International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 02 | February - 2023 Impact Factor: 7.185 ISSN: 2582-3930

Review on Machine Learning Based Hybrid Model for Suicide Attempt Prediction

Sarika H. Zile[1], Dr. Smita L. Kasar[2], Daivashala R. Deshmukh[3]

[1]M.Tech. CST Student – G. S. Mandal's Maharashtra Institute of Technology, Aurangabad, MH – IN
[2]Head of Dept. - CSE Dept – G.S. Mandal's Maharashtra Institute of Technology, Aurangabad, MH – IN
[3]Asst. Professor - CSE Dept – G.S. Mandal's Maharashtra Institute of Technology, Aurangabad, MH – IN

Abstract – Suicide attempts is a critical issue in modern society of the world. In this paper early detection and prediction of suicide attempts should be proposed to save people's life. Machine learning techniques with feature engineering or deep learning for automatic detection based on online social contents and to perform awareness are currently suicidal information of detection methods. Clinical methods based on interaction between social workers or experts and the targeted individuals. The first survey that provides a comprehensive overview of these categories' methods and performance is included in this paper. The data sources of process applications of suicide attempts and detection—such as questionnaires, electronic health records, suicide notes, and online user content and datasets—are examined. In order to make it easier to continue the analysis of the research, numerous specific tasks and datasets are described and summarized. Finally, we offer a perspective on the proposed work's future research directions and a summary of the work's limitations.

Key Words: Suicidal ideation detection, social content, features engineering, Machine learning.

1. INTRODUCTION

In today's society, suicide and attempts at suicide are extremely difficult to predict and detect. Human life's complex social, psychological, and biological factors all come together to form the proposed system. Practically speaking, it is also relatively uncommon, making suicide attempts even more difficult to predict. There are a number of data from years of research into factors that increase the likelihood of suicidal thoughts and actions

To overcome and against this possible, machine-learning based risk-algorithms such as ,Linear Regression, Feature ranking , Logistic regression , Decision tree, SVM ,KNN, Naive Bayes, ROC, Random forest, Bagging ,Voting, have been makes as an important tool to improve prediction accurately through the datasets. According to available data on Wikipedia, it says, "Machine learning (…) uses statistical techniques (…) to "learn" (e.g. progressively improve performance on a specific task) from available dataset, without being specific programmed". Machine learning can deal with (big) data sets with large numbers of records which are in the form of CSV format and many different variables and is mainly data driven in the matching form to the system. Also it can help hospitals or organizations such as radiologists make better clinical decisions randomly to carry out the required data within a short period of time.

Machine learning has been introduce to a system as a way of more improving our prediction as well as detection of suicide attempts in various forms and datasets.

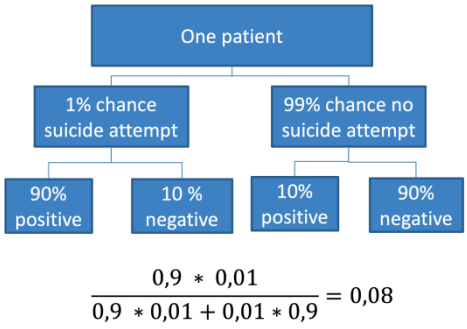

Fig 1. The suicide attempt after a positive test following Bayes' rule.

In this proposed paper, the author applied a time saving and flexible machine learning algorithm called random forest to carry out suicide data. The data were collected from the social media platform as well as by visiting the various local hospitals nearby.

Since not all self-injury entries in the available dataset refer to actual attempts at suicide, we collected this used dataset over several years of attempts at suicide across India; Each candidate record was graded by two experts. Some cases had a confirmed record of at least one suicide attempt after this expert review.

Additionally, some of the Predictors came from electronic health records, and they included: The results will be carried out with the assistance of diagnostics, medication, local hospitals, social and economic factors, and so on.

In this proposed paper Random forests were used to calculate the models to predict and detection of suicide attempts at different time points (from 1 day to 2 years) and many others.

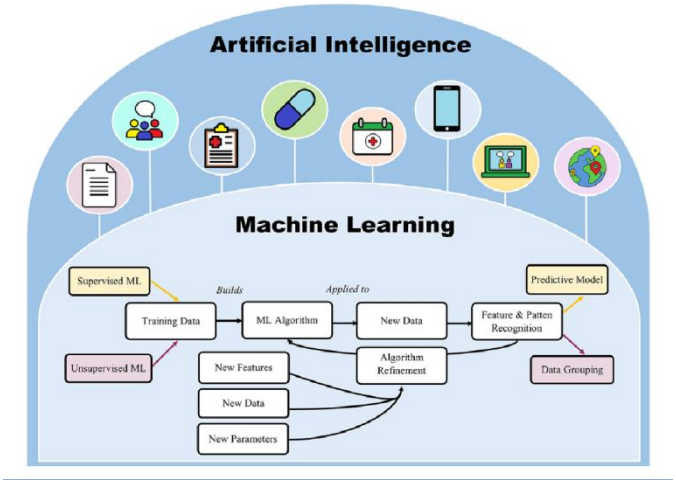

Figure 2. The relationship between machine learning and artificial intelligence.

2. Literature Review

S. Hinduja and J. W. Patchin says, Severe mental disorders without effective treatment can turn to suicidal ideation or even suicide attempts. Some online posts contain much negative information and generate problematic phenomena such as cyber stalking and cyber bullying. Consequences can be severe and risky since such lousy information is often engaged in some form of social cruelty, leading to rumors or even mental damage. Research shows that there is a link between cyber bullying and suicide [1].

J. Joo, S. Hwang, and J. J. Gallo says, Victims overexposed to too many negative messages or events may become depressed and desperate; even worse, some may commit suicide. The reasons that people commit suicide are complicated. People with depression are highly likely to commit suicide, but many without depression can also have suicidal thoughts [2]. M. J. Vioules, B. Moulahi, J. Az ` e, and S. Bringay says ,According to the American Foundation for Suicide Prevention (AFSP), suicide factors fall under three categories: health factors, environmental factors, and historical factors [3]. Ferrari et al. [4] found that mental health issues and substance use disorders are attributed to the factors of suicide. O'Connor and Nock [5] conducted a thorough review of the psychology of suicide and summarized psychological risks as personality and individual differences, cognitive factors, social factors, and negative life events. . Online communication channels are becoming a new way for people to express their feelings, suffering, and suicidal tendencies. Hence, online channels have naturally started to act as a surveillance tool for suicidal ideation, and mining social content can improve suicide prevention [6]. Suicidal ideation detection is to find out these risks of intentions or behaviors before tragedy strikes. A meta-analysis conducted by McHugh et al. [7] shown statistical limitations of ideation as a screening tool, but also pointed out that people's expression of suicidal ideation represents their psychological distress. Effective detection of early signals of suicidal ideation can identify people with suicidal thoughts and open a communication portal to let social workers mitigate their mental issues. The reasons for suicide are complicated and attributed to a complex interaction of many factors [5], [8]. To detect suicidal ideation, many researchers conducted psychological and clinical studies [9] and classified responses of questionnaires [10]. Based on their social media data, artificial intelligence (AI) and machine learning techniques can predict people's likelihood of suicide [11], which can better understand people's intentions and pave the way for early intervention. Detection on social content focuses on feature engineering [12], [13], sentiment analysis [14], [15], and deep learning [16], [17], [18]. Those methods generally require heuristics to select features or design artificial neural network architectures for learning rich representation. The research trend focuses on selecting more useful features from people's health records and developing neural architectures to understand the language with suicidal ideation better. Mobile technologies have been studied and applied to suicide prevention, for example, the mobile suicide intervention application iBobbly [19]. Applying cutting-edge AI technologies for suicidal ideation detection inevitably comes with privacy issues [20] and ethical concerns [21]. Linthicum et al. [22] put forward three ethical issues, including the influence of bias on machine learning algorithms, the prediction on time of suicide act, and ethical and legal questions raised by false positive and false negative prediction. It is not easy to

answer ethical questions for AI as these require algorithms to reach a balance between competing values, issues, and interests [20].

3. Body of Paper

- They apply novel algorithms to a diverse collection of datasets that are typically accessible in the proposed paper. Exploration, detection, hypothesis generation, and prediction of suicides are made possible by the expansion of routinely available data and stronger algorithms.

- There is a possibility that the results will be implemented, but our expectations about the kinds of problems that machine learning algorithms can solve are too high. The fact that suicide attempts are extremely uncommon and that their causes are difficult to identify is one of the main obstacles in predicting and identifying suicide attempts. Any machine learning algorithm test can have a good detection range, sensitivity, and specificity. However, if the prevalence in the population being tested is low (like in general population counting, primary care, or a general hospital system), a positive test result cannot tell us if a person is sick or has any disease. To put it another way, tests are more likely to be incorrect than accurate when the count of an event is low. As shown in fig. 1 above, this is known as the Bayes' rule.

- Additionally, the authors reported more cases of suicide attempts in recent years over a number of years of data collection and detection. Even if we set prevalence at 1 percent and specificity and sensitivity at 0.9, the likelihood that an adolescent with a "positive test" actually being a case (i.e., suicide attempt) is only 8%. That prediction is beyond the scope of any machine learning method.

- By using some machine learning it is possible to predict and detect suicide attempts when it compared with the control group of applied algorithms.

- Also some standard performance stages such as area under the curve were good across all time windows as applicable.

- This model performed better if:
    - Available timeframe to the suicide attempt was shorter and shorter.
    - Given control group was less at risk factor.

- Random forest performs significantly better than models using standard logistic regression analyses algorithms.

- Machine learning has the ability to greatly perform & improve our prediction as well as detection of suicidal behavior.

- Also, given the low prevalence of suicidal behavior, prediction will always be challenging to carry out the required outcomes.

- This process is unrealistic to think that an applied algorithm will solve this problem in the short run of time.

- Appling the more powerful field of suicide prevention and detection, machine learning should be combined with available information and the data collected consistent with the system.

• The datasets used in this proposed paper were gathered as part of a rigorous longitudinal research project by applying machine learning to a variety of medical organizations. As is common knowledge, many nearby mental health facilities and hospitals lack an appropriate and high-quality data infrastructure, making prediction even more challenging. Before algorithm-based suicide prediction can be implemented on a large scale, privacy concerns in Indian society must also be addressed.

• We believe that the model building and validation using machine learning algorithms offers hope for suicide detection and prevention. Prediction, detection, cross-validation, and bootstrapping are some of the methods used in the proposed system approach to determine whether a statistical model accurately reflects the data. When these systems are used, the required outcomes are more reliable and accurate.

While researching these methods, a significant factor in the relative lack of progress in the detection and prediction of suicidal behavior is not the incorrect use of statistics but rather the absence of a clear scientific theory. We are essentially looking in the dark side without theory.

4. CONCLUSIONS

In today's modern Indian society, suicide detection and prevention remain crucial tasks. Predicting and identifying suicidal thoughts in a timely manner is critical to ongoing suicide prevention. From a broader perspective, the proposed paper survey examines the various clinical approaches used to detect suicidal ideation, such as patient-clinician interaction and medical signal sensing.

Examples of feature engineering include word cloud visualization and dictionary-based filtering in text analysis of tabular, textual, and affective features; and representation learning based on deep learning, like CNN and LSTM-based text encoders. There are four main applications specific to a particular domain: questionnaires, suicide notes, and electronic health records from online users. Psychologists have used statistical analysis for the majority of this field's research, while computer scientists have used feature engineering-based machine learning and deep learning-based representation learning.

We provided a summary of existing tasks and further suggested new potential tasks on the basis of current research. Last but not least, we go over some of the limitations of the research that is currently being done and offer some ideas for the future, like using new learning methods, understanding interpretable intentions, detecting temporal patterns, and using proactive conversational intervention. In the not-too-distant future, online social content will most likely serve as the primary method for detecting suicidal ideation. Therefore, in the hope that suicide can be avoided, it is essential to develop new strategies for bridging the gap between automatic machine detection and clinical mental health detection in online texts.

REFERENCES

[1]S. Hinduja and J. W. Patchin, "Bullying, cyberbullying, and suicide," Archives of suicide research, vol. 14, no. 3, pp. 206–221, 2010.

[2]J. Joo, S. Hwang, and J. J. Gallo, "Death ideation and suicidal ideation in a community sample who do not meet criteria for major depression," Crisis, 2016.

[3] M. J. Vioules, B. Moulahi, J. Az ` e, and S. Bringay, "Detection ´ of suicide-related posts in twitter data streams," IBM Journal of Research and Development, vol. 62, no. 1, pp. 7:1–7:12, 2018.

[4] A. J. Ferrari, R. E. Norman, G. Freedman, A. J. Baxter, J. E. Pirkis, M. G. Harris, A. Page, E. Carnahan, L. Degenhardt, T. Vos et al., "The burden attributable to mental and substance use disorders as risk factors for suicide: findings from the global burden of disease study 2010," PloS one, vol. 9, no. 4, p. e91936, 2014.

[5] R. C. O'Connor and M. K. Nock, "The psychology of suicidal behaviour," The Lancet Psychiatry, vol. 1, no. 1, pp. 73–85, 2014.

[6] J. Lopez-Castroman, B. Moulahi, J. Aze, S. Bringay, J. Deninotti, ´ S. Guillaume, and E. Baca-Garcia, "Mining social networks to improve suicide prevention: A scoping review," Journal of neuroscience research, 2019.

[7] C. M. McHugh, A. Corderoy, C. J. Ryan, I. B. Hickie, and M. M. Large, "Association between suicidal ideation and suicide: meta-analyses of odds ratios, sensitivity, specificity and positive predictive value," BJPsych open, vol. 5, no. 2, 2019.

[8] G. Kassen, A. Kudaibergenova, A. Mukasheva, D. Yertargynkyzy, and K. Moldassan, "Behavioral risk factors for suicide among adolescent schoolchildren," Elementary Education Online, vol. 19, no. 1, pp. 66–77, 2019.

[9]V. Venek, S. Scherer, L.-P. Morency, J. Pestian et al., "Adolescent suicidal risk assessment in clinician-patient interaction," IEEE Transactions on Affective Computing, vol. 8, no. 2, pp. 204–215, 2017.

[10] D. Delgado-Gomez, H. Blasco-Fontecilla, A. A. Alegria, T. LegidoGil, A. Artes-Rodriguez, and E. Baca-Garcia, "Improving the accuracy of suicide attempter classification," Artificial intelligence in medicine, vol. 52, no. 3, pp. 165–168, 2011.

[11] G. Liu, C. Wang, K. Peng, H. Huang, Y. Li, and W. Cheng, "SocInf: Membership inference attacks on social media health data with machine learning," IEEE Transactions on Computational Social Systems, 2019.

[12] B. O'Dea, S. Wan, P. J. Batterham, A. L. Calear, C. Paris, and H. Christensen, "Detecting suicidality on

twitter," Internet Interventions, vol. 2, no. 2, pp. 183–188, 2015.

[13] H.-C. Shing, S. Nair, A. Zirikly, M. Friedenberg, H. Daume III, ´ and P. Resnik, "Expert, crowdsourced, and machine assessment of suicide risk via online postings," in Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic, 2018, pp. 25–36.

[14]F. Ren, X. Kang, and C. Quan, "Examining accumulated emotional traits in suicide blogs with an emotion topic model," IEEE journal of biomedical and health informatics, vol. 20, no. 5, pp. 1384–1396, 2016.

[15]L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, "A survey of sentiment analysis in social media," Knowledge and Information Systems, pp. 1–47, 2018.

[16] A. Benton, M. Mitchell, and D. Hovy, "Multi-task learning for mental health using social media text," in EACL.Association for Computational Linguistics, 2017.

[17]S. Ji, C. P. Yu, S.-f.Fung, S. Pan, and G. Long, "Supervised learning for suicidal ideation detection in online user content," Complexity, 2018.

[18] S. Ji, G. Long, S. Pan, T. Zhu, J. Jiang, and S. Wang, "Detecting suicidal ideation with data protection in online communities," in 24th International Conference on Database Systems for Advanced Applications (DASFAA). Springer, Cham, 2019, pp. 225–229.

[19] J. Tighe, F. Shand, R. Ridani, A. Mackinnon, N. De La Mata, and H. Christensen, "Ibobbly mobile health intervention for suicide prevention in australian indigenous youth: a pilot randomised controlled trial," BMJ open, vol. 7, no. 1, p. e013518, 2017.

[20] N. N. G. de Andrade, D. Pawson, D. Muriello, L. Donahue, and J. Guadagno, "Ethics and artificial intelligence: suicide prevention on facebook," Philosophy & Technology, vol. 31, no. 4, pp. 669–684, 2018.