

Suicidal Post Detection in Social Networks using NLP

Mukhtarkhanuly Daniyar* and Alan Abishev

Suleyman Demirel University, Almaty, Kazakhstan

Received: 22 Jul. 2018, Revised: 26 Aug. 2018, Accepted: 28 Aug. 2018

Published online: 1 Sep. 2018

Abstract: The social problem of suicide and alcoholism among youth is one of the problems, that the government currently faces. According to statistics, Kazakhstan is in the top 10 in the world in teenage suicide and alcoholism rates, as well as in several other social problems. An important stimulus in creating the aforementioned information system (IS) are the global trends in sociology, those focused on research of people with the use of internet technologies. The main methodology used for the development of the IS, is the content analysis of incoming data, because text(oral and written) reflects individual characteristics of a person like a fingerprint, as well as voice characteristics (frequency of vowels, tone, etc.), this allows for the creation of sophisticated analytics, control of psychological stability and observation of mood changes in youth. For this approach various preprocessing methods and machine learning algorithms were used.

Keywords: sentiment analysis, suicide detection, social media, machine learning

1 Introduction

Suicide is a global problem of our generation. Each year, about 800,000 suicides are committed, and about 10 million suicide attempts are registered. The World Health Organization (WHO) predicts that by 2020 the number of suicides can increase by 50%, and the number of suicide attempts will increase by 10-15 times [1].

As the WHO reports, 20 cases of suicide per 100 000 people are considered as "critical", while in Kazakhstan this number is 26.9 [2]. Kazakhstan is at the top of the Commonwealth of Independent States (CIS) in terms of the number of suicides among children and in the third place around the world. According to the General Prosecutor's Office of the Republic of Kazakhstan in 2015, 201 cases of suicide and 485 suicide attempts were recorded. The most susceptible age group is teenagers between 15-17 years old, they account for 60% of all suicides [3].

To reduce child and teenager suicide rates, the Ministry of Healthcare and Social Development initiated an implementation of a state project on the prevention of suicidal behavior. The set of measures includes raising awareness in schools, specialized trainings for teachers and medical personnel, as well as identification of adolescents "at risk" groups through monitoring of social networks.

The main goal of this research is to predict whether a student is likely to have suicidal tendencies based on his/her social network feed, and to reduce national suicide rates by notifying suicide hotlines and parents, as well as improving mental health and identifying the main issue of suicides among children. Knowing mental health problems are the bane of the 21st century, we think this is a very promising and important project that uses the power of machine learning to solve such a difficult topic as suicide.

2 Related Works

Social networks are not only a comfortable environment for communication, but also a place for self-expression. People share their feelings and record their daily lives. Since some people publish posts with suicidal content, researchers came to a conclusion that social networks can be useful for identifying people at risk.

The usage of machine learning techniques to identify people at risk became popular in the last year [4]. Since the data of the people who committed suicide is not public, in the research of [5] Ermakov Sergey and Ermakova Liana textual data was collected from writings of famous people such as writers, poets and critics, who

were close to death (the last year of life). In the document for counsellors [1], it is said that a person who prepares for a suicide, often speaks about it. According to the research, suicide is not impulsive and preparation for a suicide can last for about a year. During this time a person shows signs of his condition, sometimes with the help of various social networks.

For suicidologists it is difficult to find data that would broadly and qualitatively cover the mental state of a person who is going to commit suicide. First of all, the written notes of suicide survivors are used for this, where they describe in great detail how their consciousness changed for several months before they decided to commit such a desperate act [1]. Another valuable resource is the suicide notes - the last words of a person who has reached his last point. In a research by [6], Tong Liu, Qijin Cheng, Christopher Homan, Vincent Silenzo collected suicide notes from Twitter, relying on the Werther effect (when a person copies a famous person's suicide), thereby collecting all the people's posts on the key words about the suicide of Robin Williams.

Recently, Facebook announced its system for identifying suicidal posts [7]. If the system detects that a post is suicidal, this post is sent to the 'Help Center' and is further processed by the center's staff. The system takes into account not only what text was written, but also several important metrics: at what time the post was written, the day of the week, comments under the post, etc [8].

In the United States, Crisis Text Line (a text messaging trust service) has implemented machine learning for text and emoji analysis. The algorithm shows the most relevant risk group and prioritizes those who need immediate assistance[9].

In 2018, the Canadian government, in conjunction with Advanced Symbolics, launched a pilot project to monitor social networks for suicidal tendencies among the population [[10], [11]].

3 Data

Since the data on people who committed suicide is not public, we collected them from open source. Moreover, the Ministry of Education and Science and the Ministry of Health Republic of Kazakhstan provided local data. As for psychological analysis, the Republican Scientific and Practical Center of Mental Health helps with domain knowledge of their experts.

Each text was classified into two types. Where the 'positive' class means that a person has committed a suicide, 'negative' class means that the text does not have suicidal content.

3.1 Open source data

Approximately 180 suicide notes in Russian language were collected manually from VK social network profiles of the individuals who committed suicide. All of them were from CIS countries under the age of 18. Our domain experts checked the text, their last visit date, as well as the comments under their last post and labeled the text as positive or negative. We have built a search system by using 'VK' API. We searched for posts by keywords, such as (original Russian text and translation):

- 'Goodbye' & 'fault'
- 'why' & 'pity' & 'commit'
- 'rest in peace' | 'RIP'
- 'suicide'

3.2 Data from Ministry of Health of Republic of Kazakhstan

The Ministry of Health provided us with roughly 500 copies of suicide notes of victims aged under 18. All data was anonymized, there is only the date of birth and the date of death. Unfortunately, the data was provided in a paper format, and we have to transform it into an electronic format. As the research develops, more data will be gathered and processed.

4 Analysis of suicidal acts

Based on a thorough analysis of text corpus and the underlying themes and emotions, our psychologists outlined three potential motives for suicide:

1. Call for help – a desperate attempt to draw the attention of other people to their mental state.
2. Avoidance – inability to tolerate any further unbearable heartache, guilt or shame for a socially unacceptable act.
3. Protest – a protest against difficult family problems, often associated with expression of emotions of anger and accusation. A written note is often addressed to a specific person or group of people.

Suicide is not an impulsive act and preparation for suicide can last about a year, during which a person will show signs of his condition. We use sentiment analysis to detect this dangerous period.

5 Experimental evaluation and analysis

Suicide classification using social networks is a classic sentiment classification task. The experiments were performed using Python programming language. There are various techniques that were proposed by the

researchers [[12], [13], [14]]. All data were in JSON format with text and labels where 'True' means a post has suicide content and 'False' indicating its' absence. The corpus contains about 10000 texts with suicide content and we added approximately 12500 pieces of text that have other content such as blogs, news, ads, description of photos, posters and etc.

Firstly, all texts were lemmatized - the process of removing inflectional endings only and returning the base or dictionary form of a word, which is known as the lemma. To lemmatize words in a context of Russian language, Yandex 'MyStem' lemmatizer was used, as it demonstrated excellent performance. Afterwards, nltk Russian stopwords python library was used to remove the stopwords, therefore reducing potential noise in data.

Secondly, preprocessed texts were vectorized - the process of representation of texts in vector space for arithmetic operations on a whole data structure. Vectors are time and space efficient. For vectorization of texts bag of words and gensim word2vec word embedding methods were used. They are shown in tables 1 and 2 correspondingly.

Thirdly, vectors space with labels were fitted to machine learning models. Logistic Regression - one of the traditional machine learning algorithms with unigrams was used as a baseline.

After that, we use various improvements such as using n-grams, vectorization with word2vec, changing to non-linear models such as Random Forest and Gradient Boosting.

Table 1. Mean cross-validation with k = 5 on different classifiers with BoW

Algorithm	Log. Regression with unigram	Log. Regression with bigram	SVM with bigram	Random Forest with bigram	Gradient Boosting with bigram
Precision	73 %	74 %	75 %	80 %	80 %
Recall	70 %	74 %	74 %	74 %	74 %
F1-score	70 %	74 %	74 %	75 %	75 %

Table 2. Mean cross-validation with k = 5 on different classifiers with Word2Vec

Algorithm	Log. Regression with unigram	Log. Regression with bigram	SVM with bigram	Random Forest with bigram	Gradient Boosting with bigram
Precision	75 %	73 %	78 %	77 %	78 %
Recall	61 %	70 %	74 %	74 %	78 %
F1-score	62 %	73 %	74 %	73 %	78 %

6 Results

For the experiment we tested the following 5 algorithms:

- 1.Logistic Regression with unigrams
- 2.Logistic Regression with bigrams
- 3.SVM with bigrams
- 4.Random Forest with bigrams
- 5.Gradient Boosting with bigrams

As you can see in the 2 table, the best performing algorithm was Gradient Boosting with bigrams. To visualize the results, we use Area Under the Curve (AUC) Receiver Operating Characteristic (ROC) curve. ROC is a probability curve and AUC represents degree or measure of separability. The higher the AUC, the better the model is at distinguishing between users with suicide ideations and without.

The figure 1 shows the ROC response of different datasets, created from k-fold cross-validation. Taking all of these curves, it is possible to calculate the mean area under curve, and see the variance of the curve when the training set is split into different subsets. This roughly shows how the classifier output is affected by changes in the training data, and how different the splits generated by k-fold cross-validation are from one another[15].

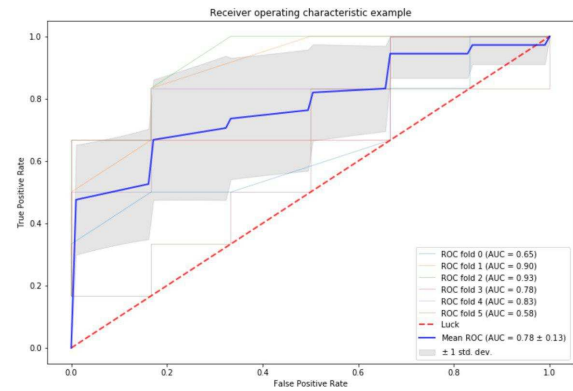


Figure 1: AUC-ROC with k-fold cross-validation.

7 Conclusion

In this preliminary research we tested different machine learning models to predict if a piece of text contains suicide content. We tried different vectorization techniques with k-fold cross-validation. The best algorithm was Gradient Boosting with word2vec that shows f1-score of 78%.

In the future work we will expand our data using more resources and try to use deep learning algorithms, we believe that they will show more accurate results. Our main goal is to build an alerting system for the government to monitor emotional state of a person to prevent possible suicide attempts or any self-inflicting injuries.

8 Acknowledgment

We thank reviewers for their comments and suggestions. We are grateful to the following members of the Practical Center of Mental Health for the psychological assessment:

Nikolay Nigay, Aigul Nashkenova, Bakhytzhana Ashirbekov, Sandygash Kudaibergenova and Kyanysh Altynbekov.

This work was funded by Suleyman Demirel University and the Ministry of Education and Science grant BR05236699 "Development of digital adaptive educational environment using large scale data analytics".

References

- [1] World Health Organization. Preventing suicide. A resource for counsellors. Geneva 2006
- [2] Oksana Lysenko. "The number of suicides among children in Kazakhstan continues to grow",
[<http://www.zakon.kz/4524024-kolichestvo-suicidov-sredi-detej-v.html>]
- [3] Dilbegim Mavloniy. "Kazakhstan ranks third in the world in the number of suicides"
[https://online.zakon.kz/Document/?doc_id=30468239]
- [4] Sidney Kennedy, Trehani M. Fonseka. How AI is helping to predict and prevent suicides, 2018.
[<http://theconversation.com/how-ai-is-helping-to-predict-and-prevent-suicides-91460>]
- [5] Ermakov Sergei, Ermakova Liana. Linguistic Approach to Suicide Detection, 2011
- [6] Tong Liu, Qijin Cheng, Christopher M. Homan, Vincent M.B. Silenzo. Learning from various labeling strategies for suicide-related messages on social media: An experimental study. 2017
- [7] Dan Muriello, Lizzy Donahue, Danny Ben-David, Umut Ozertem, Reshef Shilon. Under the hood: Suicide prevention tools powered by AI, 2018
[<https://code.fb.com/ml-applications/under-the-hood-suicide-prevention-tools-powered-by-ai/>]
- [8] Diana Kwon. Can Facebook's Machine Learning Algorithms Accurately Predict Suicide. 2017
[<https://www.scientificamerican.com/article/can-facebooks-machine-learning-algorithms-accurately-predict-suicide/>]
- [9] Crisis text line [<https://www.crisistextline.org/>]
- [10] Ottawa's Advanced Symbolics to lead AI project looking for suicide warning signs
[<https://obj.ca/article/ottawas-advanced-symbolics-lead-ai-project-looking-suicide-warning-signs>]
- [11] Advanced symbolics
[<http://www.advancedsymbolics.com/>]
- [12] Marouane Birjali, Abderrahim Beni-Hssane, Mohammed Erritali. Prediction of Suicidal Ideation in Twitter Data using Machine Learning algorithms. 2016
- [13] Dave Heller. Researcher uses machine learning to improve suicide prediction, 2017
[<https://medicalxpress.com/news/2017-03-machine-suicide.html>]
- [14] Hardik A. Patel, Cheng-Yuan Hsieh. Early Detection of Suicide Using Big-Data Analytics in Real Time. 2016
- [15] Receiver Operating Characteristic with cross-validation
[http://scikit-learn.org/stable/auto_examples/model_selection/plot_roc_crossval.html#sphx-glr-auto-examples-model-selection-plot-roc-crossval-py]



Daniyar Mukhtarkhanuly master's degree student in Mathematical and Computer Modeling at International Information Technology University in Almaty, Kazakhstan. His research interest includes Data science, Data analysis, NLP. Data scientist at Alem Research.



Alan Abishev received the BSc in Computer science at University of Southampton in Southampton, United Kingdom. His research interest includes Data science, Data analysis, Big data. Data engineer at Alem Research.