Investigating sets of linguistic features for two sentiment analysis tasks in Brazilian Portuguese web reviews

Miguel V. Oliveira
mvo.snf18@uea.edu.br
Laboratório de Sistemas Inteligentes,
Universidade do Estado do Amazonas
Manaus, Amazonas

Tiago de Melo
tmelo@uea.edu.br
Laboratório de Sistemas Inteligentes,
Universidade do Estado do Amazonas
Manaus, Amazonas

ABSTRACT

Identifying subjective sentences and classifying the polarity of subjective sentences are two important tasks in sentiment analysis. Besides being a hot topic, there is still a lack of resources to perform the mentioned sentiment analysis tasks in the Portuguese language, with its syntactic specificities. This paper describes the identified challenges and next steps in an initial study regarding the classification of subjectivity and polarity of sentences with a small set of syntactic features extracted directly from the text. Our approach reached satisfying results in experiments with two classic machine learning models in four datasets consisting of user reviews from different domains.

KEYWORDS

natural language processing, sentiment analysis, web reviews

1 INTRODUÇÃO

Com o advento da Web 2.0, milhões de pessoas publicam comentários sobre todos os tipos de produtos que são oferecidos *online*. Isto resulta em um grande número de opiniões que podem ser uma valiosa fonte de informação para a tomada de decisão de fabricantes e potenciais clientes. Os fabricantes podem obter um *feedback* imediato para melhorar a qualidade dos seus produtos e os clientes podem avaliar essas opiniões antes de efetuar a compra de um produto. Porém, o grande volume de comentários excede a capacidade humana de processar e interpretar as opiniões e, desta forma, acabam se tornando uma barreira no uso efetivo destes dados [8, 12].

Atualmente, as técnicas de análise de sentimentos (AS), também conhecidas como mineração de opinião, são popularmente aplicadas para lidar com as opiniões textuais publicadas na Web. AS é o campo de estudo que investiga as opiniões, sentimentos, avaliações, atitudes e emoções de indivíduos em relação a diversos tipos de entidades, como produtos, serviços, organizações, pessoas, eventos e seus diversos aspectos [13].

A análise de sentimentos, quando aplicada em texto, utiliza técnicas de processamento de linguagem natural (PLN) para realizar duas principais tarefas: (1) classificação de sentenças em factuais (objetivas) ou opinativas (subjetivas) e (2) identificação da polaridade dos sentimentos das sentenças subjetivas como positiva, negativa ou neutra.

In: XVII Workshop de Trabalhos de Iniciação Científica (WTIC 2020), São Luís, Brasil. Anais Estendidos do Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia). Porto Alegre: Sociedade Brasileira de Computação, 2020.
© 2020 SBC – Sociedade Brasileira de Computação.
ISSN 2596-1683

A tarefa de identificar a subjetividade consiste em classificar um dado texto como sendo subjetivo ou objetivo. Por exemplo, a sentença *"ontem eu fui na loja"* é factual, portanto, deve ser classificada como objetiva. Já a sentença *"adorei a resolução do celular"* traz uma opinião sobre um aspecto do produto e, portanto, deve ser classificada como subjetiva. Já a tarefa de identificar ou calcular a polaridade envolve classificar a opinião de um dado texto. Por exemplo, a sentença *"adorei a resolução do celular"* mencionada anteriormente tem uma opinião positiva.

Devido à relevância das duas tarefas de análise de sentimentos, existem muitos trabalhos que tratam destas duas tarefas em textos em inglês [18, 19, 23]. Porém, ainda existe necessidade de mais investigação em outros idiomas, como o português [3, 16]. A falta de ferramentas de processamento de texto e de dados anotados para experimentos figuram como desafios para a análise de sentimentos neste idioma. Além disso, conforme reportado por Pereira [16], ainda há muito espaço para o desenvolvimento de métodos de análise de sentimentos para a língua portuguesa que explorem especificidades linguísticas.

Motivados por esses desafios, nós propusemos e avaliamos um conjunto de características linguísticas dos comentários escritos em português do Brasil para lidar com as duas tarefas de classificação da análise de sentimentos, utilizando dois métodos clássicos de aprendizagem de máquina. Nós avaliamos os métodos com *datasets* de domínios diferentes e comparamos os resultados com um conjunto representativo de *baselines*. Os resultados obtidos nos experimentos iniciais foram apresentados em um artigo científico [15]. Os resultados reportados demonstram a efetividade das características extraídas diretamente do texto para a classificação baseada em aprendizagem de máquina, mas também apontam as limitações da nossa abordagem inicial.

No presente trabalho, nós apresentaremos as limitações identificadas e os próximos passos de uma investigação mais profunda do uso de *features* textuais para as tarefas de identificação de subjetividade e cálculo de polaridade. Mais especificamente, nós pretendemos investigar o uso de novas *features*, a aplicação de métodos baseados em técnicas de redes neurais e adição de novos conjuntos de dados.

O artigo está organizado da seguinte maneira. Na Seção 2, nós apresentamos os principais trabalhos relacionados com a nossa pesquisa. Na Seção 3, nós descrevemos o método já desenvolvido e também a continuidade de investigação dos próximos métodos. Os resultados já obtidos são apresentados e discutidos na Seção 4. Na Seção 5, discutiremos os próximos passos na realização desta pesquisa. Finalmente, na Seção 6, apresentamos as conclusões e trabalhos futuros.

2 TRABALHOS RELACIONADOS

Nesta seção apresentaremos um conjunto de trabalhos relacionados em três principais eixos que formam esta pesquisa.

2.1 Classificação de subjetividade e polaridade

Na literatura, há uma vasta quantidade de trabalhos que tratam das duas tarefas de análise de sentimentos para a língua inglesa. Muitos desses trabalhos lidam individualmente com os problemas, como [21] e demais outros trabalhos descritos em [18, 19, 23].

Na língua portuguesa, essas duas tarefas são tratadas por uma quantidade muito menor de trabalhos, que em geral apresentam métodos baseados em léxicos e ontologias para tratar as tarefas individualmente [16]. Porém, existem trabalhos que tratam a identificação de subjetividade e polaridade no âmbito da aprendizagem de máquina, como Moraes *et al* [14] e Britto e Pacífico [4].

2.2 Características linguísticas

A exploração de conjuntos de características linguísticas aplicadas a essas duas tarefas vem sendo tema de interesse da academia dos últimos anos, e o estudo de *features* extraídas do texto foi apontado por Cambria [5] como um dos caminhos de evolução da Análise de Sentimentos.

Em uma avaliação empírica, Chenlo e Losada [6] investigaram uma grande variedade de *features* textuais aplicadas em quatro *datasets* em inglês, a fim de lidar com identificação de sentenças subjetivas e cálculo de polaridade. Almatarneh e Gamallo [1] apresentaram uma comparação entre *features* linguísticas e abordagens baseadas em dicionário para identificar a polaridade de opiniões extremas (como extremamente negativo ou extremamente positivo), também em inglês.

O nosso conjunto de *features* proposto obteve bons resultados iniciais, mais ainda é necessário explorar novos conjuntos de características com maior abrangência em diferentes tipos de textos em Português. Até onde sabemos, não existem trabalhos na literatura que descrevam estudos aprofundados em *features* linguísticas para o Português do Brasil.

2.3 Modelos de classificação

Os modelos clássicos de aprendizagem de máquina supervisionada são comumente empregados em tarefas da análise de sentimentos [14, 24]. Em Moraes *et al* [14], os autores avaliaram Naïve Bayes e a Otimização Mínima Sequencial para a classificação de sentenças subjetivas em tuítes.

No entanto, é cada vez mais frequente que tais modelos sejam comparados com os de redes neurais. Britto e Pacífico [4] compararam uma série de modelos como Naïve Bayes, Máquinas de Vetores de Suporte, Árvores de Decisão, Redes Neurais Recorrentes e *Long Short-Term Memory* para o cálculo de polaridade em sentenças extraídas de comentários sobre aplicativos móveis.

3 MÉTODOS E MATERIAIS

Na realização desta pesquisa, nós lidamos com as duas tarefas da análise de sentimentos a nível de sentença. As duas tarefas foram tratadas como um problema de classificação binária, e para ambas, investigamos o uso de métodos clássicos de aprendizagem de máquina. As características linguísticas dos textos foram codificadas em uma representação vetorial.

3.1 Conjunto de características (*features*)

O conjunto de características investigado é composto por sete *features* e foi utilizado pelos modelos para as duas tarefas. Todas as *features* foram extraídas utilizando o modelo em Português do pacote spaCy[1], que provê informações de classes gramaticais e árvores de dependência para os textos processados.

A seguir, nós apresentamos o conjunto de *features* utilizado nos nossos experimentos e uma breve descrição:

- **Número de palavras.** A contagem total de palavras na sentença.
- **Número de adjetivos.** O número de adjetivos identificados na sentença.
- **Número de substantivos.** O número de substantivos identificados na sentença.
- **Número de advérbios.** A contagem de advérbios presentes na sentença.
- **Número de termos superlativos.** A quantidade de termos superlativos na sentença.
- **Número de termos comparativos.** A contagem de termos comparativos presentes na sentença.
- **Número de `amod`.** A contagem de relações *amod*, que está presente quando um adjetivo modifica um substantivo em uma sentença.

Durante a fase inicial da pesquisa, nós avaliamos o impacto de remover cada uma das *features* no desempenho classificadores, para todos os conjuntos de dados. Após a análise dos resultados, foi possível identificar que nenhuma das *features* se sobressai como mais relevante na decisão do classificador linear. Isso indica que a combinação do conjunto de *features* funciona melhor do que a utilização individual de uma *feature*.

3.2 Modelos de aprendizagem de máquina

Para a primeira fase da nossa investigação, nós consideramos um modelo linear e outro não-linear para a classificação. O modelo linear adotado foi a Máquina de Vetores de Suporte (MVS), um modelo bastante efetivo na classificação de texto, cujo objetivo é encontrar um hiperplano que divida as instâncias em classes [7]. Já o modelo não-linear foi *Gradient Boosting Trees* (GBT), que consiste em um conjunto de árvores de decisão otimizadas por uma função de perda [17].

3.3 *Datasets*

Quatro diferentes conjuntos de dados foram utilizados na avaliação do nosso método. Estes *datasets* variam em domínio, tamanho e balanceamento. Considerando que o desempenho dos métodos de classificação são afetados por essas características nos dados, avaliar um método em conjuntos de dados com diferentes configurações permite estimar sua robustez. A seguir, apresentamos uma breve descrição dos *datasets*.

[1] https://spacy.io/models/pt

- **TA-Restaurantes:**[2] Nós fizemos uma coleta de 350 comentários de usuários sobre restaurantes no Brasil no site TripAdvisor[3], no período de janeiro a março de 2020. Os textos dos comentários foram divididos em 1.049 sentenças e estas foram manualmente anotadas por nós. Inicialmente, nós classificamos as sentenças em objetivas e subjetivas. Em seguida, nós classificamos as sentenças subjetivas em positivas e negativas. As sentenças que eram subjetivas, mas com polaridade neutra, foram desconsideradas. Na prática, nós observamos que existem poucas sentenças que são subjetivas e neutras.
- **ReLi Corpus:** é um *corpus* composto por um conjunto de 350 resenhas de livros em português publicadas na Internet e anotado manualmente [11]. As resenhas foram igualmente divididas em sentenças objetivas e subjetivas, sendo que as sentenças subjetivas foram igualmente divididas em relação à polaridade (positivas e negativas).
- **Computer-BR:** o *corpus* é composto de 2.281 tuítes que foram coletados do período de janeiro a fevereiro de 2015 [14]. As mensagens são comentários da área de tecnologia sobre notebooks. As mensagens foram divididas em sentenças e anotadas tanto para identificar a subjetividade como a polaridade. Porém, em ambos os casos, as classes não estão balanceadas.
- **Google Play:** o *corpus* original é composto de 10.000 avaliações de aplicativos móveis publicadas na loja virtual Google Play Store[4]. As sentenças foram anotadas manualmente apenas para a identificação da polaridade [10]. Nós consideramos um subgrupo de 1630 sentenças escolhidas aleatoriamente, igualmente distribuídas entre negativas e positivas [20].

Após a análise dos resultados iniciais, nós identificamos algumas limitações envolvendo os conjuntos de dados.

A primeira delas envolve o balanceamento das classes de polaridade no *dataset* TA-Restaurantes. Nós entendemos que o fato de as sentenças negativas representarem menos de 10% do conjunto de sentenças subjetivas comprometeu a avaliação do método para a tarefa de cálculo de polaridade quando aplicada a esse conjunto.

Além disso, nós identificamos a necessidade de explorar melhor o conjunto de dados com textos extraídos do Twitter. Considerando a natureza diferente da escrita nessa plataforma, nós acreditamos que outras combinações de *features*, juntamente com o pré-processamento adequado, possam ter melhores resultados em tuítes do que em avaliações publicadas em sites de avaliação de produtos e serviços.

4 EXPERIMENTOS

Para obter os resultados iniciais utilizando essa abordagem, nós realizamos um conjunto de experimentos e comparamos os valores obtidos com um conjunto de *baselines*. A descrição de cada um destes métodos, assim como as razões para sua escolha estão descritas no artigo onde os resultados foram publicados inicialmente.

[2]Disponível em https://data.mendeley.com/datasets/hsn6g3dbsk
[3]https://www.tripadvisor.com.br
[4]https://play.google.com/store

4.1 Métricas

As métricas utilizadas para a avaliação do método foram precisão (P), revocação (R) e medida-F (F_1). Seja A o conjunto de respostas corretas, de acordo com um conjunto de referência, e seja B o conjunto de respostas produzidas pelo método que está sendo avaliado. Nós definimos precisão (P), revocação (R) e medida-F (F_1) como:

$$P = \frac{|A \cap B|}{|B|} \qquad R = \frac{|A \cap B|}{|A|} \qquad F_1 = \frac{2 \times (P \times R)}{(P + R)}$$

4.2 Baselines

Com o objetivo de avaliar os métodos propostos, nós comparamos com um conjunto de métodos que representam o estado da arte no tratamento dos problema de identificação de subjetividade e cálculo da polaridade de texto em português. Por razões de limitação de espaço, nós recomendamos a leitura dos artigos publicados em [2, 15] para um resumo completo dos métodos.

4.3 Resultados

A Tabela 1 apresenta a comparação dos resultados para a tarefa de identificação de subjetividade, em termos de precisão (P), revocação (R) e medida-F (F_1). Os maiores valores para cada métrica em cada *dataset* estão marcados em negrito. Nos experimentos, nós executamos uma validação cruzada de 10 *folds* sobre o conjunto de dados. O método Máquina de Vetores de Suporte será denominado MVS, enquanto o *Gradient Boosting Trees* será referido como GBT.

Table 1: Resultados experimentais para a tarefa de identificação de subjetividade.

	Reli			**TA-Restaurantes**			**Computer-BR**		
	P	R	F_1	P	R	F_1	P	R	F_1
EmoticonsDS	0.49	**0.97**	0.65	0.58	**0.97**	0.72	0.26	**0.99**	0.42
Stanford	0.49	0.94	0.65	0.57	0.87	0.69	0.27	0.94	0.42
Sentiment140	0.47	0.85	0.61	0.58	0.83	0.68	0.27	0.86	0.41
NRC Hashtag	0.48	0.77	0.59	0.56	0.60	0.58	0.27	0.79	0.40
SentisStrenght	0.53	0.29	0.38	0.64	0.26	0.37	0.23	0.30	0.26
SentiWordNet	0.51	0.46	0.49	0.63	0.25	0.36	0.28	0.56	0.37
MVS	0.73	0.65	0.68	**0.74**	0.78	0.76	**0.42**	0.68	**0.51**
GBT	**0.76**	0.68	**0.71**	0.71	0.91	**0.80**	0.39	0.34	0.36

De acordo com os resultados apresentados na Tabela 1, os métodos MVS e GBT alcançaram os melhores resultados gerais. Apenas o método *EmoticonsDS* conseguiu obter os melhores resultados de revocação para os três *datasets*. No entanto, este método obteve um valor mais baixo de F_1, devido à precisão nos três conjuntos de dados. Ao compararmos a média de F_1 entre os métodos MVS e GBT nos três *datasets*, nós observamos que o método MVS alcançou um resultado superior.

A Tabela 2 apresenta a comparação dos resultados para a tarefa de cálculo da polaridade, em termos de precisão (P), revocação (R) e medida-F (F_1). Os maiores valores para cada métrica em cada *dataset* estão marcados em negrito. Para esses experimentos, novamente nós executamos uma validação cruzada de 10 *folds* sobre o conjunto de dados. O método Máquina de Vetores de Suporte será denominado MVS, enquanto o *Gradient Boosting Trees* será referido como GBT.

De acordo com os resultados apresentados na Tabela 2, os métodos MVS e GBT novamente alcançaram os melhores resultados

Table 2: Resultados experimentais para a tarefa de cálculo de polaridade.

	Reli			TA-Restaurantes			Computer-BR			Google Play		
	P	R	F_1	P	R	F_1	P	R	F_1	P	R_1	F
EmoticonsDS	**0.75**	0.50	0.34	0.90	0.87	0.83	0.10	0.32	0.16	0.72	*0.49*	*0.35*
Stanford	0.49	0.46	0.33	0.85	0.11	0.08	0.44	**0.62**	0.52	0.64	0.39	0.35
Sentiment140	0.55	0.47	0.50	0.84	0.68	0.75	0.63	0.50	**0.54**	0.73	0.60	0.62
NRC Hashtag	0.51	0.40	0.45	0.78	0.31	0.43	0.59	0.47	0.52	0.45	0.31	0.35
SentisStrenght	0.69	0.21	0.32	0.91	0.23	0.37	**0.69**	0.19	0.28	**0.76**	0.24	0.36
SentiWordNet	0.56	0.27	0.36	0.86	0.15	0.24	0.58	0.27	0.35	0.51	0.17	0.25
MVS	0.48	0.52	0.49	**0.92**	0.94	0.92	0.38	0.74	0.50	0.52	**0.96**	0.67
GBT	0.57	**0.64**	**0.59**	0.90	**0.99**	**0.95**	0.44	0.44	0.43	0.69	0.68	**0.69**

gerais. No entanto, o método *Sentiment140* alcançou o melhor resultado de F_1 no *dataset* Computer-BR, composto por textos do Twitter. Nós entendemos que isto aconteceu porque o método foi desenvolvido especificamente para a tarefa de cálculo de polaridade em textos extraídos dessa plataforma.

5 PRÓXIMOS PASSOS

Considerando as limitações que foram apresentadas nas seções anteriores, nós planejamos ampliar esta pesquisa em algumas direções.

Primeiramente, nós pretendemos explorar outros conjuntos de características linguísticas para o Português, de maneira análoga à investigação de Chenlo e Losada [6] na língua inglesa. O objetivo é realizar um estudo empírico mais aprofundado para descobrir quais características textuais extraídas de comentários em língua portuguesa são mais relevantes para um modelo classificador na realização das tarefas de identificação de subjetividade e cálculo da polaridade.

Segundo, nós pretendemos construir um *dataset* mais equilibrado ao aumentar o número de instâncias negativas. Já coletamos uma grande quantidade de comentários e iremos anotar uma porção manualmente para enriquecer o conjunto de dados. Além disso, nós desejamos explorar outros conjuntos de dados públicos que estejam anotados com informações que possibilitem a avaliação para as duas tarefas da análise de sentimentos.

Finalmente, nós pretendemos investigar o uso de outros tipos de classificadores para lidar com as tarefas investigadas neste trabalho. Apesar dos dois modelos terem apresentados bons resultados, nós pretendemos incluir arquiteturas baseadas em redes neurais que estão em evidência no cenário da AS [22, 24], tais como as Redes Neurais Recorrentes, os modelos de *Long Short-Term Memory* (LSTM) e BERT (*Bidirectional Encoder Representations from Transformers*) [9], que vêm obtendo os melhores resultados para diversas tarefas no processamento de linguagem natural.

6 CONSIDERAÇÕES FINAIS

Neste artigo, nós apresentamos os próximos passos em uma investigação no uso de características linguísticas para duas tarefas de análise de sentimentos aplicadas em comentários extraídos da web.

A abordagem que apresentamos é baseada em aprendizagem de máquina e em sete características linguísticas dos comentários escritos em português. Os resultados reportados demonstram que o método é efetivo, mas indicam que ainda é necessário explorar novos conjuntos de características para melhorar o desempenho dos classificadores. Nós também planejamos investigar outros modelos de aprendizagem de máquina e de redes neurais, a fim de obter um método com melhor desempenho na classificação de sentenças subjetivas e cálculo da polaridade em Português.

REFERENCES

[1] Sattam Almatarneh and Pablo Gamallo. 2018. Linguistic features to identify extreme opinions: an empirical study. In *International Conference on Intelligent Data Engineering and Automated Learning*. Springer, 215–223.

[2] Matheus Araújo, Pollyanna Gonçalves, Meeyoung Cha, and Fabrício Benevenuto. 2014. iFeel: a system that compares and combines sentiment analysis methods. In *Proceedings of the 23rd International Conference on World Wide Web*. 75–78.

[3] Pedro Paulo Balage Filho. 2017. *Aspect extraction in sentiment analysis for portuguese language*. Ph.D. Dissertation. Universidade de Sao Paulo.

[4] Larissa Britto and Luciano Pacífico. 2019. Análise de Sentimentos para Revisões de Aplicativos Mobile em Português Brasileiro. In *Anais do XVI Encontro Nacional de Inteligência Artificial e Computacional*. SBC, 1080–1090.

[5] Erik Cambria, Björn Schuller, Yunqing Xia, and Catherine Havasi. 2013. New avenues in opinion mining and sentiment analysis. *IEEE Intelligent systems* 28, 2 (2013), 15–21.

[6] Jose M Chenlo and David E Losada. 2014. An empirical study of sentence features for subjectivity and polarity classification. *Information Sciences* 280 (2014), 275–288.

[7] Corinna Cortes and Vladimir Vapnik. 1995. Support-vector networks. *Machine learning* 20, 3 (1995), 273–297.

[8] Tiago de Melo, Altigran S da Silva, Edleno S de Moura, and Pável Calado. 2019. OpinionLink: Leveraging user opinions for product catalog enrichment. *Information Processing & Management* 56, 3 (2019), 823–843.

[9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018).

[10] Fernando Leandro dos Santos and Marcelo Ladeira. 2014. The role of text preprocessing in opinion mining on a social media language dataset. In *2014 Brazilian Conference on Intelligent Systems*. IEEE, 50–54.

[11] Cláudia Freitas, Eduardo Motta, Ruy L Milidiú, and Juliana César. 2014. Sparkling vampire... lol! annotating opinions in a book review corpus. *New Language Technologies and Linguistic Research: A Two-Way Road* (2014), 128–146.

[12] Bum Chul Kwon, Sung-Hee Kim, Timothy Duket, Adrián Catalán, and Ji Soo Yi. 2015. Do people really experience information overload while reading online reviews? *International Journal of Human-Computer Interaction* 31, 12 (2015), 959–973.

[13] Bing Liu. 2015. *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge university press.

[14] Silvia MW Moraes, André LL Santos, Matheus Redecker, Rackel M Machado, and Felipe R Meneguzzi. 2016. Comparing approaches to subjectivity classification: A study on portuguese tweets. In *International Conference on Computational Processing of the Portuguese Language*. Springer, 86–94.

[15] Miguel Oliveira and Tiago Melo. 2020. Investigating Sentences Features for Subjectivity and Polarity Classification in Brazilian Portuguese. In *Anais do XVII Encontro Nacional de Inteligência Artificial e Computacional*. SBC, Porto Alegre, RS, Brasil, 270–281. https://doi.org/10.5753/eniac.2020.12135

[16] Denilson Alves Pereira. 2020. A survey of sentiment analysis in the Portuguese language. *Artificial Intelligence Review* (2020), 1–29.

[17] J Ross Quinlan. 1996. Boosting first-order learning. In *International Workshop on Algorithmic Learning Theory*. Springer, 143–155.

[18] Kumar Ravi and Vadlamani Ravi. 2015. A survey on opinion mining and sentiment analysis: tasks, approaches and applications. *Knowledge-Based Systems* 89 (2015), 14–46.

[19] Kim Schouten and Flavius Frasincar. 2015. Survey on aspect-level sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering* 28, 3 (2015), 813–830.

[20] Milton Stiilpen Júnior. 2016. *Um arcabouço de processamento de textos informais em português brasileiro para aplicações de mineração de dados*. Master's thesis. Universidade Federal de Ouro Preto.

[21] Suge Wang, Deyu Li, Xiaolei Song, Yingjie Wei, and Hongxia Li. 2011. A feature selection method based on improved fisher's discriminant ratio for text sentiment classification. *Expert Systems with Applications* 38, 7 (2011), 8696–8702.

[22] Ashima Yadav and Dinesh Kumar Vishwakarma. 2020. Sentiment analysis using deep learning architectures: a review. *Artificial Intelligence Review* 53, 6 (2020), 4335–4385.

[23] Lin Yue, Weitong Chen, Xue Li, Wanli Zuo, and Minghao Yin. 2019. A survey of sentiment analysis in social media. *Knowledge and Information Systems* (2019), 1–47.

[24] Lei Zhang, Shuai Wang, and Bing Liu. 2018. Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8, 4 (2018), e1253.