

A language agnostic framework for analysis of political content in OSNs

Despoina Antonakaki
Dimitris Spiliotopoulos
Polyvios Pratikakis
Sotiris Ioannidis
Paraskevi Fragopoulou

Institute of Computer Science (ICS) of the Foundation for Research and Technology - Hellas (FORTH)
N. Plastira 100
Heraklion, Crete, Vassilika Vouton GR70013, Greece
despoina@ics.forth.gr

ABSTRACT

Today a significant part of social discussion regarding political affairs is taking place in Online Social Networks (OSNs). The analysis of this content can reveal interesting patterns and tendencies on the social sentiment towards political notions, institutions, events, ideologies, and public figures. Conclusions from this analysis can be of extreme importance in figuring out the major themes that dominate the online discourse and assessing the public positive and negative sentiment towards these themes. Through this analysis we can also measure the effect of certain political actions or events in the public sentiment and assess the efficacy of political campaigns in general. The main toolbox of this research includes Natural Language Processing (NLP) methods, lexicons for sentiment analysis, techniques for topic analysis (like Latent Dirichlet Allocation) and methods for sarcasm identification and correction [1].

Existing studies in this area are focusing in specific electoral events whilst analyzing content from a single language. Yet, 2017 has been characterised as a politically crucial year for Europe due to the plethora of electoral events that are happening in major European countries. Therefore, in this study we attempt to collectively analyze the OSN content regarding these events. In particular, as of April 2017 we have collected the entire corpus of tweets referring to the Dutch general elections (March) and to the French Presidential Elections (April). We are currently collecting tweets referring to the French legislative election (June), the United Kingdom general elections (June) and we also plan to include tweets from the German federal elections (September) and the Czech legislative elections (October).

The purpose of this ongoing study is to assess both the commonalities and the important differences between the political online discussions in various European regions. To accomplish this, we augment the existing analysis toolbox with language agnostic techniques and automatic translation methods that uniformly process

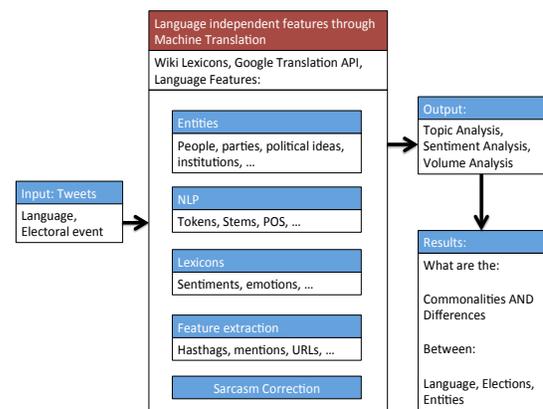


Figure 1: A schematic preview of the proposed Language Agnostic analysis workflow.

all languages and minimize potential biases. In this paper we report our preliminary results based on four electoral events in three (Dutch, French, English) languages.

CCS CONCEPTS

• **Information systems** → **Information extraction; Sentiment analysis;** • **Networks** → **Online social networks;** • **Computing methodologies** → **Natural language processing;**

ACM Reference format:

Despoina Antonakaki, Dimitris Spiliotopoulos, Polyvios Pratikakis, Sotiris Ioannidis, and Paraskevi Fragopoulou. 2017. A language agnostic framework for analysis of political content in OSNs. In *Proceedings of ACM Conference, Washington, DC, USA, July 2017 (Conference'17)*, 1 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

REFERENCES

- [1] Despoina Antonakaki, Dimitris Spiliotopoulos, Christos V Samaras, Sotiris Ioannidis, and Paraskevi Fragopoulou. 2016. Investigating the complete corpus of referendum and elections tweets. In *Advances in Social Networks Analysis and Mining (ASONAM), 2016 IEEE/ACM International Conference on*. IEEE, 100–105.

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
© 2017 Copyright held by the owner/author(s).
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00
<https://doi.org/10.1145/nnnnnnn.nnnnnnn>