

Sentiment Analysis of Tweets for Estimating Criticality and Security of Events

V. Subramaniaswamy, School of Computing, SASTRA University, Thanjavur, India

R. Logesh, School of Computing, SASTRA University, Thanjavur, India

M. Abejith, School of Computing, SASTRA University, Thanjavur, India

Sunil Umasankar, School of Computing, SASTRA University, Thanjavur, India

A. Umamakeswari, School of Computing, SASTRA University, Thanjavur, India

ABSTRACT

Social Media has become one of the major industries in the world. It has been noted that almost three fourth of the world's population use social media. This has instigated many researches towards social media. One such useful application is the sentimental analysis of real time social media data for security purposes. The insights that are generated can be used by law enforcement agencies and for intelligence purposes. There are many types of analyses that have been done for security purposes. Here, the authors propose a comprehensive software application which will meticulously scrape data from Twitter and analyse them using the lexicon based analysis to look for possible threats. They propose a methodology to obtain a quantitative result called criticality to assess the level of threat for a public event. The results can be used to understand people's opinions and comments with regard to specific events. The proposed system combines this lexicon based sentimental analysis along with deep data collection and segregates the emotions into different levels to analyse the threat for an event.

KEYWORDS

Artificial Intelligence, Human-Computer Interaction, Lexicon Based Sentimental Analysis, Security, Social Aspects, Social Media Analytics

1. INTRODUCTION

Social media websites are growing at a very fast rate throughout the world. The increase in the reach of the Internet to more people every day is also helping this. It is turning out to be one of the major industries in the world and is revolutionising several existing industries such as print media and marketing. Organisations are also benefitting from this advancement as they can now enhance their processes such as understanding customer's opinions, analysing reviews and feedbacks. However, with the increased availability of people's data online, there is an increasing concern for people's privacy and security. The advancement in social media access and increased available features has also started attracting potential hackers. This makes it possible for several crimes to occur such as selling private information, usage of such social sites to do secret transactions that are banned. Due

DOI: 10.4018/JOEUC.2017100103

to the massive size of the data available online, sometimes activities and potential threats can also go undetected. This motivated the development of Application Programming Interface that is freely available for people to use to create software that can track these crimes online. But owing to user's privacy these tools have limits to the amount of data that they can get. There are trackers inside these tools that ban people automatically the people that try to over use it. There are many undergoing projects in many universities that endlessly utilize everything they have to build an efficient product that helps law enforcement agencies to detect crimes in a better way.

In this work we focus on applying sentimental analysis to Twitter data to obtain tweets that can possibly lead to disruptions during an event. To use the Internet and obtain Intelligence, monitoring important sites is necessary. Useful data can be obtained for law enforcement purposes from social networking sites that can be processed for obtaining insights, detecting threats, making predictions and performing many other analyses. Such analyses will provide private organizations, authorities and law enforcement agencies to make decisions and obtain a better understanding of their people. Most analyses, including this analysis use the data that are publicly visible on the web. Users always have the option to choose their data to be publicly or privately visible. The trust and privacy gets breached when an organisation or analysis uses data that is wrongfully obtained. The proposed system is a data collection and text mining application for Intelligence and Law enforcement agencies.

A three step method is proposed for obtaining useful insights and the steps are as follows:

- Obtaining data from Social media sites like Twitter in the form of raw data and tweets by identifying important keywords and locations that are of interest, such as specific public events or places. The data is obtained by crawling periodically and using the APIs provided by these sites such as search and streaming APIs;
- A lexicon based on SentiWordNet was generated to obtain sentiment values for words;
- On the collected real time data in the knowledge base, lexicon based sentimental analysis is done to calculate the emotional sentimental value for every entity. Sentiment values are first calculated for individual words using their values in the lexicon and the sentiment for the entire sentence is obtained by evidence based combination function and sentiment normalisation as described by Jurek et al. (2015).

The tweets that are found to be negative are further analysed to measure their intensity and how they may influence the selected public event. By measuring the number of people who have negative opinion tweets about the event against the total number of people who tweeted about the event, we provide 3 levels of possible threat or disorder for the event that is about to take place. Thus, the relationship between the number of people who tweet with a negative sentiment about an event and the actual disruption or disorder that happens during an event can be obtained. Summarization of the steps is as follows:

- To collect and monitor real time data, using deep data collection by identifying public events or places of interest;
- To perform lexicon based sentimental analysis on the collected data to determine emotional values;
- To classify the obtained emotional values into possible levels of disorder that will happen during the event.

The reminder of the manuscript is organized as follows: The next section describes the existing related work of the literature. The section 3 explains the existing work in depth and section 4 depicts the proposed methodology along with the novel algorithms. Later in the section 5, the proposed method is evaluated for its performance in the real-time scenario. Finally, the article concludes with the summary of results and presents future work guidelines.

19 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage:

www.igi-global.com/article/sentiment-analysis-of-tweets-for-estimating-criticality-and-security-of-events/187258?camid=4v1

This title is available in InfoSci-Journals, InfoSci-Select, InfoSci-Select, InfoSci-Journal Disciplines Communications and Social Science, InfoSci-Technology Adoption, Ethics, and Human Computer Interaction eJournal Collection, InfoSci-Management Science and Organizational Research eJournal Collection, InfoSci-Journal Disciplines Business, Administration, and Management. Recommend this product to your librarian:

www.igi-global.com/e-resources/library-recommendation/?id=2

Related Content

End User Types: An Instrument to Clarify Users Based on the User Cube
Chittibabu Govindarajulu and Bay Arinze (2008). *Journal of Organizational and End User Computing* (pp. 61-81).

www.igi-global.com/article/end-user-types/3841?camid=4v1a

User Performance Testing Indicator: User Performance Indicator Tool (UPIT)
Bernadette Imuetinyan Iyawe (2017). *Design Solutions for User-Centric Information Systems* (pp. 205-229).

www.igi-global.com/chapter/user-performance-testing-indicator/173975?camid=4v1a

Gestural Interaction with Mobile Devices Based on Magnetic Field

Kamer Ali Yüksel (2014). *Research and Design Innovations for Mobile User Experience* (pp. 203-222).

www.igi-global.com/chapter/gestural-interaction-with-mobile-devices-based-on-magnetic-field/80371?camid=4v1a

Quality of Use of a Complex Technology: A Learning-Based Model

Marie-Claude Boudreau and Larry Sligman (2007). *Contemporary Issues in End User Computing* (pp. 248-272).

www.igi-global.com/chapter/quality-use-complex-technology/7039?camid=4v1a