

Detecção de Posicionamento em Dados Textuais

O estado da arte e considerações

José Lucas Silva Freitas

Universidade Federal de Campina Grande

Campina Grande, Paraíba

jsslucassf@gmail.com

ABSTRACT

The task of stance detection in textual data can be understood as given some text, automatically identifying whether the author is in favor of, against or even neutral about a certain target. The target may be an idea, a person, a company or even a product. For instance, automatically detecting a politician's positioning towards a certain government policy on his speeches. The text to be analysed can even come from a normal person posting in social media platforms like Twitter or Facebook.

In this work, many papers that address this problem were studied. Our goal being to better understand the nature of the problem, listing the techniques that were already used in dealing with it and identifying possible open areas that may be explored in future works.

The solution to this problem may be really important for various systems based on interaction with users, based on text written by humans. One of the main applications that were mentioned in

the works that we studied was the detection of fake news, mostly due to the troublesome social and political moment, in which we see ourselves nowadays.

KEYWORDS

Texto, Posicionamento, Linguagem, notícias

ACM Reference Format:

José Lucas Silva Freitas. 2018. Detecção de Posicionamento em Dados Textuais: O estado da arte e considerações. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 INTRODUÇÃO

Politicamente e socialmente, vivemos tempos conturbados. A cada dia novos temas polêmicos são trazidos para discussão da sociedade e cada vez mais, as pessoas se sentem no dever de se posicionar. Entender tais posicionamentos pode ser muito útil, para entender o espírito do posicionamento da sociedade em relação aos temas mais importantes. Além disto, entender o posicionamento das pessoas pode ter várias outras aplicações, como desenvolvimento de aplicações específicas, e mais recentemente ajudar na detecção de notícias falsas, as *fake news* [Referência].

Em reconhecimento da importância deste problema, vários trabalhos utilizando diversas técnicas para detecção automática de posicionamento

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

Conference'17, July 2017, Washington, DC, USA

© 2018 Copyright held by the owner/author(s).

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.

<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

foram publicados. A associação de linguística computacional (ACL) organiza anualmente o *SemEval*,¹ um workshop com vários desafios na área da análise de semântica computacional, dentre estes desafios, algumas tarefas relacionadas a detecção de posicionamento já foram propostas. Estudar a contribuição de todos estes trabalhos e os métodos por eles utilizados é muito importante pois possibilita o entendimento do estado da arte e a identificação de possíveis territórios de pesquisa ainda não explorados.

Iniciamos com uma visão geral da área da mineração de textos. Explorando o estado da arte no processamento de textos, mostramos de uma forma geral, como são realizadas as tarefas de processamento de texto. Então falamos um pouco mais sobre a aplicação no contexto especialmente das fake news. Prosseguimos para esclarecer as diferenças entre duas tarefas da análise semântica, a detecção de posicionamento e a detecção de sentimento. Nas próximas sessões, exploramos abordagens que usam os textos, regras, e outras informações relevantes, como as informações temporais e a estrutura das conversas (quando analisamos textos produzidos em redes sociais).

2 MINERAÇÃO DE TEXTO

- Porquê falar de text mining

Quando tentamos detectar o posicionamento dos autores em um texto, estamos implicitamente buscando extrair informações semânticas importantes e não triviais.

- O que é text mining - Framework conceitual - Problemas em aberto

- 3 FAKE NEWS E APLICAÇÕES
- 4 DETECÇÃO DE POSICIONAMENTO X DETECÇÃO DE SENTIMENTO
- 5 ABORDAGENS BASEADAS APENAS NO TEXTO
- 6 REGRAS DE CLASSIFICAÇÃO
- 7 UTILIZANDO OUTRAS INFORMAÇÕES RELEVANTES
 - 7.1 Informações temporais
 - 7.2 Estrutura das Conversas em Redes Sociais
- 8 CONCLUSÕES
- REFERENCES

¹https://aclweb.org/aclwiki/SemEval_Portal