Sentimental Statement Analysis to Prevent Suicide

Md Sharif Khan , Shakera Chowdhury , Iffat Zerin
Department of Computer Science and Engineering,
East West University, Dhaka, Bangladesh
Email: (sharif60.052, mail2shakerachowdhury, zeriniffat2)@gmail.com
ID :2015-3-60-052, 2015-3-60-051, 2015-3-60-035

***Abstract*—People tend to express their feelings and sufferings in online communities because of the anonymity of online media and social networks. To prevent suicide, natural language processing methods need to detect suicide - related posts and the suicide identification of users in cyberspace. We focus on the online community called Reddit and classify posts of users with potential suicide and suicidal risk through text - based processing features and methods based on machine learning. This paper proposes a model for detecting depressive and non-depressive posts from reddit posts. We developed a system using machine learning concepts and algorithms to detect a persons tendency to self-destruction by observing the thoughts and attitudes towards suicide and similar factors. This system consist of five algorithms that are Nave Bayes, Decision Tree, Support Vector Machine (SVM), Logistic Regression and Neural Network**

***Index Terms*—Machine learning, Nave Bayes, Decision Tree, Support Vector Machine (SVM), Logistic Regression and Neural Network, suicidal prediction, mental depression.**

I. Introduction

Suicide is one of the most relevant issues now a day. Just as it is common for people to share their personal thoughts, ideas and perspectives on social media such as facebook, twitter, blogs and other sites, sharing suicidal thoughts becomes widespread[1]. Occurrence of suicide incidents are increasing all over the world. Every year more than one million people commit suicide. Mental stress, grief and sorrow are the most common reason for suicide. The victims think that suicide is the only solution to escape from their misery. According to WHO, worldwide more people die from suicide than in warfare. Suicide rate in Bangladesh reached 8,879 or 1.13 % of total death, according to WHO data published in 2017. However, this serious mental health problem needs more attention. People who have a tendency to commit suicide are seen to share their thoughts, feelings and sorrow in the social media which indicate that automated monitoring system is needed to reduce the suicide rate or identify the people who are at risk of self-destruction[1]. Basically we tried to focus on the statements that people post on social media to express their feelings. If we can figure out through the statements of people that how much depress they are then by taking preparations suicide can be prevented. We take into account every explanation posted by users. Five different classification models are implemented for this task, such as Decision Tree, Support Vector Machine (SVM), Naive Bayes, Logistic Regression, Long Short-Term Memory. A combination of these classification models is also being tested to further improve prediction accuracy. Using scikit learning, these models are implemented using k - fold (k=10) cross - validation to prepare from the training dataset, and then predict using the dataset[2]. This research aim is to develop a system predicting the suicidal tendency observing sentimental statement from Bangla texts. Our contributions in this paper:

- Identify the peoples thoughts and attitudes towards suicide.
- To analysis the sentimental statement to prevent suicide.

II. Related Work

The importance of sentiment analysis is increasing in machine learning field. There are several work done in this topic [3]. used Support Vector Machine (SVMs) for analyzing sentimental data. They use SVMs for combining diverse source of potentially pertinent information. From Opinions.com movie review data has taken for analysis and uni-gram model also has been used. SVMs based with uni gram style feature provides a good result for real-valued favourably measures [4]. introduced an technique to extract emotion from 900 suicide note by labeling probable sentiment. They suggested a hybrid system consisting of both machine learning concepts and rule based approach. They obtain precision of 41.79% with recall of 55.03% for an f-measure of 47.50%. The overall mean f measure of all submissions was 48.75% with a standard deviation of 7%[5], recommended an automated approach consisting of lexico-syntactic rules with few models of machine learning to classify emotion in the suicide note. It results the best micro F-measure of up to 53.36%, precision 67.17%. Finally to classify suicide notes, automated text mining approach would be beneficial proposed by them [6]. bring together classification rules, meta learners, function models, decision trees, instance-based learner and lazy learners to analyze suicide notes using feature selection method. Accuracy of learning based algorithm is 78%, mental health professionals 63percent and trainees are 49%. Determining suicidal ideation of people via their public post is the main objective of text-based classification. Key term related to suicide phrase filtering [7] and keyword filtering [8], [9] are used in this type of method. Supervised learning, NLP (natural language processing) as well sentence embedding [10], word embedding [11] are implemented in this sector. Using Becks suicide intent scale to determine suicide tendency Chattopadhayay [12] recommended model. For a singular vector decomposition and text TF-IDF (term frequency-Inverse Document Frequency) are mentioned as well formed a dictionary regarding to the topic that contain suicidal ideation [13]. SVM(support Vector Machine) classification method

implemented to determine cyber suicide and a psychological dictionary formed [14]. Therefore, experiments indicates that machine learning can be used to analyze or detect sentimental issues.[1]

III. Methodology

Methodology is the research area's key term. A researcher should select a proper method to complete the research. There are also few parts of a method such as data collection, data analysis, ethics, etc. For any research study, data collection is one of the most important components. The data collected affect the study as a whole and also lead the valid or invalid results. We use the Reddit site for user statements to collect data manually. And we have applied various machine learning algorithms to analyze their emotions. We have already come to know from the above discussion that our dataset contains two classes, whether the statement is depressive or not. First we need to understand the problem definition to classify the statements, then we go to our model and evaluate the outcome. Machine learning is full of algorithms and we tried to implements some of them to find the best accuracy according to our dataset.

Therefore, surveys are applicable in terms of data collection. Collecting information about facts and attitudes of people, survey is best suited.

This is a new application of knowledge discovery and data mining to detect suicidal thinking in the content of online users. Psychological experts with statistical analysis conducted previous work in this field; this approach reveals knowledge of suicidal identification from a data analytical perspective. Our analysis reveal that suicidal individuals often use personal pronouns to display their ego. In their dialogue, they are more likely to use words that express negativity, anxiety, and sadness. They are also more likely to choose the present tense to describe their suffering and the tension of the future to describe their hopelessness and suicide plans.

This paper presents the Reddit platform and gathers a new dataset to detect suicidal ideation. Reddit's SuicideWatch BBS is a new online channel to express anxiety and pressures for people with suicidal ideation. Social volunteers respond to depression in positive, supportive ways and hopefully prevent potential suicides. This source of data is not only useful to detect suicide, but also to study how to effectively prevent suicide through effective online communication.

A. Algorithms

We have used 5 different types of machine learning algorithms in our model and as our programmable language we used Python 3.6.5 for the implementation work. Decision Tree, Bayesian Model, Logistic Regression Support Vector Machine are the classification models we implemented using the above mentioned dataset. Long Short-Term Memory is also implemented as one of the most famous deep learning methods to see how well our data fit into the model. These algorithms are good for different classifications and they have different datasets based on their own properties and performance. Naive

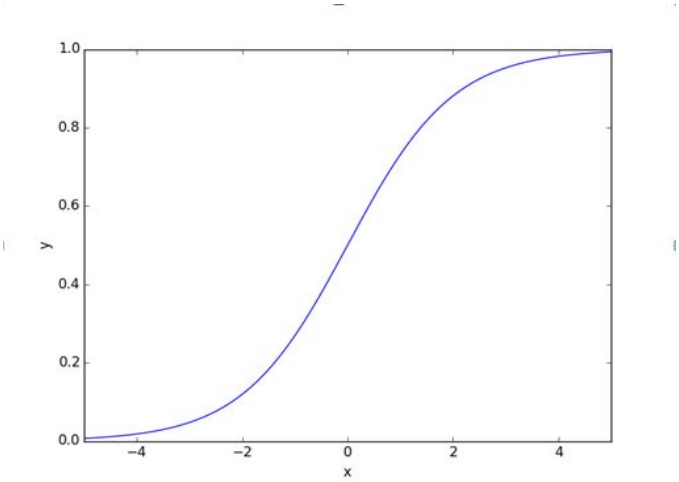


Fig. 1. Sigmoid Function

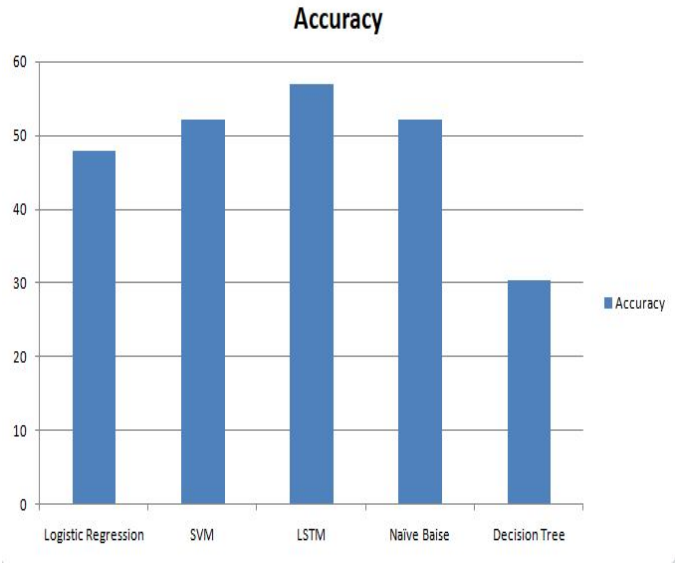


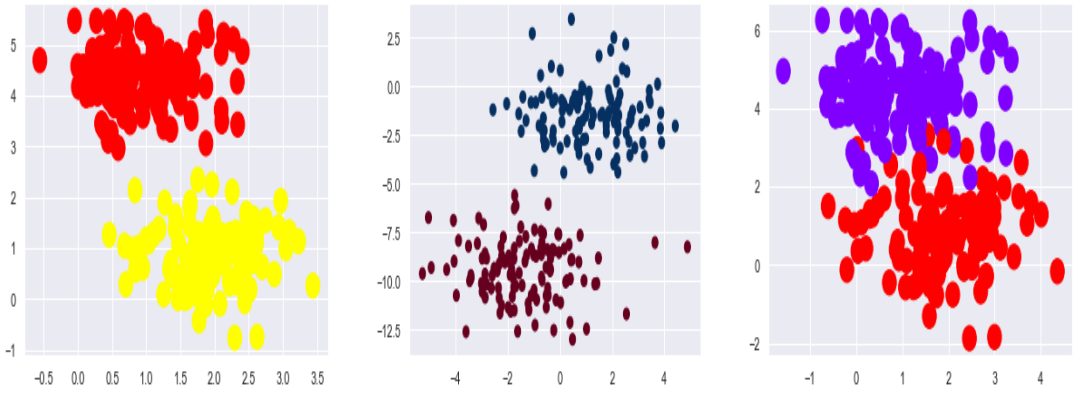
Fig. 2. Accuracy Comparison of different algorithm

Bayes, Logistic Regression and SVM are more commonly used algorithms for classification issues, as we said earlier in the Decision Tree.

B. Data and Knowledge

We collect suicidal ideation texts from Reddit and Facebook as well as some non-depressive texts. Then we check all posts manually to make sure they have been properly labeled.

Reddit Data: Reddit is an online community registered to aggregate social news and discussions online. It is made up of many topic categories and is called a subreddit for each area of interest within a topic. The content of online users includes body of text in this dataset. We collected posts from a subreddit called "SuicideWatch"(SW) (https:/www.reddit.com/r/SuicideWatch/) with potential suicide intentions. Non-suicidal data collection is entirely a user-generated content, and news aggregation and administrator



(a)Support Vector Machine (b)Naive Bayes Algorithm (c) Decision Tree

Fig. 3. Data visualization in different algorithm

TABLE I
PRECISION, RECALL AND F1-SCORE FOR LOGISTIC REGRESSION, DECISION TREE, NAVE BAYES, SVM

Box Name	Precision	Recall	f1 score
Logistic Regression	0.57	0.48	0.40
Decision Tree	0.26	0.30	0.25
Naive Bayes	0.63	0.52	0.47
SVM	0.63	0.52	0.47

posts are excluded. We will study the balanced dataset in Reddit and study the imbalanced dataset on Twitter in the following subsection to facilitate study and demonstration. The Reddit dataset includes 160 suicidal ideation samples and a number of nonsuicide texts. The dataset includes many subreddits, one from SuicideWatch and another from popular Reddit posts such as gaming, jokes, books, etc.

Facebook Data: People post their most personal thoughts on their Facebook feeds every day, entrusting information that they may never trust to an actual person to the internet. While these posts may seem meaningless to other users. However we collected rest of the data from facebook status of users which seemed depressive also some non-depressive texts. After collecting data we translated them into Bangla language to test that whether our selected algorithms will fit our dataset or not.

C. Preparing the datasets

First of all, we need to process our dataset to apply these algorithms. We used libraries of Pandas, Numpy, Scikit - learn, keras. As we deal with text data, to process the dataset, we will implement the following different ideas. Those are as follows: Count Vectors, TF-IDF, Word Embedding

1) Count Vectors: Count Vector represents a notation in the form of a matrix dataset matrix notation in which each row represents a corpus document, each column represents a corpus term, and each cell represents the frequency count of a specific term in a specific document

2) TF-IDF: -IDF shows how often a term is in a document as a whole. It attempts to assign a metric value to represent that term's presence. This is used extensively and frequently in text mining. This weight is a factual measure used in a gathering or corpus to evaluate how essential a word is to a report

Based on several types of input tokens like words, characters, n-grams, TF-IDF Vectors can be generated.

3) Word Embedding: A word embedding is a form of word and document representation with a dense vector representation. The text learns the location of the word in the space of the vector and is based on the words used around the word. There are four key steps:

- Embedding of the pre-trained word
- Create an object tokenizer
- Transform text documents into tokens ' sequence and pad them
- Create a token mapping and its embedding

We divide our dataset into 60% train data and 20% test data including 20% LSTM, RNN model validation and use kfold cross validation for our other models where k = 10. We used the embedding layer, hidden layer, input layer and output layer in the LSTM, as it is a part of the neural network. We used Sigmoid activation, glorot normal kernel initializers, kullback leibler divergence loss and 'Nadam' as an optimizer.

4) Sigmoid activation: The Sigmoid curve looks like a function in S shape demonstrated in figure 1. Sigmoid is used because it exists between 0 and 1. It is therefore particularly used for models in which the probability as an output must be predicted. Since there's only a chance of anything from 0 to 1, Sigmoid's the right decision.

IV. RESULT AND DISCUSSION

We collected a corpus of 200+ text posts from Reddit and manually labeled them based on the emotions:

- Texts containing positive emotions, such as happiness, amusement or joy we labeled them 0

- Texts containing negative emotions, such as sadness, anger or disappointment we labeled them 1.
- Then we implemented all the algorithms mentioned above using jupyter and some python libraries such as pandas, sklearn, keras.

Firstly, Logistic regression model was tested on each of the feature vectors mentioned above. It gives 47% accuracy on TF-IDF feature. Then Decision Tree model was performed and it gave 30.44% accuracy on the same feature. Nave Bayes and SVM gave same accuracy which is 52% and finally we applied LSTM which gave the highest accuracy. The accuracy comparison shown in figure 2.

Classification algorithms can also give a perfect visualization of a certain dataset. To visualize our dataset according the algorithms we have implemented some of are presented by figure 3:

Algorithms precision. recall, f1-score show in table I:

Here the first figure is for SVM , second one is for Nave Bayes and third one is for Decision Tree. After importing necessary libraries from python we can see how the data are plotted in 2D area based on X and Y axis. As there are two classes in our dataset depressive texts and non-depressive texts we could see two different color data point has been plotted in the graph to differentiate the data points.

A. Discussion

After implementation of five mentioned algorithms in our dataset its cleared that LSTM gave the best accuracy but the question is why does it fit to the dataset? Our dataset have two columns one is for labeling and the other one is for the statements of the users of reddit. LSTM basically works betters for qualitative data and it is the drawback of RNN (Recurrent Neural Network) which can predict better for long term duration. LSTM allows us to produce reliable explanations of which words in individual texts are responsible for sentiment attributes compared to the explanations obtained by using other classification algorithms that we used in our dataset. LSTM model for the prediction of sentences sentiment, demonstrating that the resulting word relevance reveals trustworthy words that support the decision of the classifier for or against a specific class and performs better. Finally it can be said that through machine learning if machine can analyze the mental condition of a person then their suicidal tendency can also be predicted which will help further to stop suicidal incidents.

V. Conclusion

In this paper we presented dataset gathered from Reddit. We presented our method of collecting and annotating the dataset. We investigated the properties and the statistics of the dataset and tried to implement some classification algorithms. The dataset in this examination is relied upon to be utilized for arrangements which utilized machine learning based statistical calculations, for example, Decision Tree, Support Vector Machines (SVM), Naive Bayes (NB), Recurrent Neural Network (RNN), Logistic Regression (LR), Long Short Time Memory (LSTM). We tried to find which of them works best and fit for the dataset and found that LSTM gave the best accuracy.

VI. Future Work

Our planned next steps include:

- Increase the size of the dataset.
- Discuss the issue of unbalanced dataset and text classification.
- Extend the generated method either automated or manually.

References

[1] S.-f. F. S. P. Shaoxiong Ji, Celina Ping Yu and G. Long, "Supervised learning for suicidal ideation detection in online user content," vol. 2018, 2018, p. 10.

[2] L. Arras, G. Montavon, K.-R. Mller, and W. Samek, "Explaining recurrent neural network predictions in sentiment analysis," 09 2017.

[3] T. Mullen and N. Collier, "Sentiment analysis using support vector machines with diverse information sources," in *Proceedings of EMNLP 2004*. Barcelona, Spain: Association for Computational Linguistics, Jul. 2004, pp. 412–418. [Online]. Available: https://www.aclweb.org/anthology/W04-3253

[4] A. Nikfarjam, E. Emadzadeh, and G. Gonzalez, "A hybrid system for emotion extraction from suicide notes," *Biomedical informatics insights*, vol. 5, pp. 165–74, 01 2012.

[5] A. Kovačević, A. Dehghan, J. A. Keane, and G. Nenadic, "Topic categorisation of statements in suicide notes with integrated rules and machine learning," *Biomedical informatics insights*, vol. 5, pp. BII–S8978, 2012.

[6] J. Pestian, H. Nasrallah, P. Matykiewicz, A. Bennett, and A. Leenaars, "Suicide note classification using natural language processing: A content analysis," *Biomedical informatics insights*, vol. 3, pp. BII–S4706, 2010.

[7] J. Jashinsky, S. H. Burton, C. L. Hanson, J. West, C. Giraud-Carrier, M. D. Barnes, and T. Argyle, "Tracking suicide risk factors through twitter in the us," *Crisis*, 2014.

[8] Y.-P. Huang, T. Goh, and C. L. Liew, "Hunting suicide notes in web 2.0-preliminary findings," in *Ninth IEEE International Symposium on Multimedia Workshops (ISMW 2007)*. IEEE, 2007, pp. 517–521.

[9] K. Varathan and N. Talib, "Suicide detection system based on twitter," in *2014 Science and Information Conference*. IEEE, 2014, pp. 785–788.

[10] T. Shen, T. Zhou, G. Long, J. Jiang, S. Pan, and C. Zhang, "Disan: Directional self-attention network for rnn/cnn-free language understanding," in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[11] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.

[12] S. Chattopadhyay, "A mathematical model of suicidal-intent-estimation in adults," *American Journal of Biomedical Engineering*, vol. 2, no. 6, pp. 251–262, 2012.

[13] E. Okhapkina, V. Okhapkin, and O. Kazarin, "Adaptation of information retrieval methods for identifying of destructive informational influence in social networks," in *2017 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA)*. IEEE, 2017, pp. 87–92.

[14] X. Huang, L. Zhang, D. Chiu, T. Liu, X. Li, and T. Zhu, "Detecting suicidal ideation in chinese microblogs with psychological lexicons," in *2014 IEEE 11th Intl Conf on Ubiquitous Intelligence and Computing and 2014 IEEE 11th Intl Conf on Autonomic and Trusted Computing and 2014 IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops*. IEEE, 2014, pp. 844–849.