

Detecção de Posicionamento em *Tweets* sobre Política no Contexto Brasileiro

William Christie, Julio C. S. Reis, Fabrício Benevenuto
Mirella M. Moro, Virgílio Almeida

¹Universidade Federal de Minas Gerais (UFMG) – Brasil

{williamchristhie, julio.reis, fabricio, mirella, virgilio}@dcc.ufmg.br

Resumo. *Opiniões compartilhadas na Web constituem um grande volume de dados. Nelas, posicionamentos são expressos direta ou indiretamente, e sua detecção identifica qual a polaridade em relação a uma ideia alvo. Neste trabalho, apresentamos a caracterização de um amplo conjunto de tweets em português sobre a corrida presidencial brasileira de 2018. Tal conjunto serve como base para a detecção automática de posicionamento através de uma abordagem semi-supervisionada. Em nossa avaliação, encontramos indícios de bots na rede. Também avaliamos três classificadores com teste estatístico pareado, e nossos resultados apresentam F-Measure acima de 94%.*

Abstract. *Opinions shared over the Web constitute big volumes of data. Moreover, they may contain stances that are expressed directly or indirectly. Hence, stance detection may help to define the polarity related to a target idea. Here, we present the characterization of a broad set of tweets in Portuguese about the 2018 Brazilian presidential race. Such a set serves as the basis for automatic stance detection through a semi-supervised approach. In our evaluation, we find clues on the presence of bots in the network. We also evaluate three classifiers with paired statistical test, and our results present F-Measure above 94%.*

1. Introdução

As redes sociais online são um ambiente propício para o compartilhamento de vários assuntos, tais como visões políticas e opiniões pessoais sobre temas polêmicos como religião e sexualidade [Reis et al. 2015]. Dentre os milhões de usuários de redes sociais online, os brasileiros são especialmente ativos, e gastam em média 3 horas por dia no acesso a estes ambientes¹. Com tamanha riqueza de conteúdo, várias aplicações têm surgido a fim de minerar as opiniões existentes. Uma das técnicas utilizadas na mineração de opiniões é a Detecção de Posicionamentos, que consiste em identificar como o autor de determinado texto se posiciona em relação a uma proposição alvo. Tal posicionamento pode ocorrer de forma neutra, favorável ou contrária sobre proposições que podem tratar de uma causa, pessoa, produto, organização ou política de governo, por exemplo.

A detecção de posicionamentos é abordada em trabalhos recentes sobre diversos assuntos, incluindo aborto, ateísmo, clima, feminismo e a corrida presidencial americana (Hillary e Trump) [Dias and Becker 2016a, Mohammad et al. 2017]. Porém, vários

¹<https://www.techtudo.com.br/noticias/2018/02/10-fatos-sobre-o-uso-de-redes-sociais-no-brasil-que-voce-precisa-saber.ghtml>

fatores contribuem para que a tarefa não seja trivial. Por exemplo, embora existam registros favoráveis e contrários nos assuntos que envolvem controvérsia, o mesmo não ocorre em situações onde uma grande mobilização se manifesta em apenas uma das polaridades, como por exemplo os temas cura gay, aborto ou preconceito. Nesses casos, “posicionar-se abertamente” pode ter impacto nos relacionamentos internos e externos à rede, já que geralmente os usuários desses sistemas podem ser identificados no mundo real. Além deste fator, o uso de técnicas como Análise de Sentimentos não é uma solução viável, visto que os posicionamentos podem ocorrer indiretamente, através de opiniões sobre alvos relacionados [Mohammad et al. 2017]. Por exemplo, a manifestação de apoio a um determinado candidato que é contra a privatização de empresas públicas poderia ser expressa apenas com palavras negativas sobre a privatização, sem que o candidato fosse efetivamente citado no texto. Deste modo, a análise poderia classificar o texto como negativo, sendo que na verdade o texto é positivo em relação ao candidato em questão.

Embora existam trabalhos sobre Análise de Sentimentos no contexto brasileiro e para o idioma português (e.g., [Araújo et al. 2016] e [Bigonha et al. 2012]), o mesmo não ocorre com a Detecção de Posicionamentos, uma vez que não encontramos trabalhos abordando as especificidades do idioma e contexto brasileiros para o assunto (busca na DBLP em março/2018). Além disso, o desenvolvimento de técnicas e ferramentas capazes de detectar posicionamentos de usuários brasileiros é de grande importância. Por exemplo, a análise político-eleitoral do contexto brasileiro pode beneficiar-se da utilização de tais ferramentas, especialmente no monitoramento da aceitação de determinado candidato por parte da população, ou no apoio em pesquisas sobre comportamento político que são regularmente realizadas por institutos como Ibope².

Neste trabalho desenvolvemos um classificador que infere automaticamente o posicionamento expresso em *tweets* escritos em português. Especificamente, apresentamos uma metodologia que pode ser utilizada independentemente de contexto, e aplicamos ao cenário político-eleitoral brasileiro. A hipótese investigada é a de que o forte viés ideológico existente em *hashtags* pode ser utilizado para etiquetar de forma automática um conjunto de *tweets* que possa ser utilizado para treinar um classificador. Para isso, nós coletamos uma amostra do *Twitter* com mensagens favoráveis ou contrárias a quatro presidenciais, estabelecidos como alvo. Então, caracterizamos esse conteúdo e treinamos um classificador capaz de explorar diferenças psico-linguísticas expressas nos textos para a categorização das mensagens. Resultados mostram indícios de uma possível ação massiva de robôs (*bots*) atuando em polaridades e alvos específicos. Além disso, obtivemos F-Measure acima de 94% para as tarefas de classificação.

O restante desse artigo está organizado da seguinte forma. A Seção 2 descreve os trabalhos relacionados. A Seção 3 detalha a metodologia utilizada para a caracterização e classificação. A Seção 4 apresenta a análise e discussão dos resultados. A Seção 5 contém as conclusões e direções para trabalhos futuros.

2. Trabalhos Relacionados

A Detecção de Posicionamentos tem sido abordada por diferentes técnicas. Por exemplo, Shenoy et al [2017] empregam técnicas linguísticas em um processo supervisionado de aprendizado de máquina. Mourad et al. [2018] utilizam a técnica de maioria de votos

²<http://www.ibope.com.br>

entre os classificadores *Random Forest*, *Linear SVM* e *Naive Bayes*. Além disso, Dias e Becker [2016b] apresentam uma metodologia baseada em regras e análise de sentimentos para rotular os *tweets* que são utilizados para treinar um classificador SVM. Em outra abordagem, Xu et al. [2017] empregam fatores latentes para a extração de características dos textos e um *lexicon* de sentimentos para dividir o corpus em duas partes e treinar um classificador SVM. Por fim, Chen et al. [2017] abordam o problema com redes neurais convolucionais e *word embedding*.

De outro modo, Mohammad et al. [2017] demonstram que o sentimento expresso em um *tweet* é benéfico, mas não suficiente para a tarefa de classificação do posicionamento. Eles baseiam-se na ideia de que uma pessoa pode expressar a mesma posição em relação a um alvo usando linguagem negativa ou positiva. Além dessas, outra proposta para detecção de posicionamento em texto de *tweets* de forma não-supervisionada é apresentada em [Dias and Becker 2016a]. O algoritmo recebe como entrada apenas o alvo e um conjunto de *tweets* a rotular. Ele é baseado em uma abordagem híbrida composta pelas etapas de rotulação automática baseada em um conjunto de heurísticas e classificação complementar baseada em aprendizado supervisionado de máquina.

Entre a diversidade de estudos publicados sobre análise de sentimento, existem os que tratam especificamente do contexto político-eleitoral. Por exemplo, o cenário da eleição presidencial americana de 2016 é estudado em [Caetano et al. 2017], onde os autores utilizam análise de sentimentos para estudar a homofilia política entre os usuários do Twitter. Já no contexto específico brasileiro, Verona et al. [2017] estudam um ponto diferente: o de poder em redes sociais avaliado através das doações de campanha para o Senado Federal. Entretanto, não existem trabalhos que abordem a detecção de posicionamentos para textos em português no contexto político-eleitoral brasileiro.

3. Metodologia

A metodologia utilizada está dividida em duas partes, de acordo com as principais contribuições do trabalho: a primeira é a coleta e a descrição dos dados; e a segunda é a construção de *datasets* e o treinamento dos classificadores. A seguir, cada etapa é descrita individualmente.

3.1. Coleta e Descrição dos Dados

A API do Twitter³ foi utilizada para coletar mensagens (*tweets*) sobre possíveis candidatos para as eleições presidenciais de 2018⁴. Neste estudo quatro presidenciáveis foram considerados em função do volume de mensagens: Alckmin, Bolsonaro, Dória e Lula.

Especificamente, definimos um conjunto de *hashtags* (Tabela 1) através de pesquisa manual para selecionar aquelas com forte viés ideológico por assunto, observando sua ocorrência em *tweets* favoráveis ou contrários à ideia alvo. Embora o número de etiquetas a favor e contra esteja desbalanceado, o total de registros coletados não é diretamente proporcional à essa quantidade. Por exemplo, o mesmo número de *hashtags* favoráveis e contrárias sobre o assunto Lula é utilizado para a coleta, porém o total de mensagens extraído é maior para o posicionamento contrário. O assunto Dória também apresenta situação semelhante, pois a maioria das mensagens coletadas está na polaridade

³<https://dev.twitter.com>

⁴<http://politica.estadao.com.br/noticias/geral,veja-quem-quer-ser-presidente-em-2018,70002054149>

Tabela 1. Hashtags utilizadas para a coleta

Candidato	Favor	Contra
Alckmin	#geraldoAlckmin, #alckmin2018, #geraldopresidente, #geraldoAlckminPresidente, #geraldoAlckmin2018	#ForaAlckmin, #alckminladrão
Bolsonaro	#BolsonaroMito, #BolsoMito, #BolsonaroPresidente	#ForaBolsonaro, #ForaBolsoLixo, #BolsonaroDitador
Dória	#doriaPresidente, #doria2018, #joaodoriapresidente	#foraDoria
Lula	#LulaPresidente, #TocomLula, #LulaPeloBrasil	#ForaLula, #LulaNaCadeia, #ForaPT

que possui menos etiquetas. Assim, o fator de maior influência no total de *tweets* coletados é a importância dada pelos usuários para cada *hashtags*, e não o número de etiquetas. Embora neste trabalho tenhamos utilizado esse conjunto de *hashtags*, a metodologia pode ser utilizada para outros conjuntos mais amplos.

Ao final da coleta, dois conjuntos de dados são obtidos para cada presidenciável pesquisado (*tweets* favoráveis, e *tweets* contrários). Em seguida, as *hashtags* são removidas do texto original para que elas não enviesem o processo de treinamento do classificador. Além do texto, cada *tweet* inclui informações adicionais sobre o usuário, como nome e identificador, além de informações sobre a mensagem, tais como *timestamp* e indicador de repostagem. Também é importante destacar que, embora a API restrinja o volume da coleta a 1% dos *tweets* públicos, isso representa um volume significativo de mensagens, visto que diariamente cerca de 500 milhões de *tweets*⁵ são compartilhados. Por fim, utilizamos os conjuntos de dados extraídos e etiquetados como favoráveis ou contrários a cada assunto para efetuar uma análise estatística descritiva da coleta.

3.2. Treinamento do Classificador

Composição dos Datasets de treino. Criamos conjuntos de dados para treinar os classificadores utilizando os registros dos dois assuntos mais comentados após a filtragem efetuada durante a caracterização. O primeiro contém 41.498 *tweets* em português com o assunto Lula, sendo 27.717 contrários e 13.781 favoráveis. Porém, para o segundo *dataset*, o volume de *tweets* contrários coletado é muito baixo (159), inviabilizando a aplicação direta das técnicas de aprendizado de máquina. Para contornar tal situação, analisamos a viabilidade de utilização das mensagens de temas distintos na elaboração do *dataset*. Assim, a Figura 1 apresenta o cruzamento efetuado entre os usuários que postam mensagens em cada assunto, de modo a verificar a interseção entre os conjuntos. Os maiores valores encontrados são entre os usuários que postam: contrários a Dória e Alckmin (11,29%), além de contrários a Lula e favoráveis a Bolsonaro (5,72%), sendo que todas as demais apresentam valores baixos. Especificamente, a interseção entre os usuários que postam a favor de Lula e são favoráveis ao Bolsonaro é apenas 2,03%, e a interseção entre Lula Contra e Bolsonaro Contra é de somente 0,14%. Assim, os valores sugerem que os usuários realmente se posicionam em grupos contrários. Nota-se ainda que as interseções podem representar pessoas indecisas ou que mudaram de opinião. No contexto brasileiro tal cenário faz sentido, uma vez que esses presidenciáveis posicionam-se em extremos antagônicos de direita e esquerda. Com isso, utilizamos os dados favoráveis ao Lula como sendo contrários ao Bolsonaro para compor o segundo *dataset*, que contém 31.366 *tweets*, sendo 13.940 contrários e 17.426 favoráveis.

Extração de características. A estruturação dos dados em forma de texto é necessária

⁵<http://www.internetlivestats.com/twitter-statistics/>

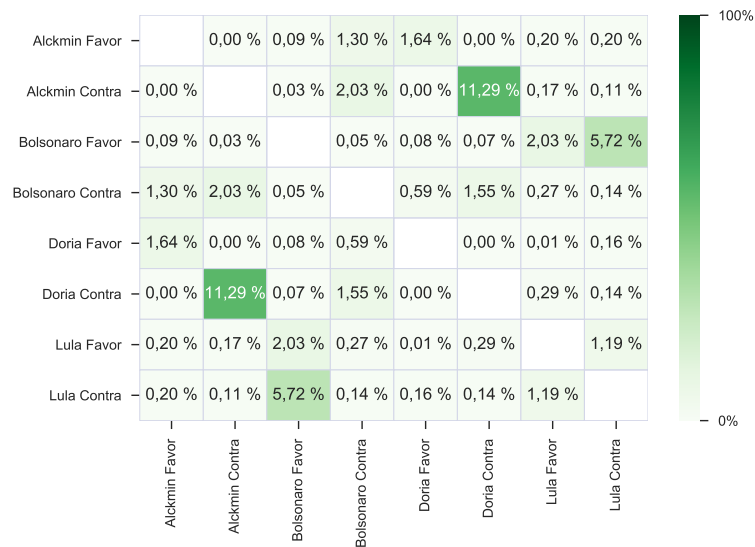


Figura 1. Cruzamento dos conjuntos de usuários que postam em cada classe

Tabela 2. Importância das características (ganho de informação)

Bolsonaro		Lula	
Importância	Característica (feature)	Importância	Característica (feature)
0,0149893	hora (hour)	0,13919312	todas pontuações (AllPunc)
0,0126505	dia da semana (dayOfWeek)	0,12782251	vírgulas (Colon)
0,0002846	lugar (place)	0,11818538	palavras de dicionário (Dic)
0,0000573	maiores que 6 letras (SixlTr)	0,10034218	outras pontuações (OtherP)
0,0000464	palavras com função gramatical (funct)	0,08829323	palavras com função gramatical (funct)

para possibilitar a utilização dos *tweets* por um classificador. Assim, utilizamos a ferramenta LIWC⁶ para a extração de *features* dos textos em português de cada *dataset*. O software efetua a análise psicolinguística do texto e fornece um vetor com suas características quantitativas morfológicas, gramaticais e psicolinguísticas. O texto é então removido do *dataset* e as colunas com o processamento do LIWC são adicionadas.

Importância das características. A fim de investigar o poder discriminativo das *features* criadas e a possível redução de dimensionalidade dos dados, criamos um *ranking* com o ganho de informação de cada característica para a predição da classe. Desse modo, filtramos os atributos para manter somente aqueles com importância acima de zero, uma vez que os demais adicionam ruído no processo classificatório. Para exemplificar o processo, os cinco melhores atributos encontrados na análise são apresentados na Tabela 2. Ao final da etapa, restam 56 atributos para o *dataset* Bolsonaro e 82 para o *dataset* Lula.

Teste estatístico. Neste trabalho utilizamos os mesmos classificadores aplicados por Mourad et al. [2018] (*Naive Bayes*, *SVM* e *Random Forest*) para comprovar a hipótese inicial de que a técnica semi-supervisionada de etiquetamento funciona para o contexto brasileiro. Assim, executamos um teste-t pareado para comparar o desempenho de tais classificadores. Para tal, a ferramenta Weka Experimenter⁷ é utilizada, uma vez que ela

⁶<https://liwc.wpengine.com>

⁷<https://www.cs.waikato.ac.nz/ml/weka/>

Tabela 3. Quantitativos de dados extraídos

Candidato	Extração			Idioma Português			% Português		Datas	
	Contra	Favor	Ambos	Contra	Favor	Ambos	Contra	Favor	Início	Fim
Alckmin	167	691	3	144	624	3	86%	90%	21/10/17	01/01/18
Bolsonaro	178	19.691	0	135	18.183	0	76%	92%	21/10/17	02/01/18
Dória	161	99	0	115	77	0	71%	78%	21/10/17	01/01/18
Lula	91.012	36.721	144	35.105	30.380	90	39%	83%	21/10/17	02/01/18

Tabela 4. Idiomas, maior número de postagens e de idiomas por usuário

Candidato	Posicionamento	# Idiomas	Maior # Postagens	Maior # Idiomas
Alckmin	Contra	3	72	2
Alckmin	Favor	5	441	2
Bolsonaro	Contra	3	21	3
Bolsonaro	Favor	19	152	4
Doria	Contra	4	26	2
Doria	Favor	4	20	2
Lula	Contra	35	11.882	33
Lula	Favor	16	1.229	6

permite fácil acesso as técnicas de aprendizado de máquina [Witten et al. 2016]. Realizamos 15 repetições entre os classificadores para cada *dataset*, utilizando a métrica F-Measure para *5-folds Cross Validation* em cada repetição, totalizando 75 medições para cada algoritmo. A possibilidade de rejeição da hipótese nula de que os classificadores são iguais é verificada com um nível de significância de 5%.

4. Resultados e Discussões

Nesta seção, apresentamos e discutimos os resultados, que estão divididos em caracterização dos dados coletados e classificação automática. Além disso, discutimos alguns pontos de ameaça à validade deste estudo.

4.1. Caracterização

Estatísticas Básicas. Caracterizamos os dados coletados para entender melhor sua composição e evitar que os classificadores deem maior importância para as características de *tweets* provenientes de postagens automatizadas e recorrentes. A Tabela 3 contém o volume coletado para cada tema, bem como o viés ideológico do conjunto de *hashtags*, visto que poucas mensagens aparecem com etiquetas dos dois posicionamentos (a favor e contra). Essas mensagens são descartadas, de modo a garantir a separação dos dados para treinamento dos classificadores. Além disso, também pode ser observado na Tabela 3 que o maior número de *tweets* aborda o assunto Lula, e a maioria dos presidentiáveis não apresenta grande diferença entre os totais de *tweets* em português e em outros idiomas. Por outro lado, somente 38% das mensagens abordando o assunto Lula possuem o idioma português, o que pode ser um indício da presença de *bots*.

Estatísticas de Idioma. Analisamos a divisão por idiomas envolvendo cada categoria (candidato e polaridade) para investigar os valores observados no assunto Lula. Então, selecionamos para cada categoria o maior número de mensagens postadas por um único usuário e o maior número de idiomas utilizados por um único usuário. Os resultados são apresentados na Tabela 4, que apresenta um grande número de idiomas distintos em algumas categorias, sendo 35 contrários ao Lula, 19 favoráveis ao Bolsonaro, e 16 favoráveis ao Lula. Investigamos manualmente alguns desses *tweets* e, aparentemente, o idioma

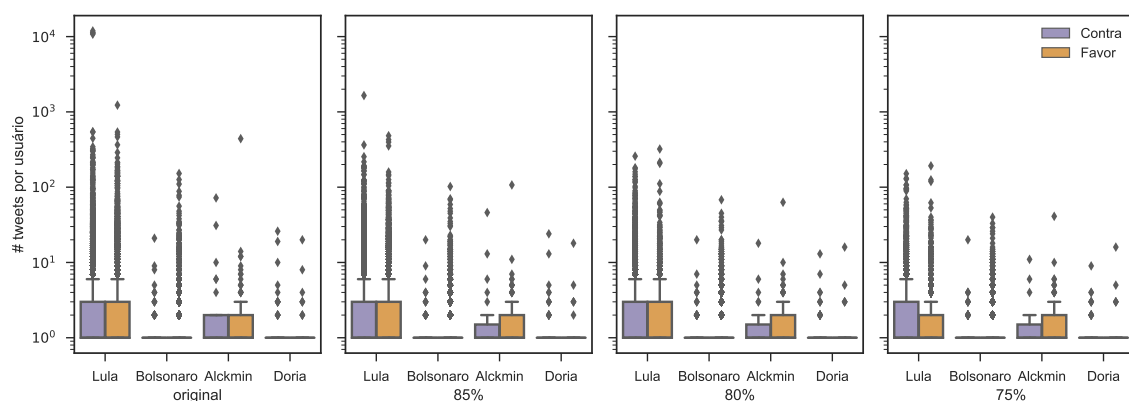


Figura 2. Distribuição do volume de *tweets* por usuário em cada classe com diferentes níveis de similaridade

de mensagens formadas por links, muitas *hashtags* e poucas palavras não é identificado corretamente. Um grande número desse tipo de mensagem foi encontrado para os temas citados, o que corrobora tal hipótese da existência de *bots* postando mensagens automaticamente. Também corrobora com a hipótese o volume de 33 idiomas distintos para um único usuário com *tweets* contrários ao Lula, e o alto volume de mensagens por usuário, sendo 11.882 *tweets* na categoria contrário ao Lula e 1.229 na favorável ao Lula.

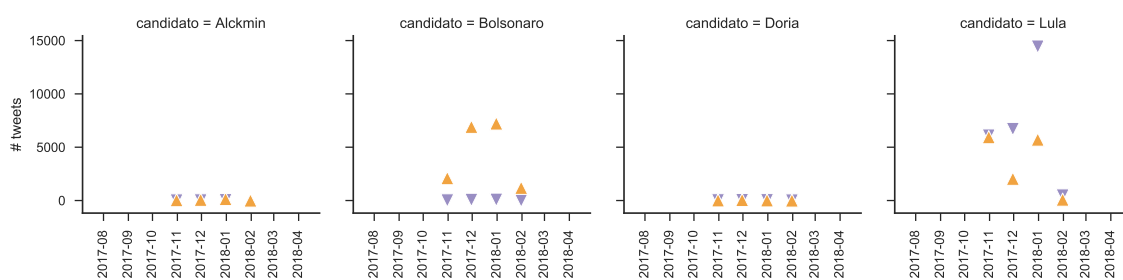
Volume por Usuário. A partir da possibilidade identificada da existência de *bots*, efetuamos uma análise mais detalhada da distribuição do volume de postagens por usuários para investigá-la. A Figura 2 apresenta as distribuições de volume de *tweets* por usuário em cada categoria. O primeiro *boxplot* representa a distribuição original dos dados coletados, no qual pode-se observar o grande número de mensagens postadas por usuários *outliers* em todos os temas. A exclusão de mensagens iguais postadas por um mesmo usuário não é suficiente para remover os *outliers*, uma vez que a maioria das mensagens possui pequenas alterações entre si.

Filtragem por Similaridade. Como a exclusão por igualdade não serve, adotamos a remoção de *tweets* através de funções de similaridade. Em [Christen 2006], várias dessas funções são testadas para a deduplicação de nomes. Além disso, essas funções são comparadas nas etapas de deduplicação de registros considerando 11 domínios distintos em [Silva et al. 2017]. Conforme os autores, a função *Jaro-winkler* está entre as que possuem o melhor desempenho. Por este motivo, neste estudo a função é utilizada para mensurar a semelhança entre dois textos.

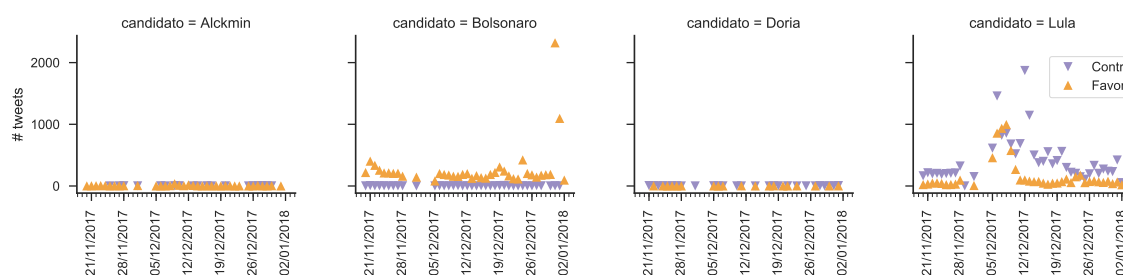
Especificamente, todas as mensagens pertencentes a um mesmo usuário e uma mesma classe são comparadas entre si utilizando-se três opções *threshold* para a definição de textos similares: 75%, 80% e 85%. Assim, os resultados das exclusões de mensagens identificadas como similares em cada um dos níveis podem ser observados nos três últimos *boxplots* da Figura 2. A remoção de *tweets* semelhantes nos três níveis de similaridade reduz o volume de mensagens dos usuários *outliers*. Entretanto, constatamos em inspeção manual que algumas mensagens similares continuam presentes nos níveis 85% e 80%. Além disso, a inspeção demonstra que muitas das mensagens semelhantes são removida no nível 75%, mas que o nível 70% remove mensagens legítimas. Deste modo, neste trabalho os *tweets* com 75% de semelhança são removidos para o restante

Tabela 5. Estatísticas do *dataset* antes e depois da remoção de *tweets* similares.

Candidato	Dataset original					Tweets duplicados removidos (75%)				
	Usuários	<i>tweets</i>	\bar{x}	75%	max	Usuários	<i>tweets</i>	\bar{x}	75%	max
Alckmin Contra	39	167	4 ± 12	2	72	39	64	2 ± 2	2	11
Alckmin Favor	122	691	5 ± 35	2	387	122	249	2 ± 4	2	41
Bolsonaro Contra	109	178	2 ± 2	1	21	109	159	1 ± 2	1	20
Bolsonaro Favor	14.832	19.691	1 ± 2	1	126	14.832	17.426	1 ± 1	1	40
Dória Contra	85	161	2 ± 3	1	26	85	114	1 ± 1	1	9
Dória Favor	61	99	2 ± 2	1	18	61	86	1 ± 2	1	16
Lula Contra	10.356	91.012	9 ± 222	3	11.877	10.356	27.717	3 ± 5	3	151
Lula Favor	7.051	36.721	5 ± 23	3	1.228	7.051	13.781	2 ± 4	2	192



(a) Distribuição mensal



(b) Distribuição diária destacando os meses de dezembro e janeiro

Figura 3. Volume de *tweets* em função do tempo

das caracterizações apresentadas.

A Tabela 5 apresenta estatísticas do conjunto de dados antes e depois da remoção das mensagens similares para analisar seu impacto. Aqui, 50% dos usuários de todas as classes possuem um único *tweet*, e por questões de simplicidade estas colunas não foram apresentadas na tabela. Observa-se uma pequena alteração no número de *tweets* da faixa de 75% dos usuários para a categoria a favor do Lula, que passa de quatro para três *tweets* com a remoção. Além disso, o total de usuários de cada classe não sofre alterações. Por outro lado, ocorre grande redução na média, desvio padrão e na quantidade de postagens do usuário que mais posta em cada categoria (coluna *max*) com a remoção de mensagens semelhantes. Como pode ser observado, o impacto da remoção de *tweets* semelhantes ocorre principalmente nos 25% dos usuários que mais postam, o que comprova o comportamento desejado de remoção de *outliers*.

Análise Temporal. Também analisamos a distribuição de *tweets* ao longo do tempo, de modo a complementar a caracterização. Assim, a Figura 3 apresenta: (a) o volume de postagens mensais em cada classe, onde é possível verificar um aumento de postagens

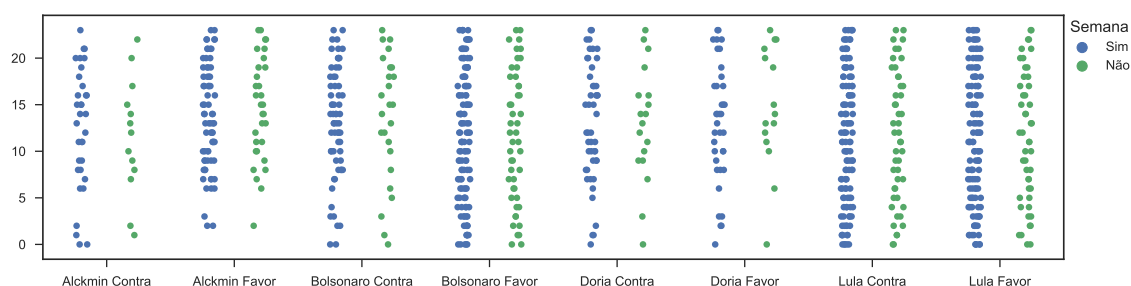


Figura 4. Tweets por horários da semana

sobre os presidenciáveis Bolsonaro e Lula nos meses dezembro de 2017 e janeiro de 2018; (b) mais detalhes sobre o período a fim de esclarecer tal comportamento, onde verifica-se um aumento dos *tweets* sobre o tema Lula em datas de eventos específicos. Como exemplo, por volta do dia 12/12/2017 houve a marcação do julgamento do recurso impetrado pela defesa de Lula na segunda instância para o dia 24/01/2018⁸, e no dia 30/12/2017 houve a chegada dos militares no RN devido à paralisação de policiais⁹.

A fim de refinar o estudo da distribuição de *tweets* no tempo, a Figura 4 apresenta a ocorrência de postagens nos horários do dia (eixo vertical). Para cada presidenciável a imagem apresenta duas colunas, uma para os dias da semana (à esquerda em azul) e outra para o final de semana (à direita em verde). Um comportamento mais intenso e bem distribuído pode ser notado para as categorias favorável e contrário ao Lula, bem como favorável ao Bolsonaro. Por outro lado, nota-se um comportamento mais disperso para as demais categorias. A frequente existência de postagens entre 3:00 e 5:00 para os três posicionamentos mais comentados pode significar que ainda existam postagens automatizadas (*bots*) no *dataset*. Entretanto, assumimos que essas mensagens não têm alto poder de enviesar o classificador, uma vez que pelo processo de remoção de mensagens semelhantes, a frequência de um mesmo *tweet* deve ser baixa. De qualquer modo, o desenvolvimento de ferramentas de detecção automática de *bots* em *tweets* escritos em português é necessária para validar os resultados.

Termos mais Frequentes. A Figura 5 apresenta uma nuvem de termos para cada posicionamento do *dataset*. Um padrão observado é a existência do termo “fora Temer” em todos os posicionamentos contrários, com exceção ao Lula, onde o termo não aparece. Outro padrão é a existência do termo “Minas Gerais” nos posicionamentos favorável e contrário a Lula, mas de modo mais intenso no posicionamento favorável. Também chama a atenção a falta de termos em destaque para o posicionamento a favor de Alckmin e a existência do termo “pais falavam” e “bullingDFuturo” para o posicionamento favorável ao Bolsonaro. Tais termos são utilizados de forma ironizada, demonstrando sarcasmo, sendo que existem inclusive comunidades sobre o assunto nas redes sociais.

⁸<https://g1.globo.com/rs/rio-grande-do-sul/noticia/julgamento-de-lula-no-caso-triplex-em-segunda-instancia-e-marcado-para-24-de-janeiro-no-trf4.ghtml>

⁹<http://www1.folha.uol.com.br/cotidiano/2017/12/1947131-com-pm-parada-exercito-chega-ao-rn-e-forcas-armadas-assumem-a-operacao.shtml>



Figura 5. Nuvem de termos

Tabela 6. Teste t pareado

Dataset	Naive Bayes	SVM	Random Forest
Bolsonaro	0,03156 ± 0,00365	0,99419 ± 0,00103	0,43404 ± 0,01210
Lula	0,64877 ± 0,00856	0,94185 ± 0,00225	0,93702 ± 0,00220

4.2. Avaliação do Classificador

O modelo criado para classificação dos dados é avaliado na Tabela 6, que apresenta o F-measure médio e o desvio padrão obtidos. O classificador SVM é estatisticamente superior aos demais classificadores para todos os *datasets* analisados, com um nível de significância de 5%. Porém, é importante observar a diferença encontrada para o classificador *Random Forest* nas duas bases de dados. Enquanto o classificador obteve um acerto de 93% para o *dataset* Lula, para o *dataset* Bolsonaro o acerto foi de somente 43%. Da mesma forma o classificador *Naive Bayes* encontrou maior dificuldade para classificar os registros do *dataset* Bolsonaro, conseguindo acertar somente 3%. Uma hipótese para o desempenho observado é o processo de elaboração do conjunto de dados Bolsonaro. Tendo em vista que os dados foram formados por características psico-linguísticas de usuários comentando sobre dois presidentes distintos, o processo de inferência dos dois classificadores foi de alguma forma prejudicado.

4.3. Ameaças à Validade

Alguns pontos precisam ser observados quanto à validade dos resultados encontrados para os classificadores. O primeiro diz respeito às suposições de existência de *bots* nas postagens da rede social. Encontramos diversas evidências da existência de postagens automatizadas, embora este não seja um dos objetivos diretos deste trabalho. Porém, uma investigação mais detalhada sobre sua existência e extensão é necessária, uma vez que a frequente repetição de postagens pode influenciar no classificador. A segunda consideração diz respeito à presença de sarcasmo e ironia. Embora neste trabalho não tenhamos abordado diretamente o assunto, detectamos indícios de sua existência em termos frequentes para a classe a favor do Bolsonaro, o que também pode influenciar nos resultados apresentados. A última consideração refere-se à elaboração do *dataset* Bolsonaro. Reconhecemos que não necessariamente um *tweet* que seja favorável ao presidente Lula também demonstre posicionamento ideológico contrário ao presidente Bolsonaro. Embora nossos estudos evidenciem a separação entre os usuários defensores de

cada candidato, as mensagens podem conter posicionamentos exclusivamente sobre um assunto, sem indicar necessariamente oposição ao outro.

5. Conclusões e Direções Futuras

Este trabalho aborda a detecção automática de posicionamentos ideológicos em textos escritos em português. Para tal, coletamos *tweets* sobre quatro presidentiáveis para a corrida presidencial brasileira de 2018. Além disso, os dados são caracterizados e possíveis mensagens postadas por *bots* são identificadas e removidas. Utilizamos o viés ideológico das *hashtags* de coleta para construir dois *datasets* etiquetados. Enquanto um contém somente dados sobre o tema Lula, o outro contém dados de dois temas. Especificamente, devido ao baixo volume de mensagens contrárias ao Bolsonaro, utilizamos as mensagens favoráveis ao Lula como sendo contrárias ao Bolsonaro. Assim, os dois *datasets* foram utilizados para treinar dois modelos de classificação usando três algoritmos clássicos de aprendizado de máquina em cada, os quais foram testados com um teste estatístico.

Os resultados obtidos com a caracterização dos dados coletados apontam indícios de *bots* atuando principalmente em postagens dos dois presidentiáveis mais comentados. De fato, são removidas 59,93% das mensagens inicialmente coletadas por serem muito similares entre si e pertencerem a um mesmo usuário e categoria. Além disso, a grande maioria delas trata do tema Lula (57,98%), sendo que do total de cada categoria são removidas 69,55% contra e 62,47% a favor. Por fim, os resultados demonstram que as técnicas de aprendizado de máquinas exploram as diferenças linguísticas existentes entre textos favoráveis e contrários ao alvo definido, possibilitando a detecção de posicionamentos, mesmo no contexto multi polarizado político-eleitoral brasileiro. Neste sentido, conseguimos um F-Measure médio de 99% para o *dataset* Bolsonaro e 94% para o *dataset* Lula, sendo que em ambos o classificador SVM obteve o melhor desempenho.

Como trabalhos futuros, pretendemos utilizar um processo de expansão automática das *hashtags* de coleta através de um processamento de verificação do poder discriminativo das *hashtags* por classe. Essas *hashtags* poderiam retro-alimentar o processo, possibilitando um novo ciclo de coletas e análises. Além disso, pretendemos analisar outros alvos sobre assuntos em outros contextos. Sobre a acurácia dos classificadores, pretendemos utilizar um conjunto de dados aleatório como *baseline* para as comparações nas análises realizadas, o que possibilitaria mensurar o comportamento dos classificadores em relação a textos que não tratam do assunto alvo. Além disso, pretendemos realizar uma avaliação da melhoria dos classificadores com a adição de novas *features*, extraídas com outras técnicas como *bag-of-words* e *Word2Vec*, dentre outras.

Agradecimentos. Trabalho parcialmente financiado por CNPq, CAPES e FAPEMIG.

Referências

- Araújo, M., Reis, J., Pereira, A., and Benevenuto, F. (2016). An Evaluation of Machine Translation for Multilingual Sentence-level Sentiment Analysis. In *Proceedings of the ACM Symposium on Applied Computing*, pages 1140–1145.
- Bigonha, C., Cardoso, T. N. C., Moro, M. M., Gonçalves, M. A., and Almeida, V. A. F. (2012). Sentiment-based influence detection on twitter. *Journal of the Brazilian Computer Society*, 18(3):169–183.

- Caetano, J. A. C., Lima, H. S. L., dos Santos Santos, M. F., and Marques-Neto, H. T. M.-N. (2017). Utilizando análise de sentimentos para definição da homofilia política dos usuários do Twitter durante a eleição presidencial americana de 2016. In *BraSNAM - Brazilian Workshop on Social Network Analysis and Mining*, pages 480–491.
- Chen, Y.-C., Liu, Z.-Y., and Kao, H.-Y. (2017). Ikm at semeval-2017 task 8: Convolutional neural networks for stance detection and rumor verification. In *International Workshop on Semantic Evaluation*, pages 465–469.
- Christen, P. (2006). A comparison of personal name matching: Techniques and practical issues. In *IEEE International Conference on Data Mining*, pages 290–294.
- Dias, M. and Becker, K. (2016a). An Heuristics-Based, Weakly-Supervised Approach for Classification of Stance in Tweets. In *IEEE/WIC/ACM International Conference on Web Intelligence*, pages 73–80.
- Dias, M. and Becker, K. (2016b). Detecção semi-supervisionada de posicionamento em tweets baseada em regras de sentimento. In *SBBDD - Simpósio Brasileiro de Bancos de Dados*, pages 40–51.
- Mohammad, S. M., Sobhani, P., and Kiritchenko, S. (2017). Stance and sentiment in tweets. *TOIT*, 17(3):26.
- Mourad, S. S., Shawky, D. M., Fayed, H. A., and Badawi, A. H. (2018). Stance detection in tweets using a majority vote classifier. In *International Conference on Advanced Machine Learning Technologies and Applications*, pages 375–384.
- Reis, J. C., Gonçalves, P., Araújo, M., Pereira, A. C., and Benevenuto, F. (2015). Uma abordagem multilíngue para análise de sentimentos. In *BraSNAM - Brazilian Workshop on Social Network Analysis and Mining*.
- Shenoy, G. G., Dsouza, E. H., and Kübler, S. (2017). Performing stance detection on Twitter data using computational linguistics techniques. *arXiv preprint arXiv:1703.02019*.
- Silva, L. S., do Amaral, D. C., and Moro, M. M. (2017). Uma avaliação de eficiência e eficácia da combinação de técnicas para deduplicação de dados. In *SBBDD - Simpósio Brasileiro de Bancos de Dados*, pages 160–171.
- Verona, L. V., Oliveira, J. O., and Campos, M. L. M. C. (2017). Métricas para análise de poder em redes sociais e sua aplicação nas doações de campanha para o senado federal brasileiro. In *BraSNAM - Brazilian Workshop on Social Network Analysis and Mining*, pages 544–554.
- Witten, I. H., Frank, E., Hall, M. A., and Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.
- Xu, X., Hu, F., Du, P., Wang, J., and Li, L. (2017). Efficient stance detection with latent feature. In *APWeb-WAIM - First International Joint Conference on Web and Big Data*, pages 21–30.