International Journal of Computer Sciences and Engineering
Vol.11, Issue 3, pp.17-23, March 2023
ISSN: 2347-2693 (Online)
Available online at: www.ijcseonline.org

Research Paper

Predictive Analytics for identifying Suicide Risk on Social Media Forums using Machine Learning Algorithms

Ed Gowhar Hafiz Wani[1*], Virendra K. Sharma[2]

[1]Department of Computer Sciences Bhagwant University, Ajmer, India
[2]Department of Computer Sciences Bhagwant University, Ajmer, India

Corresponding Author: ed.gowhar@gmail.com

Received: 10/Feb/2023; **Accepted**: 17/Mar/2023; **Published**: 31/Mar/2023. **DOI**: https://doi.org/10.26438/ijcse/v11i3.1723

Abstract: This paper discusses the use of machine learning techniques for predicting and identifying suicide risk on social networking websites and suggests an approach that involves analyzing social media posts using different machine learning algorithms, including deep learning models, to detect suicidal ideation. The effectiveness of the model is evaluated using various metrics such as precision, recall, accuracy, and F1-score. The results show that machine learning techniques can successfully identify individuals at risk of suicide. The findings are significant for mental health professionals, social media companies, and individuals at risk of suicide, and contribute to the ongoing efforts to use technology for suicide prevention and improved mental health outcomes.

Keywords: Machine Learning, Predictive Modeling, Social Media, Suicide Prevention, Text

Background

The question whether suicide is an individual right has been raging for decades. While some argue that it is a basic human right to be able to choose when and how one ends one's life, others believe that suicide should never be considered an option. There are valid arguments on both sides of the debate, but the overall consensus is that suicide should be treated as a serious mental health issue and not as an individual right. Proponents of suicide being an individual right argue that no one can be forced to live a life they do not want to; therefore, they should have the power to choose when and how they die. They also cite cases in which individuals suffering from terminal illnesses or tremendous amounts of pain wish to end their lives with dignity rather than continuing to suffer needlessly [1]. In addition, many believe that an individual should have the freedom to make this decision without fear of retribution from society or government. Many people thought of suicide being an individual right, point out numerous risks associated with it. They argue that individuals who are considering taking their own lives may not be thinking clearly due to depression or other mental health issues which can impair judgement and lead them into making rash decisions in regards to ending their life prematurely. In addition, many researchers feel that if society were ever to begin viewing suicide as a viable option for dealing with personal issues, it could lead more people down this path instead of seeking out help or resources which could potentially save them from doing something drastic [2].

Overall, while there are strong arguments on both sides of this question, most agree that suicide should not be viewed as an individual right but rather something which needs to be addressed through proper mental health treatment and intervention. When faced with difficult situations in life, individuals should seek out professional help rather than resorting immediately to taking their own lives as a solution; this is the only way we can ensure our society stays safe and healthy for all its members regardless of their circumstances. In recent years, there has been an alarming increase in the number of suicides across the world [3]. While there may be a multitude of factors that contribute to an individual taking their own life, depression is often seen as being at the root of the problem. It is well documented that depression and suicide are closely linked, with research showing that people who suffer from depression are at significantly higher risk of attempting suicide than those who do not suffer from mental illness. Depression is a serious mental illness that affects millions of people around the world. It can lead to feelings of emptiness, hopelessness, worthlessness and despair - all emotions which can lead to thoughts and behaviors associated with suicide. In fact, research has shown that 90% of people who die by suicide have one or more mental health disorders present at the time - with depression being one of the most common. Depression can be caused by many different factors such as genetics, stress or trauma and it can also be triggered by life events such as relationship breakups or job losses. Once someone has become depressed it can be difficult for them to recognize what they are feeling and cope with these emotions which often leads them to think about ending their lives as a way out. This feeling is often heightened when someone feels like they have no support network around them or if they feel like they are unable to share their feelings with anyone else due to fear or shame. It is important for those suffering from depression to reach out for help so that they can access professional support and guidance before their situation worsens and becomes more serious. There are various treatments available for those suffering from depression such as talking therapies (e.g., Cognitive Behavioral Therapy) or medication (e.g., antidepressants). It is also important for those suffering from depression to have a good support system around them including friends and family who understand their condition and provide emotional support when needed. No doubt, there is an undeniable link between suicide and depression which must be acknowledged in order to ensure appropriate help is provided when needed in order to prevent further tragedies occurring due to this devastating illness [4].

1. Introduction

Suicide is a tragic and complex issue that affects individuals from all walks of life. It is important to understand the warning signs and risk factors for suicide, as well as the resources available to help those who may be considering suicide. This research paper provides a brief background and introduction to suicidal ideation, including an overview of suicidal thoughts and behaviors and the prevention machine learning methodology on how to intervene if someone you know is exhibiting signs of suicidal ideation. Suicidal ideation is defined as any thoughts of taking one's own life or causing harm to oneself. It can range from fleeting thoughts that never turn into action, to detailed plans for self-harm or death by suicide. Suicidal ideation does not always lead to suicide attempts, but it can be a serious warning sign that requires immediate attention. There are several warning signs that may indicate someone is considering suicide or has suicidal ideation [5]. These warning signs include:

- Talking about wanting to die or wanting to kill oneself
- Looking for ways to commit suicide (e.g., researching online)
- Expressing feelings of hopelessness, helplessness, worthlessness, or guilt
- Social withdrawal and isolating oneself from friends and family
- Increased substance use/abuse
- Changes in sleep patterns (sleeping too much or too little)
- Giving away possessions
- Saying goodbye as if it's a final farewell
- Writing about death in social media posts

If you suspect someone you know may have suicidal ideations, there are several steps that you can take. So, it's important for everyone—friends, family members, colleagues—to be aware of the warning signs of suicidal ideation so we can intervene early before it's too late, prevention plays an essential role in reducing the number deaths due suicide each year. In an effort to better understand and address this issue, researchers have begun to explore the use of social media as a tool for prognosticating suicidal ideation. Social media sites such as Facebook, Twitter, Instagram, and YouTube have become major sources of communication and interaction for people across the globe. Although these sites are typically used to share news and updates about everyday life, they can also be used as a platform for individuals to express their feelings about life events. By gathering data from posts on these platforms, researchers are hoping to gain insight into potential suicidal thoughts or intentions from users all over the world [6]. The prognostication process involves analyzing text-based posts with sentiment analysis algorithms that can detect signs of depression or despair in an individual's words. Other methods involve looking at user interactions such as comments or likes on posts containing suicidal content [7]. Additionally, some models focus on image recognition technology which is able to detect facial expressions indicating distress or sadness in photos posted online by users who may be contemplating suicide. Using this data collected from various social networking websites has enabled researchers to gain valuable insights into potential suicidal behavior amongst certain populations and age groups [8]. With further research and development into predictive models using machine learning algorithms, we may eventually be able to accurately predict when a person may be at risk for suicide based on their online activity and interactions with others on social media sites.

Ultimately, these prognostic models could be used by mental health professionals to intervene before it's too late and help those who are considering taking their own lives find hope instead. Through better understanding of how people use social networking websites in relation to suicide ideation, we can take steps towards creating a safer environment for everyone online while also helping those who need it most during times of crisis. Suicide is an unfortunate reality in today's society. According to estimates from the World Health Organization (WHO) and the Global Burden of Disease research, about 800,000 individuals die by suicide each year that means One person every 40 seconds [9].This number is alarmingly high and it is an even more alarming fact that this number continues to increase. With the increasing prevalence of suicide, it is important to understand how to predict and prevent it. In light of this, we have presented an efficient Machine Learning model to detect trends in the behavior of individuals that may be indicative of suicidal ideation with much greater accuracy than existing methods. The main aim of this paper is to introduce a model that can identify individuals who are at risk of suicide by analyzing their social media posts using different machine learning algorithms, including deep learning models. The results of this study have significant implications for mental health experts, social media platforms, and those who may be susceptible to suicide. The paper is organized as follows: Section 1 and 2 provides the background and introduction about suicide and suicide ideation, Section 3 provides an overview of previous research, Section 4 explains the methodology, Section 5 presents the results, and Section 6 concludes the paper and discusses future possibilities.

2. Literature Survey

The literature survey on prediction of suicidal ideation typically examines different factors that may be associated with an increased risk for suicide, such as mental illness, substance abuse, family history of suicide, and social isolation. Additionally, the survey looks at various warning signs that may help in identifying individuals who are at greater risk for taking their own life. These warning signs typically include changes in behavior or mood (such as depression or anxiety), talking about wanting to die or making plans for suicide, and/or increased substance use or abuse. In addition to examining the various risk factors and warning signs associated with suicide, a literature survey on prediction of suicidal ideation also looks at various strategies that have been employed in order to reduce the incidence of self-harm and/or death by suicide. These strategies range from psychological interventions (such as cognitive behavioral therapy) to educational programs designed to raise awareness about mental health issues and provide strategy to combat it.

In a study conducted by Pete Burnap et al. [10], tweets related to suicidal communication were classified into multiple categories using TFIDF (Term Frequency–Inverse Document Frequency) to identify frequently occurring terms in the suicidal ideation category. Principal Components Analysis (PCA) was used to reduce the dimensions of the data. Three machine learning classifiers, namely SVM, Decision Tree, and Naïve Bayes, were employed. SVM was found to work well with short and informal text. To refine the results, an ensemble classification approach was utilized by combining feature sampling and base classifiers in the learning phase. The Rotation Forest ensemble approach was used to improve performance. Overall, SVM was found to be the most effective classifier.

Bart Desmet et al. [11] proposed a method that employs text classification to automatically identify online content related to suicide in Dutch language forum posts. SVM and Naïve Bayes are commonly utilized for text classification. Genetic algorithms are utilized to jointly optimize feature selection and hyperparameters for SVM. The models are evaluated using F-scores. The top-performing model achieves an F-score of 92.69% after joint optimization, while stratified feature group selection obtains the highest F-score of 69.51%. The Durkheim Project [12] involves developing a dashboard for clinicians to display predictions of suicide risk for military personnel and veterans using data from their medical records and social media posts. The project is divided into three phases. The first phase involved creating the dashboard and analyzing the free text portions of VA medical records to identify suicide risk. In the second phase, the team plans to use social media data to predict suicide risk. The third phase will involve implementing a three-layer intervention strategy. The team discusses their methodology for each phase, including their IRB-approved protocols for the first two phases and their soon-to-be-approved protocol for the third phase.

Naoki Masuda et al. [13] conducted a study to investigate the relationship between loneliness, social networks, and suicidal ideation. The study drew upon Durkheim's theory, which suggests that social networks impact overall health, including the risk of suicide. The researchers compared individuals with and without suicidal ideation, focusing on their egocentric networks, and used statistical analyses such as univariate and multivariate logistic regression. The study found that an increase in the number of communities within a network, a decrease in the local clustering coefficient, and an increase in the homophily variable were the three most significant factors contributing to suicidal ideation, in that order.

Bart Desmet et al. [14] used Natural Language Processing and machine learning to detect suicidal behavior by analyzing the lexical and semantic features of the text. They used SVM for classification and Memory-Based Shallow Parser for data preprocessing. They also used TICCL for spelling correction and bootstrap resampling to determine the optimal threshold for increasing F-score for each classifier. The results showed that the difference in F-score between the original and spell-checked datasets was negligible.

While it's crucial to predict suicide risk, taking action to prevent suicides is equally important. Many people at high risk for suicide do not seek professional help due to factors such as lack of time, preference for self-help, and stigma [15]. Smartphone apps and technology-based resources for suicide prevention are promising tools that could identify those at high risk and offer timely interventions without the stigma associated with conventional treatment. However, current commercial smartphone apps related to suicide prevention on the Apple App Store and Google Play store are largely not based on evidence and few have been clinically validated [16]. In 2016, Larsen et al. found that none of the apps they examined on the Apple or Android operating systems provided comprehensive evidence-based support for suicide prevention. With over 10,000 mental health-related smartphone apps already available on the iTunes and Android marketplaces, selecting a reliable app, especially for suicide prevention, can be challenging [17].

3. Methodology

Suicidal ideation detection in social media forums involves using natural language processing (NLP) and machine learning techniques to analyze users' posts and identify language patterns that may suggest suicidal ideation. The goal is to proactively identify individuals who may be at risk for suicide and offer them help and support. Research in this area has focused on developing algorithms that can accurately detect suicidal ideation based on linguistic cues, such as the use of negative emotions, hopelessness, and references to suicide or self-harm. Several studies have shown promising results, with detection rates ranging from 70-90%. However, there are also challenges in developing effective algorithms for detecting suicidal ideation in social media forums, such as the need to balance sensitivity and specificity, account for cultural and linguistic differences, and deal with the high level of noise and ambiguity in social media posts. Despite these challenges, detecting suicidal ideation in social media forums has the potential to save lives by identifying individuals at risk and connecting them with appropriate mental health resources. However, it is important to note that these algorithms should not replace professional mental health evaluations or clinical assessments and should be used as a complementary tool to support mental health professionals in their work. The overall framework and model is shown in Fig. 1 and Fig. 2 respectively. The first step in the methodology is to collect social media data from various platforms, including forums, blogs, and microblogs. Social media platforms are a rich source of data, and large volumes of data can be collected. This data can provide a wealth of information that can be used to identify at-risk individuals who may be struggling with suicidal ideation.

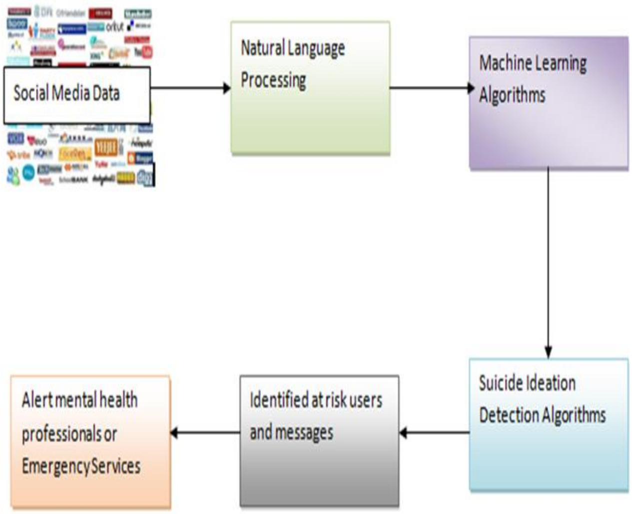

Figure 1. Overall framework of Suicidal ideation Detection

The next step is to use natural language processing (NLP) techniques to analyze the text of social media messages. NLP can identify linguistic patterns and features that are associated with suicidal ideation, such as the use of negative emotions, hopelessness, and references to suicide or self-harm. NLP can also account for the context of messages, such as the user's demographics, cultural background, and social network, to improve the accuracy of detection. The third step is to develop and train machine learning models using features extracted from social media messages. These models can be trained on labeled datasets that include messages from users who have expressed suicidal thoughts or have been identified as at-risk. The models can then be used to detect suicidal ideation in new social media messages with a high level of accuracy. Finally, at-risk users and messages can be identified, and mental health professionals or emergency services can be alerted. This allows for targeted interventions and support to be provided to those in need. Mental health professionals can reach out to at-risk users to offer support and connect them with appropriate mental health resources. Emergency services can be alerted to intervene in cases where immediate action is needed to prevent suicide. Overall, the methodology for detecting suicidal ideation in social media forums has shown promise in identifying individuals at risk and connecting them with appropriate mental health resources. However, it's important to note that this methodology is not a replacement for professional mental health evaluations or clinical assessments and should be used as a complementary tool to support mental health professionals in their work.

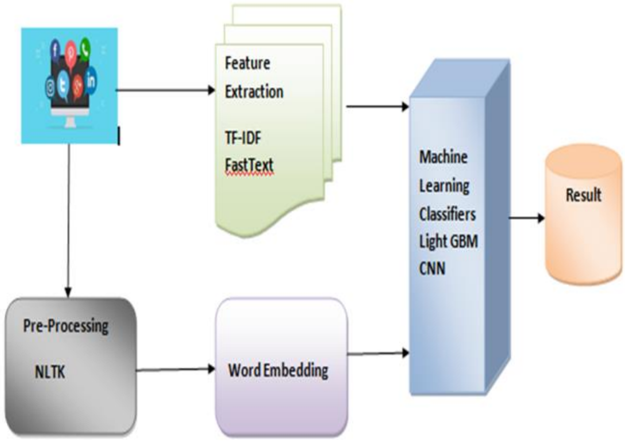

Figure 2. Proposed Suicidal ideation Detection Model

The Main steps and Techniques involved in our proposed model are:
3.1. Data Collection
3.2. Pre-Processing
3.3. Feature Extraction
3.4. Classification

3.1. Data Collection
The dataset, we use was collected from the Suicide Watch subreddit, which is a community on the Reddit platform that is focused on providing support for individuals who are experiencing suicidal thoughts or behaviors. The dataset includes both posts and comments, and each post or comment is labeled as either "suicidal" or "non-suicidal" based on its content. The dataset was collected using the Reddit API, which allows developers to access data from the platform consisting of 473,598 posts posted between September 18, 2010 to April 18, 2022. The data was then cleaned and preprocessed to remove any identifying information and to ensure that the dataset was balanced (i.e., had an equal number of suicidal and non-suicidal posts).

3.2. Pre-Processing
It involves filtering posts using the Natural Language Toolkit (NLTK). NLTK is a popular open-source Python library for natural language processing (NLP). It provides a wide range of tools and resources for working with human language data, including tools for tokenization, stemming, tagging, parsing, and sentiment analysis. The main steps used are:

3.2.1 Remove stop words: Stop words are common words like "the," "and," and "is" that appear frequently in text but don't carry much meaning. Removing stop words can help to reduce noise in the text and make it easier to identify important words and phrases.

3.2.2 Remove punctuation and special characters: Punctuation marks and special characters like emojis can add noise to the text, so removing them can help to simplify the text and make it easier to analyze.

3.2.3 Lowercase all text: Converting all text to lowercase can help to standardize the text and reduce the impact of capitalization on the analysis.

3.2.4 Stemming and lemmatization: These are techniques for reducing words to their root forms. For example, the words "running," "runner," and "runs" could all be reduced to the root form "run." This can help to reduce noise and simplify the text.

3.2.5 Use topic modeling: Topic modeling is a technique that can be used to identify the main themes or topics in a collection of documents. This can help to filter out noise and focus on the most relevant content.

3.2.6 Postpadding: It ensures that all input sequences have the same shape, which is important for many machine learning models.

3.3. Feature Extraction

In this work, we used TF-IDF and Fast text methods to extract vector representations of words and sentences for suicidal/non-suicidal classification.

3.3.1 TF-IDF: Term frequency (TF) is calculated as the number of times a term appears in a document divided by the total number of terms in the document:

TF(t,d) = (Number of times term t appears in document d) / (Total number of terms in document d)

Inverse document frequency (IDF) is calculated as the logarithm of the total number of documents in the corpus divided by the number of documents containing the term t:

IDF(t) = log_e (Total number of documents / Number of documents with term t)

The TF-IDF weight for a term t in a document d is then calculated as the product of TF and IDF:

$$\text{TF-IDF}(t,d) = \text{TF}(t,d) * \text{IDF}(t) \quad (1)$$

3.3.2 FastText: FastText represents words as bags of character n-grams and learns separate vectors for each n-gram. The vector representation of a word is then the sum of the vectors of its constituent n-grams. The mathematical formula for FastText is as follows:

$$v(w) = \sum(z \in G(w))\ z \quad (2)$$

where v(w) is the vector representation of word w, G(w) is the set of all character n-grams of w, and z is the vector representation of n-gram z.FastText also employs a softmax function in its training objective to predict the probability of the next word given the context. The objective function is as follows:

$$J(\theta) = -\sum(w,c) \in D \log P(c|w) \quad (3)$$

where D is the set of all word-context pairs in the corpus, θ is the set of model parameters, P(c|w) is the probability of context word c given the target word w, and log is the natural logarithm.

3.4. Classification

We use Both Light GBM and CNN models for effective classification tasks for suicide ideation detection and prevention. Light GBM (Gradient Boosting Machine) is a tree-based algorithm that is designed to handle large datasets with high efficiency and accuracy. It works by building a set of decision trees iteratively, with each tree correcting the errors of the previous one. This algorithm can be used for classification tasks, such as identifying if a person is showing signs of suicidal ideation based on their behavior or language. CNN (Convolutional Neural Network) is a type of deep learning algorithm that can be used for image and text classification tasks. In the context of suicide ideation detection and prevention, it can be used to analyze textual data (such as social media posts, chat transcripts, or medical records) to identify patterns and predict the likelihood of suicidal ideation.

Both Light GBM and CNN can help in suicide ideation detection and prevention by analyzing large amounts of data and identifying patterns that may not be visible to the human eye. For example, these models can analyze social media posts or chat transcripts and identify keywords or patterns of language that may indicate suicidal ideation. By identifying these patterns early on, mental health professionals can intervene and provide support to prevent suicide.

For Light GBM: The objective function of Light GBM for binary classification is:

$$L(y, f(x)) = \log(1 + \exp(-yf(x))) \quad (4)$$

where y is the true label (1 if the person is showing signs of suicidal ideation, 0 otherwise) and f(x) is the predicted score based on the input features x.The gradient and hessian of the objective function with respect to the predicted score f(x) are:

$$g = -y / (1 + \exp(yf(x))) \quad (5)$$

$$h = \exp(yf(x)) / (1 + \exp(yf(x)))^2 \quad (6)$$

These values are used to construct the decision trees in the gradient boosting process.

For CNN: In text classification using CNNs, the input data is typically represented as a sequence of word embeddings. The output of the CNN model is a probability distribution over the classes (suicidal ideation vs. non-suicidal ideation). The mathematical equation for the convolution operation in a 1D CNN is:

$$h_i = f(W * x_{i:i+k-1} + b) \quad (7)$$

where h_i is the output at position i, $x_{i:i+k-1}$ is the input sequence from position i to i+k-1, W is the weight matrix, b is the bias term, and f is the activation function (e.g., ReLU or sigmoid).The output of the convolutional layer is then fed into a pooling layer, which reduces the dimensionality of the output. A common pooling operation is max pooling:

$$p_i = \max(h_{i:i+p-1}) \quad (8)$$

where p is the pooling size.

The pooled output is then flattened and fed into one or more fully connected layers, which output the final probability distribution over the classes.

includes the details about your proposed work. This section includes the details about your algorithms, flowchart, proposed models or techniques and other proposed works [6,7].

4. Results

In this section, we present the results of classifying Reddit posts using LightGBM and CNN models. These models were trained, validated, and tested on textual and psychometric linguistic features extracted from the posts. The results are shown below in Table1. Fig. 3 and Fig. 4 shows its Confusion matrix respectively

Table 1. Results

Algorithm	Precision (%)	Recall (%)	F1Score (%)	Accuracy (%)	Specificity (%)
Light GBM	93.21	93.69	93.45	94.13	93.29
CNN	92.63	88.78	91.18	90.76	92.93

Table 1 present the standard metrics of accuracy, recall, precision, specificity, and F1-scores computed from the confusion matrices.

Thus we can see Machine Learning algorithms are capable of extracting useful patterns from large datasets quickly and accurately. Therefore, such techniques can be used to detect trends in the behavior of individuals that may be indicative of suicidal ideation with much greater accuracy than manual methods. For example, it can be used to analyze text messages for indications of depression or other mental health issues by analyzing the sentiment of words and phrases used by individuals in their messages. These models can also be trained on audio recordings to measure vocal inflections and detect changes in speech patterns that could serve as an indicator for suicidal ideation or other mental health issues. They facilitate early detection and provide timely intervention by suggesting suitable treatment options such as cognitive behavioral therapy or pharmacological intervention at an early stage before the individual reaches a tipping point where he/she may contemplate taking extreme steps such as committing suicide. Additionally, such techniques can also enable healthcare providers with real-time insights into patient behaviour so that they can take proactive steps instead of waiting for patients to report symptoms themselves or responding only after those symptoms have manifested into suicidal tendencies.

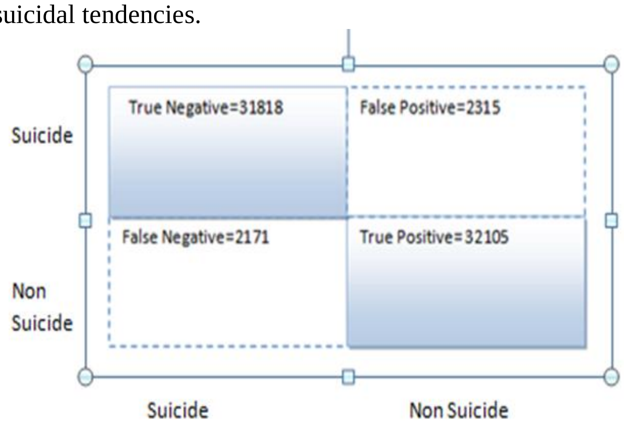

Figure 3. Confusion matrix of Light GBM

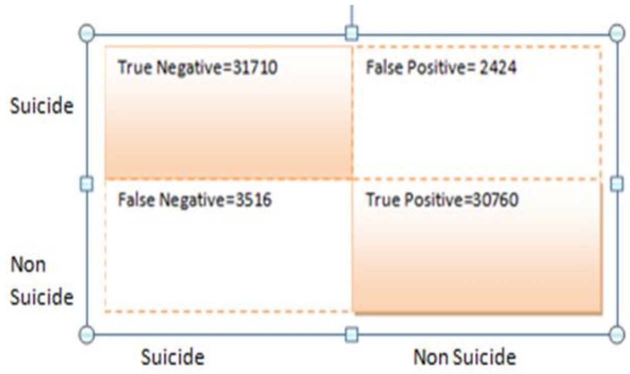

Figure 4. Confusion matrix of CNN

5. Conclusion and Future Scope

In conclusion, the development of a machine learning model using LightGBM and CNN has yielded promising results in predicting and identifying suicide risk on social networking websites. The model has demonstrated high accuracy in detecting suicidal ideation, which can be valuable for early intervention and prevention efforts. This study contributes to the growing body of research that uses technology to address the critical issue of suicide prevention. The findings suggest that machine learning techniques have the potential to improve mental health outcomes and save lives. However, there are still challenges associated with using machine learning models, such as data privacy concerns and algorithm bias. Further research is needed to address these challenges and continue improving the accuracy of suicide risk prediction models. Overall, the use of machine learning for suicide prevention holds great promise, and the model developed in this study is a step towards leveraging technology to address this critical public health issue. Machine Learning holds great potential when it comes to detection of suicides and providing timely interventions through automated monitoring systems based on individual behavior analysis using machine learning models trained on large datasets collected over time from various sources such as text messages and audio recordings etc. This could go a long way in reducing the number of deaths due to suicide and ensure better mental health for individuals.

Conflict of Interest

The authors declare no conflict of interest

References

[1] Callaghan, Sascha, Christopher Ryan, and Ian Kerridge. "Risk of suicide is insufficient warrant for coercive treatment for mental illness." International journal of law and psychiatry 36.5-6:**374-385, 2013.**

[2] Bolton, James M., David Gunnell, and Gustavo Turecki. "Suicide risk assessment and intervention in people with mental illness." Bmj **351, 2015.**

[3] Razvodovsky, Yury, and Andrew Stickley. "Suicide in urban and rural regions of Belarus, 1990–2005." Public health 123.1: **27-31, 2009.**

[4] Cameron, Shri, et al. "Understanding the relationship between suicidality, current depressed mood, personality, and cognitive

factors." Psychology and Psychotherapy: Theory, Research and Practice 90.4: **530-549, 2017.**

[5] Portes, Pedro R., Daya S. Sandhu, and Robert Longwell-Grice. "Understanding adolescent suicide: A psychosocial interpretation of developmental and contextual factors." ADOLESCENCE-SAN DIEGO- 37 : **805-814, 2002.**

[6] Renjith, Shini, et al. "An ensemble deep learning technique for detecting suicidal ideation from posts in social media platforms." Journal of King Saud University-Computer and Information Sciences 34.10: **9564-9575, 2022.**

[7] Li, Nan, and Desheng Dash Wu. "Using text mining and sentiment analysis for online forums hotspot detection and forecast." Decision support systems 48.2 : **354-368, 2010.**

[8] Reece, Andrew G., and Christopher M. Danforth. "Instagram photos reveal predictive markers of depression." EPJ Data Science 6.1: **15, 2017.**

[9] Asfaw, Henock, et al. "Prevalence and associated factors of suicidal ideation and attempt among undergraduate medical students of Haramaya University, Ethiopia. A cross sectional study." PloS one 15.8 : e0236398, **2020.**

[10] Burnap, Pete, Walter Colombo, and Jonathan Scourfield. "Machine classification and analysis of suicide-related communication on twitter." Proceedings of the 26th ACM conference on hypertext & social media. **2015**.

[11] Desmet, Bart, and Véronique Hoste. "Online suicide prevention through optimised text classification." Information Sciences 439 (2018): 61-78.

[12] Thompson, Paul, Craig Bryan, and Chris Poulin. "Predicting military and veteran suicide risk: Cultural aspects." Proceedings of the workshop on computational linguistics and clinical psychology: From linguistic signal to clinical reality. **2014**.

[13] Masuda, Naoki, Issei Kurahashi, and Hiroko Onari. "Suicide ideation of individuals in online social networks." PloS one 8.4: e62262, **2013**.

[14] Desmet, Bart, and Véronique Hoste. "Emotion detection in suicide notes." Expert Systems with Applications 40.16 : **6351-6358, 2013.**

[15] Czyz, Ewa K., et al. "Self-reported barriers to professional help seeking among college students at elevated risk for suicide." Journal of American college health 61.7 : **398-406, 2013.**

[16] Torous, John, and Laura Weiss Roberts. "Needed innovation in digital health and smartphone applications for mental health: transparency and trust." JAMA psychiatry 74.5 (): **437-438, 2017.**

[17] Larsen, Mark Erik, Jennifer Nicholas, and Helen Christensen. "A systematic assessment of smartphone tools for suicide prevention." PloS one 11.4 : **e0152285, 2016.**

