Topic Model for Identifying Suicidal Ideation in Chinese Microblog

Xiaolei Huang[1], **Xin Li**[1], **Lei Zhang**[2], **Tianli Liu**[3], **David Chiu**[4], **Tingshao Zhu***[5,6]

[1]Computer Network Information Center, Chinese Academy of Sciences (CAS), China,
`{xlhuang,lixin}@cashq.ac.cn`
[2]Jinan University, China, `zl.ramiel@gmail.com`
[3]Inst. of Population Research, Peking University, China, `tianli.liu@pku.edu.cn`
[4]Dept. of Math and Computer Science, University of Puget Sound, USA,
`dchiu@pugetsound.edu`
[5]Inst. of Psychology, CAS China, `tszhu@psych.ac.cn`
[6]Inst. of Computing Technology, CAS China

Abstract

Suicide is one of major public health problems worldwide. Traditionally, suicidal ideation is assessed by surveys or interviews, which lacks of a real-time assessment of personal mental state. Online social networks, with large amount of user-generated data, offer opportunities to gain insights of suicide assessment and prevention. In this paper, we explore potentiality to identify and monitor suicide expressed in microblog on social networks. First, we identify users who have committed suicide and collect millions of microblogs from social networks. Second, we build suicide psychological lexicon by psychological standards and word embedding technique. Third, by leveraging both language styles and online behaviors, we employ Topic Model and other machine learning algorithms to identify suicidal ideation. Our approach achieves the best results on topic-500, yielding $F_1 - measure$ of 80.0%, Precision of 87.1%, Recall of 73.9%, and Accuracy of 93.2%. Furthermore, a prototype system for monitoring suicidal ideation on several social networks is deployed.

1 Introduction

Suicide is a severe health problem worldwide, which is one of leading causes of youth death in the world, especially in China. In the latest report (Organization and others, 2014) from World Health Organization (*WHO*), over $800,000$ people committed suicide in 2012, including $120,730$ Chinese; and it is very likely that the data is underestimated. Indeed, there are many more people who attempt suicide every year. Instead of calling health services or seeking for help in-person, choosing social networks is a preferable choice for some suicide because of privacy and facilitating sharing similar experiences among peers (Luxton et al., 2011).

Social network sites (*SNS*), such as Twitter, Sina Weibo, have become popular platforms for people to express themselves. Sina Weibo is a Chinese leading social network akin to Twitter. According to the latest Sina Weibo User Activity Report (`http://data.weibo.com/report/reportDetail?id=215`), Weibo now has more than 70 million active users per day, and over 160 million active users per month. It becomes a great platform for sharing opinions, emotions, and even to breaking news or public events. Recent work (Fu et al., 2013) showed that *SNS* not only enhanced our connections with others, but also facilitated selective self-presentation of undesirable behaviors, such as suicide.

The association between social media and suicide has drawn public attention recently, since several actual suicidal cases were reported in Sina Weibo, e.g., (`http://news.sina.com.cn/zl/zatan/2014-12-02/18032759.shtml`). However, new approaches towards online suicide ideation monitoring and prevention are still under development. (Fu et al., 2013) suggested that diffusion of microblogs about one's suicidal ideation or behaviors on social networks might serve as an early indication of a person's mental state. These indicators include one's writing through style, format, selection of specific words, and general

structure. It would therefore be desirable to build an appropriate suicide-monitoring system, to identify people who gave expressed suicidal ideation on *SNS* and provide follow-up support and services.

In this paper, we propose to detect suicide ideation in Chinese *SNS* and explore the possibility of using Topic Model (Blei et al., 2003). In particular, we first collect and evaluate suicidal microblogs by psychological standards. Second, we construct our psychological lexicons using word embedding techniques and explore the differences of online behaviors between suicide and none-suicide Weibo users. Then, in order to avoid the adverse outcomes that high dimensions of lexicon feature, which could weaken both efficiency and accuracy of classifiers, we model features of microblog in social networks and utilize the popular unsupervised model, Topic Model. Finally we design, develop, and test a model that can effectively identify suicidal ideation in *SNS*.

To summarize, our research has three main contributions: first, we use word embedding technique to construct psychological lexicons to enable utilization of suicide online behaviours; second, we employ Topic Model with lexicon knowledge and hybrid approaches for suicidal ideation identification on real-world datasets; third, a real-time application of suicide ideation prototype monitoring system is deployed online.

2 Related Work

Psychologists' researches on suicide cases in social networks started in recent years. Research (Fu et al., 2013) found that social media not only can spread suicidal ideation very quickly, but it also can be used to identify suicidal ideation in its early stages. Psychologists (Jashinsky et al., 2013) implied that evaluating suicidal risk factors in Twitter can be used to prevent suicide. Undoubtedly, previous research (Li et al., 2014b; Guan et al., 2014; Li et al., 2014b) also have dug interesting patterns for suicide prevention, however, these patterns can not be deployed on large population to provide timely service.

"Sentiment Analysis"(Pang and Lee, 2008) has been researched on various corpus for years, such as product reviews (Wang et al., 2010), movie reviews (Whitelaw et al., 2005). The core method in most of previous works (Pestian et al., 2010; Pestian et al., 2012) in this field is using $N-Gram$ (Brown et al., 1992) to model clinical suicide note. Another promising approach to learn sentiment from microblogs is *LDA* (Blei et al., 2003), a unsupervised algorithm that takes documents as a mixture of topics. It can discover latent semantics in documents and compute documents into a low-dimensional topic distributions. The potentiality of using topic models also has been applied in sentiment analysis (Mei et al., 2007; Lin and He, 2009).

Mental Health problems have been attracted much attention from researchers all over the world. Wu *et al.* (Wu et al., 2005) mined depressive symptoms from psychiatric consultation records. Researches on depression in *SNS* have been studied in single depression case (Wang et al., 2013), multimedia content of depressive microblogs (Lin et al., 2014), depression research in Twitter (De Choudhury et al., 2013) or mining emotion labels from social texts (Yu and Ho, 2014). Researchers (Resnik et al., 2013) applied *LDA* on essay's depression judgement among college students and compared its effectiveness with *LIWC*.

Given much research have been done on suicide ideation and sentiment analysis, and they have produced much promising results, which inspire us to investigate efficient methods to better understand and identify suicidal ideation on *SNS*.

3 Data Collection

We trained model on data collected through our Java-based crawler from Sina Weibo. We collected publicly reported suicide cases from 2011 to 2014, and spent another 6 months collecting and tagging their data. Based on evaluation criteria (Rudd et al., 2006), six experts first summarized and evaluated 12 warning signs of suicide, such as threatening to hurt or kill themselves. Those experts were trained to ensure the lowest biased tagging. They spent significant time in assessment and diagnosis of suicide risk. Before tagging data, the experts were tested by tagging 50 microblogs independently; and the test's interrater agreement coefficient[1] is 0.819.

---

[1]Kendall's W: a statistic, ranged from 0 to 1, can be used for assessing agreement among raters. The higher score, more agreements among raters are reached.

For each piece of microblog, only if it was voted by more than half of the experts, it would be tagged as suicidal microblog. During tagging process, each suicidal microblog has three levels: there is suicide warning sign, but no suicide plan; microblog indicates suicide plan, but author is not going to commit it; microblog indicates the author is going to commit suicide. Since we only focus on binary classification in this study, all three levels will be consider as suicide. Finally, 664 suicidal microblogs were obtained from over 30,000 microblogs.

Table 1: The composition of experiment data

All	Suicide	Non-suicide
7,314	664	6,650

To perform 10-fold cross validation, we randomly sampled 6,650 microblogs from a Weibo User Pool (WUP) (Li et al., 2014a) with 1.06 million active users' microblogs, which share the same time interval. Statistics of the data is illustrated in Table 1.

Figure 1: Detail Descriptions of One Microblog

Microblogs were segmented and tokenized using *Ansj* (Sun, 2014), a Chinese segmentation tool. *URLs* were removed by regular expression rules. In Fig. 1, we present an example of microblog. For each microblog, we extracted both content related features and meta features (i.e., time, like, *etc*).

3.1 Traditional & Simplified Chinese Conversion

Currently, there are two types of Chinese encoding: Simplified Chinese and Traditional Chinese. Words have the same meaning but different encodings, so computer program treats them differently. We converted Traditional Chinese into Simplified Chinese in our research.

4 Lexicons Construction

Primarily, we took advantage of existing sentiment dictionary, the latest version of *HowNet* (Dong and Dong, 2003), which is a Chinese-based emotional-words resource for sentiment analysis. *HowNet* is designed for general sentiment analysis.

Intuitively, words with similar contexts or co-occurrence may share similar meanings (Turney and Pantel, 2010). In order to extended our existing suicide lexicons and take advantage of words contextual information, we run word2vec (Mikolov et al., 2013) over 100 millions microblogs from WUP by the following steps: first, we segmented and tokenized corpus, and trained vector features for each word; top 5 semantic synonyms were chosen empirically for each word in our existing lexicons (each synonym was chosen if at least half of our experts reached agreement); we elaborately collected words into our extended suicide dictionary. Examples of suicide dictionary are shown in Table 2.

Table 2: Suicide Lexicon Table

Type	Number	Example
Suicide Words	3453	insomnia, stilnox
Suicide Phrases	3763	disapear+myself

We categorized suicide words and phrases according to 12 suicide warning signs, which differs from categories in previous work (Gao et al., 2013). Details of 12 potential suicide warning signs can be found in (Rudd et al., 2006), such as "Acting reckless" . Because when we associate each word or phrase with one or more warning signs, words or phrases become more interpretable and might describe more details of suicide.

We also extended our suicide lexicons by adding references such as I, me, mom and so on. (Li et al., 2014b) found that self-reference was used more common among suicide than none suicide's microblog. We found that suicidal ideation words (*i.e.* death, depression or estazolam) always co-occur with some particular words, such as "I", "psychologist", or "medicine". Furthermore, suicide pre-

fer to mention their families or other suicidal victims. The statistics of reference comparison between suicide and none suicide is shown in Table 3.

Table 3: Statistics of Reference Comparison

Type	Self-reference	Other-reference
Suicide	71%	29%
None Suicide	26%	22%
Example	I, myself	Dad, mother

5 Modeling

5.1 Knowledge-based Modeling

In order to take advantage of domain knowledge from psychology, the extended suicide lexicons was also used. These features are based on its subjectivity, sentiments or categories. It contains both positive and negative words. In addition, we added reference including both self-reference and other-reference as an independent category in the modeling process. For any input sentence, we count the numbers of positive, negative, suicide words, reference words according to our lexicon resources.

5.2 Syntactic Features

Syntactic features contains dependency relation, Part of Speech (POS) tagging, *etc*. Considering that some types of words or their POS(e.g., adverbs, adjectives, *etc*) are likely to convey sentiments, we obtained POS features by counting the numbers of words with the following POS tags: adjective (VA), adverb (AD), noun (NN/NR/NT), verb (VV/VE/VC), pronoun (PN) and preposition (P). Those tagging signs are from Chinese Penn Treebank Tag Set[2].

Table 4: Syntactic Features Comparison Table

Type	Suicide POS	None Suicide POS
Adjective	2.10%	1.63%
Adverb	18.50%	11.52%
Noun	20.56%	37.21%
Verb	29.95%	26.75%
Pronoun	10.36%	4.21%
Preposition	2.77%	2.98%
Total	84.24%	84.30%

---

[2]http://www.cis.upenn.edu/\~chinese/posguide.3rd.ch.pdf

We applied Stanford Parser (Toutanova et al., 2003) to acquire the statistics of POS. The statistics of data is illustrated in Table 4. Two observations were derived from the Table 4: in suicide microblogs, the users prefer to mention self-reference or other people than in none suicidal microblogs; more adverb, more verb and adjective appear in suicidal microblogs may suggest people with suicidal ideation would like to express more about their emotions or behaviors in their microblogs.

5.3 Topic Modeling

Another approach we used in our experiment is Latent Dirichlet Allocation (*LDA*) (Blei et al., 2003). It can generate predefined topics over the "bag of words" and infer topic distributions in new documents. We are interested in incorporating sentiment dictionary with topic models to make topics more interpretable. Part of LDA-induced topics are shown in Table 5.

Table 5: Part of LDA-induced topics related to suicide

Themes	Topic Words
Depression	me, depression, leave, bye
Stress & Negative	death, I won't love fear of death, to die
Anxiety	long, desperate, take medicine
Family	Mother, Father
Sadness & Hopeless	dead, don't, one day pain, past, wrong
Reference	me, we, myself, you

5.4 Topic Model with More Layers

In this paper, we hypothesize that non-sentiment words around implicit sentiment words could be affected, which may be interpreted as sentiment-propagation from word-to-word. Derived from *LDA*, sentiment associated the topic will also be reflected by its associated sentiment words. Motivated by these observations, we implemented a new approach, which adopts sentiment dictionaries into the topic model.

As illustrated in Fig. 2, the basic idea is that each suicide microblog may contain multiple topics, and each topic may associate with one or several suicide keywords. From the topic perspective, a topic associating with sentiment words could be identified as sentiment topic. Thus, words, associated with sen-

timent words within the same topic, convey some sort of sentiment, which could be viewed as the process of sentiment propagation. Each microblog has several topics labeled by sentiment words following different multinomial distributions. Therefore, a suicidal-sentiment layer can be extracted from annotating each sentiment word and computing the sentiment polarity that is associated with words and topics in documents.
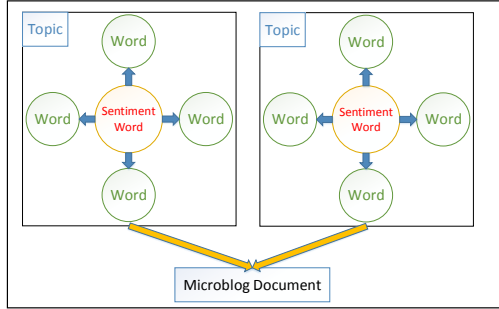

Figure 2: Basic Idea of Sentiment Propagation

While training model, psychological dictionaries are incorporated into topic modeling. The sentiment layer in our approach associates with both topic and word. Each topic associates with more sentiment words. From this perspective, the process of computing the sentiment layer could be viewed as mining for sentiments from documents and labeling suicidal words to topics on behalf of psychologists. For example, as illustrated in Fig. 3, if one microblog is about "Insomnia" and "Dysthymia", then intuitively, the topics within that microblog could be annotated by keywords, which are associated with psychological dictionaries. In general, the sentiment layer can be viewed as describing a group of words that represent a psychological mental state.
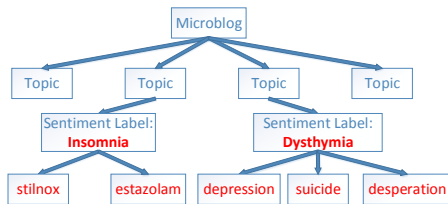

Figure 3: Extend sentiment words with dictionary

Our proposed algorithm is presented in Algo-

rithm 1. We first load the map of sentiment words with their initial polarity, $Lexicon_p$, from lexicons. We scan each microblog and annotate sentiment words within microblog. Then the labeled word will be enriched with more sentiment words, which is measured by initial sentiment polarity. Next a matrix of the data's topic multinomial probabilities, $TopicProb$, and the map of topic alphabet, $Topic_{alphabet}$, are inferenced. K is the number of topics.

Algorithm 1 Process of recomputing Topic Model

Ensure:

    Matrix of Topic Distributions, $TopicProb$;

1: Build and load $Lexicon_p$;
2: Scan each microblog and label sentiment words;
3: Label word with psychological lexicons;
4: Recompute Topic Probabilities, $TopicProb$;
5: Recompute Topic alphabets, $Topic_{alphabet}$;
6: Iterate $Topic_{alphabet}$, calculate $Polarity_k$ for each topic
7: Normalize $Polarity_k$ matrix;
8: **for** each topic multinomial probability in $TopicProb$ **do**
9:     Recompute probability using $Polarity_k$;
10:     Update topic probability in $TopicProb$;
11: **end for**
12: **return** $TopicProb$;

$$Polarity_k = \log\left(e^{\left(\sum_{k,w} P_{wk}\right)} e^{\frac{N_{ws}}{N_w} - \frac{\sum_{k,w} \widetilde{P}_{wk}}{e^{\frac{N_{ws}}{N_w}}}} \right) \quad (1)$$

The $Polarity_k$ encodes the sentiment-topic polarity for the kth topic of the microblog. P_{wk} is the initial weight of negative word w appears in k^{th} topic, and $\widetilde{P}_{wk}$ refers to initial weight of the positive word, respectively. According to (Kay et al., 1987), positive information would reduce the influence of negative emotion. However, researchers (Martin et al., 1993) found that negative effects bring longer and deeper impacts than positive effects. We thus use $e^{\frac{N_{ws}}{N_w}}$ to simulate the cumulative impact of negative sentiment impact, where N_{ws} refers to the frequency of the word w appearing in the kth topic, and

N_w refers to the number of words associated with k^{th} topic.

Given the normalized sentiment matrix at step 7, we incorporate it with original topic multinomial distribution in step 8. We recompute the topic multinomial distribution to simulate the sentiment-diffusion process as shown back in Fig. 2.

5.5 Meta Features within Microblogs

Although previous work (Lin et al., 2014) reported detecting depression from pictures in microblog, in our dataset, microblogs rarely contain pictures, thus we mainly focused on text content. We also observed that the number of critics, like, retweet surged after the suicide was reported publicly. Thus, we took three meta features into consideration: *posting type*, *posting time* and *social relationship*. Detail descriptions can be found in Fig. 1.

5.5.1 Posting Type

The posting type refers to one microblog's origin, either *original creation* or *retweet*. In our research, it appears that suicidal users prefer to post original suicide notes instead of retweeting. The comparisons in experimental data are shown in Fig. 4.
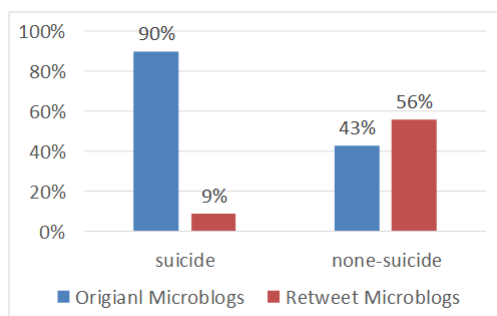

Figure 4: Comparisons in Microblog's Type

5.5.2 Posting Time

We also found that temporal features matter. Empirically, we separated 24 hours into four periods: 23:00 to 06:00, 07:00 to 13:00, 14:00 to 18:00, and 18:00 to 23:00. The result shows that suicidal microblogs are posted more frequently during 23:00 to 06:00 and less in the morning, which is in contrast to none suicidal microblogs. Specific details can be gleaned from Fig. 5. A plausible explanation

might be that some suicidal users suffer from insomnia, which hypnotic pills appear in their microblogs, such as tranquillizers or stilnox.
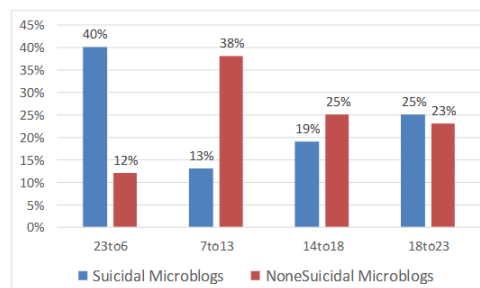

Figure 5: Comparisons in Microblogs' Time

5.5.3 Social Relationships

Microblogs also contain much useful information like social relationships, which connect microblogs to microblogs, or microblogs to users, as shown in Figure 6.
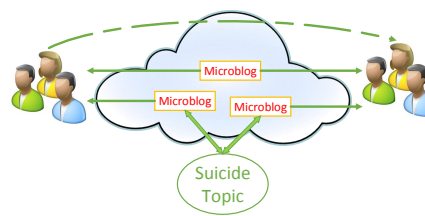

Figure 6: Social Relations within Microblogs

There is one major social relationship existed in microblog: mention. People use "@" to mention individuals or group of individuals. In addition, retweeting microblogs also generate mention behavior. According to our study, we found that several users mentioned other suicide before committing suicide. From this perspective, it could be explained by part of suicidal ideation diffusion and suicide mimic engagement among social networks (Fu et al., 2013). We employed binary value to indicate whether the microblogs have relationship with other suicide or suicide related subjects.

6 Experiment and Discussion

6.1 Experiment Approach

In the experiment, we run *LDA* (implemented by Mallet (McCallum, 2002)) to infer *k*-topic probabilities and alphabet associated with each topic. Mallet's parameters were set with default values, and stoplist was extended by Chinese punctuations. We trained *SVM* classifier by using *LibSVM* (Chang and Lin, 2011) package. The *SVM* classifier in our experiments used a *RBF* kernel and was trained by default parameter values. Weka (Hall et al., 2009), an useful machine learning tools, was also employed in our experiment for training and testing.

We run 10-fold cross validation to avoid evaluation bias. In Table 6, we list all features that were selected for training classification models in our experiment.

Table 6: Summarization of Features in Experiment

Feature	Compute Method
Knowledge-based Features	Count * Lexicon polarity
Syntactic Features	Count
Topic Model	Topic Distributions
Advanced Topic Model	Topic Distributions
Posting Type	Binary Value
Posting Time	Intervals
Social Relationships	Binary Value
N-gram	Count

The performance of classification are measured by "Precision", "Recall", "$F_1 - measure$", "Accuracy". "Precision" refers to the ratio of true suicidal microblogs against all microblogs predicted as suicidal. "Recall" refers to the fraction of suicidal instances retrieved by trained models. "Accuracy" refers to all predictions match their labels regardless whether they are suicidal microblogs or not. "$F_1 - measure$" is defined as follows 2.

$$F_1 - measure = 2 * \frac{precision * recall}{precision + recall} \quad (2)$$

6.2 Comparison between Lexicons and *LDA*

The Lexicon approach uses psychological lexicons, described in Section 3, to extract lexicon features and train the classifiers. We run *LDA* with number of topics from 100 to 1000 topics with increment of 100 (*i.e.* 100, 200, ..., 1000). Comparison of classification performance is presented in Table 7.

Table 7: Comparison between Prior-knowledge-based Features and *LDA* approach

Topics	F_1	Precision	Recall	Accuracy
100	31.2%	74.9%	19.7%	92.1%
200	47.4%	**86.5%**	32.7%	93.4%
300	53.1%	82.5%	39.2%	93.7%
400	48.8%	78.3%	35.4%	93.2%
500	56.8%	68.7%	40.4%	93.6%
600	52.1%	79.6%	38.7%	93.5%
700	59.0%	80.7%	46.5%	94.1%
800	59.5%	80.7%	47.1%	94.2%
900	**60.3%**	80.2%	**48.3%**	**94.3%**
1000	58.1%	79.9%	45.6%	94.0%
Lexicon	54.2%	85.4%	39.6%	93.9%

Table 7 shows that more topic features can improve the performance of classification. The highest $F_1 - measure$ is 60.3% with 900 topics in *LDA*. Although in the low topic dimensions *LDA* performs poorer than Lexicon approach, *LDA* could perform better than Lexicons when assigned a high topic value.

6.3 Experiment with advanced Topic Model

To test our proposed method in Section 5.4, we also conducted experiments with the same topic number and default parameter settings as in Section 6.2, and the results are shown in Table 8. Table 8 shows that compared with *LDA* in Table 7, our approach works better.

Table 8: Cross-validation performance on Topic Model after adding psychological lexicons

Topics	F_1	Precision	Recall	Accuracy
100	44.1%	95.9%	28.6%	93.4%
200	62.4%	89.6%	47.9%	94.8%
300	67.9%	93.2%	53.4%	95.4%
400	74.2%	96.8%	60.1%	96.2%
500	**76.2%**	94.6%	**63.9%**	**96.4%**
600	75.1%	**98.8%**	60.5%	96.3%
700	74.0%	95.0%	60.5%	96.1%
800	67.3%	84.0%	56.2%	95.1%
900	64.6%	76.1%	56.2%	94.4%
1000	61.8%	72.3%	53.9%	93.9%

The best performance is on 500 topics, with $F_1 - measure$ at 76.2%, Recall at 63.9%, Accuracy over 96%. The results indicate that it is feasible to predict and even prevent the suicides through analyzing microblogs.
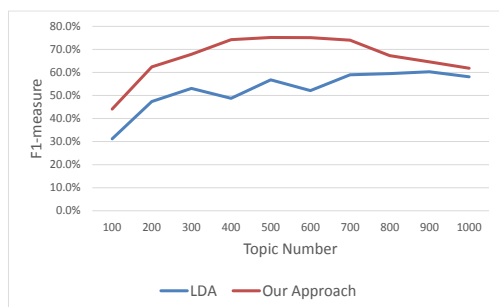
Figure 7: Comparisons in $F_1 - measure$ between *LDA* and Advanced Topic Model

Compared with previous approaches both in lexicons and *LDA*, we obtained around 25% improvement in F1-measure. Fig. 7 presents a more detailed comparison between *LDA* (in blue) and our approach (in red) on $F_1 - measure$.

6.4 Results with Meta Features

To further improve the performance, we add meta features described in Section 5.5 with topic 500. We run *SVM* and several classifiers in Weka (Hall et al., 2009): Logistic, J48 classifier, Random Forest(*RF*), Random Tree(*RT*), Decision Table(*DT*). All classifiers are trained and tested with default parameters, and performances are presented in Table 9.

Table 9: Cross-validation performance of Different classifiers

	F_1	Precision	Recall	Accuracy
SVM	76.8%	**96.8%**	63.7%	**96.5%**
Logistic	53.0%	59.2%	48.0%	92.3%
J48	**80.0%**	87.1%	**73.9%**	93.2%
RF	71.3%	98.2%	56.0%	93.6%
RT	67.7%	71.0%	64.6%	94.4%
DT	74.6%	92.0%	62.7%	96.1%

Clearly, *J48* attains the best $F_1 - measure$ (80.0%) and Recall (73.9%). Compared with Table 8, we acquired a better performance after integrating meta features.

We found that there are still about nearly one fourth suicidal microblogs that were not identified correctly. There might be several reasons: first, the complexity and ambiguity of language on the Internet, especially the *SNS*; second, the psychological lexicon is quite limited.

This research has a number of potential applications. The trained model can be used to build a suicide monitoring system to help professionals execute suicide intervention in time. If this system was effective in detecting suicide ideation in microblogs from *SNS*, it might also help psychologists investigate how linguistic and behavioral patterns are correlated with suicide thoughts and provide them with advanced decision support.

7 Conclusion

In this paper, for the purpose of identifying suicidal ideation of microblogs on social networks, first, we build a suicidal domain lexicons and develop hybrid approaches combined both contextual and meta information for suicidal ideation identification; second, we run Topic Model for feature selection with less than $1,000$ dimensions left, and achieve more than 38% accuracy increased over lexicon approach; furthermore, we deploy a real-time engine to detect suicide ideation in microblogs for monitoring suicide, which might be helpful for professional organizations to assess people's suicide ideation; finally, from psychological perspective, we found that writing styles and time variations are highly correlated with suicidal ideation. A prototype system[3] has been deployed online to detect suicide ideation of *SNS* in real-time.

The performance of model is limited by several issues as follows: first the model has been trained and tested on small size data sample; second, we need to try more advanced machine learning algorithms. Our future work is undertaken in two directions: improving performance and mining latent social relationships.

8 Acknowledge

---

[3]It can be visited at `http://ccpl.psych.ac.cn/suicide/`.

References

David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022.

Peter F Brown, Peter V Desouza, Robert L Mercer, Vincent J Della Pietra, and Jenifer C Lai. 1992. Class-based n-gram models of natural language. *Computational linguistics*, 18(4):467–479.

Chih-Chung Chang and Chih-Jen Lin. 2011. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27.

Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting depression via social media. In *ICWSM*.

Zhendong Dong and Qiang Dong. 2003. Hownet-a hybrid language and knowledge resource. In *Natural Language Processing and Knowledge Engineering, 2003. Proceedings. 2003 International Conference on*, pages 820–824. IEEE.

King-wa Fu, Qijin Cheng, Paul WC Wong, and Paul SF Yip. 2013. Responses to a self-presented suicide attempt in social media: A social network analysis. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*, 34(6):406.

Rui Gao, Bibo Hao, He Li, Yusong Gao, and Tingshao Zhu. 2013. Developing simplified chinese psychological linguistic analysis dictionary for microblog. In *Brain and Health Informatics*, pages 359–368. Springer.

Li Guan, Bibo Hao, and Tingshao Zhu. 2014. How did the suicide act and speak differently online? behavioral and linguistic features of china's suicide microblog users. *arXiv preprint arXiv:1407.0466*.

Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H Witten. 2009. The weka data mining software: an update. *ACM SIGKDD explorations newsletter*, 11(1):10–18.

Jared Jashinsky, Scott H Burton, Carl L Hanson, Josh West, Christophe Giraud-Carrier, Michael D Barnes, and Trenton Argyle. 2013. Tracking suicide risk factors through twitter in the us.

Stanley R Kay, Abraham Flszbein, and Lewis A Opfer. 1987. The positive and negative syndrome scale (panss) for schizophrenia. *Schizophrenia bulletin*, 13(2):261.

Lin Li, Ang Li, Bibo Hao, Zengda Guan, and Tingshao Zhu. 2014a. Predicting active users' personality based on micro-blogging behaviors. *PLoS ONE*, 9(e84997).

Tim MH Li, Michael Chau, Paul SF Yip, and Paul WC Wong. 2014b. Temporal and computerized psycholinguistic analysis of the blog of a chinese adolescent suicide.

Chenghua Lin and Yulan He. 2009. Joint sentiment/topic model for sentiment analysis. In *Proceedings of the 18th ACM conference on Information and knowledge management*, pages 375–384. ACM.

Huijie Lin, Jia Jia, Quan Guo, Yuanyuan Xue, Jie Huang, Lianhong Cai, and Ling Feng. 2014. Psychological stress detection from cross-media microblog data using deep sparse neural network.

David D Luxton, Jennifer D June, and Julie T Kinn. 2011. Technology-based suicide prevention: current applications and future directions. *Telemedicine and e-Health*, 17(1):50–54.

Leonard L Martin, David W Ward, John W Achee, and Robert S Wyer. 1993. Mood as input: People have to interpret the motivational implications of their moods. *Journal of Personality and Social Psychology*, 64(3):317.

Andrew Kachites McCallum. 2002. Mallet: A machine learning for language toolkit.

Qiaozhu Mei, Xu Ling, Matthew Wondra, Hang Su, and ChengXiang Zhai. 2007. Topic sentiment mixture: modeling facets and opinions in weblogs. In *Proceedings of the 16th international conference on World Wide Web*, pages 171–180. ACM.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*, pages 3111–3119.

World Health Organization et al. 2014. *Preventing suicide: A global imperative*. World Health Organization.

Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*, 2(1-2):1–135.

John Pestian, Henry Nasrallah, Pawel Matykiewicz, Aurora Bennett, and Antoon Leenaars. 2010. Suicide note classification using natural language processing: A content analysis. *Biomedical informatics insights*, 2010(3):19.

John P Pestian, Pawel Matykiewicz, Michelle Linn-Gust, Brett South, Ozlem Uzuner, Jan Wiebe, K Bretonnel Cohen, John Hurdle, and Christopher Brew. 2012. Sentiment analysis of suicide notes: A shared task. *Biomedical informatics insights*, 5(Suppl 1):3.

Philip Resnik, Anderson Garron, and Rebecca Resnik. 2013. Using topic modeling to improve prediction of neuroticism and depression. In *Proceedings of the 2013 Conference on Empirical Methods in Natural*, pages 1348–1353. Association for Computational Linguistics.

M David Rudd, Alan L Berman, Thomas E Joiner, Matthew K Nock, Morton M Silverman, Michael

Mandrusiak, Kimberly Van Orden, and Tracy Witte. 2006. Warning signs for suicide: Theory, research, and clinical applications. *Suicide and Life-Threatening Behavior*, 36(3):255–262.

Jian Sun. 2014. Ansj Chinese Lexical Analysis System. Website. `http://ansj.org/`.

Kristina Toutanova, Dan Klein, Christopher D. Manning, and Yoram Singer. 2003. Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network. In *North American Chapter of the Association for Computational Linguistics*.

Peter D. Turney and Patrick Pantel. 2010. From Frequency to Meaning: Vector Space Models of Semantics. *Computing Research Repository*, abs/1003.1.

Hongning Wang, Yue Lu, and Chengxiang Zhai. 2010. Latent aspect rating analysis on review text data: a rating regression approach. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 783–792. ACM.

Xinyu Wang, Chunhong Zhang, Yang Ji, Li Sun, Leijia Wu, and Zhana Bao. 2013. A depression detection model based on sentiment analysis in microblog social network. In *Trends and Applications in Knowledge Discovery and Data Mining*, pages 201–213. Springer.

Casey Whitelaw, Navendu Garg, and Shlomo Argamon. 2005. Using appraisal groups for sentiment analysis. In *Proceedings of the 14th ACM international conference on Information and knowledge management*, pages 625–631. ACM.

Chung-Hsien Wu, Liang-Chih Yu, and Fong-Lin Jang. 2005. Using semantic dependencies to mine depressive symptoms from consultation records. *Intelligent Systems, IEEE*, 20(6):50–58.

Liang-Chih Yu and Chun-Yuan Ho. 2014. Identifying emotion labels from psychiatric social texts using independent component analysis. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 837–847, Dublin, Ireland, August. Dublin City University and Association for Computational Linguistics.