

An Analytical Study for Predicting Suicidal Tendencies Using Machine Learning Algorithms

Mou Ghosh^{#*}, Susama Chatterjee[#], Debarupa Das[#], Kaushik Chanda[#], Dr. Sandip Roy[#]

[#]Department of Computational Science, Brainware University, Barasat, Kolkata, India

*moughosh217@gmail.com

*Corresponding Author

Abstract

Suicide, one of the most important concerns in the world, is causing one death in every 40 seconds, according to WHO. Since the last 45 years, having the suicide rate increasing by 45% around the world, is raising a question with the adaptability of fast life style and complexities related to this smart technological era. Especially in a developing country like India the increasing problem of suicidal tendency among young generations and middle- aged persons raises a question mark for the well-being in this smart city oriented highly technological scenario. Among many reasons of suicide depression is a big factor behind many deaths in today's scenario. This chapter not only highlights the factors responsible for suicidal tendency, but also performs a study on several models and machine learning based applications for anticipating the suicidal tendency of a person by remote monitoring, which is very essential to prevent the attempt of such a devastating crime. Priority has been imposed upon depression as a cause of suicidal tendencies (Esteban A, 2019).

Keywords: Depression, Deep Learning, Suicidal tendency, Smart Self-immolation prediction

1. Introduction

Suicide is derived from a Latin word *suicidium* that has its meaning as “to kill oneself”. It has become a leading reason that has acquired the position of 10th while considering the death term in the worldwide section. It has been found that millions of people died due to suicide attempt that has represented the death rate due to suicide attempt in global region as 16 people per 100000. According to the “World Health Organisation (WHO)”, the attempt of suicide is more than twenty times frequently than the complete suicides.

It has been found that the statistics show that 1,00,000 death occurs due to suicide in the world-wide region out of which 20 percent belongs to the country India but however the population of India is 17 percent of that of the population of world. According to the report of NCRB it has been found that the earners that belong to the daily wagers commit more suicide that is 26,589 in number among the males that is 92,114 in number. The second position goes to the people who would be self employed and it is 12,175 in number along with the people who are unemployed that is 10,687 in number. In case of the females the number of suicide attempt is 42,391 in the year 2018 and out of which 22937 were housewives and then 4790 are students that is followed by 3535 belongs to the daily earners of the wage. It is 11 in case of the transgenders and their case is 3 belonging to daily earners and 1 belonging to unemployed people as well as salaried or professional people and 6 people belonging to others in case of the Asian region (S.Biswas, 2020). It has been found that in the report that was released before the “World Suicide Prevention Day” the rate of suicide in India has pegged 16.5 people out of 100000 (Colic, 2018).

2. Depression as a major cause of suicidal attempt

Depression is a common illness in today's world, not less than 263 million people worldwide affected with some form of psychological disorder. It should not be mixed up with usual mood fluctuations, empty feelings or short-lived emotional outbursts that occur as a response to challenges faced in routine life style, but it becomes a major issue in case of their health when there is a long lasting occurrence along with an intensity that would be severe or moderate in nature and hence it has become an issue. This condition is known as depression (Bhat,2017).

Although depression has high severity, it is one of the most treatable of mental disorders. Effective treatments for depression do exist but several barriers to effective mental health care make it quite inaccessible for almost 76% to 85% of low-income and middle-income families from economically backward countries. There barriers exist in several ways such that the mental health become effective that would include lack in the resources along with the lack in the providers of the trained health care along with the mental disorders that would be associated with the stigma socially. The barrier related to the care effectively comprises people from distant region that include people with the problem of mobility along with people who did not prefer to leave their own home and hence suffering from disorders such as agoraphobia or anxiety. The assistance on the basis of the Internet that include physicians primarily along with doctors so that it becomes easier to treat the larger patients so that the queue related to waiting in case of the treatment related to the psychology could be minimized effectively (Birjali,2017).

At present, cases of depression, mental health issues, and suicides caused due to depression globally and there is rise from all countries towards addressing and finding solutions to the problems caused due to mental health issues so that the "World Health Assembly" has passed their resolution in 2013 related to the response in coordinated and comprehensive manner (Kaushik Chanda,2020) (Dr. Sandip Roy,2020).

Suicide is a fatal issue causing the loss of many lives in world and as per WHO this problem is increasing day by day and if this trend persists, then by 2030 in every 20 seconds one life may be destroyed forever .To consider the trend that occur recently in case of the suicides as well as factors related to the risk in India. MCA has associated the attributes like gender, age group and profession. Among many causes, depression is one important reason of suicide. Prediction and prevention of suicide ideation in today's world is one of the biggest concerns because of its effect in untimely death of different age group people specially in young generation. Several research works have been undergone and different systems have been implemented in this context. Some researchers have used supervised learning algorithms for predicting suicidal tendency whereas some have applied unsupervised learning methods to predict and prevent suicide. Ji et al have reviewed applications of different machine learning methods in suicidal idea detection and prevention where data deficiency, annotation bias, lack of proper understanding the intention and data imbalance have been identified as the limitations of the existing approaches to suicide prediction. Depending upon the machine learning methods used to detect suicide, we have divided our analysis into three division such as Unsupervised Learning, Supervised Learning and Reinforcement learning (Ji,2019) (Sandip Roy,2020).

2.1. Categories of Suicides

(1) Egoistic suicide:

This type of suicide occurs whenever one thinks that they have been isolated socially and hence have a feeling that they are not valued in their society so they decide to destroy oneself and hence the suicide is actually associated with self-centered people that lack the altruistic feelings and hence cut off usually from the main society stream.

(2) Altruistic suicide:

This type focuses on the people or a group that are close along with the intimate relation and hence result in over integration among the individuals related to the group socially such as Dannie's warriors or Sati system.

(3) Anomic suicide:

This type focuses on the breakdown certainly in the equilibrium socially that include bankruptcy suicide attempt or after the case of winning the lottery. The situation occurs suddenly due to certain action.

(4) Fatalistic suicide:

This type would focus on the society overregulation where one had to become a slave or servant such that they decide to commit suicide and hence the barren women would be committing suicide and hence it is considered the example of the fatalistic one.

2.2. Many of the symptoms and signs of suicide

- Sudden change in the sleeping and eating habits
- Sudden withdrawal from family friends as well as regular activities.
- Rebellious behaviour, violent actions or running from the situation.
- Use of alcohol and drug
- Unusual negligence in the appearance of their personality.
- Change in marked personality.
- Persistency in the boredom and difficulty in the concentration along with the decline of the quality of their work.
- Complaints related to the symptoms physically that would be related to various emotions that include headache, stomachache or fatigue.
- Sudden loss in the interest related to pleasurable activities.
- Not interested in praise or any kind of rewards.

2.2.1. One who has decided to attempt suicide would

- Continuous complain of considering oneself as worst one or feeling bad inside.

- Giving hints verbally along with statements such as “being problem no more”, “no matter as it would be of no use”, “ I would not be meeting again”
- Putting their affairs in orders such as giving their possessions or cleaning their own room or throwing their own things or belongings
- Suddenly being very cheerful after depression
- Thoughts related to hallucinations.

2.3. Diagnosis and Treatment of Suicide

2.3.1 .Diagnosis

If one has the following depression symptoms or they are carrying thoughts of suicide then the doctor would be taking the examination medically by asking certain questionnaire that would include certain questions such as (Rajesh Bose,2020):

2.3.1.1. Do you have any depression symptoms or history related to mental illness?

2.3.1.2. How long you have thoughts related to suicide?

2.3.1.3. Do you have a plan related to suicide?

2.3.1.4. Are you taking any medication— and, if so, what kind?

2.3.1.5. Are you acquainted with drugs or alcohol?

2.3.2. Treatment

If one experiences certain depression symptoms and carry thoughts about suicide the therapist or doctor might recommend medication, psychotherapy or change in lifestyle that would reduce the suicide risk.

2.3.2.1. Therapy related to talk or psychotherapy such that one would work with therapist to know the reason behind their suicidal thoughts.

2.3.2.2. Therapy related to family or education such that one would work with the loved ones so that they become better quickly and could improve their dynamics related to family.

2.3.2.3. Substance using the disorder treatment, such that an increment in the usage of drugs or alcohol.

2.3.2.4. Lifestyle changes, so that it would include stress management and improve their habits of sleeping as well as eating along with exercising so that a social network could be build and times would be related to interest as well as hobbies.

2.3.2.5. Medications so that there is treatment to the depression that would be related to

the suicide.

2.3.2.6. It might include various antipsychotic, antidepressants, anti-anxiety medications.

3. Machine learning methods

The most desired as well as the effective way by which suicide ideation could be prevented by treatment or early detection and would comprise the attempt potentially such that they could include the factors related to risk that would result in attempting the suicides successfully. Online ways of communication are considered an way by which people would express their tendency to attempt suicide . The paper would focus on the approach so that it become easy to understand the ideation of suicide using the online content that would be generated by the user such that the suicide could be detected at an early stage using the supervised learning The rich knowledge could be revealed by analysing the preference of the analysed language of user along with the description of the topic so that an system related to the early warning could detect the tendency of suicide(Dey, R. K,2020). as the negativity is more in case of individuals having suicidal thoughts there would be hopelessness and anxiety among them. It might include friends and family and hence they would cover topics related to social and personal issue which could be detected by gaining sets of information that include linguistic, syntactic, statistical embedded words and the six classifiers could be compared that include the four supervised classifier traditionally along with the neural networks. The study related to experimentation would demonstrate the approach practicability and feasibility that would provide the ideation suicidal benchmark related to platform in online such as Twitter, Reddit Suicide Watch (Gradus,2020).

The rate of high attempts in nonfatal suicide among the adolescents that would clinically provide prediction of risk in practical such that it would be a challenge so screening related to consuming of time that would include implementation scale such that the trained counsellors could provide the crisis tailored management. The Word Count and the “Simplified Chinese Linguistic Inquiry” and would use analysis of the users text psycholinguistic in the slot of next 1 month to prior consultation (Hosseinifard,2013).

Data would be preprocessed using the technology Python that would analyse the statistics and perform them in R so that the random approach of forest could be accepted using the community of Machine Learning and implemented easily along with the robustness(Sarddar, D.,,2020)(Bose, R.,2020). The performance of model would be evaluated with the help of AUC along with metrics recall and precision. The study could be calibrated and evaluated by inspecting the plots of calibration and considering the scores of Brier that would reflect the prediction probabilistic accuracy and would range from 0 to 1(Sharmistha Dey,2020)(Biswas, S,2020).

Table 1: (Types of Prediction Methodology)

| SL No. | Objective | Method | Performance | Features | Drawback(s) |
|--------|---|--|---|--|--|
| 1. | To study machine learning use in depression analysis . | -Syntax analysis of Twitter & Weibo Posts by different age groups to find certain keywords and their relevance -Semantic Analysis to find the general emotion of paragraph Classification of posts according to the symptoms of depression | -6.09 percent of the Weibo posts would be reflected in stigmazing the attitudes in deopression | SVM | -Requirement of social media post data verification to achieve perfection in prediction |
| 2. | To identify Suicide prone individuals with specialized crisis management including direct message assessment & one-to-one counselling | -Annotation of social media posts on suicide note for training dataset -Automatic identification of posts involving suicidal thought & behavior. | i) SVM Precision-0.87 Accuracy-0.86 F-measure-0.84 Recall-0.81 ii) DT Precision-0.79 Accuracy-0.78 F-measure-0.75 Recall-0.73 iii)RF Precision-0.86 Accuracy-0.82 F-measure-0.79 Recall-0.74 | SVM, DT, RF, LR with 10-fold cross validation | Consideration of only a microblog group, no access to media platforms like school bulletin board, online suicide group, online self help group -No consideration of factors like posting time, frequency etc. other than the text |

| | | | | | |
|----|--|---|--|--|--|
| 3. | To predict individual risk of suicide attempting | -dataset was from at least 5 hospitals- they have performed computation using python with tensor flow | Comparatively low accuracy rate | Neural Network Model | Data privacy issue |
| 4. | Mobile application for suicide prediction | -e health Application -approach for intelligent health (i-health) applicationn for development knowledge execution and application of data miningtechnique | not mentioned | Knowledge execution - unsupervised clustering | use of randomize data for complex dataset. |
| 5. | Suicide prevention by Feature Extraction | -medical data modelling -tested with more than 100 occurrences. | High risk performance is poor(34%) Moderate risk performance is 45% | -feature extraction with linearmodel -K nearest neighbour algorithm | -earning trust from the medical practitioners is a challenge -probability of matching between manual strategy and ML Strategy |
| 6. | Randomized machine learning technique | -They test five machine learning techniques with three feature sets | AUC 74% | -sparse logistics regression and decision tree | - Single retrospective cohort has been used -Use of randomize method for complex dataset |
| 7. | To detect suicide risk and predict future suicide attempts | Classification by random forest for statistical modeling. | Recall-95% Precision-77% | Anatomical Therapeutic Classification, Level V (ATC, Level 5) | -no integration of potential precipitating events like job loss -need experiment |

| | | | | | |
|-----|---|--|--|---|--|
| | | | | | with larger dataset |
| 8. | Design of twitter based event detection for suicide. | system developed based on PHP, mysql and twitter API | not mentioned | Text mining | lack of incorporation of public tweet |
| 9. | To predict the individuals with the ideation of suicide in the general Population | -Random selection of individuals from non suicide ideators to avoid class imbalance problem- Classification of testing dataset to predict the suicide ideation | Accuracy in test set- 78.3%, Accuracy in total population- 82%, sensitivity- 77%, specificity- 79.2% | -RF in training & recursive feature elimination | -Requirement of more study of public health & clinical data, biomarkers to predict more critical suicide risk like self-harm and suicide attempt. |
| 10. | Detection of suicide attempters among suicide ideators | -Resampling of suicide ideators by "Synthetic Minority Over Sampling Technique (SMOTE)" to obtain data of suicide attempters and non suicide attempters | Accuracy 88.9% | RF in training & recursive feature elimination via 10-fold cross validation | Dataset contains very simple scale & questions, no comparison with other algorithms Use of only class balanced data, no actual data with bias classification |

4. Conclusion

People might lack in online communication system due to lack of internet connection. Hence an approach related to scalability would lead to machine learning along with longitudinal data such that it would broaden the screening risk for the non-factual suicidal attempt among the adolescents. This chapter provides an analytical overview of the existing literature. Detection as well as prevention is a very necessary factor for an issue like suicide, especially when depression works behind the death. With the blessings of technologies like machine learning or data analytics, clinical management of suicide across diagnostics and evaluation has witnessed a paradigm shift in recent times. This chapter divided the present works according to several categories and discusses the pros and cons. But increasing the precision efficiency keeping the data privacy is till now a challenge to the researchers.

References

1. Amin, I. and Syed, S., 2017. Prediction of Suicide Causes in India using Machine Learning. *Journal of Independent Studies and Research (JISR)*,15(2).
2. Bhat, H.S. and Goldman-Mellor, S.J., 2017. Predicting adolescent suicide attempts with neural networks. *arXivpreprintarXiv:1711.10057*.
3. Birjali, M., Beni-Hssane, A. and Erritali, M., 2017. Machine learning and semantic sentiment analysis based algorithms for suicide sentiment prediction in social networks. *Procedia computer science*, 113,pp.65-72.
4. Colic, S., Richardson, D.J., Reilly, P.J. and Hasey, M.G., 2018, July. Using machine learning algorithms to enhance the management of suicide ideation. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 4936-4939).IEEE.
5. Gradus, J.L., Rosellini, A.J., Horváth- Puhó, E., Street, A.E., Galatzer-Levy, I., Jiang, T., Lash, T.L. and Sørensen, H.T., 2020. Prediction of sex-specific suicide risk using machine learning and single-payer health care registry data from Denmark. *JAMA psychiatry*, 77(1), pp.25-34.
6. Hosseinfard, B., Moradi, M.H. and Rostami, R., 2013. Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. *Computer methods and programs in biomedicine*, 109(3), pp.339-345. Walsh, C.G., Ribeiro, J.D. and Franklin, J.C., 2018. *investigation*, 15(11), p.1030.
7. Ji, S., Pan, S., Li, X., Cambria, E.,Long, G. and Huang, Z., 2019. Suicidal Ideation Detection: A Review of Machine Learning Methods and Applications.*arXivpreprintarXiv:1910.12611*.
8. Ji, S., Yu, C.P., Fung, S.F., Pan, S. and Long, G., 2018. Supervised learning for suicidal ideation detection in online user content. *Complexity*, 2018.
9. Jung, J.S., Park, S.J., Kim, E.Y., Na, K.S., Kim, Y.J. and Kim, K.G., 2019. Prediction models for high risk of suicide in Korean adolescents using machine learning techniques. *PLoS one*,14(6).
10. Morales, M., Dey, P., Theisen, T.,Belitz, D. and Chernova, N., 2019, June. An investigation of deep learning systems for suicide risk assessment. In *Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology* (pp.177-181).
11. Predicting suicide attempts in adolescents with longitudinal clinical data and machine learning. *Journal of child psychology and psychiatry*, 59(12), pp.1261-1270.
12. Rakesh, G., 2017. Suicide prediction with machine learning. *American Journal of Psychiatry Residents' Journal*, 12(1),pp.15-17.
13. Ryu, S., Lee, H., Lee, D.K. and Park, K., 2018. Use of a machine learning algorithm to predict individuals with suicide ideation in the general population.*Psychiatry*
14. Ryu, S., Lee, H., Lee, D.K., Kim, S.W. and Kim, C.E., 2019. Detection of suicide attempters among suicide ideators using machine learning. *Psychiatry investigation*, 16(8), p.588. Walsh, C.G., Ribeiro, J.D. and Franklin, J.C., 2017. Predicting risk of suicide attempts over time through machine learning.*Clinical Psychological Science*, 5(3), pp.457-469.
15. Tran, T., Phung, D., Luo, W., Harvey, R., Berk, M. and Venkatesh, S., 2013, August. An integrated framework for suicide risk prediction. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp.1410-1418).
16. Varathan, K.D. and Talib, N., 2014, August. Suicide detection system based on Twitter. In *2014 Science and Information Conference* (pp. 785-788). IEEE.
17. Sharmistha Dey, Indranil Sarkar, Srabanti Chakraborty, Sandip Roy,"DEPRESSION DETECTION USING INTELLIGENT ALGORITHMS FROM SOCIAL MEDIA CONTEXT - STATE OF THE ART, TRENDS AND FUTURE ROADMAP",Xi'an DianziKejiDaxueXuebao/Journal of Xidian University · August 2020, <https://www.researchgate.net/publication/343384461>
18. K.Chanda, P. Bhattacharjee, S. Roy and S. Biswas, "Intelligent Data Prognosis of Recurrent of Depression in Medical Diagnosis," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2020, pp. 840-844, doi: 10.1109/ICRITO48877.2020.9197843.

19. Raktim Kumar Dey, Debabrata Sarddar, Indranil Sarkar, Rajesh Bose, Sandip Roy, "A Literature Survey on Sentiment Analysis Techniques involving Social Media and Online Platforms", <https://www.researchgate.net/publication/341597914>, Article in International Journal of Scientific & Technology Research · June 2020.
20. Esteban A. R'issola, Diana Ram'irez-Cifuentes, Ana Freire, and Fabio Crestani, "Suicide Risk Assessment on Social Media", Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology, pages 167–171, Minneapolis, Minnesota, June 6, 2019. © 2019 Association for Computational Linguistics.
21. Ahona Ghosh, Sandip Roy, Suparna Biswas, "Mathematical Modelling For Decision Making of Lockdown during COVID-19", <https://www.researchsquare.com/article/rs-42770/v1>
22. Rajesh Bose, Sandip Roy, "A Novel Hybrid Approach for Diagnosis of Mental Health Condition Applying Intelligent Data Analysis 1", September 2020, International Journal of Advanced Trends in Computer Science and Engineering 9(5):1-10
23. Dey, R. K., Sarddar, D., Sarkar, I., Bose, R., & Roy, S. A Literature Survey on Sentiment Analysis Techniques involving Social Media and Online Platforms. International Journal Of Scientific & Technology Research, 1(1). INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 9, ISSUE 05, MAY 2020
24. Biswas, S., Ghosh, A., Chakraborty, S., Roy, S., & Bose, R. (2020). Scope of Sentiment Analysis on News Articles Regarding Stock Market and GDP in Struggling Economic Condition. International Journal, 8(7).
25. Dr. Sandip Roy, Dr. P. S. Aithal, Dr. Rajesh Bose, "A Novel Hybrid Approach for Diagnosis of Mental Health Condition Applying Intelligent Data Analysis", International Journal of Advanced Trends in Computer Science and Engineering Available Online at <http://www.warse.org/IJATCSE/static/pdf/file/ijatcse152952020.pdf>
<https://doi.org/10.30534/ijatcse/2020/152952020>, <http://www.warse.org/IJATCSE/static/pdf/file/ijatcse152952020.pdf>
26. Sandipan Biswas, Indranil Sarkar, Rajesh Bose and Sandip Roy, Prasenjit Das, "Examining the Effects of Pandemics on Stock Market Trends through Sentiment Analysis", <https://doi.org/10.37896/jxu14.6/138>
27. Sarddar, D., Dey, R. K., Bose, R., & Roy, S. (2020). Topic Modeling as a Tool to Gauge Political Sentiments from Twitter Feeds. International Journal of Natural Computing Research (IJNCR), 9(2), 14-35.
28. Bose, R., Dey, R. K., Roy, S., & Sarddar, D. (2020). Sentiment Analysis on Online Product Reviews. In *Information and Communication Technology for Sustainable Development* (pp. 559-569). Springer, Singapore.