Linguistic Approach to Suicide Detection

Ermakov Sergei[1], Ermakova Liana[1,2]

[1] Perm State National Research University, Russia
[2] Institut de Recherche en Informatique de Toulouse, France
{choufler,liana.ermakova.87}@gmail.com

Abstract.
Suicide is a major, preventable public health problem. Particularly the problem is critical for young people. In Russia every year thousands of teenagers commit suicide. In most of the cases it can be prevented if a risky state is detected. Nowadays internet becomes a major way of communication, mainly in the text form. Therefore we suggest a method to detect a tendency to suicide based on text messages. Our main approach is to study indicators of such condition and based on it use machine learning approach to build a classifier that could determine, whether the person is about to commit a suicide. Our experiments are based on the analysis of texts of Russian writers for past 100 years that committed suicide.

Keywords: sentiment analysis, machine learning, suicide.

1   Introduction

Suicide poses a serious public health problem worldwide. Every year around 1 million deaths and 10 million attempts occurs due to that [1, 2]. In Russia this problem is particularly urgent among young people. According to official number provided by investigating committee, it was more than one thousand deaths in 2012 [3]. Around 10 thousands of teenagers attempt to commit suicide. This mainly happens due to their social disadaptation. For the past 50 years this number has grown for around 60% [4]. This problem also is present in the army, police and jails, and other places where people experience strong stresses. Suicide can be prevented and most people who feel suicidal demonstrate warning signs. The only problem is that these signs should be noticed. However, ordinary person could easily miss it. Thus, for the major part it is a duty of psychologists to recognize that dangerous behavior. However, the number of this kind of staff in schools is constantly reducing, e.g. in 2011-2012 school year, 70% of schools did not have psychologists at all (this is more 8% to the previous year). In the police and army there is one specialist for several thousand soldiers / policemen. There is an extremely urgent need in the tools that could help to analyze the emotional condition of big groups of people.

Internet communication and self-expression are becoming more and more popular. Almost every young person uses social networks, blogs and forums. The mean of

expression is mainly text. That gives us an idea that these text messages could give us a valuable source for automatic analysis in order to help solving suicide problem.

2   Techniques used to detect and prevent suicides

First step to prevent suicide is to detect the signs that indicate it. There are numerous studies on that topic [2, 4–7]. Most of the researches agree on the fact, that suicide is related to many emotions: hostility, despair, shame, guilt, dependency and hopelessness. The general psychological state commonly assumed to be associated with suicide is a state of intolerable emotion or unbearable despair.

The common indictors are:

— sadness;
— loss of interest or pleasure;
— disturbance in sleep and appetite;
— difficulty in concentrating;
— thoughts and talks of suicide;
— isolation;
— looking for a way to kill oneself;
— talking about being trapped or in unbearable pain;
— talking about being a burden to the others;
— talking about having no reason to live or feeling hopeless;
— displaying extreme mood swings;
— extreme behavior;
— exposure to suicidal behavior of others, such as family members or peers [4, 6].

The risk factors are:

— history of previous suicide attempts;
— family history of suicide;
— family violence;
— history of depression or other mental illness;
— alcohol or drug abuse;
— stressful life event or loss;
— exposure to the suicidal behavior of others;
— incarceration [4, 6, 8].

Person, who has set up his mind to commit suicide usually talks about this. Suicide is not always impulsive: people tend to prepare themselves to the suicide, according to the researches it could take around one year. During this period of time they give some type of verbal or behavioral message about their ideations of intent to hurt themselves [4].

Thus, we could attempt to determine using text analysis:

— emotional state of the person;
— specific verbal suicidal signs.

3 Analysis based on machine learning

First part of the analysis is to analyze emotions expressed in the text. This is one of the applications of sentiment analysis task [9]. There are various approaches proposed by researchers in this field [9–12]. One of the traditional sentiment analysis methods is using machine-learning techniques to perform classification: Naïve Bayes, SVM, neural networks, etc. Most of classifiers use words, word combinations and their characteristics as features. In our research we decided to start with a most simple two-way classification: does the text express potential suicide or not. To perform classifications we used WEKA toolkit [13]. As a classifier, SVM was chosen as it proved to be one of the most efficient in sentiment classification [14]. We started from using word N-grams (unigrams, bigrams and trigrams) as features. Then in order to improve the efficiency of determining emotional state we proposed to use following advanced features:

— share of all punctuation signs in the size of the text;
— number of specific "mood" signs used: three dots, exclamation marks, question marks;
— average and maximum length of the words;
— share of each part of speech (obtained from MyStem annotation);
  various phonetic characteristics (e.g. individual letters, bigrams, trigrams and quantitative features like word length, maximum sequential consonant length in a word, etc.) [15].

The second part of analysis is about specific verbal signs that could indicate suicide. This part was also partly covered by words and word combinations as features, such as burden, death, pain, despair, etc.

4 Experiments

We classified individual texts into two classes: texts written by a person committed suicide and not.

All experiments were based on the manually retrieved corpus of Russian texts. Information about people who has committed suicide is not public. Therefore we had to choose people, who are well-known. The most accessible texts are those produced by writers, poets, critics, etc. The list of them was extracted from research [16]. To make an experiment clear we chose only native speakers. We retrieved the list of 49 persons, excluding those whose suicide is doubtful. All texts were written in XIX – XXI c. As per study described above, we used texts, written at the time most close to death (last year of life). This corpus contained 469 texts (around 10 texts per person). We added the same number of texts, taken from the writers, who have not committed

suicide, including texts about death. Hence, full corpus for experiments contained 938 texts. The examples of each category of texts are given below:

Example 1. A text written by an author committed suicide (V. Mayakovsky).

```
Как говорят - "инцидент исперчен", любовная лодка разби-
лась о быт. Я с жизнью в расчете и не к чему перечень
взаимных болей, бед и обид. Счастливо оставаться.
                    Владимир Маяковский.
```

Example 2. A text written by a random author who did not committed suicide (A. Pushkin).

```
Я памятник себе воздвиг нерукотворный,
К нему не зарастет народная тропа,
Вознесся выше он главою непокорной
        Александрийского столпа.

Нет, весь я не умру — душа в заветной лире
Мой прах переживет и тленья убежит —
И славен буду я, доколь в подлунном мире
        Жив будет хоть один пиит.
```

Example 3. A text about death written by an author who did not committed suicide (A. Akunin).

```
Ну все, пора, меня уже зовут.
Увидимся позднее - не мешайте:
Мне надо что-то вспомнить напоследок.
        Но что? Но что?
            Ума не приложу.
Все спуталися мысли. Все, пора.
Что будет за последним окоемом,
        Спешу узнать.
                Вперед!
            Царевич Смерть,
Приди в кроваво-красном облаченьи,
Подай мне руку, выведи на свет.
Где буду я стоять, простерши руки,
Как ангел, как судьба, как отраженье
        Себя самой.
        Другого - не дано.
```

For a baseline we chose word unigrams as features.

Before classification all the words in the texts were stemmed using MyStem application [17].

There are several methods used to evaluate the models. The basic ones are

— precision,
— recall,

where $precision = \dfrac{true\ positive}{true\ positive + false\ positive}$ $recall = \dfrac{true\ positive}{true\ positive + false\ negative}$

F-measure, which is a harmonic mean between precision and recall

$$F1 = 2 * \frac{precision * recall}{precison + recall}$$

Receiver Operating Characteristic (ROC-Area) [18], which is an area under the curve of tradeoff between true positive and false positive. We used F1 and Roc-Area to compare results of the runs.

Our experiments consisted of build several models with different feature set and perform 10 folds cross-validation using WEKA package.

We evaluated the following models:

— only unigrams (baseline);
— combination of unigrams and bigrams;
— combination of unigrams, bigrams, trigrams;
— combination of unigrams, bigrams, trigrams and set of quantitative features;
— full model (combination of unigrams, bigrams, trigrams and set of quantitative features and phonetic features).

The last model showed the best results, all the results are presented in the Table 1

Table 1. Two-way classification results for Russian (SVM, 10 folds cross-validation)

Feature set	Class	Precision	Recall	F-measure	ROC Area
Unigrams	Positive	0.633	0.671	0.651	0.607
	Negative	0.641	0.677	0.659	0.607
Unigrams + bigrams	Positive	0.642	0.691	0.666	0.621
	Negative	0.663	0.699	0.680	0.621
N-grams	Positive	0.653	0.710	0.680	0.693
	Negative	0.671	0.722	0.696	0.693
N-grams and quanti-	Positive	0.779	0.792	0.785	0.789

tative features	Negative	0.783	0.804	0.793	0.789
Full combination of features	Positive	0.805	0.821	0.813	0.819
	Negative	0.817	0.834	0.825	0.819

Compared to baseline results (unigrams) we increased our F-measure average value from 0.65 to 0.82, ROC-area increased from 0.6 to 0.819. Though our task was not purely sentiment classification, we compared our results with recent researches in two-way sentiment classification, which is about 0.8-0.85 [19]. We can conclude that our best effort showed satisfactory results.

5   Conclusion and Future Work

In this paper, we presented an approach to suicide detection in texts using linguistic characteristics and machine learning. With the model we built we obtained F-measure 0.82 with the ROC-area 0.819. We observed that using proposed feature set we were able to improve almost for 20% compared to the baseline, and 13% compared to the N-grams model. This shows, that in our area not only words are important, but also the punctuation signs, word lengths, and phonetic characteristics. We used only two-way classification, which means that if text is detected as positive, this directly indicates, that person is in trouble. However, psychologists determine several levels of threat, and the lowest of them could contain a lot of false-positives, meaning that low level of signs could be normal for the most of the people. In the future work we would like to split two classes to 5, but this would require an assessment of texts made by professional psychologists, to determine a real level of danger. We also will try our approach on English texts, which will require manual collection like it was done for Russian.

As for the Russian language we try to cooperate with the medical organization to retrieve more texts for training our model, and also possibly to determine other features for the model, that are not obvious to a non-professional psychologists.

Our final goal is to build software that will help to monitor emotional state of people in order to prevent tragedies.

6   References

1.   Wasserman, D., Cheng, Q., Jiang, G.X.: Global suicide rates among young people aged 15-19. World psychiatry. 4, 114–120 (2005).
2.   Paci, W.R.O.W.: Towards Evidence-Based Suicide Prevention Programmes. World Health Organization, Western Pacific Region (2010).
3.   СК: В России резко возросло число детских самоубийств, http://www.rg.ru/2012/10/24/deti.html.
4.   World Health Organization: Preventing suicide : a resource for counsellors.

5. Gomez, D.D., Blasco-Fontecilla, H., Sukno, F., Ramos-Plasencia, M.S., Baca-Garcia, E.: Suicide attempters classification: Toward predictive models of suicidal behavior. Neuro-computing. 92, 3–8 (2012).

6. Early detection of depression and prevention of suicide, http://www.philstar.com/science-and-technology/147031/early-detection-depression-and-prevention-suicide.

7. Hjelmeland, H.: Cultural Context Is Crucial in Suicide Research and Prevention, (2011).

8. Костин, Р.А., Шамкова, С.В.: Суициды в молодежной среде Санкт-Петербурга: социологический анализ : монография. Санкт-Петербургский гос. университет сервиса и экономики (2007).

9. Pang, B., Lee, L.: Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval. 2, 1–135 (2008).

10. Jia, L., Yu, C.T., Meng, W.: Improve the effectiveness of the opinion retrieval and opinion polarity classification. In: Shanahan, J.G., Amer-Yahia, S., Manolescu, I., Zhang, Y., Evans, D.A., Kolcz, A., Choi, K.-S., and Chowdhury, A. (eds.) CIKM. pp. 1415–1416. ACM (2008).

11. Eguchi, K., Lavrenko, V.: Sentiment Retrieval using Generative Models. In: Jurafsky, D. and Gaussier, É. (eds.) EMNLP. pp. 345–354. ACL (2006).

12. Mukherjee, S., Bhattacharyya, P.: Feature Specific Sentiment Analysis for Product Reviews. CoRR. abs/1209.2352, (2012).

13. Witten, I.H., Frank, E., Hall, M.A.: Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, Amsterdam (2011).

14. Chetviorkin, I.I., Loukachevitch, N.V.: Sentiment analysis track at ROMI P 2012, (2013).

15. Ermakov, S., Ermakova, L.: Sentiment Classification Based on Phonetic Characteristics. In: Serdyukov, P., Braslavski, P., Kuznetsov, S., Kamps, J., Rüger, S., Agichtein, E., Segalovich, I., and Yilmaz, E. (eds.) Advances in Information Retrieval. pp. 706–709. Springer Berlin Heidelberg (2013).

16. Чхартишвили, Г.: Писатель и самоубийство: Писатель и самоубийство. Захаров (2007).

17. Segalovich, I.: A fast morphological algorithm with unknown word guessing induced by a dictionary for a web search engine. Proceedings of MLMTA. (2003).

18. Powers, D.M.W.: Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation. School of Informatics and Engineering, Flinders University, Adelaide, Australia (2007).

19. Piao, S., Tsuruoka, Y., Ananiadou, S.: Sentiment Analysis with Knowledge Resource and NLP Tools. The International Journal of Interdisciplinary Social Sciences. 4, 17–28 (2012).