes que aparecem na base de dados. Para isso, foi realizado todos os préprocessamentos propostos na seção anterior e as palavras foram pontuadas de acordo com sua frequência.

Para a classe de texto preconceituoso as palavras que se repetiram várias vezes foram: racista, negro, preto, cabelo, gay, peso e branco. Por outro lado, as palavras de postagens neutras mais frequentes são: racismo, preconceito, Brasil, luta, homofobia, cor e mulher. Essas palavras também vão ser usadas como características.

## 4.4. Classificação

A primeira etapa da classificação foi a definição das características dos textos que foram usadas. Para realizar uma avaliação mais abrangente foram definidos diferentes conjuntos de características:

- Todas as Palavras (TP): todas as palavras do texto;
- Palavras Frequentes da Classe Preconceituoso (PFCP): palavras que foram mais vezes utilizadas nas postagens consideradas preconceituoso;
- Palavras Frequentes da Classe Neutra (PFCN): palavras que foram mais vezes utilizadas nas postagens consideradas neutras;
- **Dicionário 1 (Dic1)**: palavras consideradas preconceituosas do dicionário de [Pasqualotti 2008];
- **Dicionário 2** (**Dic2**): palavras consideradas preconceituosas do dicionário de [Tulkens et al. 2015];
- **Dicionário 3 (Dic3)**: palavras consideradas neutras do dicionário de [Tulkens et al. 2015].

O primeiro conjunto de características utilizado foi o TP. Para isso, não foi realizado nenhum processamento no texto. Depois foram criados diferentes conjuntos de características utilizando o método MFD [Pinheiro et al. 2015] nos conjuntos PFCP e PFCN. Esse método propõe a utilização de X palavras mais relevantes de cada postagem como característica, onde o valor de X é um inteiro, os autores mostraram que o valor

ideal de X sempre está entre 1 até 10 para textos grandes. Como estamos trabalhando com textos extraídos do twitter ou sentenças, vão ser testados os valores de 1 a 3.

Por exemplo, quando o valor de X for definido como 1, vai ser recuperado o termo mais importante de cada texto para compor o vetor de características. Dessa forma, teremos os seguintes conjuntos de características:

- **TP**: com todas as palavras;
- MFD1: com uma palavra mais importante de cada texto e as palavras dos dicionários;
- MFD2: com duas palavras mais importantes de cada texto e as palavras dos dicionários;
- MFD3: com três palavras mais importantes de cada texto e as palavras dos dicionários;

Após a definição das características foram utilizadas técnicas de aprendizagem de máquina para identificar os textos com indícios de preconceito. Mais especificamente foram utilizados cinco classificadores de diferentes tipos: Redes Bayseanas, KNN, SVM, Árvore de Decisão e Rede Neural (RBF) (seção 2.2). Todos os classificadores foram utilizados usando a ferramenta WEKA e suas configurações padrões [Hall et al. 2009].

#### 4.5. Resultados

Os classificadores foram avaliados utilizando a base de dados descrita na seção 4.1 e as métricas tradicionais da literatura [Baeza-Yates and Ribeiro-Neto 2013]: **Precisão**: avalia a quantidade de instâncias que foram classificadas corretamente; **Cobertura**: avalia a porcentagem instâncias de uma determinada classe que não foi classificada como pertencente a essa classe; *F-Measure*: É uma média harmônica entre Precisão e Cobertura, como mostrado na Fórmula 1.

$$F - Measure = 2 \times \frac{Precis\~ao \times Cobertura}{Precis\~ao + Cobertura}$$
 (1)

Além disso, no processo de avaliação foi aplicado o método de validação cruzada com 10 k-fold [Arlot et al. 2010] que divide o conjunto de dados em treinamento e testes, possibilitando uma melhor avaliação do classificador.

A Tabela 1 apresenta os resultados detalhados por algoritmo de classificação e grupo de característica utilizados. Os algoritmos que alcançaram os melhores resultados no geral foram os de redes bayseanas e SVM. Esse resultado já eram esperado visto que esses algoritmos tem se destacado na literatura de classficação de texto. Por outro lado, o RBF teve os piores resultados.

Em relação as características os grupos de características que tiveram melhor desempenho foram o MFD2 e MFD3. O MFD2 obteve resultados maiores para os algoritmos RBF, SVM e árvore de decisão, enquanto o MFD3 alcançou maiores resultados para redes bayseanas e KNN. É importante destacar que esses grupos de características englobam 2 ou 3 palavras de cada postagem e os termos pertencentes aos dicionários propostos, tratando-se assim de uma abordagem híbrida que utiliza dicionário e algoritmos de aprendizagem de máquina.

**Tabela 1.** Resultados dos Algoritmos

Algoritmo	Características	Precisão	Cobertura	F-Measure
Redes Bayseanas	TF	78,80	78,60	78,60
RBF	TF	59,30	46,40	49,50
SVM	TF	79,10	79,00	79,00
KNN	TF	69,70	59,50	58,20
Árvore de Decisão	TF	74,30	74,10	74,20
Redes Bayseanas	MFD1	77,20	77,40	77,30
RBF	MFD1	62,20	61,30	61,70
SVM	MFD1	80,20	80,40	80,20
KNN	MFD1	78,50	78,60	78,50
Árvore de Decisão	MFD1	79,60	79,20	79,30
Redes Bayseanas	MFD2	79,70	79,80	79,40
RBF	MFD2	65,60	61,10	63,30
SVM	MFD2	82,10	82,20	82,10
KNN	MFD2	79,20	78,20	78,40
Árvore de Decisão	MFD2	79,50	79,40	79,40
Redes Bayseanas	MFD3	80,00	80,00	80,00
RBF	MFD3	63,20	61,10	62,30
SVM	MFD3	80,80	80,80	80,80
KNN	MFD3	80,40	78,00	79,20
Árvore de Decisão	MFD3	79,30	79,20	79,20

A tupla (algoritmo, características) que alcançou o melhor resultados foi a utilizando o algoritmo SVM e o conjunto de característica MDF2. Portanto, essa configuração foi utilizada para identificação de sentenças preconceituosas em textos de redações.

# 5. Identificação de Sentenças com Indícios de Preconceito em Redações

Para avaliar a ferramenta proposta no contexto educacional foi realizado uma análise qualitativa utilizando redações extraídas do Banco de Redações da UOL que são corrigidas seguindo os parâmetros estabelecidos pelo ENEM. Foram selecionadas cinco (5) redações para o experimento. Como as redações com comentários preconceituosos são eliminadas da base, foram inseridos textos racistas nelas. Contudo, para tornar o experimento mais realista foram utilizadas redações com temas que existem uma probabilidade maior de ter comentários preconceituosos <sup>1</sup>.

Para cada redação, foram incluídas 3 sentenças com preconceito. É importante destacar que essas sentenças não estão contidas na base de dados extraída do twitter, para não causar problemas de *overfitting*. Com isso, o banco de dados ficou com 15 sentenças preconceituosas e 70 neutras.

Das 15 frases preconceituosas inseridas nas redações 14 foram classificadas corretamente e apenas 1 não. Por outro lado, 15 sentenças que não deveriam ser marcadas como preconceituosas foram, deixando assim 56 sentenças neutras.

 $<sup>^{1}</sup> https://educacao.uol.com.br/bancoderedacoes/propostas/carta-convite-discutir-discriminacao-na-escola.htm \\$ 

Para analisar esse resultados foram consultados dois professores que trabalham com correção de redações. A opinião deles é que a ferramenta é um ótimo auxílio para lidar com o problema relacionado a competência 5 do ENEM (respeitar os direitos humanos). Contudo, ela não deve ser utilizada para uma avaliação automática e sim para indicar para os professores as frases que tem indícios de preconceito, mas a decisão final continuaria com o avaliador. Ou seja, seria uma abordagem semi-automática.

# 6. Considerações Finais

Este artigo apresentou uma análise de diferentes algoritmos para identificação de textos com indícios de preconceitos para o idioma português. Na proposta foi utilizada uma abordagem híbrida combinado dicionários léxicos e algoritmos de aprendizagem de máquina.

Foi realizado um experimento para avaliar diferente combinações de características e algoritmos de aprendizagem de máquina para o problema citado. O melhor resultado encontrado combinou como características duas palavras de cada texto da base de dados e os termos dos dicionários propostos. O SVM foi o melhor algoritmo avaliado, atingindo uma f-measure de 82,10%.

Como trabalhos futuros sugere-se: (i) a ampliação da base de dados para utilização de algoritmos de *deep learning*; (ii) obtenção de redações reais que contém sentenças com comentários preconceituosos, para uma melhor avaliação educacional; (iii) criação de uma ferramenta visual para que professores possam utilizar.

#### Referências

- Agarwal, S. and Sureka, A. (2016). But i did not mean it!—intent classification of racist posts on tumblr. In *Intelligence and Security Informatics*, pages 124–127. IEEE.
- Aggarwal, C. C. and Zhai, C. (2012). *Mining text data*. Springer Science & Business Media.
- Arlot, S., Celisse, A., et al. (2010). A survey of cross-validation procedures for model selection. *Statistics surveys*, 4:40–79.
- Baeza-Yates, R. and Ribeiro-Neto, B. (2013). *Recuperação de Informação-: Conceitos e Tecnologia das Máquinas de Busca*. Bookman Editora.
- Bermingham, A., Conway, M., McInerney, L., O'Hare, N., and Smeaton, A. F. (2009). Combining social network analysis and sentiment analysis to explore the potential for online radicalisation. In *Social Network Analysis and Mining*, 2009. ASONAM'09. International Conference on Advances in, pages 231–236. IEEE.
- Berry, M. W. and Castellanos, M. (2008). Survey of text mining II, volume 6. Springer.
- Coutinho, E., Moreira, L., Paillard, G., and de Lima, E. T. (2016). Análise do sentimento de mensagens de chats em uma turma de graduação de um curso de educação à distância. In *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*, volume 5, page 1019.
- Dosciatti, M. M., Ferreira, L. P. C., and Paraiso, E. C. (2013). Identificando emoçoes em textos em português do brasil usando máquina de vetores de suporte em solução multiclasse. *ENIAC-Encontro Nacional de Inteligência Artificial e Computacional. Fortaleza, Brasil.*

- Gitari, N. D., Zuping, Z., Damien, H., and Long, J. (2015). A lexicon-based approach for hate speech detection. *International Journal of Multimedia and Ubiquitous Engineering*, 10(4):215–230.
- Greevy, E. and Smeaton, A. F. (2004). Classifying racist texts using a support vector machine. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 468–469. ACM.
- Guia, A. R. (2012). Guia do participante. *Instituto Nacional de Estudos e. Pesquisas Educacionais Anísio Teixeira. Brasília.*
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The weka data mining software: an update. *ACM SIGKDD explorations newsletter*, 11(1):10–18.
- INEP (2017). Inep instituto nacional de estudos e pesquisas educacionais—enem. Disponível em: http://gl.globo.com/educacao/enem/2017/noticia/enem-2017-chega-a-38-milhoes-de-inscritos-a-quatro-\dias-do-fim-do-prazo.ghtml. Acesso em: 12 Julho 2017.
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1):1–167.
- Martin, J. H. and Jurafsky, D. (2000). Speech and language processing. *International Edition*, 710:25.
- Nascimento, P. C. (2014). *DICIONÁRIO DE POLARIDADES PARA APOIO A ANÁLISE DE SENTIMENTO*. PhD thesis, Universidade Federal do Rio de Janeiro.
- Nassif, L. F. d. C. (2012). Técnicas de agrupamento de textos aplicadas à computação forense.
- Panceri, S. and de Menezes, C. (2015). Apoio a mediação pedagógica em um debate de teses utilizando técnicas de processamento de texto. In *Simpósio Brasileiro de Informática na Educação*, volume 26, page 977.
- Pasqualotti, P. R. (2008). Reconhecimento de expressões de emoções na interação mediada por computador.
- Pinheiro, R. H., Cavalcanti, G. D., and Ren, T. I. (2015). Data-driven global-ranking local feature selection methods for text categorization. *Expert Systems with Applications*, 42(4):1941–1949.
- Rezende, S. O., Marcacini, R. M., and Moura, M. F. (2011). O uso da mineração de textos para extração e organização não supervisionada de conhecimento. *Revista de Sistemas de Informação da FSMA*, 7:7–21.
- Tan, S. (2005). Neighbor-weighted k-nearest neighbor for unbalanced text corpus. *Expert Systems with Applications*, 28(4):667–671.
- Tulkens, S., Hilte, L., Lodewyckx, E., Verhoeven, B., and Daelemans, W. (2015). The automated detection of racist discourse in dutch social media. *Mirror*.
- Vapnik, V. (2013). The nature of statistical learning theory. Springer science & business media.