Detection of Suicide Related Posts in Twitter Data Stream

[1]Mr. Vishal Shinde, [2]Mr.Vaibhav Khare, [3]Mr.Atul Patil, [4]Mr.Pravin Shinde

[1]Asst.Professor, [2, 3,4]UG Student, [1,2,3,4]Computer Engg. Dept. Shivajirao S. Jondhle College of Engineering & Technology, Asangaon, Maharashtra, India.

[1]mailme.vishalshinde@gmail.com,[2]kharevaibhav519@gmail.com,[3]atulp082@gmail.com, [4]spravin1852@gmail.com

Abstract- Suicidal ideation detection in online social networks is an emerging research area with major challenges. Recent research has shown that the publicly available information, spread across social media platforms, holds valuable indicators for effectively detecting individuals with suicidal intentions. The key challenge of suicide prevention is understanding and detecting the complex risk factors and warning signs that may precipitate the event. In this paper, present a new approach that uses the social media platform Twitter to quantify suicide warning signs for individuals and to detect posts containing suicide-related content. The main originality of this approach is the automatic identification of sudden changes in a user's online behavior. To detect such changes, its combine natural language processing techniques to aggregate behavioral and textual features and pass these features through a martingale framework, which is widely used for change detection in data streams. Experiments show that the text-scoring approach effectively captures warning signs in text compared to traditional machine learning classifiers. Additionally, the application of the martingale framework highlights changes in online behavior and shows promise for detecting behavioral changes in at-risk individuals.

Keywords: Twitter; Social-network Services; Real-time Systems; Pragmatic; Information analysis; Feature extraction.

I. INTRODUCTION

Currently, a social network is an important source of exchange information and communication now a days 900 social media sites available on the internet.so many people exchange their views and day to day activity. Millions of people are using twitter and post their views.

As per the research of the world health organization (WHO) 800,000 people tack an action of suicide attempts every day. According to this calculation of suicide attempts has appreciated to world health Organization member states and commits to decrease rate of suicide. American Foundation for Suicide Prevention (AFSP) has found, the major point of suicide Prevention categories risk factors.1) Health Factors 2) Environmental Factor 3) Historical Factors as a Result, It showed the Grapple and Notation on Facebook. And Social Media Preceding Proposed Detection Methods rely heavily on manually annotated speech, which can limit their effectiveness due in part to the varying forms of suicide warning signs in at risk Individuals.

Sometimes users post harmful or suicide-related tweets on their account on twitter or user's text messages to their friend about making suicide so its main work to detect and identify that post of users. Admin keeps watch on users posts so that users cannot do anything wrong with their life.

The Main Motto of this paper find answer from field of Psychology and others is to detect sudden behavioral changes.

II. AIMS AND OBJECTIVE

a) Aim

The main Aims of this set-up to Decreased suicide attempts. Many People Share their day to day routine and whatever Happened in Their Life. With this system Cybercrime easily Identify the Suspect.

b) Objective

- The Objective of Track and Identify Accurate Suspect.
- Use of This System Evaluate Performance and Decreased the Rate of Suicide.

III. LITERATURE SURVEY

Paper 1:- Prediction of Suicidal Ideation in Twitter Data using Machine Learning algorithms.

Using Twitter data that can automatically analyze the sentiments of these tweets. Then investigate a tool of data mining to extract useful information for classification of tweets collected from Twitter based on machine learning classification algorithms.

Experimental results show that the method for detecting the suicidal acts using Twitter data and the machine learning algorithms verify the effectiveness of performance in terms of recall, precision, and accuracy on sentiment analysis.[2]

Paper 2:- Suicide-Related Text Classification with Prism Algorithm

The aim is to contribute to the research relating to suicide communication on social media.

The output of the study shows that the Prism algorithm has exceeded the other machine learning algorithms with an F-measure of 0.84 for the target classes. This result, to the best of knowledge, is the highest performance that has been achieved in classifying social media suicide-related text. [3]

Paper 3:- Suicidal Trend Analysis of Twitter using Machine Learning and Neural Network.

Support Vector Machine (SVM) is one of the best machine learning algorithms for text analysis and neural network is also well known for performance in complex cases. In the case of the neural network, there used three types of weight optimizers namely Limited-memory BFGS, Stochastic gradient descent and an extension of stochastic gradient descent which is Adam to attain maximum accuracy. [4]

IV. EXISTING SYSTEM

Past work in this field has been conducted by psychological authority with statistical analysis; this approach reveals knowledge on suicidal ideation from a data analytic perspective. To detect suicidal ideation, extract several informative sets of features, counting statistical, syntactic, and linguistic, word embedding, and topic features, and compare six classifiers, counting four traditional supervised classifiers and two neural network models. (1) identifies informative features from a number of perspectives, counting statistical, syntactic, linguistic, word embedding features, and topic features; (2) compares with different classifiers from both traditional and deep learning perspectives, such as support vector machine, Random Forest, gradient boost classification tree (GBDT), XG Boost, multilayer feed-forward neural net (MLFFNN) and long short-term memory (LSTM).[4]

V. COMPARATIVE STUDY

Table 1: Comparative Study of Existing System

SR NO.	PAPER TITLE	AUTHOR NAME	TECHNOLOGY	ADVANTAGE	DISADVANTAGE
1	Prediction of Suicidal Ideation in Twitter Data using Machine Learning algorithms	M. Birjali, Abderrahim Beni-Hssane & Mohammed Erritali Morocco	Machine Learning Technology	Works relatively well when there is clear margin of separation between classes.	Does not perform very well, when the data set has more noise i.e. target classes are overlapping.
2	Suicide Related Text Classification with Prism Algorithm	Fatima Chiroma, Han Liu, Mihaela Cocea	Prims algorithm	Prism algorithm has outperformed the other machine learning algorithms with an F-measure of 0.84 for the target classes (Suicide and Flippant)	In cases where number of features for each data point exceeds the number of training data sample, the SVM will underperform.
3	Suicidal Trend Analysis of Twitter using Machine Learning and Neural Network	Nabia Shahreen,Mahfuze Subhani ,Md Mahfuzur Rahman	Machine Learning Technology	Relatively memory efficient	Kernel function Choosing Techniques is difficult.
4	Suicidal Behaviour Detection on Twitter Using Neural Network	Ghelmar Astoveza, Randolph Jay P. Obias, Roi Jed L. Palcon, Ramon L. Rodriguez, Bernie S. Fabito, Manolito V. Octaviano Jr (October 2018)	Machine Learning Technology	Artificial neural networks have numerical strength that can perform more than one job at the same time.	Realization of the equipment is dependent Because of networks require processors with parallel processing power, in accordance with their Structure.

VI. PROBLEM STATEMENT

If left untreated, suicidal ideation may lead to the devastating consequence of an actual suicide attempt.

Attempted suicide is a devastating experience for both the individual and that person's loved ones. [9] Besides a successful suicide attempt, other concerns of untreated suicidal thinking include:

- Conflicts in relationships
- Loss of relationships
- Missed work or job loss
- Lowered social function
- Financial concerns

VII. PROPOSED SYSTEM

In investigating the problem of suicidal ideation detection in online social bites, with a focus on understanding and detecting the suicidal thoughts in online user content. Suicidal Person Is Think about Suicide Ideas or tricks and after that, he/she Post about Any suicidal related tweets on their Tweeter Account. The system is to verify the Word like Kill, die, death, end Life, etc. Words like this are stored in the twitter database when post a comment admin or server check that content in a database to check whether it is a match or not. If content match with server data it assumes that the user has some mental stress and user supposed Words like this are stored in the twitter database when post a comment admin or server check that content in the database to check whether it is a match or not. It's mainly used to identify people who are stressed or look like in depression By Watching various Comments Post by Users. Admin keeps watch on users posts so that users cannot do anything wrong with their life. [5]

VIII. ALGORITHM

Step 1: Start

Step 2: User Open social site

Step 3: User login

Step 4: Enter user id and password

User_id = root;

Password = pass;

If (user_id==id && password==pass)

{

System.out.prinln("user id and password is correct and user login successfully ");}

Else If (! (user_id == id && password == pass))

{System.out.prinln("User need to login again ");}

Else {

System.out.prinln("Users Identity is   not varified , User need to register again ");}

Step 5: Successfully user login

Step 6: Access Account

Step 7: Checks friend's comments or posts

Step 8: User make post on twitter

K = words stored in database or   server

K = sad , stress , die ,etc.

P= positive post

N = negative post

If (user_post==P){

System.out.prinln("User is happy ");}

Else If (user_post==N){

System.out.prinln("User seen in stress ");}

Else If (user_post==K){

System.out.prinln("User seem in danger ");}

Step 9: Admin detect risk

For (int i =1; i<=number;i++){

If (user_post==K){

System.out.prinln("User tweet doubtful tweets multiple time ");

System.out.prinln("User is in risk , need to warn user ");}}

Step 10: Stop

IX. MATHEMATICAL MODEL

Point wise Mutual Information (PMI):

$$PMI\ (w,c)\ =\ \log(\frac{P\ (w,c)}{(w)P(c)})$$

Where

w =word, c =class

P (w, c) = probability that word w occurs in class c.

P (w) = frequency of the word across all classes

P(c) = frequency of class c.

This is used to identify the distress symptoms of users.

Suicide Prevention Assistant (SPA):

$$SPA\ =\ fsymptoms\ +\ fswear\ +\ fintensifiers\ +\ ffirst\_pronouns$$

Where

$fsymptoms$ = sum of PMI score of every word in tweet that appears in symptom lexicon.

Frequency of symptoms= $fswear$, $fintensifiers$, $ffirst\_pronouns$.

Unified Strangeness Measure:

Let X = {x1; x2; x3 . . . xn}

Where

X= sequence of unlabeled m-dimensional data points with new data points xi.

xn is the tweet that come very first in the data stream

USM is defined as follows:

$$USMi(X,x_n) = \sum_{k=1}^{m} |x_{ik} - \mu(\{x_{ik} \ldots x_{(n-1)k}\} \cup \{x_{nk}\})|$$

Where, m = mean

i : 1 . . . n; m = number of (user-centric and postcentric) behavioral.

xik = value of the k-th feature for tweet xi.
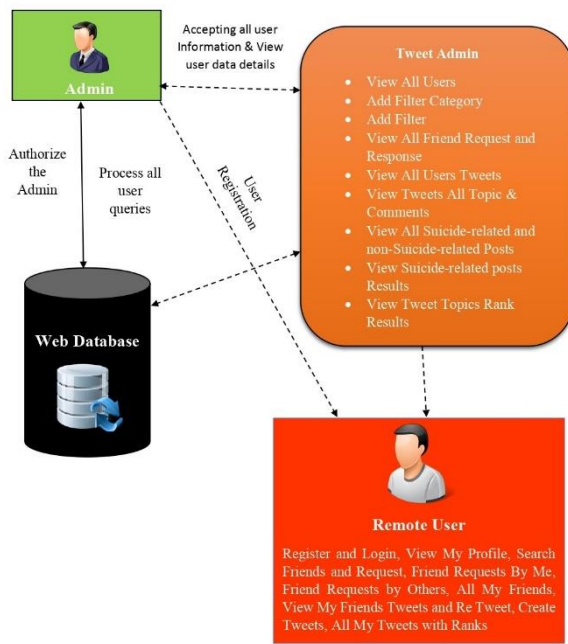
X. SYSTEM ARCHITECTURE

Fig.1: System Architecture

XI. ADVANTAGES

- Very fast Suicide warning signs in online behavior.

- The system is effective due to a general framework for detecting suicide-related posts in social networks.

- It is used for a mixture of the sentence to better understanding.

- Less memory used and it is well designed.

XII. DESIGN DETAILS

Fig 2: Administrator

XIII. CONCLUSION

Thus, we have tried to implement the paper "M. Johnson Vioules, B. Moulahi, J. Aze, S. Bringay", "*Detection of Suicide-Related Posts on Twitter Using Data Stream*", IEEE 2018 and according to the implementation, the conclusion is for the verification of suicide-related posts.Where mainly the youth and almost all of the people from the whole world is suffering from mental stress because different different reasons, it is very important to have a reality check how a person really stressed about is. This project will be very beneficial for most of the people who are not so comfortable opening up about their stress to friends or family. It will help them to get out of from there stress.

REFERENCE

[1] Shaoxiong Ji, Celina Ping Yu, Sai-fu Fung, Shirui Pan, Guodong Long "Supervised Learning for Suicidal Ideation Detection in Online User Content", 2018.

[2] Marouane Birjali, Abderrahim Beni-Hssane & Mohammed Erritali Morocco "Prediction of Suicidal Ideation in Twitter Data using Machine Learning algorithms", Morocco (2016).

[3] Fatima Chiroma, Han Liu, Mihaela Cocea (July, 2018), Suicide Related Text Classification with Prism Algorithm, Proceedings of the 2018 International Conference on Machine Learning and Cybernetics, Chengdu, china. (July,2018).

[4] Nabia Shahreen, Mahfuze Subhani, Md Mahfuzur Rahman (September, 2018), Suicidal Trend Analysis of Twitter using Machine Learning and Neural Network, International Conference on Bangla Speech and Language Processing (ICBSLP), Bangladesh, (September, 2018).

[5] M. Moreno, L. Jelenchick, K. Egan, et al., "Feeling bad on Facebook: depression disclosures by college students on a social networking site," Depression Anxiety, 2011.

[6] J. F. Gunn and D. Lester, "Twitter postings and suicide: Ananalysis of the postings of a fatal suicide in the 24 hours prior to death," Suicidologi, 2012.

[7] M. T. Lehrman, C. O. Alm, and R. A. Proano, "Detecting distressed and non-distressed affect states in short forum texts," in Proc. 2012 Workshop Lang. Social Media, Montreal, QC, Canada,2012.

[8] M. De Choudhury, S. Counts, and E. Horvitz, "Predicting postpartum changes in emotion and behavior via social media," in Proc. SIGCHI Conf. Human Factors Comput. Syst., New York, NY, USA, 2013.