

# O Discurso de Ódio Homofóbico no *Twitter* a partir da Análise de Dados

Vinícius S. dos Santos<sup>1</sup>, Felipe da R. Henriques<sup>1</sup>, Gustavo Guedes<sup>1</sup>

<sup>1</sup>Programa de Pós-graduação em Ciência da Computação - PPCIC  
Centro Federal de Educação Tecnológica Celso Suckow da Fonseca – CEFET/RJ  
Caixa Postal - 20271-204 - Rio de Janeiro - RJ - Brasil

vinicius.soares@aluno.cefet-rj.br,  
{felipe.henriques,gustavo.guedes}@cefet-rj.br

**Abstract.** *This study aims to analyze Twitter messages with homophobic content. Initially, we investigate and summarize the terms in the texts through text mining techniques via Bag of Words (BoW). Next, we evaluate sequences of two adjacent words, the Bigrams. After the initial exploratory analyses, we use clustering techniques in the extracted terms, allowing the assessment of the differences between textual elements in the resulting groups. Finally, this study uses text mining techniques, cluster analysis, and data visualization applied to a current issue in society, allowing for the discovery of information contained in the collection of extracted texts and explaining cases of homophobia on Twitter.*

**Resumo.** *O objetivo deste estudo é analisar mensagens do Twitter com conteúdo homofóbico. Inicialmente, nós investigamos e resumimos os termos contidos nos textos utilizando técnicas de mineração textual, por meio do Bag of Words (BoW). Após as análises exploratórias iniciais, utilizamos técnicas de agrupamento nos termos extraídos, permitindo avaliar as diferenças entre os elementos textuais nos grupos resultantes. Por fim, este estudo utiliza técnicas de mineração de texto, análise de agrupamentos e visualização de dados aplicados a um tema atual da sociedade, permitindo a descoberta de informações contidas nos textos extraídos e explicando casos de homofobia no twitter.*

## 1. Introdução

O Supremo Tribunal Federal (STF) do Brasil, em 13 de junho de 2019, julgou como crime de racismo casos de agressões contra pessoas de orientação sexual lésbicas, gays, bissexuais, transexuais e travestis (LGBT) [UOL, 2020]. Na ação denominada “*Homofobia e omissão legislativa*”, lista-se o julgamento e as alegações da suprema corte. Dentre os trechos da determinação, destaca-se o enquadramento de atos homofóbicos e transfóbicos como formas contemporâneas do racismo [Brasil, 2019].

Mesmo diante dessas ações, jornais e portais de notícia alertam que o discurso de ódio está se dissipando nas redes sociais. Segundo o portal *UOL Notícias*, um perfil homofóbico chamado “*HomofobiaSIM*” teve rápida ascensão na rede social *Twitter*, contando com mais de 15 mil seguidores. O perfil, como seu próprio nome sugere, incitava violência contra homossexuais. A reportagem ainda destaca que o perfil chegava ao ponto de compartilhar que homossexuais mereciam morrer (e.g., “*Um viadinho [sic] brasileiro foi espancado, aí que está o erro, ele deveria ter sido é morto, acabe com Aids, mate um gay*”) [UOL, 2010].

Outra notícia que traz à tona casos de homofobia foi relatada pelo jornal *O Estado de São Paulo*. Um usuário do *Twitter* com mais de 7 milhões de inscritos em suas redes sociais, contratado pelo Governo Federal para divulgação do programa de Reforma do Ensino Médio no ano de 2017, teve mensagens antigas denunciadas por diversos usuários da rede por conter discurso de ódio. Dentre essas mensagens, destaca-se uma de cunho homofóbico, divulgada na mesma matéria jornalística, conforme exibido na Figura 1.



**Figura 1. Homofobia no Twitter - fonte: [S.PAULO, 2017]**

O *Twitter* é definido como um serviço de microblogue, por meio do qual pessoas e empresas podem se comunicar por mensagens [Honey and Herring, 2009]. O significado do termo microblogue é definido pelo tipo de serviço que é oferecido, e que segundo a própria companhia consiste em: “*as pessoas publicam mensagens, que podem conter fotos, vídeos, links e texto. Essas mensagens são publicadas em seu perfil e enviadas a seus seguidores, podendo ser encontradas por meio de busca na rede*” [Twitter, 2020]. A característica que destaca o *Twitter* das outras redes sociais é que as mensagens trocadas limitam-se até 280 caracteres, chamadas pelo microblogue de *tweets*, que permitem que os usuários compartilhem suas “*últimas histórias, ideias, opiniões e notícias em tempo real, escolhendo entre mensagens públicas ou um grupo selecionado*” [Twitter, 2020].

Em 2017, o *Twitter* definiu novas regras para combater o abuso, violência e discurso de ódio. Em constante atualização, a conduta de propagação de ódio define que não é permitido “*promover violência, atacar diretamente ou ameaçar outras pessoas com base em raça, etnia, nacionalidade, orientação sexual, sexo, identidade de gênero, religião, idade, deficiência ou doença grave*” [Twitter, 2020].

A aplicação dessas regras, entretanto, pode requerer um grande esforço manual para revisar cada mensagem compartilhada na rede. MacAvaney et al. [2019] destacam que as ferramentas de identificação e classificação automáticas podem acelerar o processo de curadoria dos textos, e ainda que os recursos humanos poderiam ser utilizados em análises mais aprofundadas de temas específicos ainda não automatizados. Desse modo, o *Twitter* poderia impedir que a mensagem de cunho homofóbico fosse compartilhada ainda no momento da sua postagem, impedindo sua disseminação na rede.

Uma limitação à classificação automática de textos para este contexto de pesquisa é a incipiência de alguns léxicos para discurso de ódio homofóbico no idioma português. Cita-se como exemplo a ferramenta *Hatebase*, um vocabulário multilíngue, mantido por funcionários e voluntários em mais de 90 idiomas e mais de 200 países. Para construção do *Hatebase* [2020], foi desenvolvido o *Hatebrain*, uma engenharia de processamento de linguagem natural que realiza análises linguísticas em conversas públicas para estimar uma probabilidade de conteúdo com discurso de ódio. Todos os dados estão disponíveis

para consulta por meio do seu sítio eletrônico ou sua interface de programação de aplicativos (em inglês, *Application Programming Interface* - API). Atualmente, nessa ferramenta, existem 28 termos relacionados à homofobia no idioma português.

Em janeiro de 2021, o *Twitter* lançou sua nova API para acadêmicos. Conforme notícia veiculada em seu sítio eletrônico, a nova ferramenta permite coletar até 10 milhões de *tweets* mensalmente, o que significa algo vinte vezes maior que a anterior [Twitter, 2020]. Uma das vantagens abordadas pela empresa, além do acesso gratuito para pesquisadores, é a busca por dados históricos antes restritos a usuários pagantes. O *Twitter* disponibiliza guias técnicos e metodológicos, com o objetivo de facilitar a construção das consultas e análises dos conteúdos extraídos. Vale destacar que a justificativa para uso dos dados provenientes do *Twitter* nesse estudo é a relação dos fatos abordados sobre homofobia nas matérias jornalísticas e os recursos disponibilizados pela rede social aos pesquisadores para a construção de suas pesquisas.

Dada a problemática supracitada, o objetivo geral deste trabalho é analisar o discurso de ódio homofóbico em textos em português. Para isso, inicialmente são utilizadas técnicas de mineração de textos para resumir e avaliar os termos contidos nos textos, análise de agrupamentos, permitindo avaliar as diferenças entre cada elemento textual nos grupos propostos e visualização de dados. Para alcançar o objetivo geral, os seguintes objetivos específicos são ilustrados:

- consultar e construir uma base de dados de mensagens do *Twitter* em português brasileiro;
- aplicar o léxico do *Hatebase* para filtragem dos conteúdos relacionados;
- analisar os conteúdos das mensagens extraídas do *Twitter*;

As próximas seções estão estruturadas da seguinte forma: na Seção 2, são apresentados os trabalhos relacionados, necessários para contextualizar e permitir o entendimento do presente estudo. Na Seção 3, é detalhada a metodologia adotada. Na Seção 4, são apresentadas as análises exploratórias e a proposta de agrupamento. Na Seção 5, são apresentadas as conclusões e a proposta de próximos passos.

## 2. Trabalhos Relacionados

Com o intuito de promover o entendimento dos principais conceitos abordados nessa proposta, são apresentados a seguir os trabalhos relacionados.

Uma maneira clássica de se identificar o discurso de ódio é usar uma abordagem baseada em palavras-chave, conforme descrito em MacAvaney et al. [2019]. Para os autores, com base em uma ontologia ou dicionário, textos que contenham potencial discurso de ódio podem ser identificados. Um exemplo considerado pelos autores é o vocabulário *Hatebase*. Eles ponderam que apenas o uso do termo presente no texto não é suficiente para definir a presença do discurso de ódio. Além disso, segundo eles, a utilização dessa abordagem exclusivamente não permite identificar discurso de ódio em textos que não contenham quaisquer palavras-chave de incitação ao ódio, como é o caso do uso de linguagem figurada.

Nascimento et al. [2021] abordaram o tema da detecção de mensagens homofóbicas em português, no *Twitter*, usando análise de sentimentos. Os autores construíram a base de dados consultando a interface do *Twitter* com quatro palavras-chave de

cunho homofóbico (e.g., “*sapatão*”, “*viado*”, “*traveco*” e “*gay*”). Após isso, vetorizaram as mensagens utilizando TF-IDF e treinaram um modelo de Regressão Logística aplicados ao TASH-PT, um banco de dados para análise de sentimentos na língua portuguesa com dados coletados do *Twitter*. Os autores relatam que para construção do TASH-PT foram avaliados 2.787 *tweets*, sendo 888 positivos, 881 negativos e 1.018 neutros. Para validação do modelo, os autores fizeram uso de cinquenta e sete julgadores humanos. Para avaliar a qualidade do modelo de classificação adotado, tais julgadores responderam se cerca de 160 mensagens classificadas pela Regressão Logística eram ou não mensagens homofóbicas. Embora ainda incipiente, os resultados observados foram: acurácia de 61,48%, precisão de 66,67%, revocação de 62,16% e *F-Measure* de 64,33%.

Quatro modelos de aprendizado de máquina para classificação foram estudados em Silva and Roman [2020]. Os autores avaliaram modelos gerados pelos algoritmos *Naive Bayes*, *Support Vector Machine*, *Bayesian Logistic Regression* e *Multilayer Perceptron Neural Network* (MLP), com diversas configurações. Os dados provenientes do *Twitter* foram utilizados para aplicação dos modelos para língua portuguesa. O modelo mais eficiente foi o que utilizou o *Support Vector Machine*, com acurácia de 88,36% e *F-measure* de 71,3%.

Em Burnap and Williams [2015], utilizou-se a abordagem híbrida para um classificador de aprendizado de máquina supervisionado e modelagem estatística aplicados a um conjunto de dados de mensagens do *Twitter*. Eles propuseram uma combinação de classificadores probabilísticos, *rule-based* e *spatial-based*, com um meta-classificador conjunto, *voted ensemble* com máxima probabilidade. Segundo os autores, o meta-classificador utilizou como base os modelos *Bayesian Logistic Regression* (BLR), *Random Forest Decision Tree* (RFDT) e *Support Vector Machine* (SVM). Os resultados obtidos foram: precisão de 95%, revocação de 95% e *F-measure* de 95%.

Uma abordagem híbrida também foi apresentada em Silva Neto [2017]. O autor combinou técnicas de aprendizado de máquina e dicionários léxicos para detecção de preconceito em textos no idioma português. De acordo com o autor, o modelo que obteve melhor resultado utilizou Redes Bayesianas considerando a frequência de palavras em unigramas, denominado *Token-N-Grams*. Os resultados obtidos foram: precisão de 83,60%, revocação de 83,80% e *F-measure* de 83,60%.

Técnicas de agrupamento de textos foram estudadas em Nassif [2011] e aplicadas à Computação Forense. Para o autor, o método que melhor se adaptou aos dados coletados foi o agrupamento por técnicas hierárquicas e ainda o algoritmo K-médias (*kmeans*) foi um dos que apresentou a menor perda de convergência para mínimos locais. O autor conclui que o algoritmo em questão é bastante escalável por possuir complexidade computacional linear.

Nesse estudo, considerou-se a análise dos conteúdos textuais de mensagens extraídas do *Twitter*. A aplicação de técnicas de mineração de texto e agrupamentos é utilizada para extrair informações dos conteúdos das mensagens a fim de se verificar a homofobia. Esse estudo adota o uso de léxicos homofóbicos extraídos do *Hatebase*, além de uma proposição de técnica de agrupamento de textos, com o objetivo de avaliar a diferença do uso de palavras para cada grupo.

### 3. Metodologia

Nesta seção detalha-se a metodologia para o alcance dos objetivos. Para tal, utiliza-se como inspiração o modelo de mineração de dados que foi proposto por Han et al. [2012].

#### 3.1. Coleta e Pré-filtragem dos Dados

A pesquisa inicia-se com a construção da Base de Dados, que aqui é representada por uma coleção de *tweets*. Esta etapa do processo prevê o acesso aos dados e, para isso foi utilizada a API do *Twitter*, em <https://developer.twitter.com/en/apps>. Em seguida, para acessar a *url* em questão e coletar os dados necessários, foi utilizada uma linguagem de programação de alto nível, o *Python* [Van Rossum and Drake, 2009].

A API possibilita buscas a quaisquer mensagens desde que se faça uso de filtros, como por exemplo a identificação da mensagem (id). Utilizou-se a seguinte estratégia para levantar as mensagens:

- localizar a identificação da mensagem do ex-deputado federal Jean Wyllys (em seu perfil: *jeanwyllys\_real*) na qual o mesmo informa seu asilo (*i.e.*, mensagem com id 1088491273227522049);
- a partir dessa mensagem, utilizar o léxico referente a homofobia, extraído em consulta ao sítio *Hatebase*, para filtrar os comentários que contivessem um ou mais termos existentes no *Hatebase*;
- filtrar apenas as mensagens em português e referentes ao ano de 2019.

#### 3.2. Pré-processamento

A etapa de pré-processamento do texto consiste em converter os documentos coletados em uma estrutura que seja compatível com a análise proposta. Para isso, foram considerados os seguintes passos:

- remoção de conteúdos irrelevantes na análise como menções, urls, emojis, imagens e outros;
- transformação de todo conteúdo textual para *lower case*;
- remoção de palavras irrelevantes (*i.e.*, *stopwords*), como artigo, preposição, numeral e outros. Para isso, foi utilizada a biblioteca NLTK do Python.
- por fim, houve a redução de todas as palavras ao seu radical, a fim de estabelecer a unicidade de cada termo.

##### 3.2.1. Transformação dos Dados

Nesta etapa, cada termo foi separado com o intuito de construir uma matriz de termos e documentos (TF-IDF). Este método tem por objetivo contar a frequência de uso de palavras para cada documento (*i.e.*, mensagem do *Twitter*), e estimar a importância do termo no texto, conforme abordado em Santana [2020].

##### 3.2.2. Mineração dos Dados

Nesta etapa o objetivo foi identificar as palavras mais frequentes por meio do uso da nuvem de palavras. Além disso, foi utilizada a análise de bigramas (*i.e.*, palavras em pares) com o objetivo de verificar a incidência das mesmas para o entendimento da homofobia.

### 3.2.3. Modelo de Agrupamento

Nesta etapa, o estudo se propôs a identificar como os textos podem ser agrupados em termos similares (similaridade). Assim, pode-se apresentar uma possível evidência de como os termos homofóbicos podem ser separados, o que pode fomentar estudos futuros.

## 4. Análises e Resultados

O levantamento dos dados na etapa de coleta evidencia um total de 2.597 mensagens únicas oriundas das respostas à mensagem inicial do ex-deputado federal Jean Wyllys. Deste total, conta-se 827 usuários únicos, ou seja, 827 pessoas responderam à postagem do ex-deputado utilizando um ou mais termos do *Hatebase*. Pode-se concluir com esses dados iniciais que, em média, esses usuários responderam 2,9 vezes à mensagem original.

### 4.1. Análise exploratória

A análise dos termos mais utilizados nas 2.597 mensagens coletadas tem o objetivo de apresentar uma visão geral sobre o teor das mensagens. A nuvem de palavras ilustrada na Figura 2 apresenta uma visão geral sobre os termos mais utilizados. A interpretação subjacente é que, quanto maior for o tamanho da palavra em relação ao conjunto de palavras, maior sua frequência nos textos coletados das mensagens do *Twitter*. Pode-se perceber que as palavras mais frequentes são “viado” e “bicha”. Também é possível notar algumas outras palavras com menor ocorrência, como “biba” e “gay”.

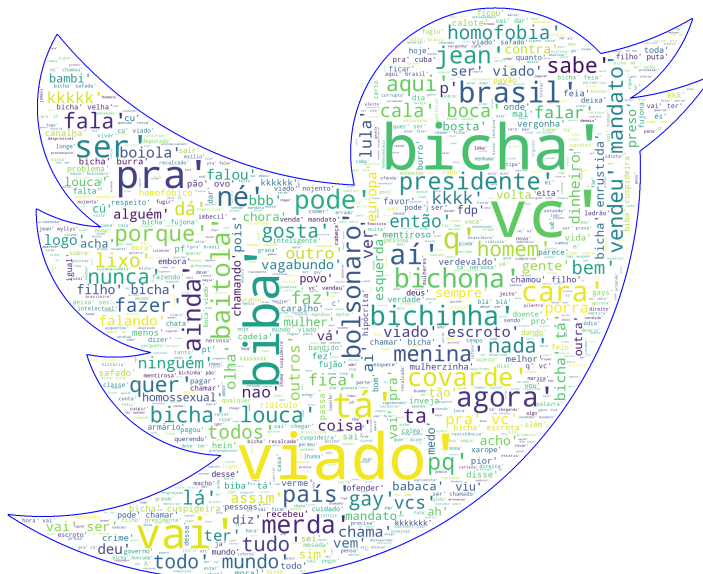


Figura 2. Nuvem de Palavra de termos frequentes.

A Figura 3 ilustra os dez termos mais frequentes dentre todos presentes nas mensagens. O objetivo desse gráfico é observar, em termos percentuais, a ocorrência de cada termo. Nota-se que, por exemplo, o termo bicha ocorre ao menos uma vez em 3,5% de todos os documentos. Os termos com maior ocorrência fazem parte do conjunto de termos do *Hatebase*. Ainda observando a Figura 3, pode-se perceber a necessidade de incluir alguns termos dentre as palavras não significantes, como o “pra” e “tá”, por exemplo. Vale

destacar que essa análise preliminar é útil para a mineração de textos por permitir uma visão macro dos termos.

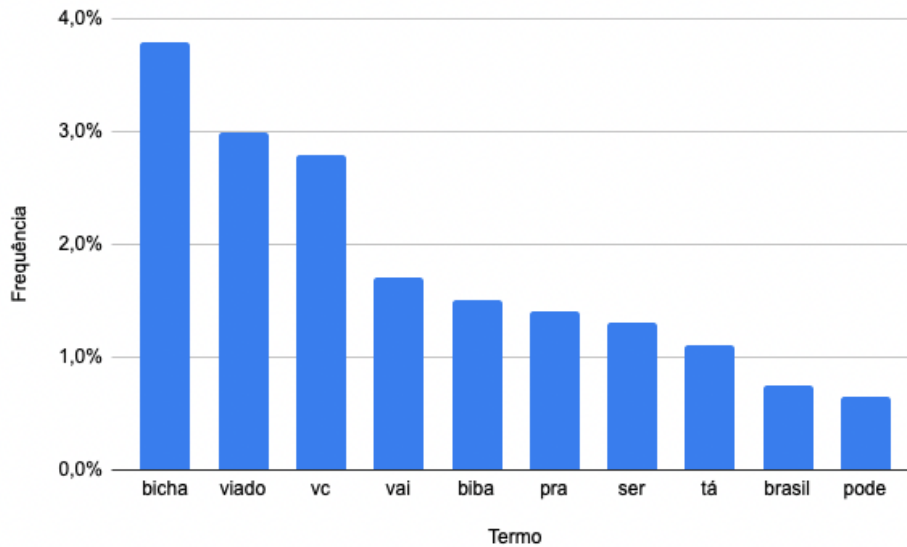


Figura 3. Os 10 Termos mais frequentes.

Seguindo a metodologia utilizada, o passo seguinte é analisar as palavras termo a termo, ou seja, pares de termos que compõem cada conteúdo textual e como estas se relacionam. A Figura 4 ajuda a visualizar e compreender os bigramas. A análise de bigramas ilustra os 20 termos mais frequentes e suas relações. Essa análise permite observar como a palavra “bicha” se relaciona ao mesmo tempo com a palavra “burra” e “enrustida”, o que ressalta que a ofensa também se refere às capacidades racionais do sujeito ou a maneira que o mesmo se reconhece enquanto sujeito.

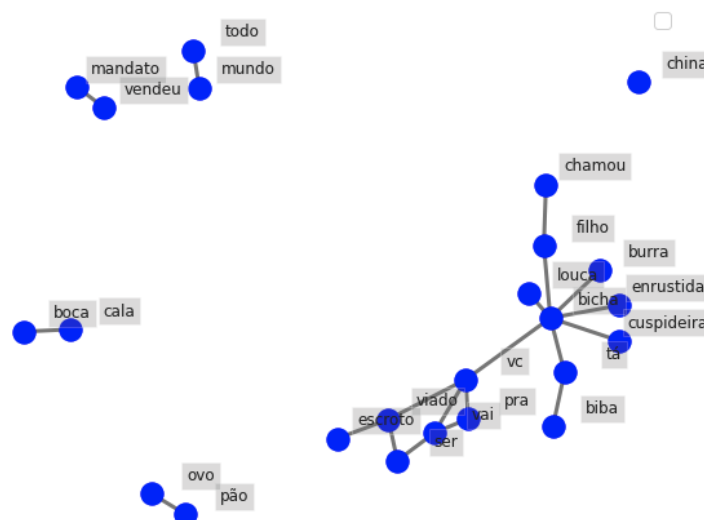


Figura 4. Análise de Bigramas dos 20 termos mais frequentes.

## 4.2. Análise de Agrupamento

A análise de agrupamento nos permite avaliar secções dentre os termos/palavras. O método avalia o quão similar estão os termos para cada grupo. Marutho et al. [2018] abordaram a análise de agrupamento para conteúdos textuais. Eles adotaram o Método do Cotovelo como métrica para identificação do melhor número  $k$  de grupos. A Figura 5 ilustra a curva do Método do Cotovelo e o tamanho adequado de grupos (i.e., 5). Um número maior poderia ser adotado, porém, como relatado em Marutho et al. [2018], não haveria ganhos expressivos e poderia tornar a análise mais difícil de se interpretar.

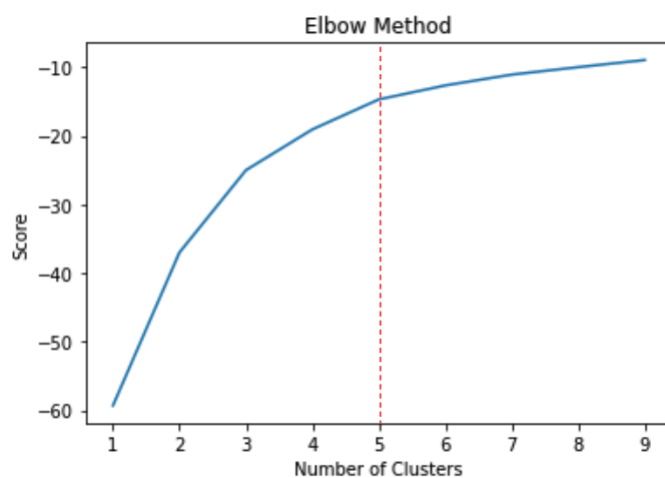


Figura 5. Análise de Agrupamento - Método do Cotovelo (*Elbow*).

Com base no número de grupos definidos, ou seja, cinco, a ideia é analisar a ocorrência de termos nesses grupos. A Figura 6 permite avaliar se o recorte de cinco grupos parece razoável *a priori*. Também percebe-se que alguns termos se misturam e a classificação dos grupos, embora tenha respeitado a análise do cotovelo, pode carecer de uma segunda abordagem para avaliar o quão boa foi a decisão.

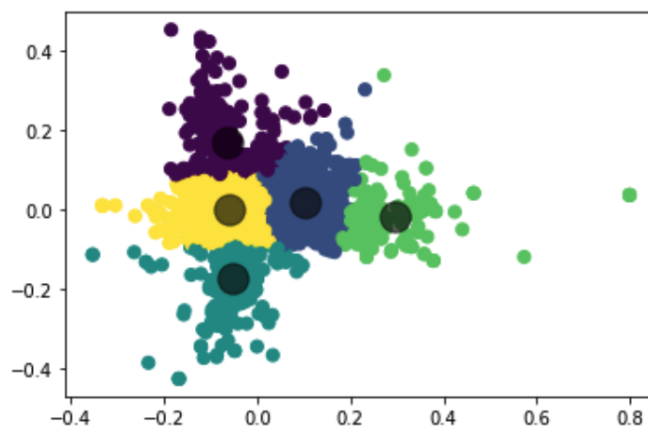
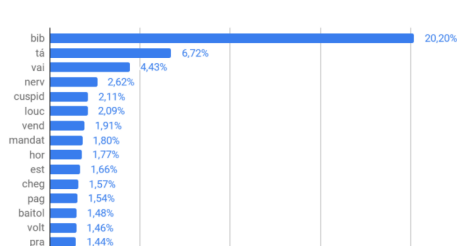


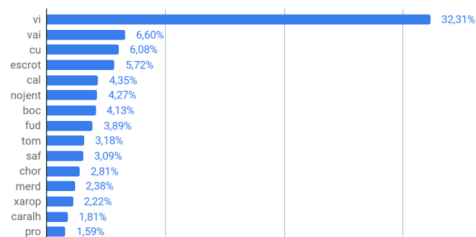
Figura 6. Visualização dos grupos.



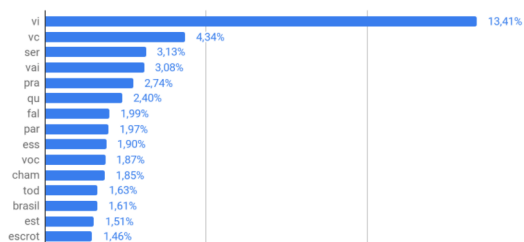
As Figuras 7 (a)-(e) referem-se as diferenças textuais presentes em cada grupo proposto. Por exemplo, no grupo 1 o termo “bib” ou “biba” é o mais frequente, com aproximadamente 20% de representação em relação a todos os termos extraídos dos textos desse grupo, seguido de palavras associadas a eventos que envolveram o político em questão. Enquanto que para o grupo 2, a palavra homofóbica “viado”, representada pelo termo tratado “vi”, apresenta uma frequência relativa de 33% em relação aos termos extraídos, além de insultos e palavras de baixo calão, nesse caso, indicando um conteúdo mais ofensivo. No grupo 3, o mesmo termo homofóbico representa aproximadamente 13% de frequência em relação aos termos dos textos agrupados, seguido de outros termos menos relevantes ou associados ao país. No grupo 4, o termo homofóbico está em 9,6% de frequência, seguido de outros termos menos relevantes ou associados ao contexto do evento de pedido de asilo político do referido ex-deputado. No grupo 5, pode-se notar que alguns termos presentes se relacionam com termos homofóbicos, como por exemplo o “enrust”, o que foi evidenciado na análise dos bigramas, apresentada na Seção 4.1.



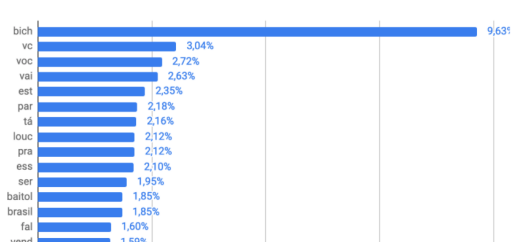
(a) Grupo 1



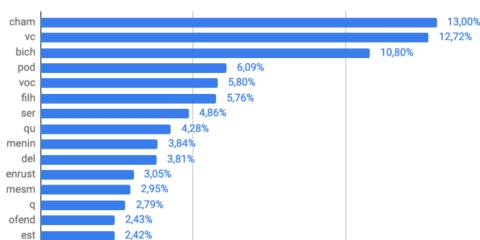
(b) Grupo 2



(c) Grupo 3



(d) Grupo 4



(e) Grupo 5

Figura 7. Termos mais frequentes nos cinco grupos. (clusters).

A análise dos grupos nos permite perceber que de fato há diferenças nos tipos de homofobia, seja pelos termos presentes no léxico, seja nos qualificadores textuais como adjetivos e substantivos. Com base no caso do grupo 2, grupo em que mais termos mostram um nítido discurso de ódio, pode-se avaliar que alguns grupos tendem a ser mais agressivos que outros.

As discussões reunidas com base nessa amostra podem ajudar futuros trabalhos a identificarem as relações dentro desses grupos, abordando novas proposições e subclassificações com relação à homofobia textual.

## 5. Conclusões

O discurso de ódio homofóbico trata-se da ação ou incitamento ao ódio, à discriminação ou à hostilização de um indivíduo, motivado por um preconceito contra pessoas de orientação sexual lésbicas, gays, bissexuais, transexuais e travestis (LGBT). Nesse cenário, este trabalho apresenta uma análise exploratória em respostas à uma mensagem do ex-deputado federal Jean Wyllys no *Twitter*, em conteúdos textuais, no idioma português, utilizando técnicas de análise de textos e léxico. O que se busca é a aplicação de tais técnicas em um tema de bastante relevância na sociedade, propiciando uma visão que permita trazer à luz casos de homofobia em redes sociais e fomentar futuros trabalhos no tema.

A análise permitiu ilustrar que o conteúdo homofóbico não é composto apenas do termo homofóbico, os textos direcionados ao ex-deputado contém o uso de outros elementos textuais acompanhados ao termo. Pelo uso de bigramas, podemos avaliar que o insulto ao sujeito ocorre de diversas maneiras, atacando sua sexualidade, sua capacidade intelectual ou seu posicionamento político.

Como limitação, evidenciamos a manutenção de termos que podem ser removidos no desenvolvimento de outras pesquisas. Assim, os próximos trabalhos podem melhorar a escolha de palavras irrelevantes com variações de palavras comumente utilizadas na internet, como por exemplo: “vc” ao invés de “você”. Nesse caso, poderiam ser utilizados conjuntos de dados que fossem formados por palavras informais (*i.e.*, NetSpeak), conforme o conjunto proposto por Nascimento et al. [2019] e analisado em Nascimento et al. [2021].

Um próximo passo importante também é buscar as classificações dos textos por meio de análise de sentimento, como conteúdos negativos, positivos e neutros. A abordagem permitirá o uso de modelos de aprendizado de máquina para classificação de textos sem classe definida.

## Referências

- Brasil (2019). *Direito Constitucional - Ação Direta de Inconstitucionalidade por omissão*. Supremo Tribunal Federal. <http://www.stf.jus.br/arquivo/informativo/documento/informativo931.htm>. Acessado em 14 de setembro de 2020.
- Burnap, P. and Williams, M. L. (2015). Cyber hate speech on twitter: An application of machine classification and statistical modeling for policy and decision making. *Policy & Internet*, 7(2):223–242. Acessado em 21 de setembro de 2020.

- Han, J., Kamber, M., and Pei, J. (2012). *Data mining concepts and techniques*, third edition.
- Hatebase, I. (2020). How it works? [https://hatebase.org/how\\_it\\_works](https://hatebase.org/how_it_works). Acessado em 13 de setembro 2020.
- Honey, C. and Herring, S. (2009). Beyond microblogging: Conversation and collaboration via twitter. pages 1 – 10. Acessado em 21 de setembro de 2020.
- MacAvaney, S., Yao, H.-R., Yang, E., Russell, K., Goharian, N., and Frieder, O. (2019). Hate speech detection: Challenges and solutions. *PLOS ONE*, 14(8):1–16. <https://doi.org/10.1371/journal.pone.0221152>. Acessado em: 2020-09-21.
- Marutho, D., Hendra Handaka, S., Wijaya, E., and Muljono (2018). The determination of cluster number at k-mean using elbow method and purity evaluation on headline news. In *2018 International Seminar on Application for Technology of Information and Communication*, pages 533–538.
- Nascimento, R., Santos, G., Carvalho, F., and Guedes, G. (2021). Avaliando contribuições na substituição de termos informais em classificação de texto de redes sociais com netspeak-br. In *Anais do X Brazilian Workshop on Social Network Analysis and Mining*, pages 181–186, Porto Alegre, RS, Brasil. SBC.
- Nascimento, R., Santos, L., and Guedes, G. (2019). Netspeak-br: Um léxico sobre expressões criadas na língua portuguesa brasileira para a internet. In *Anais do Symposium in Information and Human Language Technology*, pages 71–75, Online Event, Brazil. SBC.
- Nassif, L. F. (2011). *Técnicas de Agrupamento de Texto aplicadas à computação forense*. Acessado em 13 de setembro de 2021.
- Santana, R. (2020). *Como preparar Dados de Texto para Machine Learning*. <https://minerandodados.com.br/como-preparar-dados-de-texto-para-machine-learning>. Acessado em 13 de setembro de 2021.
- Silva, A. and Roman, N. (2020). Hate speech detection in portuguese with naïve bayes, svm, mlp and logistic regression. In *Anais do XVII Encontro Nacional de Inteligência Artificial e Computacional*, pages 1–12, Porto Alegre, RS, Brasil. SBC.
- Silva Neto, S. R. d. (2017). Uma abordagem computacional para identificação de indício de preconceito em textos baseada em análise de sentimentos. In *Proposta de Dissertação Curso de Pós-Graduação em Informática da UFAL*, Maceió, AL, Brasil. Acessado em 21 de setembro de 2020.
- S.PAULO, O. E. D. (2017). Youtuber contratado pelo governo fez posts racistas e homofóbicos no Twitter. <https://emails.estadao.com.br/noticias/gente/youtuber-contratado-por-temer-fez-posts-racistas-e-homofobicos-no-twitter-confira,70001670052>. Acessado em 16 de setembro de 2020.
- Twitter, I. (2020). Perguntas frequentes de novos usuários. <https://help.twitter.com/pt/new-user-faq>. Acessado em 13 de setembro de 2021.
- UOL (2010). *Grupo gay condena perfil homofóbico que já tem 15 mil adeptos no Twitter*. <https://noticias.uol.com.br/cotidiano/ultimas-noticias/2010/11/18/grupo-gay-condena-perfil-homofobico-do-twitter-que-ja-tem-15-mil-adeptos.htm>. Acessado em 16 de setembro de 2020.

- UOL (2020). *Entenda o que significa cada uma das letras desta bandeira de luta.* <https://www.uol.com.br/universa/amp-stories/lgbtqi-entenda-as-letras>. Acessado em 13 de setembro de 2021.
- Van Rossum, G. and Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA.