International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET207476

Analysis of Active Learning for Social Media to Support Crisis Management

Mayuri S. Waghmare[1], Prof. Tarun Yengantiwar[2]
[1]M.Tech Scholar, Department of Computer Science and Engineering.V.M.Institute of Engineering and Technology, Nagpur, Maharashtra, India
[2]Project Guide Dept. of Computer Science and Engineering. V.M.Institute of Engineering and Technology, Nagpur, Maharashtra, India

Article Info
Volume 7 Issue 4
Page Number: 291-294
Publication Issue :
July-August-2020

Article History
Accepted : 15 Aug 2020
Published : 25 Aug 2020

ABSTRACT

People use social media (SM) to describe and discuss different situations in which they are involved, such as crises. It is therefore useful to exploit SM content to support crisis management, especially by revealing useful and unknown information about real-time crises. Therefore, we propose a new active online multi-prototype classifier called AOMPC. It identifies relevant data related to a crisis. AOMPC is an online learning algorithm that runs on data streams and contains active learning algorithms for actively querying the label of obscure unnamed data. The number of queries is controlled by a consistent budget strategy. In general, AOMPC allows for somewhat labeled data streams. AOMPC evaluated using two types of data: (1) synthetic data and (2) SM data from Twitter related to two crises, the Colorado floods and the Australian pushfires. To provide a complete estimate, a complete set of known measurements was used to examine the quality of the results. Furthermore, a sensitivity analysis was conducted to show the effect of the parameters of AOMPC on the accuracy of the results. AOMPC's comparative study was conducted against other available online learning methods. Tests to handle emerging, somewhat labeled data streams showed AOMPC's excellent behavior.

Keywords : Information Access, Social Networks, Twitter, Online Learning, Active Learning, Crisis Management

I. INTRODUCTION

Internet based social media websites such as Facebook (www.facebook.com) and Twitter (www.twitter.com) represent a growing facet of modern experience. These sites boast large numbers of users and their influence is increasingly being experienced in clinical practice [1].Social media and forum websites have been examined as a resource for healthcare service users. Internet based social media, such as Twitter, has also received attention for its potential in supporting social activism where its role can be seen in linking groups together and coordinating activity [2][3].

Social networks can provide a range of benefits to members of an organization[4]:

Support for learning: Social networks can enhance informal learning and support social connections within groups of learners and with those involved in the support of learning.

Support for members of an organization: Social networks can potentially be used my all members of an organization, and not just those involved in working with students. Social networks can help the development of communities of practice.

Engaging with others: Passive use of social networks can provide valuable business intelligence and feedback on institutional services (although this may give rise to ethical concerns).

Ease of access to information and applications: The ease of use of many social networking services can provide benefits to users by simplifying access to other tools and applications. The Facebook Platform provides an example of how a social networking service can be used as an environment for other tools.

Common interface: A possible benefit of social networks may be the common interface which spans work / social boundaries. Since such services are often used in a personal capacity the interface and the way the service works may be familiar, thus minimizing training and support needed to exploit the services in a professional context. This can, however, also be a barrier to those who wish to have strict boundaries between work and social activities.

Examples of popular social networking services include:

Facebook: Facebook is a social networking Web site that allows people to communicate with their friends and exchange information. In May 2007 Facebook launched the Facebook Platform which provides a framework for developers to create applications that interact with core Facebook features

MySpace: MySpace is a social networking Web site offering an interactive, user-submitted network of friends, personal profiles, blogs and groups, commonly used for sharing photos, music and videos..

Ning: An online platform for creating social Web sites and social networks aimed at users who want to create networks around specific interests or have limited technical skills.

Twitter: Twitter is an example of a micro-blogging service. Twitter can be used in a variety of ways

including sharing brief information with users and providing support for one's peers.

Note that this brief list of popular social networking services omits popular social sharing services such as Flickr and YouTube.

II. IMPLEMENTATION

OSN System Construction Module

In the first module, we develop the Online Social Networking (OSN) system module. We build up the system with the feature of Online Social Networking System, Twitter. Where, this module is used for new user registrations and after registrations the users can login with their authentication. Where after the existing users can send messages to privately and publicly, options are built. Users can also share post with others. The user can able to search the other user profiles and public posts. In this module users can also accept and send friend requests. With all the basic feature of Online Social Networking System modules is build up in the initial module, to prove and evaluate our system features. We present the proposed framework for the analysis and real-time detection of Crisis Management on Twitter. First, we introduce the real-time detection problem. Then, we define our online proxy measurements (behavior features) for Crisis Management signs. Finally, we describe the approach we implement for detecting behavioral change points[1].

Data collection

we use dataset. The aim of these normalization techniques is to remove the features that are too specific. For each of these two models, we run experiments on counts of unigrams as well as unigrams and bigrams. Due to the absence of publicly

available datasets for the evaluation of Crisis Management in social media, we used the Twitter streaming API to collect tweets. We used the Twitter streaming API to collect tweets containing key phrases generated from the APA's list of risk factors and AAS's list of warning signs related to Crisis Management.

Active Online Multiple Prototype Classifier (Aompc)

Due to the fact that SM data is noisy, it is important to identify relevant SM items for the crisis situation at hand. The idea is to find an algorithm that performs this classification and also handles ambiguous items in a reasonable way. Ambiguous denotes items where a clear classification is not possible based on the current knowledge of the classifier. The knowledge should be gained by asking an expert for feedback. The algorithm should be highly self-dependent, by asking the expert only labels for a limited number of items. Therefore, we propose an original approach that combines different aspects - such as online learning and active learning - to build a hybrid classifier, AOMPC.AOMPC learns from both, labeled and unlabeled data, in a continuous and evolving way[1]. In this context, AOMPC is designed to distinguish between relevant and irrelevant SM data related to a crisis situation in order to identify the needs of individuals affected by the crisis. AOMPC relies on active learning. It implies the intervention of a user in some situations to enhance its effectiveness in terms of identifying relevant data and the related event in the SM stream of[5]. The user is asked to label an item if there is a high uncertainty about the classification as to whether it is relevant or irrelevant. The classifier assigns then the item (be it actively labeled or unlabeled) to the closest cluster or uses it to create a

new cluster. A cluster - in this case - represents either relevant (i.e., specific information about the crisis of interest) or irrelevant information (i.e., not related to the crisis)[6].

Multiple Prototype Classification and LVQ Classification

A prototype-based classification approach operates on data items mapped to a vector representation (e.g., vector space model for text data). Data points are classified via prototypes considering similarity measures. Prototypes are adapted based on items related/similar to them. A Rocchio classifier is an example of a single prototype-based classifier. It distinguishes between two classes, e.g., "relevant" and "irrelevant". In real world-scenarios, due to the nature of the data, it is often not possible to describe the data with a single prototype-based classifier.

III. SYSTEM ANALYSIS

Currently, there are systems with crowd-sourcing platform characteristics, but no procedure (like active learning) is available to directly involve emergency management personnel in filtering relevant information. Data is noisy.

We propose a Learning Vector Quantization (LVQ) like approach based on multiple prototype classification. The classifier operates online to deal with the evolving stream of data. The algorithm, named active online multiple prototype classifier (AOMPC), uses unlabeled and labeled data which are tagged through active learning. An original online learning algorithm, AOMPC, is proposed to handle data streams in an efficient way. It is a multi-prototype LVQ-like algorithm inspired by our previous work. As part of AOMPC, an active learning strategy is introduced to guide AOMPC towards accurate classification, and in this paper towards sub event detection. Such a strategy makes use of budget and uncertainty notions to decide when and what to label. AOMPC is evaluated on different data: synthetic datasets (synthetic numerical data, generated microblogs, which are geo-tagged) and real-world datasets collected from Twitter related to two crises, Colorado Floods in 2013 and Australia Bushfires in 2013. The choice and the use of all these datasets was motivated by their diversity. That allows to thoroughly evaluate AOMPC because these datasets have different characteristics. A sensitivity analysis based on the different AOMPC parameters and datasets is carried out. A comparison of AOMPC against well- known online algorithms is conducted and discussed.

IV. CONCLUSION

This paper presents a streaming analysis framework for distinguishing between relevant and irrelevant data items. It integrates the user into the learning process by considering the active learning mechanism. We evaluated the framework for different datasets, with different parameters and active learning strategies. We considered synthetic datasets to understand the behavior of the algorithm and real-world social media datasets related to crises. We compared the proposed algorithm, AOMPC, against many existing algorithms to illustrate the good performance under different parameter settings. The algorithm can be extended to overcome many issues, for instance by considering dynamic budget, dynamic deletion of stale clusters, and generalization to handle non-contiguous class distribution.

V. REFERENCES

[1]. Daniela Pohl, Abdelhamid Bouchachia SMIEEE, and Hermann Hellwagner SMIEEE.

[2]. Active Online Learning for Social Media Analysis to Support Crisis Management (IEEE, VOL. , NO. , —- 2015.)

[3]. Deen SR, Withers A, Hellerstein DJ. Mental health practitioners' use and attitudes regarding the Internet and social media. J Psychiatr Pract. 2013;19:454–63

[4]. Segerberg A, Bennett WL. Social media and the organization of collective action: using twitter to explore the ecologies of two climate change protests. Commun Rev. 2011;14:197–215.

[5]. F. Abel, C. Hauff, G.-J. Houben, R. Stronkman, and K. Tao, "Semantics + Filtering + Search = Twitcident. Exploring Information in Social Web Streams," in Proc. of the 23rd ACM Conf. on Hypertext and Social Media. ACM, 2012, pp. 285–294

[6]. U. Ahmad, A. Zahid, M. Shoaib, and A. AlAmri, "Harvis: An integrated social media content analysis framework for youtube platform," Information Systems, vol. 69, pp. 25 – 39, 2017

[7]. G. Backfried, J. Gollner, G. Qirchmayr, K. Rainer, G. Kienast,G. Thallinger, C. Schmidt, and A. Peer, "Integration of Media Sources for Situation Analysis in the Different Phases of Disaster Management: The QuOIMA Project," in Eur. Intel. and Security Informatics Conf., Aug 2013, pp. 143–146.

Cite this article as : Mayuri S. Waghmare, Prof. Tarun Yengantiwar, "Analysis of Active Learning for Social Media to Support Crisis Management", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 7 Issue 4, pp. 291-294 , July-August 2020. Available at
doi : https://doi.org/10.32628/IJSRSET207476
Journal URL : http://ijsrset.com/IJSRSET207476