

Suicidal Tendency Detection using Social Media Data

Sachit Nair¹ • Harisha S¹ • Riddhi Yogesh Phade¹ • Alan Shaji¹

Abstract Mental health is an often neglected area in our lives, tragically many people end up suffering and taking their own lives just because they didn't receive proper help at the right moment. This project proposes a solution that can hopefully provide help to those who are in dire need of it. We will use Machine learning (NLP in particular) to identify if a particular individual is showing signs of a mental health crisis. If it identifies such a person with a high enough confidence, then the system would alert a mental health professional who can then quickly provide the much-needed help and

Sachit Nair, Harisha S, Riddhi Yogesh Phade , Alan Shaji have contributed equally to this work

Significance Statement This paper is an attempt to help those who may be considering self-harm by detecting their distressed state and providing them help before they are able to do damage to themselves. Thus, the aim of this paper is to prevent possible self-harm attempts and suicides, thereby helping many victims of mental health issues.

Email Harisha S
harisha.s2020@vitstudent.ac.in

Sachit Nair
sachit.nair2020@vitstudent.ac.in

Riddhi Yogesh Phade
riddhiyogesh.phade2020@vitstudent.ac.in

Alan Shaji
alan.shaji2020@vitstudent.ac.in

guidance. This project is intended to make sure no other life is thrown away and to alleviate suffering.

Keywords Suicide - Social Media - Prediction - Machine Learning

Introduction

Death by suicide is one of the leading causes of death among young adults. Suicides generally do not happen out of the blue, usually, the victim displays prolonged periods of distress. During this phase, there are some significant changes in behavior and actions.

If we were somehow able to use this information to our advantage and recognize if a person is at risk of suicide, we could give them the help they require. This is the problem we attempt to solve.

Using machine learning techniques, we can perform sentiment analysis and predict if a person is distressed based on the way they chat or write posts. We can then do many things like pass this information to a certified mental health professional or conduct further analysis to be even more certain.

Over 800,000 people have died due to suicide every year worldwide. India has a suicide rate of 10.3 Of which 71% of people are below the age of 44 years. That is 71% of people attempting suicide in India are of working age and active on social media apps for daily communication. Early detection of such suicidal thoughts might help in preventing suicide attempts.

However, there is still lack of effective methods to identify an individual's will to attempt suicide. Researchers have examined people's behaviour and tendencies to determine suicidal ideation in two main perspectives. The first one is a traditional method of clinical counseling which can be burdensome for individuals or use of questionnaires to assess the individual's mental health (4). The second technique is through social media posts and texts. The suicidal ideation is identified through assessing and analysing data over the social media using machine learning algorithms. Studies have shown that young people trust the Internet to find out what's wrong with them. They also seek help for online counseling or just chatting with strangers. (5,6). With increasing users of social media apps like Facebook and Twitter and people creating new trends, people post notes about their problems and depression through which suicidal ideation could be detected.

Detecting ideas for suicidal attempts in people through social media and machine learning algorithms might help professionals intervene earlier. Such an idea has been applied on real dataset and has received a great feedback from psychiatrists. For example, Chiang et al. developed an early warning system for detecting suicidal ideation based on Facebook data to identify users with potential of suicidal ideation. (7)

Previous suicide Detection methods:

Features used for detection	Machine learning algorithm used	Social media app	Performance summary
Vocabulary (8)	Naïve Bayes	Twitter	Performance: Accuracy is 0.6315 in Leave One Out validation (LOO) and 0.6327 in 10-CV.

Word bags, Polarity dictionary, LSA topic model, Named entities (9)	LIBSVM	Dutch-language forum	Performance: F1 is 0.93 for relevant messages, 0.70 for severe messages.
Demographic features, Emotion labels, (10)	Logistic regression	Twitter	Performance: F1 is 0.53.
N-gram, Word vectors, Document vectors (11)	Random Forest, SVM, CEM, deep learning, ensemble models	Microblogging and movie reviews domain	Performance: The F1-scores are 0.7302, 0.6379, 0.7532, 0.7181, 0.8120 and 77.41 in the six datasets, respectively.

Linear Support Vector Classifier produces the optimal separating hyperplane that separates the data such that it maximizes the distance between it and the nearest data point of each class. It is used for classification problems, and works well on unseen examples. (Gunn 17)

A multilayer perceptron Classifier is used to perform classification of the texts as suicidal and non-suicidal. The MLP network takes a feature vector extracted from the object to be classified as input, and outputs the class of the object. (Ruck et al. 2)

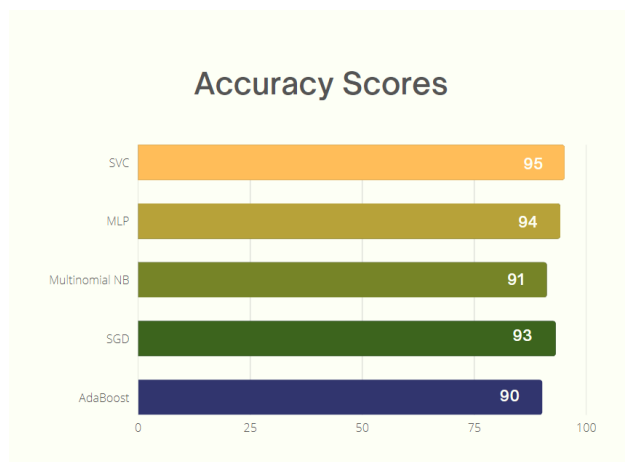
Multinomial Naive Bayes is a supervised learning technique utilized for automatic text classification. It outperforms the Multivariate Bernoulli event model, and is computationally very efficient, and simple to implement. (Kibriya et al. 1)

Stochastic Gradient Descent Classifier implements a SGD learning routine alongside multiple loss functions and penalties for classification. It is an optimization algorithm that is used to identify function parameters that minimizes the cost function.

The AdaBoost algorithm is an ensemble method which utilizes weak classifiers to make the final prediction. It produces a series of stumps, where every subsequent tree is built to correct the errors in the previous model.

The dataset used in the project is sourced from social media websites, Twitter and Reddit. A consolidated dataset containing the text and it's classification was made with more than 240,000 data records.

After training models with Linear Support Vector Classifier (SVC), Multi Layer Perceptron (MLP), Multinomial Naive Bayes, Stochastic Gradient Descent Classifier (SGD) and AdaBoost, these are the accuracy scores for each model:



Tabulation

Classification Models		Precision	Recall	F1 score
Linear SVC	0	0.95	0.95	0.95
	1	0.95	0.95	0.95
	accuracy			0.95
MLP	0	0.95	0.94	0.95
	1	0.94	0.95	0.94
	accuracy			0.94
MultinomialNB	0	0.95	0.85	0.90
	1	0.86	0.95	0.91
	accuracy			0.90
SDG	0	0.93	0.93	0.93
	1	0.93	0.93	0.93
	accuracy			0.93
Adaboost	0	0.87	0.93	0.90
	1	0.92	0.86	0.89
	0			0.90

Result Analysis

After using the datasets and training models using the above 5 techniques, it is clear that Linear Support Vector Classifier gives the maximum accuracy of 95%.

The reason for this is that different machine learning techniques have different sensitivity levels to different datasets. This is known as the bias-variance trade-off. It is a measure of how sensitive an algorithm is to a particular dataset. A higher variance value means the algorithm is more sensitive to the specific dataset and a lower variance means vice versa.

It is calculated as:

$$\text{Error}(\text{Model}) = \text{Variance}(\text{Model}) + \text{Bias}(\text{Model}) + \text{Variance}(\text{Irreducible Error})$$

after training the model with the dataset.

So we can conclude on the technique we have to use to classify if a given person is suicidal or not

Future scope

The findings of this paper may be of use in the development of many applications of varied scope. In this case, the model was applied to detect distressed emotions in a user via their messages, but with slight modifications, it may be applied to detect other emotions as well.

Some applications of this include

Automatic moderation of messages, for example in applications where children interact, no inappropriate messages must be sent.

Chatbot that can hold a conversation keeping emotions in mind. Leading to a richer conversation.

References

1. Gunn, Steve R. "Support vector Machines for Classification and Regression." *ISIS Technical report*.
2. Kibriya, Ashraf M., et al. "Multinomial Naïve Bayes for Text Categorization Revisited."
3. Ruck, Dennis W., et al. "Feature Selection Using a Multilayer Perceptron." *n Journal of Neural Network Computing*, vol. 2, no. 2, 1990, pp. 40-48.
4. .Weber A.N., Michail M., Thompson A., Fiedorowicz J.G. Psychiatric emergencies: Assessing and managing suicidal ideation.
5. Colombo G.B., Burnap P., Hodorog A., Scourfield J. Analysing the connectivity and communication of suicidal users on twitter.
6. Daine K., Hawton K., Singaravelu V., Stewart A., Simkin S., Montgomery P. The power of the web: A systematic review of studies of the influence of the internet on self-harm and suicide in young people.
7. Chiang W.C., Cheng P.H., Su M.J., Chen H.S., Wu S.W., Lin J.K. Socio-health with personal mental health records: Suicidal-tendency observation system on Facebook for Taiwanese adolescents and young adults;
8. Abboute A., Boudjeriou Y., Entringer G., Azé J., Bringay S., Poncelet P. Natural Language Processing and Information Systems, Proceedings of the International Conference on Applications of Natural Language to Data Bases/Information Systems, Lecture Notes in Computer Science, Montpellier, France, 18–20 June 2014. Volume 8455. Springer; Cham, Switzerland: 2014. Mining Twitter for Suicide Prevention;
9. Desmet B., Hoste V. Online suicide prevention through optimised text classification
10. Coppersmith G., Ngo K., Leary R., Wood A. Exploratory Analysis of Social Media Prior

to a Suicide Attempt; Proceedings of the
Third Workshop on Computational
Linguistics and Clinical Psychology.

11. Araque O., Corcuera-Platas I.,
Sánchez-Rada J.F., Iglesias C.A. Enhancing
deep learning sentiment analysis with
ensemble techniques in social applications.