Prediction Model For Preventing Suicide Attempts Using Machine Learning

Jayashree Agarkhed
Department of Computer Science and Engineering
PDA College of Engineering
Kalaburagi, India
Email: jayashreeptl@yahoo.com

Monika Reddy
Department of Computer Science and Engineering
PDA College of Engineering
Kalaburagi, India
E-mail: monikareddy3564@gmail.com

2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT) | 978-1-6654-1480-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICECCT52121.2021.9616815

Abstract— **Suicide is that the act of inflicting damage to oneself with the intention of inflicting death. Suicides are often derived back to numerous reasons like depression, monetary standing, mental standing, legal status, encompassing conditions etc. Self-destructive tendency in folks could be a serious concern that isn't restricted to a selected state or country. The worldwide statistics show that suicide could be an important contributor to the world's fatality rate. To boot, suicide is one in every of the twenty primary causes for death within the world. Thus it's necessary to prevent this menace by developing correct prediction systems supported offered information. Primarily the analysis of suicide information is finished to spot important attributes contributory towards suicide try and predict future such makes an attempt with important exactness. Here the person might specific their thoughts of committing suicide within the kind of fugitive thoughts, suicide arrange and role taking part in. this method is to seek out these dangerous intentions or behaviors before tragedy strikes. The scope of the project is to investigate the pattern of suicide cases and predict the causes of future suicides by exploitation machine learning algorithms.**

Keywords—Fatality rate, Fugitive thoughts, Machine learning, Self-destructive, Suicide

I. INTRODUCTION

Self destruction has been characterized as a purposeful demonstration that outcomes in one's own passing. Around 800, 000 individuals end it all on the planet every year, from 135000 (17%) of the units in the Republic of India, a country with 17. 5% of the number of inhabitants on the planet [1]. Somewhere in the range of 1987 and 2007, the collected self destruction rate went from 7.9 to 10.3 per 100,000 with self-destruction rates in southern and eastern Indian states [2]. As per the National Crime Records Bureau (NCRB), the provinces of Tamil Nadu, West Bengal, Andhra Pradesh, Maharashtra and Karnataka have recorded different self-dispensed passing's deliberately in the course of recent years and incorporate 56.2% of detailed general suicides in country [3]. Uttar Pradesh, purportedly a thickly populated state (16. 5% portion of the populace) with a lower extent of self - caused passing's, represent under 3.6% of revealed suicides in the nation, yet research laborers feel it could make this so because of appraisals of self destruction cases inside this space [4].

Therefore, it is important to create a predictive model for people at high risk of suicide to predict suicide attempts and causes of suicide such as hopelessness, stress, depression, anxiety, etc. in Indians and learn about trends and changes in suicides rates and assess and find reasons for growing suicides rate and make a report that can be used in finding a solution.

Machine learning was originally used to build faster search engines like Google, to find signal detection and plenty of other engineering features. The article in JAMA emphasizes how machine learning is important in health care in the 21st century [5]. Electronic learning (ML) methods enable computer learning of advanced classifiers that can improve predictive accuracy using large data sets. This prediction model analyzes the mean and mode of suicide cases and predicts the causes of suicides by using algorithms Support Vector Machine.

The prediction model for preventing suicide attempts analyzes suicidal speculation and human emotions so that this can be used in predicting suicide attempts. In [6] the author discussed on Affective computing and emotional analysis to recognize the emotions in human its one among the most important factor in human growth and development. Emotional processing is a prime in closely related polarity function. The spontaneous opportunity to capture the general public's feelings is very challenging in the world. The discovery of human behavior [7] looks at different behavioral patterns. This paper analyzes body signal, speech and gestures, text data and facial expressions of behavioral detection. This project analyzes human behavior.

In the previous edition Lakshmi Vijayakumar's research analyzes various suicidal articles published in the IJP (Indian Journal of Psychiatry). Research reveals that suicide rates are much higher in India than ever before. Another researcher studied on social, public and mental health responses which is important in preventing and predicting suicide attempts among gay, lesbian, bisexual and transgender (LGB) youths were tested and in the existing system only a few behaviors are analyzed such as age, gender and body image, etc.,. But the current system collects data from patients and doctors and analyzes the data collected, which prompts the processing of this problem.

Therefore the proposed model called Prediction model for preventing suicide attempts using machine learning. This model predicts the suicide attempts in human by collecting the data from the patient of all ages and gender by questionnaires, by doctor's report and by heredity that is patient's history, family, friends and surrounding environment etc, to make an accurate prediction model using and ML algorithm.

II. RELATED WORK

Just about all throughout the earth, self destruction costs are really perhaps the key issues. Typically the quantity of folks who-else ending all of it is usually expanding annually. Expected to be as a result of an assortment regarding causes, around 700, 000 men and women approved on while trying self destruction. Do it yourself

destruction is noticeable an illness by simply the WHO (World Health Organization), in addition to as per a study 17 % regarding worldwide self devastation victims incorporate Indians.

In [8] this paper study theoretically derived risk of suicide ideation to predict which factor is much important in suicide ideation. They used randomized data n=97, the data consist of patient who already tries to harm themselves. Here they followed N for one year and then assessment in 4 months interval on their daily factors. This paper concluded that the factors will differ and their emotions those who likely to reattempt from their past attempt.

In [9] is a survey of huge data to identify the suicide attempts and its resource in the common people, here they used random forest for training data to invent a risk model. The author has gone through 2500 survey questions to identify suicide attempts. This paper concluded that 1.8% of their population has 10% or greater risk of attempts.

In [10] this they predicted the self-disaster among sex-specific and non-fetal people. They used cohort study and got the result by developing classification tree and random forest to predict suicide attempts with the help of patient details. In this the author found that poisoning diagnosis and stress disorder is the self disaster among men and women. The conclusion of this paper is among top 5% of the data 44.7% are men and 43.2% are women who attempted suicide.

In [11] the author described that obsessive-compulsive disorder patients has more risk of attempts as compared in common people. In this 959 data used and elastic net model is used to predict attempts, variable used in this are social and clinical data. Here total 10.8% samples were of suicide deaths and got a result in the form of curve under area is 0.95, the author used only obsessive-compulsive disordered samples.

In [12] the purpose of this paper was to examine the relationship between aggressive and the suicide attempts behavior in military. In this they conducted longitudinal study with 944 military mans which are on duty. The person who reported aggressive behavior is more near to the suicide ideation, here they improved the role of aggression to predict suicide in service mans. They developed interventions to reduce risk within the service mans.

In [13] the inventor makes reference for an exertion among those fighters to dismiss the idea associated with self - harm during the Military Study to Figure out the Risk plus Resilience of commission rate individuals (Army STARRS) he researched whenever an example associated with U. S. opportunities dismissing SI age group. Still alive plus afterward sent regulatory records for upward to forty -- five months to become recounted the authoritatively recorded self destruction putting forth attempt (SA). Different overview factors detailed generally minimal then it was entirely expected to create a prescient model that knew 100% of the principal stage peril test (I. e., the third component, everything being equal) with the most noteworthy local army hazard (contributing 45 SAs inside the all out example). we have the likelihood to examine the possible utilization of this way to deal with separate future SI doubters with the most noteworthy state army hazard.

Inside [14] described suicides and its factors who are in the general hospital the result obtained in this that 6.9% among all are women and the admitted person due to suicide are younger then admitted for other reasons.

Inside [15] Journal of Self-employed Studies and Computer Research, the sum of suicides inside a couple regarding areas predominantly inside India and afterwards make use of those to be able to appraise the number of suicides with an conclusion goal to end up being saved in the upwards and coming. These kinds of self destruction critiques can help or perhaps help government inside making major determinations related with the location experiencing the do it yourself destruction. The crucial trait in the observe talks about the tiny part of typically the general population of which surveys pressure regarding the most portion on account regarding self destruction.

Inside [16] the creators explored the final results of medication reliance, sorrow, nervousness, mental cases, breaks, and immaturity injury in self - dangerous practices, self destruction attempts and self - harm however now not fatalities in network revision and they expected event of self - ruinous insight themselves, attempted self destruction, and self - perpetrated throb and no lethal outcomes have been 41%, 19% and 14% individually. Self - antagonistic center got expected with drug reliance, inordinate PCL: SV scored, and White race. The ability that hoping to hurt oneself must be expected on account of breaks, sadness, and youth injury. Self - hurt is anticipated because of break, alarm, PCL: SV score, and junior injury.

III. METHOD

The proposed model is “Prediction model to prevent suicide attempts using Machine Learning” the architecture of model is shown in figure 1, This figure is a block diagram of a prediction model and its working process initially the data is collected in data collection process and then that data is split into two parts to test the data and to train the data, for training data feature selection is done features such as: depressed mood, stress level, anxiety, education, age, gender, reason of unemployment etc., then classifier algorithm is applied to get accurate analyses then the test data is used to test the suicide prediction.

The inputs used in this model are patient Questionnaires, Below are the few questionnaires:

- **Name, Age, Gender:** The details of patient to predict their future or suicide ideation.
- **Marital status:** weather the person is married or unmarried or separated and its reason.
- **Economic Status:** The person’s income.
- **Education details**: Educated or uneducated, if educated then weather they faced any failure in their exam.
- **Heath Status:** Any heath related issues.
- **Family Details:** This consists of family history and details.

Data Collection: Data is collected from the online data science platform. Patient’s Data: Collection of data from patients and the Psychiatrist, the data consist of the details of patients like age, gender, employment, marital status, total

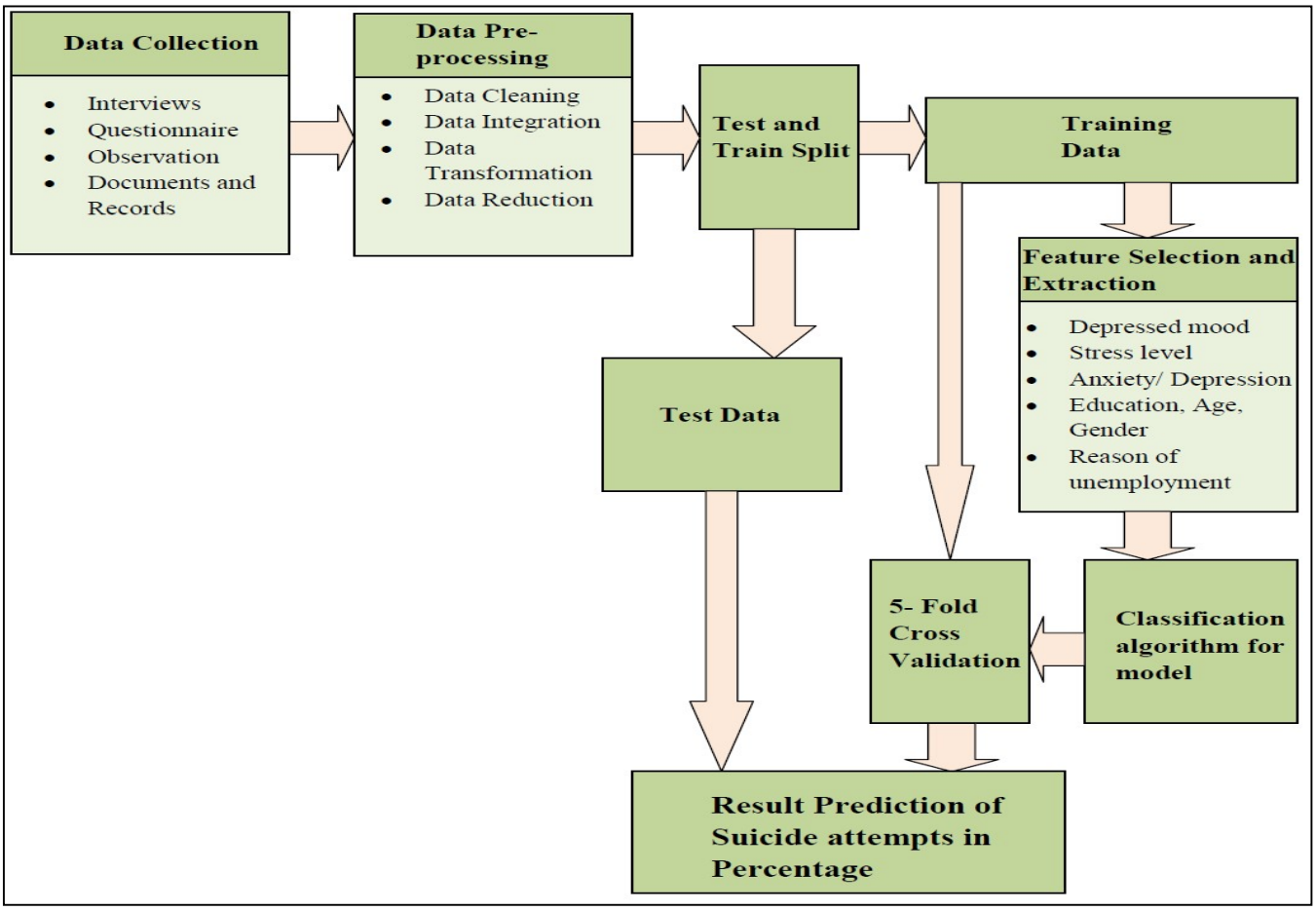

Fig. 1. Block Diagram of Proposed System

number of suicide attempts and its cause etc. Throw interview, Questionnaires, Observation and documents/ records. After the collection of patient's data the collected data is again pre-processed with the doctor's data with the help of doctors.

***Data Pre-processing*:** Data Pre-handling is usually the way regarding preparing information together with the aim to be able to obtain the wonderful results. Pre-handling of realities comprises of different strides alongside measurements purifying, filling in clean fields, encoding factors and so on.; Data Pre-Processing is a essential advance on the grounds that jam-packed insights can interrupt with impacts and may even prompt incorrect ramifications. Since many regions of the facts set are obscure factors, there's a new need to help to make or code ahead of utilizing a proper relapse model. Typically the records are eliminated through filling inside clear information implemented with the guideline of express elements. Sham factors usually are embedded for just about all fields with precise factors to alter to binomial elements (i.e. 0 or 1). Pre-handling of the info remembers various advances for request that individuals can acquire typical and complete details. The realities managing steps are since per the next:

- **Data Cleaning:** The initial action to handling the past data is to simple current realities. Discover realities that contains missing qualities and erase that data or put various qualities like the normal of the segments and so on The reason for insights purging is to get rid of missing, jumbled and conflicting data from the data assortment.
- **Data Integration:** Add or blend insights from more than one information bases joined. Subsequent to storing up every one of the insights clashes sooner or later of which all the data is done.
- **Data Transformation:** Data is changed the utilization of standardization or collection procedures

Splitting data into test and train: Separation of details into tests in addition to show measurements is usually partitioned into analyze sets and educate units in order that typically the teach records established can be equipped with a urge pair of rules in addition to at last applied to are planning on self-destructive airs inside investigate data established. This transformed directly into accomplished to analyze the prescient ability of the type on realities of which got not employed for assessment. twenty percent of the facts is treated along of evaluations so that it doesn't leave us all creating almost zero train records which often could save an individual the model coming from fitting and evaluating the measurements properly.

Feature Extraction and selection: Within device contemplating, the trademark is the trademark or house of a situation this is beneath assessment. Highlight extraction is really a method that will involves eliminating large quantity records from the information set.

Since everything choices probably will not assist our mode plus should bring regarding your decision of appropriate, valuable, informational plus independent capacities will be the substantial point to persuasive type

and relapse. Selecting the appropriate aspects is additionally remarkably fundamental for device group of rules considering program gear. Emphasize Elimination Algorithm (RFE) have been never genuinely out the handiest alternative. A section in the choices usually are recorded while typically the calculation introductions typically the area unit postponing body (weight), wretchedness, work - just like express (can't locate artworks) and thus on

Classification algorithm for model: Inside machine contemplating elements, SVM is several other fundamental and is also thought about to be able to be the finest class set regarding rules for examining or anticipating beliefs. In fashion, a couple of - degree method is expected to be able to produce a SVM model.

Cross validation: Cross - approval is a strategy used to explore the prescient by and large execution of factual models. Cross - approval got finished to check whether the factual examination become right or exact or presently don't just as to check for particularly right data.
Subsequent to preparing the data, check data had been utilized to are expecting self destruction attempts

A. Prediction Model using Support Vector machine(SVM)

In device acquiring information on systems, SVM is some other imperative and is viewed as the most extreme a triumph set of rules for assessing or foreseeing values. In loved, a - phase technique is needed to produce a SVM model.

- At first, the example realities is recorded or test conscious to district measurements which is often full-size in correlation with the top components of the genuine realities.
- The particular subsequent advance is to find the suitable hyper plane with an incredibly enormous distance to order the data effectively

This SVM algorithm is used to find out two classes suicide attempts and non-suicide attempts, these two classes are divided by the best boundaries known as hyper planes to train and predict the future attempts. Equation (1) shows, the person who attempt suicide SA is the fraction of truly attempt suicide t(SA) by sum of truly attempt suicide t(SA) and falsely attempt suicide f(SA) and equation (2) shows, the person with suicidal tendency NSA is the fraction of truly not attempt suicide t(NSA) by sum of truly not attempt suicide t(NSA) and falsely not attempt suicide f(NSA).

The person who attempt suicide:

$$SA = (t(SA))/(t(SA)+f(SA)) \quad (1)$$

The person who not attempt:

$$NSA = (t(NSA))/(t(NSA)+f(NSA)) \quad (2)$$

In this model, assumed value x=30% that is accuracy of the suicide ideation. Equation (3) implies, if accuracy x is less than or equal to predicted value than suicide attempt SA and if accuracy x is greater than or equal to predicted value then NSA Non-suicide attempt.

Here if particular person prediction value that is i

$$\text{Suicide Attempt (SA)} \geq \text{Accuracy (x)}$$
$$SA \geq 30\% \quad (3)$$

$$\text{Non-suicide Attempt (NSA)} \leq \text{Accuracy (x)}$$
$$NSA \leq 30\% \quad (4)$$

B. Algorithm used in model

Algorithm: Prediction model for preventing suicide attempts

Input: Data set D, Questionnaires Q
Output: Predict future SA
1 Start model
2 Load D into a model
3 Split D to D_{train} and D_{test}
4 To evaluate a model
5 Set assumed average value i for SA
6 Set features and target to D_{train}
7 Initiate objects for SVM and fit the D_{train} to the model
8 Add input as Q to the model
9 Make a prediction on D_{test} calculate percentage
10 Print percentage
11 Stop

In this algorithm of Prediction model for preventing suicide attempts the inputs are collected dataset and questionnaires and predict whether a person attempt suicide or not. In this D_{train} is the data used to train and analyse and D_{test} is the data used to test using SVM.

Algorithm: Prediction Model for Preventing SA/ NSA using Machine Learning

Input: Data set
Output: Suicide attempt SA, Non suicide attempt NSA.
1 **Start**
2 def docviewrecords(dataset, userinputs)
3 vectordata ← fetchdat_fromdatabase(features)
4 patientdata ← getpatientdetails(Questionnaires)
5 pcount == 0
6 ncount == 0
7 Assume x == 30%
8 **for** patientdata in vector data
9 val = compare_vectordata_patientdata(parameter, pval, vval)
10 if val> 0
pcount←pcount+1
11 else
ncount←ncount+1
12 percentage← (pcount/ no_of_parameter)/ 100
13 **end**
14 **if** percentage ≥ x
then print (SA)
15 **else**
print (NSA)
16 **Stop**

Initially we start the model by giving input from data set, in step 2 'def docviewrecords' is a method name with two inputs dataset and userinputs. Step 3 fetches the previous results from the dataset and features. Step 4 fetches the

answers list of the patient for the list of questions. Step 5 and 6 are variables to store matched and unmatched values between patient's answers and dataset results. In step 7 we assumed a value for suicide ideation as 30%. In step 8 to 11 looping is done for each value provided by the patient with the existing answer of previously stored dataset. Step 12 converts the vector result to percentage for simplified and generalized result. In step 14 and 15 if the percentage is great then suicide tendency value then it prints SA else NSA.

IV. RESULT AND DISCUSSION

Prediction Model For Preventing Suicide Attempts Using Machine Learning is to predict causes of suicide in India based on the existing dataset of India's registered suicide cases which is obtained from NCRB and all type of age groups from 15 to 60 years of people and the measuring variables used in this are age, gender, marital status, health education, etc. To predict the suicide attempts in future.

In order to get the accurate result of proposed model we compared the model with existing models with respect to the method or algorithm used in their model in Table 1. In table gives the results in which this proposed model is greater than the existing model in table.

TABLE I. ANALYSIS OF ALGORITHMS' ACCURACIES

Author	Algorithms	Accuracy (%)
jun su jung, s [17]	Artificial Neural Network	77.5
	Random Forest	77.8
	Support Vector Machine	78.7
	Linear Regression	77.9
Seo-Eun [18]	Random Forest	75
Imran Amin [19]	Artificial Neural Network	77.5
	Support Vector Machine	81.5
Mrs. B. Ida Seraphim [20]	Decision Tree	91.67
	Mlp Classifier	90
J. Luo [21]	Convolutional Neural Network	83
Proposed Model	**Support Vector Machine**	**92**

The figure 2 is the graph which represents the suicide rates based on their marital, economic, education and health status in percentage separately. In this we can come to a conclusion that Health status has more number of suicide attempts compared to other status due to mental illness many people harm themselves and now a days due to this covid pandemic the person who face covid are also going to depression and mentally ill and harming them self .

Marital status has 5.5%, Economic status 8.9%, Health status 18.5 % and Education status 7.6%.

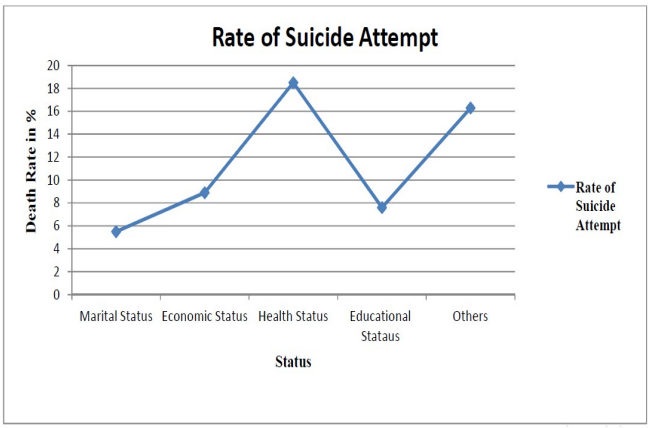

Fig. 1. Rate of Suicide Attempts

This model is used to reduce future attempts so, machine learning is very useful to train the data and classify the data and predicts future cases, this model uses ML algorithm to get correct accurate result. Support vector machine makes the correct hyper plain between SA and NSA to get correct accuracy.

V. CONCLUSION

As there is increase in suicide rate and there is no correct model to predict a particular individuals future suicide attempts. So, this particular proposed tool can be used to predict the particular person he/she commit suicide or not with the help of features used in this model within India. Machine learning is used to analyze and predict self-harm among any individual by asking the data or questions from the patients and analyzing the data. By this model we can decrease the death rate due to suicide by detection before the crises arise. This system / model can be used in clinically in future we can all departments like Crime department, Hospitals, IT sectors, Psychology department, military, etc... in single module and keep single admin to handle this module.

ACKNOWLEDGEMENT

The authors appreciate for the helpful support from the department of computer science and the authors also wants to thank the International Conference on Electrical, Computer and Communication Technologies for giving opportunity for presenting my research.

REFERENCES

[1] C. G. Walsh, J. D. Ribeiro, and J. C. Franklin, "Predicting Risk of Suicide Attempts Over Time Through Machine Learning," *Clinical Psychological Science*,vol.5,no.3, pp: 457-469, 2017. DOI: 10.1177/2167702617691560.

[2] M. Simon, E-S. Chang, P. Zeng, and X. Dong, "Prevalence of Suicidal Ideation, Attempts, and Completed Suicide Rate in Chinese Aging Populations: A Systematic Review," *Archives of Gerontology and Geriatrics*,vol.57no.3,pp.250–256,2013.DOI:10.1016/j.archger.2013.05.006.

[3] V. Arya, A. Page, J. River, G. Armstrong, and P. Mayer, "Trends and Socio-Economic Determinants of Suicide in India: 2001–2013," *Social Psychiatry and Psychiatric Epidemiology*, pp. 1–10, 2017.

[4] S. Kumar, A. K. Verma, S. Bhattacharya, and S. Rathore, "Trends in Rates and Methods of Suicide in India," *Eg tian Journal of Forensic*

Sciences, vol.3, no.3, pp. 75–80, 2013. DOI: 10.1016/j.ejfs.2013.04.003.

[5] Depression And Exposure To Suicide Predict Suicide Attempt Sonali Nanayakkara M.D. , Diane Misch M.D. , Laura Chang D.O. , David Henry Ph.D. First published: 15 August 2013.

[6] 15-year computer-record study of adolescents exposed to peer suicide CP Tang *, SF Hung, CC Lee, TP Ho, PWL Leung February 2019.

[7] Applying machine learning on health record data from general practitioners to predict suicidality Kasper Mens, Elke Elzinga, Mark Nielen, Joran Lokkerbol, Rune Poortvliet, Gé Donker, Marianne Heins, Joke Korevaar, Michel Dückers, Claire Aussems, Marco Helbich, Bea Tiemens, Renske Gilissen, Aartjan Beekman, Derek de Beurs.

[8] Modeling the suicide behavior cycle: Understanding repeated suicide attempts among individuals with borderline personality disorder and history of attempting siucide. Kuehn, K.S., King, K.M., Linehan, M.M.,& Harned, M.S.(2020)

[9] Identification of suicide attempt risk in national US survey using machine learning, Angel Garciala Garza,BA; Carlos Blanco, MD, Phd; Mark Olfson, MD, MPH; etal Melanie M Wall, Phd JAMA Psychitary 2021.

[10] Predicting sex-specific Non-Fatal Suicide Attempt Risk using machine learning and Data from Danish National Registries Jaimie L Gradus, Anthony J Rosellini, Erzebet Horvath-Puho, Tammy Jiang, Amy E Street, Isaac Galatzer-Levy, Timothy L Lash, Henrik T American Journal of Epidemiology.

[11] Predictors of suicide attempt in patients with obsessive-compulsive disorder: an explratory study with machine learning analysis .neusa Aita Agne, caroline Gewehr Tisott, Pedro ballester, Ives Cavalcante Passos and Ygor Arzeno Ferrao.Publisge by Cambridge university Press(2020).

[12] Predicting Suicide Ideation in Milirary: The Independent Roles of Aggression Amanda R. Star phd, Yvonne allard MPH, Amy Adler phd, Robin Toblin phd.

[13] Predicting suicide attempts among soldiers who deny suicidal ideation in the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS) Samantha L Bernecker, Kelly L Zuromski, Peter M Gutierrez, Thomas E Joiner, Andrew J King, Howard Liu, Matthew K Nock, Nancy A Sampson, Alan M Zaslavsky, Murray B Stein, Robert J Ursano, Ronald C Kessler Behaviour research and therapy 120, 103350, 2019.

[14] Suicide attempts and related factors in patients admitted to a general hospital: a ten-year cross-sectional study (1997-2007) Jesús Alberdi-Sudupe, Salvador Pita-Fernández, Sonia M Gómez-Pardiñas, Fernando Iglesias-Gil-de-Bernabé, Jorge García-Fernández, Gonzalo Martínez-Sande, Sara Lantes-Louzao & Sonia Pértega-Díaz BMC Psychiatry.

[15] Journal of Independent Studies and Research Computing Volume 15 Issue 2 July-December 2017.

[16] Predictors of suicidal ideation, suicide attempts, and self-harm without lethal intent in a community corrections sample Author links open overlay panelTracy D.GunterabJohn T.ChibnallaSandra K.AntoniakbRobert A.Philibert bNancyHollenbeckb.

[17] jun su jung, sung jin park, eun young kim “Prediction models for high risk of suicide in korean adolescents using machine learning techniques” 2019- journals.plos.org.

[18] Seo-Eun Cho, Zong woo Geem, Kyoung sae Na “Prediction of suicide among 372,813 individuals under medical check-up” journal of Psychiatric research December 2020.

[19] Imran Amin “Prediction of suicide causes in India using Machine Learning” Research gate 2020.

[20] Mrs. B. Ida Seraphim , Subroto Das , Apoorv Ranjan “A Machine Learning Approach to Analyze and Predict Suicide Attempts” IJERTV10IS040294 Volume 10, Issue 04 (April 2021)

[21] Luo, et al., Exploring temporal suicidal behavior patterns on social media: insight from Twitter analytics, Health Inform. J. 26 (2) (2019) 738–752,1460458219832043.