Suicidal Ideation Prediction Using Machine Learning

Dr.N.Shanthi
Professor
Department of Computer Science and Engineering
Kongu Engineering college (Autonomous) Perundurai
Erode - 638 060
shanthi.cse@kongu.edu

M.Muthuraja
Assistant Professor
Department of Computer Science and Engineering
Kongu Engineering college (Autonomous) Perundurai
Erode - 638 060
muthraja.cse@kongu.edu

C.Sharmila
Assistant Professor
Department of Computer Science and Engineering
Kongu Engineering college (Autonomous) Perundurai
Erode - 638 060
sharmila.cse@kongu.edu

Jagadeesh S
UG Student
Department of Computer Science and Engineering
Kongu Engineering college (Autonomous) Perundurai
Erode - 638 060
jagadeeshs.19cse@kongu.edu

Karthick R
UG Student
Department of Computer Science and Engineering
Kongu Engineering college (Autonomous) Perundurai
Erode - 638 060
karthickr.19cse@kongu.edu

Bharanidharan M
UG Student
Department of Computer Science and Engineering
Kongu Engineering college (Autonomous) Perundurai
Erode - 638 060
bharanidharanm.19cse@kongu.edu

Abstract— **In the present world, accidents and health complications account for the majority of fatalities. The majority of deaths after accidents are caused by suicide due to depression and natural catastrophes. The widespread use of the Internet has given people a new means of communicating their feelings. It is also a platform with a massive amount of content where users may read other users' opinions, which are divided into several sentiment groups and are becoming more and more important in decision-making process. This paper contributes to the classification that is useful to examine the data in the form of the quantity of tweets where comments are extremely unstructured and either negative or positive or somewhere in between these two. To do this, we first pre-processed the data, then extracted the adjectives with meaning from the tweet, and last utilised machine learning-base classification methods, specifically. When compared to the current system, the accuracy of the TFIDF, N-gram, and LinearSVC algorithms for suicide prediction with tweets including suicidal thoughts was improved to 95 percent. Such testing and observation may help in both individual and population-wide prevention. By establishing a baseline for suicide identification on online social networks, such as Twitter, the experimental work suggests the viability of the approach adopted. In the end, we evaluated the classifier's performance in terms of accuracy.**

Keywords- Suicidal Prevention, Machine Learning, Tweets , Depression

979-8-3503-4821-7/23/$31.00 ©2023 IEEE

I. INTRODUCTION

One of the biggest risk factors for suicide thoughts and attempts is depression. Therefore, understanding characteristics that contribute to he emergence of suicidal thoughts in depressed teenagers should help direct treatments to stop suicide acts. Suicidal planning and attempts can result from suicidal thoughts of suicidal ideas and plans are, in fact, the biggest causes of suicide and excellent indicators of future suicidal action, according to several research. Suicidal thoughts can affect people of all ages for a variety of causes, such as shock, anger, guilt, melancholy, and anxiety[1]. Despite the fact that the majority of people who suffer suicidal thoughts may not actually attempt suicide, long-term depression may result in suicide if sufficient therapy is not obtained. Instead, a lot of people decide to announce their intention to kill themselves on social media. Early detection of warning signals or risks may be the most effective method of avoiding suicide since mental illness may be recognised and treated. Applying pertinent annotation rules to massive amounts of data and training machine learning (ML) models both help improve categorization accuracy[2]. On this work, we employ machine learning models to predict the precision of individuals with suicide ideation based on the textual information they offer in the social media.

II. EXISTING SYSTEM

Because the rate of suicide deaths has risen alarmingly in recent years, particularly among young teens who are either in high school or college and adults who work in environmental stresses, many researchers have written papers on suicidal ideation. These reasons include depression, stress, mental exhaustion, and negative experiences[3,4]. Finding out if the candidate's tweets include suicidal thoughts is the goal of content-based suicide categorization. Such tactics include phrase filtering and phrase filtering for phrases relating to suicide. In this sector, artificial intelligence techniques, particularly supervised learning techniques and natural language processing (NLP), are also used. Twitter tweets are divided into sentences in the current structure by taking into account 62 keywords and employing recommended extra keywords of n-grams [5-6] as search key phrases.

Kasper et al analyzed the potential of machine learning to forecast future suicide ideation using population-based statistics. The top 5 significant predictors of probable suicidal ideation and suicide attempt were reported by the authors [7]. Damien Lekkas et al used an ensemble machine learning model to predict suicidal thoughts within the past month in a dataset of teenagers from Instagram who had previously experienced lifetime suicidal thoughts. According to the authors, existing machine learning techniques can be used to foresee suicidal thoughts [8]. By using Twitter data from the last two years as a goal of early detection through sentiment analysis and supervised learning approaches, ER Kumar et al. demonstrate various techniques to understand suicidal ideation through online user contents, in particular. In order to talk about their issues or get knowledge on related topics, several persons use social discussion platforms. Tweets are collected automatically in the proposed research and old data may be considered for classification in future [9].

Deep learning and machine learning-based classification techniques implemented to Reddit social media can be used to identify suicidal ideation early used a machine learning algorithm to try to create a model that could predict people who had suicidal thoughts in the wider population. The authors employed LSTM-CNN model for classification. This approach is combined with other algorithms [10]. This paper proposes the machine learning approach for predicting suicidal ideation at an early stage.

III. PROPOSED SYSTEM

The architecture of the proposed work is shown in Figure 1. Suicide_detection.csv is available in kaggle which is used as the dataset in the proposed work. The dataset consists of 232074 values with two outputs as suicide or non_suicide. The data is collected from the "suicidewatch" platform. Initially the text in the dataset is preprocessed. After preprocessing N-gram modeling is done for Term Frequency - Inverse Document Frequency (TF-IDF) text classification. After this the TF-IDF values are given to different classifiers for classification. The classifiers identifies the given text belongs to suicide or non-suicide. The following section describes the various modules proposed in the paper.

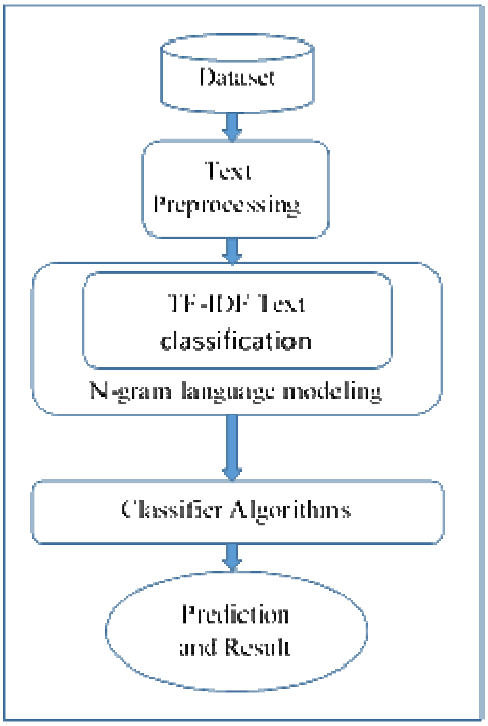

Figure1. *Architecture of proposed system*

IV. MODULE DESCRIPTION

A. Text preprocessing in NLP

Data preprocessing is a crucial stage in creating a machine learning model, and the effectiveness of the preprocessing determines the outcomes. Preprocessing text is the initial stage in NLP's model-building process. Following are the different text preparation steps:

B. TF-IDF Text Classification

Term Frequency - Inverse Document Frequency (TF-IDF), is a technique that is used to determine the meaning of sentences made up of words. It overcomes the limitations of the Bag of Words approach, which is effective for text categorization. Using TF-IDF, a word's relevance for a text is more properly described while also taking into account how that document relates to other papers in the same corpus. This is accomplished by calculating the frequency with which a word appears in a given document as well as the frequency with which the identical term appears in texts across the corpus. The following justifies this:

- A phrase with a high frequency of occurrence in a text has greater significance for that text, which raises the probability that the text seems to be about or connects to that specific word.
- Finding the right document in a group may be challenging if a phrase often used in many documents either applies to all documents or has no application at all. We won't be able to eliminate a document or a small group of documents from the entire collection, in any scenario.

Every word in every document in our dataset receives a TF-IDF score. Additionally, for each word, the TF-IDF value rises with each instance of the term in a document while progressively falling with each instance of the word in other documents

TF formula:

$$tf(w,d) = log(1 + f(w,d))$$

C. Ngram Language Modelling

Language modelling is a method for calculating the likelihood of any word sequence. Many different applications, including speech recognition and spam filtering, employ language modelling. In reality, the construction of several cutting-edge NLP models is driven mostly by language modelling.

Language modelling techniques come in two varieties:

- Statistical language modelling, often known as language modelling, is the process of creating probabilistic models that can predict the following word in a sequence based on the words that came before. Examples include language modelling using N-grams.
- Neural Language Modelling: Neural network approaches are outperforming conventional methods for difficult tasks like voice recognition and machine translation, both when used as independent language models and when included into larger models. Using word embeddings, one may simulate a neural language model.

The continuous collection of n elements from a particular set of text or voice is known as a "n-gram." Depending on the application, the objects might be letters, words, or base pairs. The N-grams are usually gathered from a corpus of text or voice. An N-gram language model forecasts the likelihood of a certain N-gram in any language's word order. A decent N-gram model can predict the value of p(w|h), or the next word in the sentence.

Examples of N-grams include unigrams and bi-grams, such as "This article," "article is," "is on, on NLP." Calculating p(w|h), where is the potential word, is necessary. In the example above, for instance, suppose that we wanted to determine the likelihood that the final word would be "NLP" given the terms that came before it.

D . Algorithms used

Linear SVC

A technique called as Linear Support Vector Machine (Linear SVC) looks for a hyperplane to optimise the distance between samples that are categorised. Using a kernel function, data may be entered and then transformed into the format needed for processing. The term "kernel" is employed because the Support Vector Machine's window for manipulating the data is provided by a collection of mathematical functions. A Linear SVC's (Support Vector Classifier) goal is to split or categorise the data you supply by producing a "best fit" hyperplane. In order to determine what the "predicted" class is like, you can then give certain characteristics to your classifier after acquiring the hyperplane

Random forest

The training phase of the random forests ensemble learning technique, which is used for classification, regression, and other tasks, involves the construction of a substantial number of decision trees. The result of using a random forest to solve classification issues is the class that most of the trees select. For regression tasks, the mean or average estimate of each individual tree is returned. Random decision forests correct for the tendency of decision trees to overestimate to their training. Gradient improved trees are more precise than random forests, albeit they frequently beat them. But how effectively they operate can be affected by data aspects.

XGBoost

Extreme Gradient Boosting, often known as XGBoost. By combining several weak classifiers, the ensemble modelling approach known as "boosting" aims to create a powerful classifier. It is accomplished by employing weak models in sequence to develop a model. First, a model is created using the training set of data. The new model is then created in an effort to fix the previous model's flaws. Models are added in this manner until either the full training data set is properly predicted or the optimum number of models have been added.

Logistic regression

In a binary classifier, the result is predicted via logistic regression. Rather of giving the precise values of 0 and 1, it provides the probabilistic values that fall between 0 and 1. It could be either Yes or No, 0 or 1, true or false, etc. As a result, the result must be a discrete or categorical value.

V. PERFORMANCE METRICS

A. Precision

Precision is mostly employed when predicting the positive class is necessary since false positives are more expensive than false negatives, like in spam filtering or medical diagnosis. When the classes are severely unbalanced, the precision score is a helpful indicator of the accuracy of the forecast. It reflects the ratio of genuine positives to the total of true positives and false positives mathematically. Precision Score = TP / (FP + TP)

B. Recall

The recall score assesses the model's accuracy in correctly predicting positives out of true positives. Contrary to precision, which measures the percentage of correct positive predictions among all positive predictions made by models, the former is a different measure. If you were attempting to detect good reviews, for example, the recall score would be the proportion of positive reviews that your machine learning model correctly recognised as positive. It evaluates how well our machine learning model can discriminate between all true positives and all false positives inside a dataset. The greater the recall score, the better the machine learning model is at identifying both positive and negative samples. Recall Score = TP / (FN + TP)

C. F1 score

The F1 score, which is based on the accuracy and recall scores, serves as a representation of the model score. The F-score is a machine learning model efficiency metric that weighs precision and recall equally when determining how accurate the model is (it does not need that we know the whole number of observations). It serves as an alternative for accuracy measurements. It is typically used to provide summaries of the output quality of the model as a single value. When aiming to maximise accuracy or recall score, the performance of the model suffers, this is a useful model measurement to use. F1 Score = 2* Precision Score * Recall Score/ (Precision Score + Recall Score)

D. Accuracy

In machine learning, based on the input, or training, data, machine learning model accuracy is the statistic used to discover which model is best at recognizing correlations and patterns among variables in a dataset.

TABLE I. ACCURACY OF DIFFERENT CLASSIFIERS

Method	**precision**	**recall**	**f1-score**	**accuracy**
LinearSVC	**0.95**	**0.93**	**0.93**	**0.94**
Random Forest Classifier	0.88	0.84	0.86	0.85
Logistic Regression	0.94	0.92	0.93	0.93
XGBoost Classifier	0.93	0.88	0.90	0.90

VI. CONCLUSION

In today's culture, there is still much that needs to be done to prevent suicide. A crucial and successful method of preventing suicide is the early identification of suicidal thoughts. This paper examines current techniques for preventing suicidal thoughts from a comprehensive angle that includes techniques like textual content analysis. We have used LinearSVC, Random Forest Classifier, Logistic Regression, XGBoost Classifier algorithms, among them LinearSVC is best. This lead to the suggested proposed model achieving accuracy of 94% with 95% precision, 93% recall, and 94% F1-score using features and classifier approaches to training datasets.

REFERENCES

[1] Haque, Rezaul, Naimul Islam, Maidul Islam, and Md Manjurul Ahsan. "A comparative analysis on suicidal ideation detection using NLP, machine, and deep learning." *Technologies* 10, no. 3 (2022): 57.

[2] Metzler, Hannah, Hubert Baginski, Thomas Niederkrotenthaler, and David Garcia. "Detecting potentially harmful and protective suicide-related content on twitter: machine learning approach." *Journal of medical internet research* 24, no. 8 (2022): e34705.

[3] Huang, Yating, Chunyan Zhu, Yu Feng, Yifu Ji, Jingze Song, Kai Wang, and Fengqiong Yu. "Comparison of three machine learning models to predict suicidal ideation and depression among Chinese adolescents: A cross-sectional study." *Journal of Affective Disorders* 319 (2022): 221-228.

[4] Chatterjee, Moumita, Poulomi Samanta, Piyush Kumar, and Dhrubasish Sarkar. "Suicide Ideation Detection using Multiple Feature Analysis from Twitter Data." In *2022 IEEE Delhi Section Conference (DELCON)*, pp. 1-6. IEEE, 2022.

[5] Wang, Shirley B., Walter Dempsey, and Matthew K. Nock. "Machine learning for suicide prediction and prevention: advances, challenges, and future directions." In *Youth Suicide Prevention and Intervention*, pp. 21-28. Springer, Cham, 2022.

[6] Ji, Shaoxiong, Shirui Pan, Xue Li, Erik Cambria, Guodong Long, and Zi Huang. "Suicidal ideation detection: A review of machine learning methods and applications." *IEEE Transactions on Computational Social Systems* 8, no. 1 (2020): 214-226.

[7] van Mens, Kasper, C. W. M. de Schepper, Ben Wijnen, Saskia J. Koldijk, Hugo Schnack, Peter de Looff, Joran Lokkerbol et al. "Predicting future suicidal behaviour in young adults, with different machine learning techniques: A population-based longitudinal study." Journal of affective disorders 271 (2020): 169-177.

[8] Lekkas, Damien, Robert J. Klein, and Nicholas C. Jacobson. "Predicting acute suicidal ideation on Instagram using ensemble machine learning models." *Internet interventions* 25 (2021): 100424.

[9] Rajesh Kumar, E., K. V. S. N. Rama Rao, Soumya Ranjan Nayak, and Ramesh Chandra. "Suicidal ideation prediction in twitter data using machine learning techniques." *Journal of Interdisciplinary Mathematics* 23, no. 1 (2020): 117-125.

[10] Tadesse, Michael Mesfin, Hongfei Lin, Bo Xu, and Liang Yang. "Detection of suicide ideation in social media forums using deep learning." *Algorithms* 13, no. 1 (2019): 7.