International Arab Conference on Information Technology (ACIT'2016)

Prediction of Suicidal Ideation in Twitter Data using Machine Learning algorithms

Marouane Birjali[1], Abderrahim Beni-Hssane[1],and MohammedErritali[2]
[1]LAROSERI Laboratory, Department of Computer Sciences, University of ChouaibDoukkali, Faculty of Sciences, El Jadida, Morocco
[2]TIAD Laboratory, Department of Computer Sciences, University of Sultan MoulaySlimane, Faculty of Sciences and Technologies, BéniMellal, Morocco

Abstract:*The rise of social network and the large amount of data generated by them, has led researchers to study the possibility of their operation in order to identify the hidden knowledge. Suicide is a serious mental health problem that demands our attention, and to control and prevent it is not an easy task. In this paper, we propose a suicidal ideation detection system, for predicting the suicidal acts using Twitter data that can automatically analyze the sentiments of these tweets. Then we investigate a tool of data mining to extract useful information for classification of tweets collected from Twitter based on machine learning classification algorithms.Experimental results show that our method for detecting the suicidal acts using Twitter data and the machine learning algorithms verify the effectiveness of performance in term of recall, precision and accuracy on sentiment analysis.*

Keywords:*Twitter; machine learning; suicide; tweets; sentiment analysis;*

1. Introduction

Nowadays, the social network is more than ever a means of communication tools for information exchange. It offers a considerable information at a great speed and services adapt more to the needs of users. These can consult the Internet to create social communications, social interaction (between individuals or groups of individuals), and content creation [1].

During the last years, the Internet has yet seen a wider scope through the development of social media. Based on easy communication techniques and accessible to all, the media promote social interaction through the Internet. Offering free access, social media has greatly promoted the mass and have triggered public debate on the Internet. Many social networks exist and there are more than 900 social media sites available on the internet [2]. Millions of people are using Twitter and ranked as one of the most visited sites with the average of 58 million tweets per day [3].

In this context, social network like Twitter and Facebook are increasingly associated with phenomena such as harassment, bullying or even suicide. It is therefore very important to detect potential victims at the earliest in order to strengthen suicide prevention on the web. Indeed, we can cite as an example the case of two American rappers Freddy E. [4] and Capital Steez [5] are given the death commenting live on their actions their Twitter accounts.

This research emphasized on suicide. Suicide is one of the top 20 causes of death worldwide [6]. The word suicide, so pronounced, is not to be taken with simplicity and lightly. It could be the cry for help from somebody, and if the signs are recognized early, lives could be saved. Suicide is preventable and suicide prevention should be responsibility for everyone.

Recent studies have demonstrated that Twitter could be used to prevent online suicide. The topics discussed and the terms used by depression and suicidal are well known. For example, these people are often victims of harassment or cyberbullying. Twitter would thus implement a real-time monitoring with respect to various risk factors. In addition, this study demonstrated that the number of suicidal tweets was highly correlated with the actual suicide rate.

2.RelatedWorks

The social network has attracted the attention of the research community that is trying to understand, among others, their structure and user interconnection and interaction between users. People tend to express their feelings and talk about their activities of daily life through Twitter.

Applying machine learning methods for the identification of suicide has grown in recent years. Linguistic Inquiry and Word Count Version 2007 (LIWC2007) evaluates different words or emotional, cognitive and structural English expressions presented in oral and written sentences of individuals. Liwccouldbe used to identify the trend of an emotional post. Ramirez-Esparza, et al. [7] worked on linguistic markers and for discussion of depression by gathering information both in depressed individuals and non-depressed existing Internet forums using bulletin board systems (BBS). They also found that depressed people who wrote in English were more likely to report medical problems.

Sentiment analysis has been handled as a Natural Language Processing task at many levels of granularity. There has been a wide range of research done on sentiment analysis, from rule-based, bag-of-words approaches to machine learning techniques. Starting from being a document level classification task Turney [8] it has been handled at the sentence level Hu and Liu [9] and more recently at the phrase level Wilson et al. [10].

The social network like Twitter, on which users post real-time reactions to and opinions about "everything", is a new and different challenge. Some of the first results and recent analysis of Twitter sentiment data. Two main research areas of mining opinion operate either on the document level [11]. Both classification methods at the document level and at the level of the sentence are generally based on the identification of opinion words or phrases.

In this work, we compute similarity of a tweet in data set and training data. We use several machine learning algorithm to build our work. In this context, we focus on pessimistic, bad thoughts and the fact of thinking by suicide. We evaluate the whole system and present results for predicting the orientations on Twitter data.

3. Methodology

The needs for this project can be divided into four parts: the requirements related to the construction of an associated vocabulary to the theme of suicide, retrieving the tweets from Twitter, and needs related to the automatic classification using machine learning algorithms and the requirements for a languistic analysis of these sentiments to improve our results.

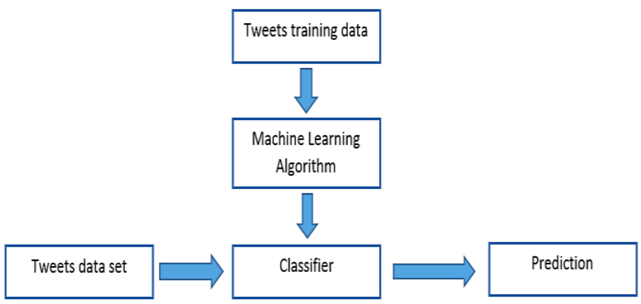


Figure 1. Process of methodology to define a machine learning of tweets.

Before embarking our work, we bring to define a vocabulary associated with different themes suicide-related critical (depression, fear, harassment, etc.). This vocabulary should be divided into different categories and sub-categories, so that can easily identify the degree of threat from the tweet. For example, in the case of harassment via a tweet, it is necessary to identify the recipients of tweets (for it is they who may take action). In contrast, for other categories of vocabulary, it is the people who posted the tweet, which must be identified.

3.1. Dataset and data collection

The collection of tweets is a second part in our project. In the present study, the API Twitter4J is used for collecting the data. Twitter4J is a Java library for the Twitter application programming. It is integrated Java application with all the Twitter service. We collected the tweets using the search word. The wire mechanism is used to collect sequential data from Twitter. The token and access keys obtained, which is required for the extraction of the tweets and extracted the real time tweets from the twitter website. The extraction is repeated for multiple times to obtain more number of tweets.

3.2. Machine learning classification

The machine learning is a set of computer algorithms that automate the construction of our classification function from a set of data is called the training set. Ma-chine learning algorithms for the automatic detection of strongly concerning suicidal tweets have been developed by supervised or unsupervised learning [12] [13]. Super-vised learning approach is done after studying the characteristics that possibly possessed by tweets at particular class of suicidal acts. This approach generally named after classification technique. This classification will divide the existing tweets into training and testing tweet. By using the training tweets and our vocabulary, this approach builds analysis model to be able to determine in which tweet is a high risk that the authors have pass the act of suicide or suspect's tweets safely is to say, without however the author of the tweet goes to the act.

Several machine learning techniques including Maximum Entropy, Support Vector Machines (SVM) and Naive Bayes are used to classify reviews. [14]. A number of features that can be used for classification are feeling term presence, term frequency, negation, n-grams and Part-of-Speech [15]. These principal features can be used to extract the sentiment value of words, sentences and tweets. For our work, we aimed to study the contents of tweets, associated with the user profiles is an effective feature set for accurately classifying user profiles hence, we do not consider a particular users' tweet. On the other hand, we targeted our vocabulary associated tweets.

3.3. Suicidal acts analysis

In our work, we used the twitter dataset and analyzed it. This analyses labeled datasets where the pre-processor is applied to the raw sentences which make it more appropriate to understand. Further, the different machine learning techniques trains the dataset with feature vectors and then the languistic analysis offers a large set of synonyms and similarity which provides the polarity of the content. The complete description of the work has been described in next sub sections and the block diagram of the same is graphically represented in Figure 2.

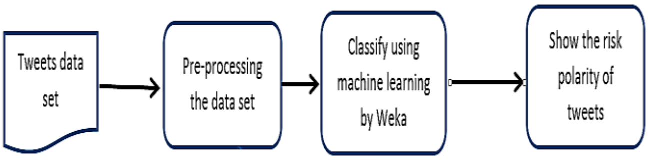


Figure 2. Block diagram detailing our work

4. Experimental evaluation and analysis

The experiment was performed using the Weka tool. The choice of this tool is due to the fact that it is widely used in the field of machine learning and data mining, and it is also easy to handle. It is compatible with the data format we chose.

This figure presents statistics to the results of the classification by Weka. Tweets suspects at risk, the percentage of suspects tweets at risk compared to suspects tweets safely.

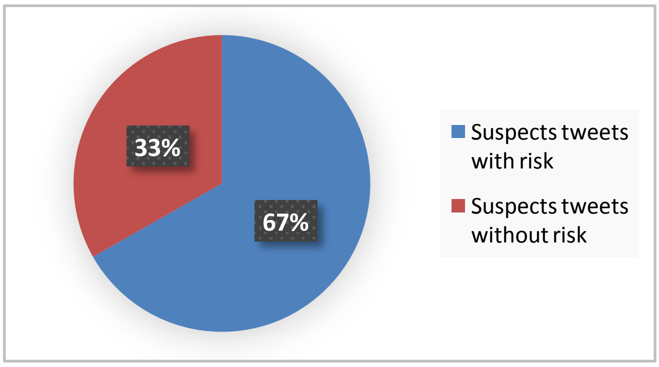


Figure 3. Suspects tweets with risk and without risk of suicidal acts

As already mentioned above, we will look at the classification. In weka, one can find all the matching algorithms classify in the tab. After loading the data into the tool, applying different functions for this study. Weka provides four options for evaluating the performance of the model learned in this experiment, we will consider only two of them. First, apply the algorithms on the entire data (use training set). Then we will use the cross validation. The figure below shows the testing quality of the sentiment analysis of machine learning algorithm. Using Twitter API, tweets related to products are collected. A dataset is created using 335 twitter posts of electronic products. Dataset splits in to a training set tweets and a test set. The following tables show the various classification algorithms and the results we have achieved in terms of accuracy:

Table 1. Cross-validation of performance on different classifiers for suspects tweets with risk.

Algorithm	**CART**	**IB1**	**Naive Bayes**	**SMO**	**J48**
Precision	66,7%	63%	61,00%	70%	75,4%
Recall	58,7%	50,5%	76,1%	51,4%	72,4%
F-measure	62,4%	55,8%	67,8%	59,3%	63,8%

Figure 4 present the correctly clustered instances are represented by crosses, incorrectly clustered once represented as squares. By changing the color dimension to other attributes, you can see their distribution within each of the clusters.

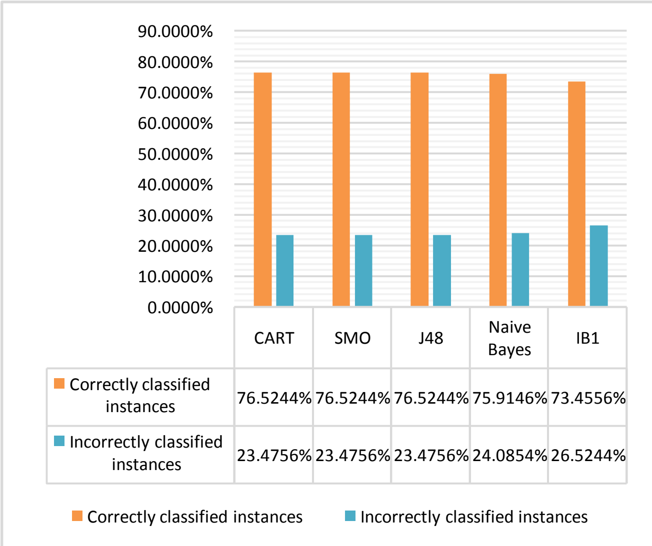


Figure 4. Performance of different classifiers

The confusion matrix shows misclassified tweets. In the figure above, the tiles are misclassified tweets and those of cross-ranked. And the figure X shows how to check suspicious tweets at risk.

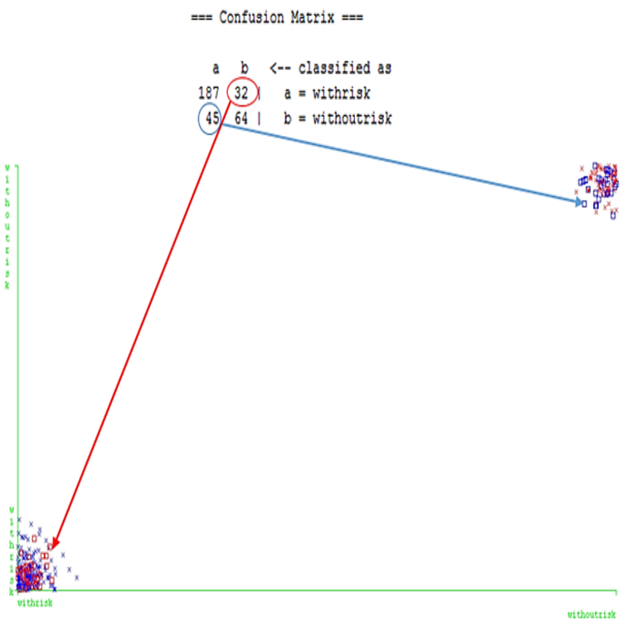
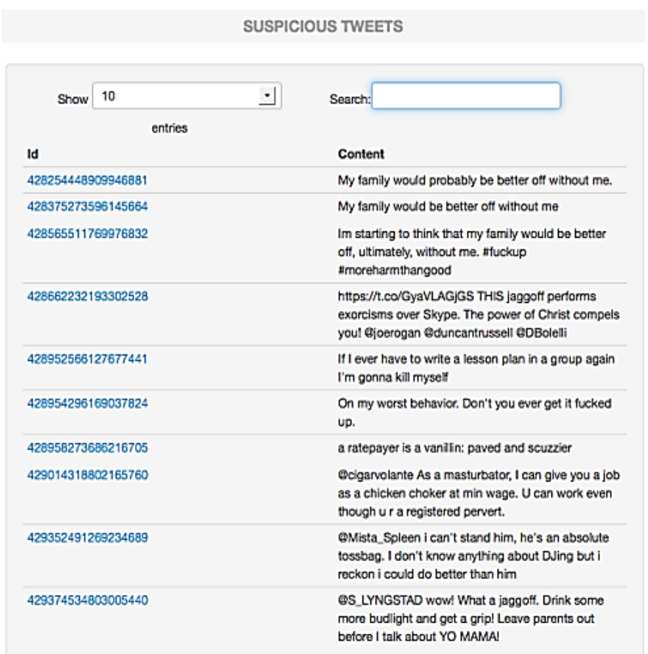


Figure 5. Visualization errors classifier



SUSPICIOUS TWEETS

Show 10 entries Search:

Id	Content
428254448909946881	My family would probably be better off without me.
428375273596145664	My family would be better off without me
428565511769976832	Im starting to think that my family would be better off, ultimately, without me. #fuckup #moreharmthangood
428662232193302528	https://t.co/GyaVLAGjGS THIS jaggoff performs exorcisms over Skype. The power of Christ compels you! @joerogan @duncantrussell @DBolelli
428952566127677441	If I ever have to write a lesson plan in a group again I'm gonna kill myself
428954296169037824	On my worst behavior. Don't you ever get it fucked up.
428958273686216705	a ratepayer is a vanillin: paved and scuzzier
429014318802165760	@cigarvolante As a masturbator, I can give you a job as a chicken choker at min wage. U can work even though u r a registered pervert.
429352491269234689	@Mista_Spleen i can't stand him, he's an absolute tossbag. I don't know anything about DJing but i reckon i could do better than him
429374534803005440	@S_LYNGSTAD wow! What a jaggoff. Drink some more budlight and get a grip! Leave parents out before I talk about YO MAMA!

Figure 5. Viewing tweets suspects at risk

5. Conclusions and Future Work

As part of this work, we present our method potentially based on different machine learning for using the social network Twitter as a preventive force in the fight against suicide.ssIn our future work, we plan to further improve and refine our techniques in order to enhance the accuracy of our method. Thereafter, we involved to test multilingual WordNet for tweets and to orient this work in a big data environment.

References

[1] Statistic Brain. Twitter Statistics, Retrieved from http://www.statisticbrain.com/ twitter-statistics/, 2014.

[2] Po-Wei Liang, Bi-Ru Dai, "Opinion Mining on Social Media Data," IEEE 14th International Conference on Mobile Data Management, pp. 91–96, 2013.

[3] Statistic Brain. Twitter Statistics, Retrieved from http://www.statisticbrain.com/ twitter-statistics/, 2014.

[4] R. Li, K. H. Lei,R. Khadiwala, Chang, "TEDAS: A Twitter-based Event Detection and Analysis System," icde, pp.1273-1276, 2012 IEEE 28th International Conference on Data Engineering, 2012.

[5] https ://twitter.com/Freddy_E

[6] https ://twitter.com/CapitalSTEEZ_

[7] Zhao, D. &Rosson, M.B., (n.d), Retrieved from http://research .ihost.com/cscw08-socialnetworkinginorgs/papers/zhao_cscw08_workshop.pdf

[8] Nairan Ramirez-Esparza, Cindy K. Chung, EwaKacewicz, and James W. Pennebaker, "The Psychology of Word Use in Depression Forums in English and in Spanish: Testing Two Text Analytic Approaches", Association for the Advancement of Artificial Intelligence (www.aaai.org).

[9] P. Turney, "Thumbs up or thumbs down Semantic orientation applied to unsupervised classification of reviews," Proceedings of the Association for Computational Linguistics.

[10] Hu M. and Liu B., "Mining and Summarizing Customer Reviews", KDD '04 Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, Pages 168-177.

[11] Wilson T., Wiebe J. and Hoffmann P., "Recognizing Contextual Polarity in Phrase-

Level Sentiment Analysis", In the Advanced Research and Development Activity (ARDA).

[12] Wu, Yuanbin, Qi Zhang, Xuanjing Huang, and Lide Wu. Structural opinion mining for graph-based sentiment representation. In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP-2011). 2011.

[13] B. O'Dea, S. Wan, P. J. Batterham, A. L. Calear, C. Paris, and H. Christensen, "Detecting suicidality on Twitter," Internet Interventions, vol. 2, pp. 183-188, 2015.

[14] Pang, Bo and Lee, Lillian. Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval Vol. 2, No 1-2. 2008.

[15] G. Vinodhini and R. Chandrasekaran, "Sentiment analysis and opinion mining: A survey," International Journal, vol. 2, no. 6, 2012.

[16] Y. Mejova, "Sentiment analysis: An overview," Comprehensive exam paper, available on http://www.cs.uiowa.edu/ymejova/publications/CompsYelenaMejova.pdf [2010-02-03], 2009.

[17] Yuan Wang, .lie Zhang and JulitaVassileva, "A User-Centric Approach for Social Data Integration and Recommendation," Human-Centric Computing (HumanCom), 2010 3rd International Conference, pp. 1- 8,2010.

[18] Maxim Grinev, Maria Grineva, Alexander Boldakov, Leonid Novak, AndreySyssoev,DmitryLizorkin, "Sifting Micro-blogging Stream for Events of User Interest," Proceedings of the third ACM conference on Recommender systems, pp. 385 - 388, 2009.

[19] Akram AI-Kouz, Ernesto William De Luca, SahinAlbayrak, "Latent Semantic Social Graph Model for Expert Discovery in Facebook," IICS 2011, pp. 128- 138, 2011.

[20] Leacock C. and M. Chodorow, Combining Local Context and WordNet Similarity for Word Sense Identification, In WordNet: An Electronic Lexical Database, C. Fellbaum, MIT Press, 1998.