(IJIEST) 2023, Vol. No. 9, Jan-Dec **e-ISSN: 2454-9584; p-ISSN: 2454-8111**

Developing a Prediction Model Based on Machine Learning Tools and Techniques to Avert Suicide and Suicidal Tendencies of Vulnerable Individuals

Aditi Singh

Department of Psychology

Manipal University, Jaipur

[1]*Date of Receiving: 19 November 2022; Date of Acceptance: 14 January 2023; Date of Publication: 12 February 2023*

ABSTRACT

In today's environment, depression and suicide rates are rising. The most frequent reasons of suicide thoughts in humans are stress, anxiety attacks, depressive illnesses, and other conditions. Elevated psychological strain can set off suicidal thoughts and is the main driving force for suicide attempts. However, the statistics employed in conventional suicide prediction systems take into account the tenuous link between individuals' psychological stress and suicidal ideation. The goal of this research study is to advance the field by developing various machine learning algorithms, such as Random Forest [RF] and Support Vector Machine [SVM], to analyze and predict suicidal ideation in individuals based on the six most significant psychological stress-causing domains and the messages they send to others. Next, an analysis and comparison are conducted between the Random Forest [RF] and SVM methodologies' accuracy. Out of them, the Random Forest has outperformed the Support Vector Machine in terms of accuracy.

INTRODUCTION

People in today's environment are exposed to psychological stress for a variety of reasons, including peer pressure, busyness, mental instability, and excessive obligations.

In the United States of America [USA], adults between the ages of 15 and 19 have suicide as one of the top three causes of mortality. It was discovered that 2.0% to 37.4% of US military personnel suffered from major depression [1]. Veterans of the US military have been reported to suffer from 2.0% to 17.0% from post-traumatic stress disorder (PTSD) [2]. Individuals who had been deployed several times, exposed to combat, high-stakes scenarios, unusual health conditions, and recurrent failures were more likely to experience mental health issues and suicide thoughts [3]. In contrast to American civilians, active army personnel have made more suicide attempts, per a survey [4-6].

Previous research has linked psychological stress with suicidal thoughts. Early detection of suicidal thoughts and subsequent suicidality prevention are crucial in today's reality [7].

Advances in technology are driving a rapid evolution in artificial intelligence (AI). The field of big data is likewise changing. Scholars have devoted a great deal of effort to developing computer algorithms that can handle enormous data sets. Artificial intelligence and computing are combined with machine learning to provide accurate disease prediction and detection.

[1] *How to cite the article:* Singh A. (2023) Developing a Prediction Model Based on Machine Learning Tools and Techniques to Avert Suicide and Suicidal Tendencies of Vulnerable Individuals; *International Journal of Inventions in Engineering and Science Technology*, Vol 9, Issue 1, 21-28

In the modern world of stress and many illnesses, a model that can predict suicidal ideation is desperately needed. They have the potential to save many lives.

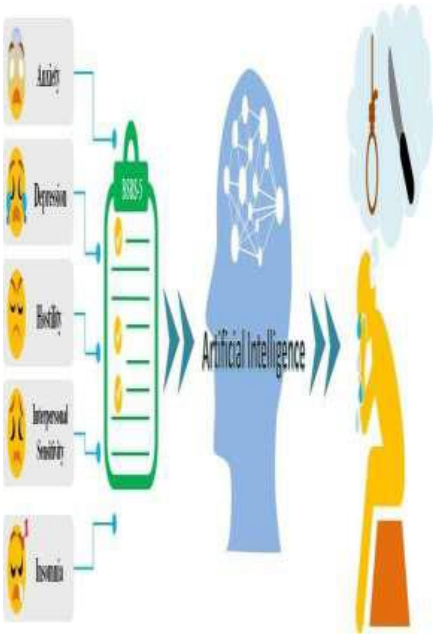

Figure 1. Diagrammatic representation

The main cause of suicide attempts is suicidal thoughts, which can be brought on by a lot of mental stress. However, conventional statistical methods only find a weak correlation between suicidal ideation and psychological stress. Machine learning may be able to more accurately anticipate suicidal thoughts [8].

In this study, we employ various machine learning algorithms on a large sample of text messages and tweets to predict the frequency of suicidal thoughts using psychological stress characteristics.

A schematic diagram of the recommended methodology used in this study is presented in Figure 1. Based on the respondents' surveys, a probabilistic classifier that maintains a binary value from a machine learning algorithm can determine whether or not the respondents had suicidal thoughts. Because machine learning can recognize suicidal thoughts automatically, it offers a useful tool for suicide prevention and early identification.

REVIEW OF LITERATURE

There are frequently more accurate methods for forecasting suicidal thoughts and behaviors. Researchers found that there hasn't been any advancement in assessing a person's risk of engaging in suicidal thoughts or behaviors over the preceding fifty years.

In addition to using recognized risk indicators, we continued to get better results in the forecasting of these outputs [9]. Using a range of machine learning approaches, Dr. Andrew Littlefield and associates aim to gain a deeper understanding of suicide markers. Machine learning differs from earlier methods in that it prioritizes prediction over explanation when identifying the factors that contribute to suicide. Algorithms for machine learning are able to look at a lot of variables and determine the intricate connections between them. Consequently, it can integrate large volumes of data and does not require an explanatory theory-based approach. Consequently, it can concurrently assimilate the enormous volumes of data and does not require an explanatory theory-based approach.

Using the MSRC's Common Data Elements (CDE), Dr. Littlefield and his associates want to pinpoint military personnel who are more likely to consider or carry out suicide acts. While a number of earlier studies have employed CDE data to search for explanatory correlations, none have employed all of the variables to forecast the risk of suicide. By employing this technique, Dr. Littlefield and his colleagues hope to identify the most effective categorization algorithms that can be used to identify individuals who are at risk in the future.

Here, machine learning is used in six distinct ways to identify and categorize individuals who are contemplating suicide.

The following groups will be targeted: those who have made a suicide attempt in the past, those who have a plan for suicide, those who are actively considering suicide, those who are at a higher risk of making a deadly suicide attempt. To test if machine learning can more accurately identify people who are at risk of suicide, it will compare the algorithms discovered by machine learning to those found by employing traditional methods.

Following a careful examination of the CDE, these ideal feature procedures can be employed as screening instruments and enhance military personnel's mental health [10]. We're hoping that better identification will enable us to stop suicides.

PROPOSED MODEL

The proposed system has used machine learning models to predict the output. This model has utilized a dataset containing a huge set of messages and tweets of different people from different parts of the world. This dataset is given to two machine learning models and the suicide ideation among the people can be predicted. The accuracy of both the machine learning algorithms is calculated and measured.

A. Data Pre-Processing

We use the technique of normalization of Min-Max scaling [11] in which the input data is normalized into the interval (0-1) which helps us in solving the occurrences of dynamic ranges that are distinct for the six different input variables. On the original data, minimum-maximum normalization will be used to perform the linear transformation.

Normalized(X')=X -mini_mum(X)/maxi _mum(X)-mini_mum(X)

where the lowest and maximum values of the input feature x are indicated by the variables mini_mum(X) and maxi_mum(X). The normalized data is represented by the symbol X′.

Ten-fold cross-validation is employed in this study.

Suicidal thoughts are predicted using two sets of data: the training set and the testing set. There are two distinct classes in the training and testing sets used to predict the high level of suicidal ideation. Class 0 refers to people who have never considered suicide, whereas class 1 is made up of people who have considered suicide. There is a glaring difference in the sample between classes 0 and 1. This issue is resolved by applying the synthetic minority over-sampling method, or SMOTE [12]. The training data for the two predictions is pre-processed by SMOTE, in turn.

B. Algorithms for Machine Learning

1. The Random Forest (RF)

Another name for random forests is random choice forests. It's an algorithm for group learning. They function by putting a lot of samples through training. In our random forest technique, the bootstrap aggregating training process is employed. When dividing a node in a decision tree, a subset with a random set of features is taken into account.

The trees in the random forest approach are not manicured because each individual tree is unrelated to any other tree. The final prediction result is obtained from the majority-votes model from many DTs. In RF, the benefits of bagging and feature selection are merged.

The decision tree number is the hyperparameter that has to be adjusted.

The RF pseudocode is

1. Select 'k' features at random from the total of 'm' features that are available, where k is significantly more than m.
2. Using the split point that is best among the other available "k" attributes, determine the node "n."
3. Divide the node into additional little child nodes by splitting it at its optimal place.
4. Continue doing steps 1 through 3 until the desired number of nodes, "li," is reached.
5. To create an RF, go through steps 1 through 4 "N" times to get a "N" number of trees.

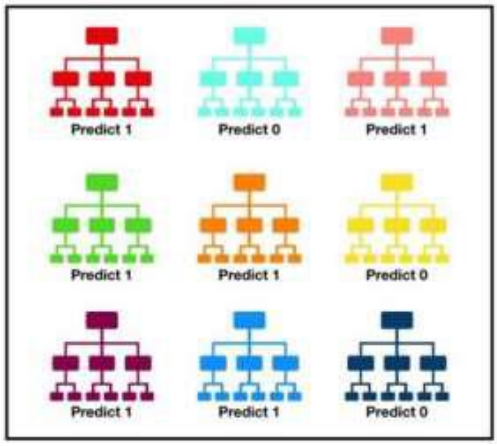

Figure 2. Depiction of Random Forest

2. Support vector machine

A Linear support vector machine (Linear SVM) with one linear kernel is used in our proposed approach (Linear SVM). A 6-dimensional vector is represented for the data point, and we use a hyperplane to separate them. This linear SVM is used to create the maximum margin hyperplane, which maximizes the distance between it and the training data point which is nearest to any class [class 0 or class 1].

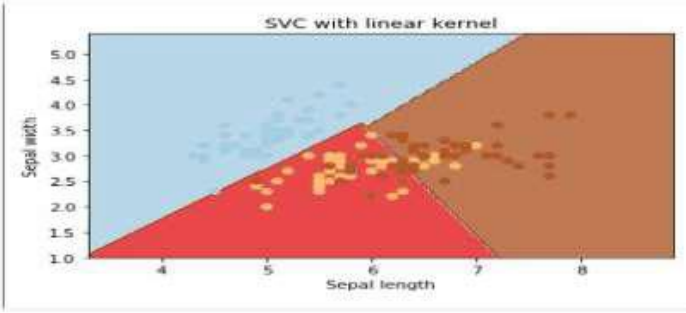

Figure 3. Graph of SVM

```
from sklearn.svm import SVC # "Support vector classifier"
classifier = SVC(kernel='linear', random_state=0)
classifier.fit(x_train, y_train)
```

Figure 4. ML implementation of SVM

EXPERTIMATED OUTCOMES

A. Assessing the evaluation

By calculating the accuracy, precision, specificity, sensitivity, F1 score, and area under the precision-recall (PR) curve, the models produced by applying various methods are assessed.

Equations for True positive (Trp), True negative (Trn), False positive (Frp), and False negative (Frn) determine accuracy, sensitivity, specificity, confusion matrix, and precision.

(Trp+Trn)/(Trp+Trn+Frp+Frn) equals accuracy.

Accuracy is among the most basic performance metrics. It is the proportion of observed values to all observed values that may be reasonably predicted [14–16].

B. Matrix of confusion

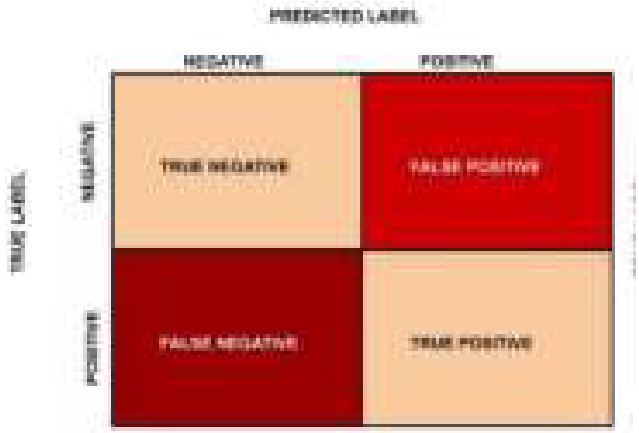

Figure 5. Confusion matrix depiction

EVALUATION AND CONVERSATION

By comparing the two machine learning algorithms' performance with the respondents' questionnaire scores, the techniques' accuracy, specificity, sensitivity, precision, and F1 score under a curve are assessed [13]. It is determined how well the support vector machine and random tree perform. The results of the suicide ideation 1 forecast are better than the results of the suicide ideation 2 prediction.

The learning curves for the two suggested methods for suicidal thoughts are shown in the following figures. when the tree depth or number of support vector machine iterations is altered. The random forest F1 score steadily improves, as indicated by the learning curves. The curves of the two suggested methods exhibit good convergence.

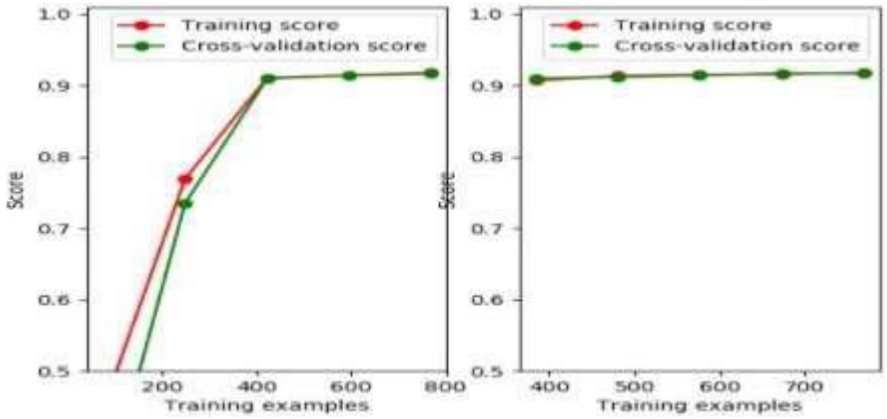

Figure 6. The learning curve of the Decision tree

Figure 7. The learning curve of Random Forest

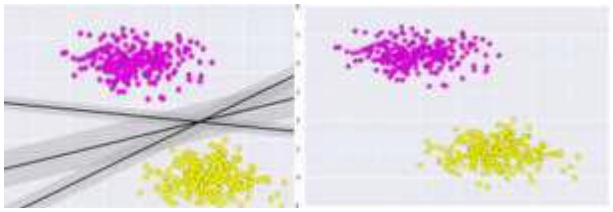

Figure 8. SVM depiction

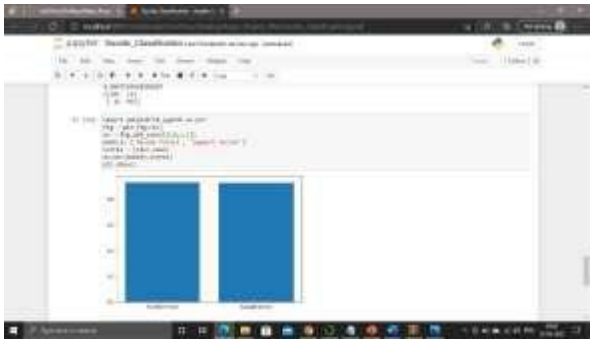

Figure 9. Graph comparing accuracies of Random Forest and SVM

CONCLUSION & FUTURE WORK

The suggested research project has trained machine learning algorithms on a range of psychological stress indicators in order to predict suicidal thoughts. In order to avoid suicide, the suggested model can greatly assist in identifying military personnel who are very susceptible to it. Future iterations of this suggested approach could benefit from the application of numerous additional machine learning methods, such as decision trees, logistic regression, naive bayes, and many more.

REFERENCES

1. A. Karthika, N. Muthukumaran, R. Joshua Samuel Raj, 'An Ads-Csab Approach for Economic Denial of Sustainability At tacks in Cloud Storage', International Journal of Scientific & Technology Research, Vol. 9, Issue. 04, pp. 2575-2578, April 2020.
2. Potluri, Sirisha, et al. "Cloud Manufacturing Service: A Secure and Protected Communication System." Cloud Security: Techniques and Applications (2021): 171.
3. R. Joshua Samuel Raj, T. Sudarson Rama Perumal, N. Muthukumaran, 'Road Accident Data Analytics Using Map - Reduce Concept ',International Journal of Innovative Technology and Exploring Engineering, Volume. 8, Issue. 11, pp. 1032- 1037, September 2019.
4. K. Lakshminarayanan, N. Muthukumaran, Y. Harold Robinson, Vimal Shanmuganathan, Seifedine Kadry and Yunyoung Nam, "Deep Learning-Based Hookworm Detect ion in Wireless Capsule Endoscopic Image Using AdaBoost Classifier", Computers, Materials & Continua, vol. 67, no.3, pp. 3045–3055, 2021.
5. A. Karthika, N. Muthukumaran, " An ADS-PAYG Approach using Trust Factor against Economic Denial of Sustainability At tacks in Cloud Storage", Wireless Personal Communications, Vol. 122, No. 1, pp. 69–85, January 2022.
6. J. Rethna Virgil Jeny, Chetan Anil Joshi, " Flexible Dat a St reaming In St ream Cloud" , International Journal of Innovative Research in Science, Engineering and Technology, Vol. 2, Issue 4, April 2013, ISSN: 2319-8753, pp:1127-1131
7. VP. Anubala, N. Muthukumaran and R. Nikitha, 'Performance Analysis of Hookworm Detection using Deep Convolutional Neural Network', 2018 International Conference on Smart Systems and Inventive Technology, pp. 348-354, 2018, doi:10.1109/ICSSIT.2018.8748645.
8. Niha, K., Amutha, S., Banu, A. , (2021), A Convolutional Neural Network based System to Detect Plant Disease, Webology, 18, pp. 944–962.
9. R. Kabilan, N. Muthukumaran, "A Neuromorphic Model for Image Recognition using SNN", 2021 6t h International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2021, pp. 720-725.
10. Chang, H. A., Tzeng, C. K, Wang, T. S., Kao, Y. C., Yeh, H. W.& Huang, S. Y. N. S., Chen, (2016). Forensic psychiatric evaluation for military absenteeism in Taiwan. Journal of the American Academy of Psychiatry and the Law Online, 44(3), 352-358.

11. J. Rebekah, D. C. J. W. Wise, D. Bhavani, P. Agatha Regina and N. Muthukumaran, "Dress code Surveillance Using Deep learning," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 394-397, Doi: 10.1109/ICESC48915.2020.9155668.
12. S. Gayathri, D. C. J. W. Wise, P. B. Shamini and N. Muthukumaran, "Image Analysis and Detect ion of Tea Leaf Disease using Deep Learning," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 398-403, Doi: 10.1109/ICESC48915.2020.9155850.
13. J. Bhola, G. Dhiman, T. Singhal, and G. S. Sajja, "A Novel Technique on Autism Spectrum Disorders Using Classification Techniques," pp. 40–53, 2021, Doi: 10.4018/978-1-7998-7460-7. ch003.
14. G. P. Devaraj, R. Kabilan, J. Z. Gabriel, U. Muthuraman, N. Muthukumaran and R. Swetha, "Design and Analysis of Modified Pre-Charge Sensing Circuit for STT-MRAM," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, March 2021, pp. 507-511.
15. G. S. Sajja, M. Mustafa, K. Phasinam, K. Kaliyaperumal, R. J. M. Ventayen and T. Kassanuk, "Towards Application of Machine Learning in Classification and Prediction of Heart Disease," 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), 2021, pp.1664-1669, doi: 10.1109/ICESC 51422.2021.9532940
16. Shahin, M.Ahmed, B.Hamida, S.T.B. Mulaffer, F.L.Glos, M., & Penzel, T.(2017). Deep learning and insomnia, assisting clinicians with their diagnosis. IEEE journal of biomedical and health informatics, 21(6), 1546-1553.
17. Gautham A Nagendran, Harshmeet Singh, R Joshua Samuel Raj and N. Muthukumaran, "Input Assistive Keyboards for People with Disabilities: A Survey," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, March 2021, pp. 829-832.