Egyptian Informatics Journal 24 (2023) 291–302

Contents lists available at ScienceDirect

Egyptian Informatics Journal

journal homepage: www.sciencedirect.com

Full length article

Detecting suicidality on social media: Machine learning at rescue

Syed Tanzeel Rabani [a], Akib Mohi Ud Din Khanday [a], Qamar Rayees Khan [a], Umar Ayoub Hajam [a], Ali Shariq Imran [b,*], Zenun Kastrati [c]

[a] Department of Computer Science, BGSBU, Rajouri, India
[b] Department of Computer Science (IDI), Norwegian University of Science and Technology (NTNU), Norway
[c] Department of Informatics, Linnaeus University (LNU), Sweden

ARTICLE INFO

Article history:
Received 6 January 2023
Revised 3 April 2023
Accepted 7 April 2023
Available online 19 April 2023

Keywords:
Suicidal ideation
Social media
Feature engineering
Machine learning
Ensemble learning

ABSTRACT

The rise in technological advancements and Social Networking Sites (SNS) made people more engaged in their virtual lives. Research has revealed that people feel more comfortable posting their feelings, including suicidal thoughts, on SNS than discussing them through face-to-face settings due to the social stigma associated with mental health. This research study aims to develop a multi-class machine learning classifier for identifying suicidal risk levels in social media posts. The proposed Enhanced Feature Engineering Approach for Suicidal Risk Identification (EFASRI) is used to extract features from a novel dataset collected from Twitter and Reddit platforms. Three machine learning algorithms, i.e. Support Vector Machine (SVM), Random Forest (RF) and Extreme Gradient Boosting (XGB) were employed for classification. The study demonstrates significant improvements in the precision, recall, and overall accuracy compared to previous research that used classical feature extraction mechanisms. The best-performing algorithm, Extreme Gradient Boosting (XGB), achieved an overall accuracy of 96.33%. The findings imply that different features contain different levels of information, and the right combination of the features supplied to the machine learning algorithms may improve the prediction results.

1. Introduction

A detailed statistic provided by World Health Organisation (WHO) mentions that, on average, suicide occurs every forty seconds [1]. The report further says that approximately 800,000 people commit suicide every year, and the number of suicide attempts is 20 times more than a completed suicide [2]. However, compared to the deaths due to various ailments, suicide is much underreported. Thus, the global estimate of suicide mortality is approximated to one million deaths per year [3]. The above statistics also reveal that suicide is the leading cause of death among youngsters, particularly among women. Suicide is not about an "all or nothing" situation. The famous book related to suicide [4] indicates that it follows a proper process and pattern wherein suicidal ideation takes the first place followed by a suicidal attempt which then matures into the completed suicide. Suicidal ideation does not always lead to suicide, but it poses a significant threat to individuals who may then attempt suicide. Suicidality gets noticed when one often talks about it with his/her caretakers or when psychiatrists/psychologists interact with the individuals and enquire about their thinking and mood. By analyzing various warning signs, caregivers can uncover the risk factors associated with suicidality and take the necessary steps for prevention. American Foundation for Suicide Prevention [AFSP] [5] identified various risk factors and warning signs related to suicidality to help potential suicidal individuals. They categorized the risk factor into three main classes. These factors are related to health (mental health, persistent pain), environment (stress, molestation, etc.) and family history (previous attempts of suicide etc.). The National Institutes of Health (NIH) further lists some of the indicators/warning signs of suicidal ideation, as shown in Fig. 1.

Many suicidal deaths can be prevented by understanding how people communicate their distressrelated thoughts. Early understanding of the risk factors and warning signs can decrease the threshold for suicide and help prevent many deaths. However,

* Corresponding author.
E-mail address: ali.imran@ntnu.no (A.S. Imran).

Peer review under responsibility of Faculty of Computers and Information, Cairo University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.eij.2023.04.003
1110-8665/

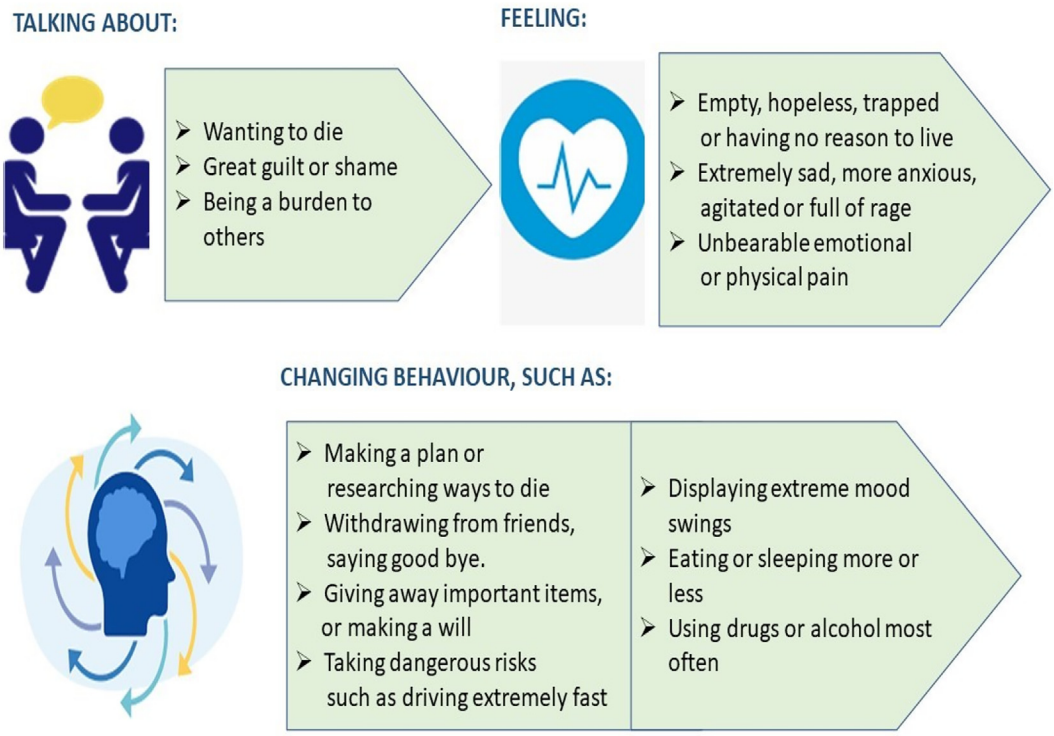

Fig. 1. Warning Signs of Suicide.

the issue with suicidality is that people usually don't cooperate with clinicians due to the social stigma associated with mental illness. As stigma plays a deadly role in suicidality, clinical interventions for at-risk individuals at a large scale become almost impossible. It is reported that 36% of individuals who die due to suicide leave a note behind [6]. Researchers found that suicide notes indicate that when the suicidal attempt of any individual fails, there is a high probability that they will still go for another attempt with more accuracy [7]. It is also found that these suicide notes talk more about shame and apology suggest if any alternative was there, they might have readily accepted that. However, most of the suicide notes are found when suicide is completed or at least attempted [7]. Many efforts were made to screen the patients, but it was a challenging task to counsel the individuals to come for evaluation in a stigmatized society [8,9]. Research has revealed that people feel more comfortable discussing their daytoday happenings on Online Social Networks (OSN) without worrying about social stigma [10]. Moreover, recent research has also found that monitoring social media can provide an alternative and excellent opportunity for uncovering the warning signs associated with the posts of at-risk individuals suffering from suicidality [11–13]. Thus social media, if appropriately mined, can act as a tool to prevent potential suicidal victims from taking the extreme step and also offer the necessary support. Moreover, the histories of persons on SNS who commit suicide can help understand the risk factors associated with suicide. Researchers have put a lot of effort into identifying the patterns in the language of social media [14,15] including mental health [16,17]. Some studies trained machine learning models to separate suicidal content from non-suicidal content [16,18]. But the model that will separate the content reflecting suicidal ideation from depression and other low-risk posts requires a highlevel feature engineering mechanism. The need for such a feature engineering mechanism is that the posts reflecting suicidal ideation and other stress-related content seem very identical in its language. In this paper, significant attention has been made to employing novel hybrid feature engineering mechanisms that will be used to train the machine learning models to help differentiate the highrisk posts from other categories. The major contributions that this research article makes to the already existing literature are as follows:

1. Novel dataset is collected from Twitter and Reddit. The dataset was annotated with the scheme developed in consultation with psychiatrists and psychologists.
2. We developed a methodology that could help differentiate the posts into three categories of risk: high risk, moderate risk, and no risk.
3. We developed a novel hybrid feature engineering mechanism for the extraction of the most relevant features.
4. We trained three learning algorithms, viz., support vector machine (SVM), random forest (RF), and extreme gradient boosting (XGB), with the features extracted through our proposed feature engineering mechanism.

The rest of our article is divided into various sections. Section 2 discusses the related work done in the domain of the detection and prevention of suicidality. Section 3 discusses the proposed methodology, which contains various subsections: data collection, data pre-processing, proposed novel feature engineering mechanism and finally, training of machine learning algorithms. Section 4 elaborates on the results and discussion. Section 5 compares our work with the previous research. Section 6 provides the conclusion and future scope of the work.

2. Literature Review

Mental health issues like depression and suicidality have usually been examined through psychological battery tests and clinical procedures [6,19]. The stigma associated with mental illness made researchers move towards informal sources like social media to understand language patterns of suicidal posts that could enhance the interpretation of suicidal ideation in a better way.

2.1. An overview of research on questionnaires, topic analysis, and machine learning techniques

The study by [11,20] indicated that questionaries play a vital role in detecting the mental state of a patient. However, due to the boom in social networking sites, people feel comfortable and express their feelings freely on social media. Various studies [20,21] discuss several scales used to predict depression on social media. More studies [22,23] were conducted to analyse the topics, which potential suicidal individuals usually like to discuss on social media, and the behaviour of suicidal individuals was also evaluated through these studies. The advancement in natural language rocessing and machine learning techniques has made it possible to process the semantic information from these social media posts to extract the various features, which can help automate the prediction of suicidal content [24,21]. Most of the research focussed on using binary classification mechanisms for identifying suicidal content through various popular algorithms like Support vector machine, Decision trees and ensemble learning algorithms [25–28]. Deep learning methods have also been used to help predict suicidal ideation [8].

2.2. Feature extraction mechanism for detecting suicidal ideation

One of the most important parts of building a machine learning model is to find the most relevant features that can help in differentiating suicidal content from non-suicidal content [18]. Alada et al. [9] did a sentimental analysis of Reddit posts and differentiated the suicidal content from non-suicidal ones using features extracted through term frequency-inverse document frequency and word count. O'Dea et al., [27] trained the classifier that helped automate and replicate the accuracy of human coders. Feature extraction was performed using unigrams, TF-IDF and Filter based techniques. Vioules et al., [29] introduced a new approach that measures the warning signs of suicide, detects the posts that contain content related to suicide, and automatically identifies the sudden changes in user behavior. The researchers developed the behavioral features to measure the level of risk of a person concerned with his online behavior on Twitter. Two groups of behavioral features, namely user-centric features and post-centric features were established. Chadha et al., [30] discuss about the performance evaluation of various machine learning algorithms for identifying and differentiating suicidal content from non-suicidal content on Twitter. The researchers manually selected the 112 features through a survey involving the doctors and patients of a psychiatric hospital. BOW and TFIDF weighting schemes were used to discard the irrelevant features. Abboute et al. [31] developed a list of terms for lexical feature extraction using nine suicidal topics Okhapkina et al. [32] also developed a lexicon of terms using term-frequency inverse document frequency technique that helped in differentiating the suicidal content from non-suicidal ones. In a study of suicide on Weibo, Cheng [33] largely employed Simplified Chinese Language Query and Word Count (SC-LIWC) to count how many times each type of word appeared in users' postings. He then used logistic regression to look at the relationship between SC-LIWC features and five suicide risk factors.

Very few studies used a hybrid feature model where features extracted through different techniques were combined and supplied to machine learning techniques for better prediction. Sawhney et al. [18]collected numerous sets of features such as statistical features, Linguistic Inquiry and Word Count (LIWC) features, term frequency-inverse document frequency, and topics probability features. These features proved to be successful for the binary classification of tweets with stable precision and recall. Shing et al. [34] also retrieved numerous features using the technique of Bag of words, topic modeling, and linguistic inquiry and word count. Mbarek et al. [35]developed a suicide user profile detection model using a set of features that included a linguistic, emotional, facial, timeline, and public features. Mbarek et al. [36], in their other study, applied different machine learning algorithms to solve the suicidal user prediction problem using a rich set of features like Emotional features, Temporal features and Account features that have been effective in detecting suicidal users. The feasibility of their method was studied on people who committed suicide, and the results were shown to be in line with expectations.

2.3. Critical review of the existing literature

Detection of suicidal ideation through social media is a growing area of research, with efforts directed towards building an intelligent mechanism through training classifiers using various algorithms and features. The ensemble approach has shown promise in improving prediction results by overcoming overfitting. However, current research has primarily focused on training classifiers on smaller datasets and tuning various parameters, with less attention given to feature engineering. Moreover, Previous studies have mainly employed feature extraction techniques such as term-frequency inverse document frequency and Bag of Words. In addition, various feature selection techniques such as univariate selection, feature importance, and correlation matrix have been used to eliminate irrelevant attributes. In contrast, our study emphasizes building a large real-world suicide dataset, utilizing a hybrid feature engineering mechanism to extract relevant features and training machine-learning models using these features.} The most relevant articles related to our work, along with their contributions and the features used are summarised in Table 1.

The literature survey indicated that work done in the field of data mining for predicting suicidal ideation on social media and its prevention is minimal that needs a lot of effort. Data scarcity is also a big problem due to the privacy and ethical issues related to this research. Moreover, the above literature mostly focuses on binary classification and uses ordinary feature extraction techniques. Our work made a novel effort to collect big data related to suicidal tweets using the API of Twitter and Reddit and also focused on a feature extraction mechanism for the extraction of rich features. Then three machine learning algorithms were trained to classify the tweets into three classes of distress using the methodology discussed in Section 3.

3. Methodology

The methodology adopted in this research article for identifying and classifying social media posts into three levels of concern consists of four major steps as shown in Fig. 2. In the first step, relevant data is extracted from SNS. The second step is about annotating the posts into three levels of risk based upon the annotation scheme that is devised in consultation with mental health experts. The third step involves preprocessing the posts to remove irrelevant and redundant information and proposing the feature extraction mechanism to extract the relevant features for training the multi-class machine learning model. The last step is about classifying the posts and evaluating the model using different metrics.

3.1. Data Collection and Exploration

The data was collected from two famous SNS: Twitter and Reddit through their APIs. We neither collected any identifiable human data from the social posts nor saved any such information. A random identifier was assigned to each post. Twitter API was used to collect the tweets using the phrases or words as used in previous research [11,14,26] and other words suggested by the mental

Table 1
Most Relevant and Recent Work Related to the Detection of Suicidal Ideation on Social Media.

Study	Contributions	Data Used	Features Used
[27]	Developed Binary classifier for suicide detection using various Machine learning techniques	Twitter	TFIDF,BOW
[37]	Developed Multi-class machine learning classifier for classification of suicide related communication on Twitter	Twitter	TFIDF
[38]	Developed Binary classifier for suicide detection using Machine and ensemble methods.	Twitter	Manual features (34 known keywords used by suicidal persons)
[18]	Develop and design new features to improve classification of suicidal content.	Twitter	LIWC Features, Topics, TFIDF and part of speech
[30]	Developed machine learning classifier for suicidal identification.	Twitter	Manual features (Suicidal Keywords),TFIDF and BOW
[29]	Quantification of suicidal warning signs for detection of distress and suicide related content.	Twitter	Textual and Behavioural features
[39]	Prediction of suicidal ideation using Deep learning and Machine learning models.	Reddit	TFIDF and Word2Vec

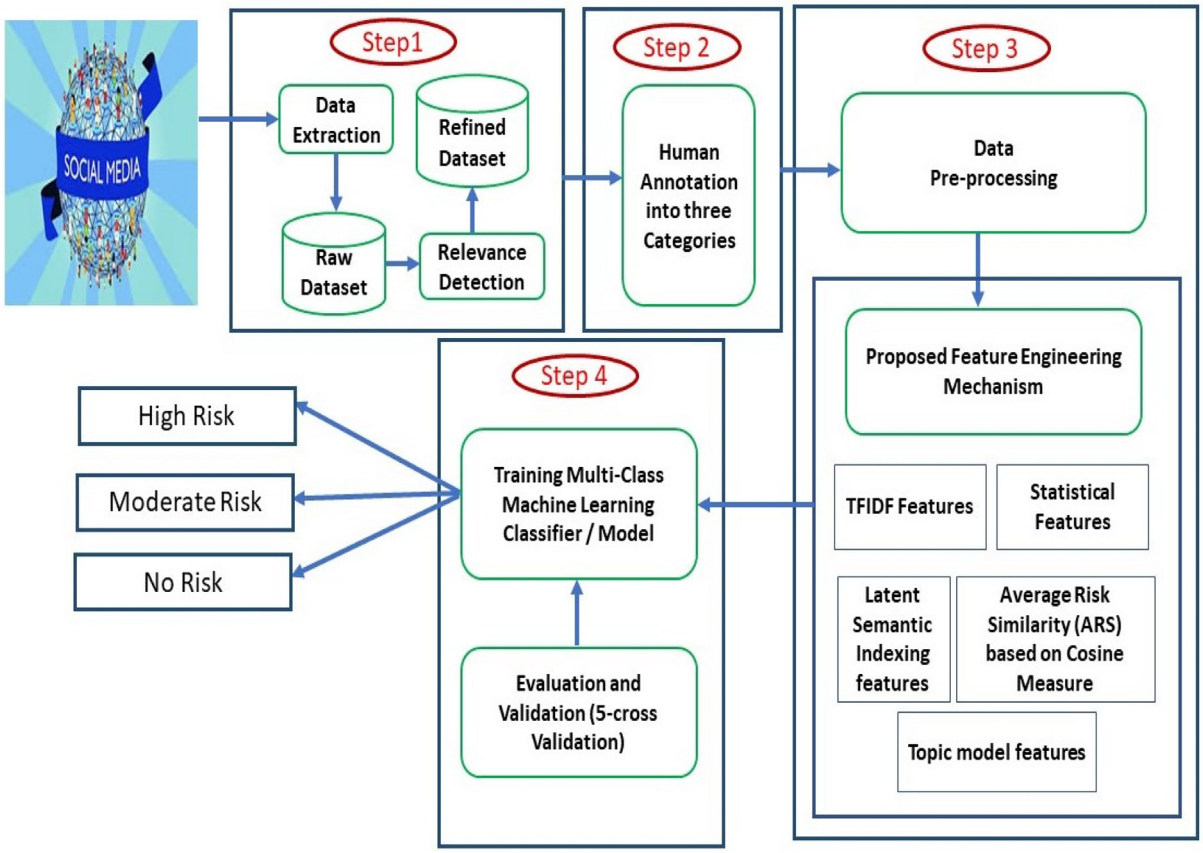

Fig. 2. Proposed methodology for suicidality detection.

health experts over the period November 8, 2019, to February 26, 2022. Some of those keywords for extraction of data from Twitter are as listed: "want to die; tired of myself; ending my life; be dead; suicidal; feeling empty; feeling suicidal; feeling alone; feel anxious; I feel helpless; unworthy life; suicide plan; cutting my wrist; Insomnia; fucking life; depression; pills depressed; diagonised schizophrenia; diagonised bipolar; MDD; never wake up; better off dead; go to sleep forever; tired of living; We also studied two subreddits, r/SuicideWatch, and r/depression. To extract the suicidal posts -and also those posts that don't directly indicate suicidality but point to any of the risk factors related to the suicide, we selected postings from r/SuicideWatch and r/depression. We expected that all r/SuicideWatch entries were suicidal and that the vocabulary of r/depression would be the most comparable to that of r/SuicideWatch as indicated by the previous research [30,31], when several control subreddits were utilized for language comparison, implying that subreddits can be used as labels also. We used the API of Reddit and Python Reddit API Wrapper (PRAW) to retrieve the data. PRAW can only be used when we authenticate ourselves. For that, we need to develop an application and get the authentication details like client id, client secret, and user agent. Some of the sample tweets and Reddit posts are shown in Table 2.

The combined dataset collected through Twitter and Reddit was analyzed through Word Cloud in Fig. 3 to get the frequent words used by individuals. The words like "life", "feel", "commit", "want", "depress", "think", and "depression", were found in abundance. These words reflected that many individuals who post their feelings on social media want people to listen to their online cry before taking this harsh step of suicide. The extracted data also contains tweets about suicidal awareness, talking about killing oneself, reporting the third person, and using suicide in another way, e.g., suicide door. The next step was to manually annotate the data by

Table 2
Sample Tweets/ Reddit Posts reflecting the Suicidalaity

Suicidal Posts	Type of Post
"I just want to end my life so badly. My life is completely empty and I don't want to have to create meaning in it. Creating meaning is pain. How long will I hold back the urge to run my car head-first into the next person coming the opposite way. When will I stop feeling jealous of tragic characters like Gomer Pile for the swift end they were able to bring to their lives?"	Reddit Post (r/SuicideWatch)
"I've been suicidal for a while now. I've made 4 attempts this year. And yet, nobody seems to get it. I can't afford to be hospitalized. I'm an emotional wreck. I keep breaking down, crying, and I don't know why. I hate myself for this."	Reddit Post (r/SuicideWatch)
"My relationship is complicated and painful, but I don't want to end it. While my girlfriend and I were broken up, I slept with her boyfriend. For three years I've been trying to make it up to her, but after a while, I realized I don't think I ever will. I'm not sure she'll ever want to touch me again. She hasn't wanted to touch me since that happened, no matter how hard I try, no matter what I do or say. She does it for me, but she looks bored the whole time and gets it over with asap. I feel worthless. She's my soulmate. I can't live in this horrible world without her, but I don't think she wants to be sexual with me. I'm so lonely. I wish I could just turn off the part of my brain that feels."	Reddit Post (r/depression)
"I'm 21 y/o working student, I want to kill myself; life is unbearable. I have no skills no talents and I'm absolutely boring person so I spend most of my free time alone. I really want to find a girlfriend but I see no way how could I find one when I'm useless like this. I used to go away from sadness by listening to music but these days I am bored of music and just want to go to sleep and never wake up."	Reddit Post (r/depression)
"I am 19 years old. I've had anxiety my entire life and depression since I was 12, and it's gotten worse and worse over the years. Add misophonia to that and it makes life torture. It's especially bad now because I don't know what the fuck I'm doing in life. The only reason I have to live is for my friends and family and because I'm too scared to die (which is ironic because I wish I was dead most of the time, or at least I wish I'd never been born). But life is scary, too. Anyway. If things don't get better in the next 5–10 years, then I fucking quit. Or maybe I'll pussy out like the pathetic piece of a shit coward I am. I don't know."	Reddit Post (r/depression)
"I am gonna attempt suicide tonight. Bye Everyone. My last wishes are for my siblings to finish school and become emergency responders"	Tweet
"For the love of God leave me fucking alone. Plz leave flowers on my grave. Fuck y'all."	Tweet
"Confession: I am not ok! I am trying but so far I am failing"	Tweet

Fig. 3. Word Cloud of the suicidal dataset.

developing the annotation scheme in consultation with psychiatrists and Psychologists.

3.2. Data Annotation

A total number of 19915 tweets and reddit posts were collected. Out of these posts, the annotators labelled 7025 posts in the no-risk category, 6590 in the moderate risk category and 6300 in the high-risk category. The annotation was performed in consultation with mental health experts and the person's active on social media. Human annotation scheme was devised in such a way that annotators were asked to indicate about a post whether it will fall in any one of the three categories of high risk, moderate risk or no risk categories. All the posts indicating killing oneself or wishing to be dead were asked to be put in the high-risk category. The moderate-risk category includes those posts that talk about general anxiety and depression. No risk category contains posts about news about suicide, suicidal awareness and campaigning etc.

Annotators were asked to select the post against the high risk and no risk categories and in case of ambiguity, put the post in moderate-risk category, the default level. We used the Kappa coefficient to check the inter-annotator agreement between various posts.

For high-risk category:

- Observed agreement: 0.92
- Expected agreement: ((0.70252̂) + (0.29752̂)) = 0.5164
- Kappa coefficient: (0.92–0.5164)/ (1–0.5164) = 0.645

For moderate-risk category:

- Observed agreement: 0.95
- Expected agreement: ((0.65902̂) + (0.34102̂)) = 0.4533
- Kappa coefficient: (0.95–0.4533)/ (1–0.4533) = 0.673

For no-risk category:

- Observed agreement: 0.98
- Expected agreement: ((0.63002̂) + (0.37002̂)) = 0.3897
- Kappa coefficient: (0.98–0.3897)/ (1–0.3897) = 0.781

Therefore, the Kappa coefficient for the high-risk, moderate-risk, and no-risk categories are 0.645, 0.673, and 0.781, respectively that indicates a substantial agreement.

3.3. Pre-processing and Feature Engineering

Data extracted through social media contains very noise. The established methods [32] like tokenisation, stop word removal, and lemmatization was applied to filter the data to use it for machine learning. Moreover, the language of suicidal ideation lacks lexical and syntactic patterns. Therefore, there is a need for hand engineering to analyze a set of features. Feature engineering is proposed to differentiate between various levels of distress. Various features that were used in our model are as under:

- **Statistical features**: As per our analysis, the posts that users generate vary in length. Therefore, the length of a post is calculated to use as a feature to train the machine learning model.
- **Term FrequencyInverse Document Frequency (TFIDF)**: TFIDF is used to measure the importance of words in the whole corpus. TFIDF is defined below.

$$tfidf(w) = freq(w) * \log \frac{N}{|t \in D : w \in t|} \quad (1)$$

Where w refers to the word feature, N is the total number of posts, t is the document and D is the document set.

- **Latent semantic indexing features**: The features generated through TF-IDF had many inherent problems. As the dataset grows, dimensionality increases. Moreover, the sparsity also increases through the approach of the n-gram technique used on the dataset. We used Singular Value Decomposition (SVD) [33] to generate features in five ranks of 50, 100, 150, 200, and 250 new features by finding the semantic relations between the features generated through the TF-IDF scheme. Among the five ranks, the most differentiating features were used.
- **Average Risk Similarity**: Average risk similarity (ARS) is a feature that has been engineered for our multi-class classification problem on the basis of cosine similarity. The feature has been built on a per-document basis. We used the following hypothesis to engineer the ARS feature.
"*On average, high-risk suicidal messages are going to have a higher average cosine similarity with other high-risk messages than moderate and no-risk ones*".
Cosine similarity measures the cosine of the angle between two vectors. The cosine measure is the metric whose range is between −1 to 1. The smaller the angle between vectors, the more similarity. When the angle between vectors is zero, the similarity becomes highest as ($\cos\theta = 1$). Let a and b be two feature vectors having the dimension of n. the feature vectors are defined as. $a = \{a1, a2, a3 \cdots an\}$ and $b = \{b1, b2, b3 \cdots bn\}$ The cosine similarity between these vectors is defined as:

$$CS = \frac{(a * b)}{(|a| * |b|} \quad (2)$$

where,

$$a * b = (a_1 * b_1 + a_2 * b_2 + \ldots.. + a_n * b_n) = \sum_{i=i}^{n} a_i b_i$$

$$|a| = \sum_{i=i}^{n} a_i^2$$

$$|b| = \sum_{i=i}^{n} b_i^2$$

CS is the Cosine Similarity and $|a|$ and $|b|$ define the magnitude of two feature vectors a and b respectively.

- **Topic Model Features**: The suicidal and non-suicidal posts talk about different themes. Moreover, the language of high risk, moderate risk, and no risk posts also differ in their probabilities of expressing a particular topic. These themes/ topics if analyzed can help in differentiating between these three categories. Latent Dirichlet Allocation (LDA) was used to identify 50 latent topics from the corpus. The LDA function of the topic model package with a parameter k was used to implement the LDA model. It was a challenge for us to select the optimal number of topics (k) that are very well segregated and can provide meaningful information for classification purposes. Thus, we selected k with different values and picked up the one that gave us the higher coherence value. The k value of '50' marked the end of the rapid growth of the topic coherence. So, the 50 topics were found to be optimal. Table 3 shows the topic coherence when k was changed. Topic modelling assumes that mixtures of topics create a document. The topics generate various words based on the probability distribution.Fig. 4 Shows the inter-topic strength between those 50 generated topics. It is easy to observe from the figure that the oval shaped figures refer to the generated topics and the lines connecting the topics represent the strength of the relationship between them. Thicker lines indicate a stronger relationship between topics, while

Table 3
Number of Topics vs Coherence value

Number of Topics	Coherence Value
5	0.4581
10	0.5143
15	0.5480
20	0.5722
25	0.6099
30	0.6298
35	0.7491
40	0.7520
45	0.7693
50	0.7828
55	0.7835
60	0.7843
65	0.7848

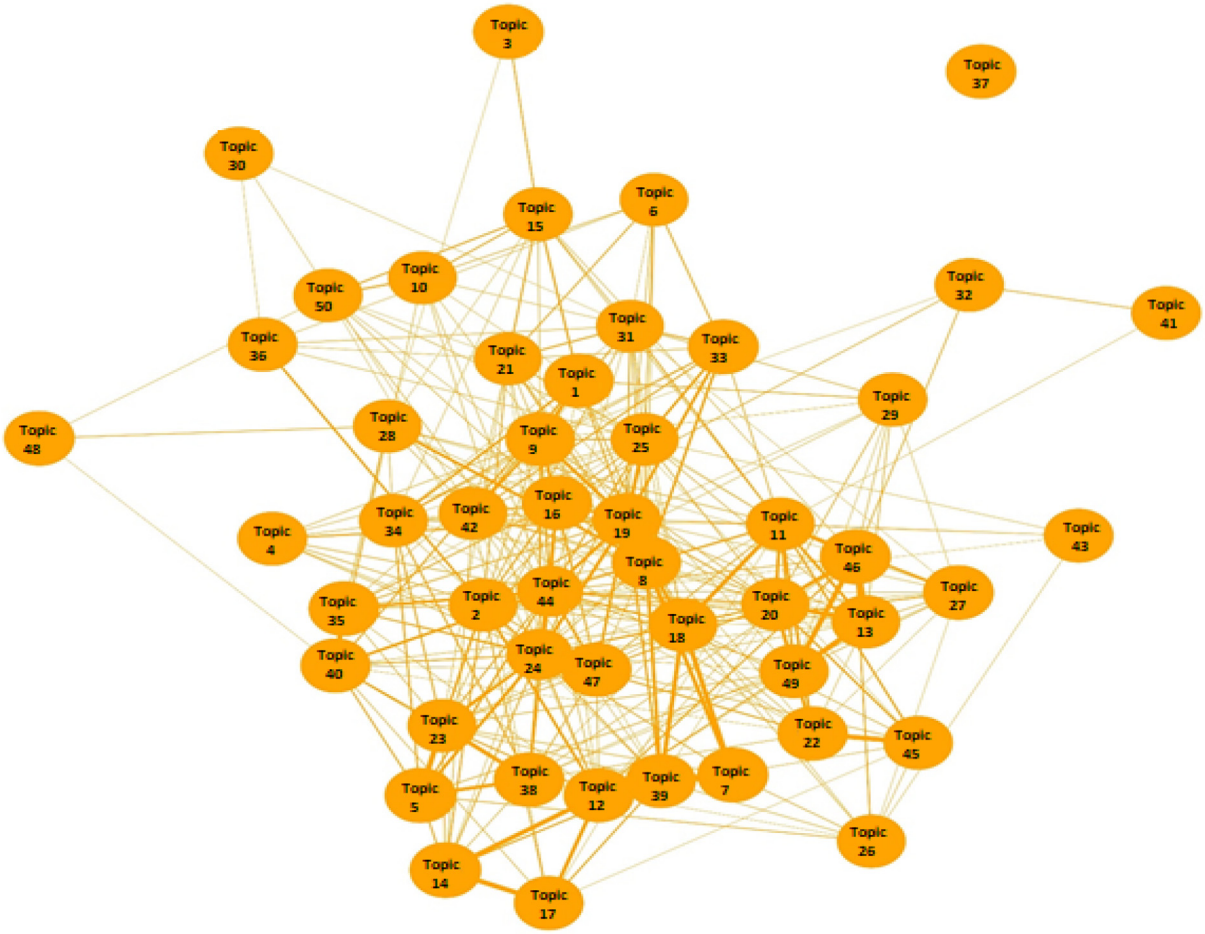

Fig. 4. Strength of relation between various generated topics.

thinner lines indicate a weaker relationship. The strength of the relationship between topics is determined by the degree of overlap between the words that make up each topic. If two topics share many common words, then they are likely to be strongly connected, and the line between them will be thicker. On the other hand, if two topics share few common words, then they are likely to be weakly connected, and the line between them will be thinner.

3.4. Classification

Suicidal ideation Detection is treated as a multi-class machine learning problem. The dataset we used to train our model consists of only two columns; the title of our text and the label. The problem is in the same way as formulated by [18]. On a corpus consisting of a set of posts/tweets $\{p_i\}_i^n$ and labels $\{l_i\}_i^n$, training is provided in such a manner that the model learns from the data consisting of a set of all the engineered features and the corresponding labels that are provided in a supervised setting. The supervisory function guides the model as under:

$$l_i = Fun(p_i) \tag{3}$$

$l_i = 2$ in case of p_i representing the high risk of suicide. The value of $l_i = 1$ and 0 in case of p_i representing the moderate risk and no risk of suicide respectively.We focused on the 'No Free Lunch' theorem of machine learning that suggested that no algorithm can work well for all problems. As a result, we tried top three well known machine and ensemble learning algorithms for text classification [40,41] viz. Support vector machine, Random forest and Extreme gradient boosting algorithm to train our multi-class classification model. The model is further validated through 5-cross validation technique. The cross-validation approach reduces the bias and variance as the majority of the dataset is used for training the model, and most of the data is also used for testing the model. The empirical evidence suggests that 10-fold cross-validation and 5-fold cross-validation is generally preferred, but it is not a thumb rule as k can take any value.

The various metrics that were used to evaluate our classification results are confusion matrix, Accuracy, Precision, Recall and F-Measure.

- **Confusion Matrix:** The confusion matrix presents the four values i.e.True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) in a matrix form. True positive is a value, where the classification model correctly predicts the positive class, while true negative is a value, where the classifier accurately predicts the negative class. The false positive is a value where the classifier incorrectly predicts the positive class. Similarly, a false negative is a value where the classifier incorrectly predicts the negative class. An ideal classifier should have more True positive and True negative values. The perfect classifier will have an off-diagonal value equal to zero, while all the values lie on the main diagonal, i.e. (FP = 0, FN = 0).

- **Accuracy:** It is one of the common metrics used to evaluate the machine learning classifiers. It measures the ratio of the total number of corrected predicted instances over the total number of predictions.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Accuracy is a good measure, but when the dataset is imbalanced, accuracy does not provide more information about the classifier's performance.

- **Precision:** It is a metric that determines the ratio of true positives over the total predicted positives (TPR).

$$Precision = \frac{TP}{TPR} \quad (5)$$

- **Recall:** It is a metric that determines the ratio of true positives over total actual positives (TAP). Recall is preferred to select the best model when the high cost is associated with false-negative, e.g. In disease prediction like in suicidal risk identification, the consequences can be very high if at-risk person is misclassified.

$$Recall = \frac{TP}{TAP} \quad (6)$$

- **F1-score:** It is a metric that is used to measure the performance of the classifier/model when that model needs a balance between Precision and Recall & also when the dataset is imbalanced, having a large number of actual negative values. Usually, for learning models, false positives and false negatives provide an important role. The F1 score tries to give more weight to these values and contribute in minimizing the impact of true negative values.

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (7)$$

The overall procedure that we used to detect level of suicidal ideation in the social media posts is shown as a pseudocode named as Algorithm 1 Enhanced feature engineering approach for suicidal risk identification (EFASRI).

Algorithm 1: Enhanced Feature Engineering Approach for Suicidal Risk Identification (EFASRI)

1: **Require:** Filtered Posts (Tweets & Reddit) ($P_{in}.csv$), Classifier_Name, Classifier_Hyperpar
2: **Ensure:** No Risk (PNR), Moderate Risk (PMR) and High Risk (PHR)
3: **ST4733p33ART**
4: **for** i from 1 to n (Total Number of Posts) **do**
5: C [i] = P_{input}[i] $ Label//Adding labels
6: Text.csv = c [i]//CSV file containing posts and corresponding labels
7: **end for**
8: Pro = tokens (Text.csv)//Tokenization and other text standardization's
9: Pro = tokens_tolower (Pro) // Lower casing
10: Pro = tokens_remove (Pro)// Hand Engineered stop word removal
11: Pro = tokens_stem (Pro)// stemming
12: Text2.csv = Pro// Processed file
13: Processedcorpus = tm_map((tolower(Text), removewords (Text), remove_punctuation(Text), preserve_intra_word_dashes = true, removenumbers (Text). stemDocument(Text), stripwhitespace(Text))// processed corpus retained for topic modeling

a (*continued*)

Algorithm 1: Enhanced Feature Engineering Approach for Suicidal Risk Identification (EFASRI)

14: **for** i from 1 to n **do**
15: P_length[i] = nchar (Text2[i])// length of Post
16: **end for**
17: Tokens = tokens_ngrams (Text2, n = 1:3)//ngram features up to 3 grams of whole dataset
18: Tokens.dfm = dfm (Tokens)//make document feature matrix
19: dfm_trimmed = dfm_trim (Tokens.dfm, min_docfreq, min_termfreq)// trimming the tokens having less importance
20: TFIDF_features = dfm_trimmed. Tfidf// Extracting TFIDF Features
21: incomplete.cases <– which (! complete. cases (TFIDF_features))
22: TFIDF_features [incomplete. cases,] <–rep (0.0, ncol (TFIDF_features))// replacing incomplete cases
23: LSA_features = SVD (TFIDF_features, nv = 50,100,150,200 250)// Most relevant features extracted by Dimensionality Reduction using Singular value decomposition of LSA
24: Train.similarities = cosine (LSA_Features)// Finding similarity of doc's based on Cosine measure
25: **for** i from 1 to n row **do**
26: ARS [i] = mean (train. similarities [i, High Risk index)// Finding Average suicide similarity based on cosine measure
27: **end for**
28: topic_model_features <–LDA (processedcorpus.dtm, k = 50, method ="Gibbs)//topic modeling features on preprocessed document term matrix
29: Optimal_Feature Set = LSA_features + P_length + ARS + t opic model features
30: ClASSIFIER(Classifier_Name, Classifier_Hyperparameters, CV = 5,Optimal_Feature Set)// Training classifier using our feature engineering approach With hyperparameters tuned and 5 cross-validation technique
31: **END**

4. Results and Discussion

Twitter and Reddit APIs were used to collect 19915 suicidal and non-suicidal posts using different keywords and phrases used in previous literature like [42,26] and also through various other terms defined in our manual library that were collected in consultation with mental health experts. The data from two subreddits r/depression and r/SuicideWatch were also extracted to get those posts that indicate the suicidal ideation and emotional state of the potential suicidal users in more detail. After collecting a large enough dataset, we pre-processed the dataset to remove the redundancy and noise using various established methods [43]. Thereafter, the proposed feature engineering mechanism was used to extract the most relevant features for classifying suicidal tweets into three classes of distress. The feature extraction mechanism consists of TFIDF, Latent semantic features, length of the post, topic model features, and average risk similarity based on the cosine similarity measure. The initial experiments were conducted by splitting the dataset into 80:20 and 70:30 because empirical results [44] show the performance of algorithms increases on such splitting. We found that the model's performance increased slightly on increasing the training data. The technique of splitting the dataset into training and testing folds suffer from inherent

problem of bias and variance. In machine learning, it is not always that the model that has fit the training data will also work for the real data. For that, we used the k-fold cross-validation technique to get assured that the model gets the correct patterns of the data and not take too much noise. The model was validated through a 5-fold cross-validation mechanism. Python was selected to implement machine learning algorithms. The various important packages used in coding are Skitlearn, pandas, and NLTK. We applied three well-known machine and ensemble learning algorithms i.e., SVM, Random Forest, and XGB to train the suicide prediction model. The configuration parameters of these machine-learning algorithms are highlighted in Table 4. Among the three machine learning algorithms, XGB-EFASRI outperformed the other two algorithms with an overall accuracy of 0.9633.The main reason XGB-EFASRI performed better than SVM and Random Forest is because XGB handled both linear and nonlinear relationships between features. In suicidal detection problem, there were complex relationships between the features and the target variable, which XGB captured better than SVM and Random Forest. Moreover, XGB uses gradient boosting, which involved training multiple models sequentially and combined their predictions. This helped reduce the variance of the model and improved its accuracy. The confusion metrics generated in our experimentation are shown in Fig. 5–7. On Analysing these figures, it can be seen that the rows represent the actual classes of the test data, while the columns represent the predicted classes. The values in the cells of the confusion matrix indicate the number of instances that were classified as belonging to a particular class. In our experimentation, confusion matrix generated has three rows and three columns, corresponding to three classes of suicide risk: Label '0', '1', 2' refers to the No Risk, Moderate Risk, and High Risk, respectively. Fig. 5 depicts the confusion matrix generated by applying SVM-EFASRI. Out of 1426 instances that truly belong to the No Risk class, the model correctly predicted 1275 and incorrectly predicted 41 as Moderate Risk and 110 as High Risk. Out of 1106 instances that truly belong to the Moderate Risk class, the model correctly predicted 905 and incorrectly predicted 76 as No Risk and 125 as High Risk. Out of 1451 instances that truly belong to the High-Risk class, the model correctly predicted 1229 and incorrectly predicted 154 as No Risk and 68 as Moderate Risk. Fig. 6 depicts the confusion matrix generated by applying RFEFASRI. Out of 1384 instances that truly belong to the No Risk class, the model correctly predicted 1357 and incorrectly predicted 1 as Moderate Risk and 26 as High Risk. Out of 1157 instances that truly belong to the Moderate Risk class, the model correctly predicted 1083 and incorrectly predicted 22 as No Risk and 52 as High risk. Out of 1442 instances that truly belong to the HighRisk class, the model correctly predicted 1366 and incor-

Table 4
Hyper-parameter tuning of various Machine Learning algorithms

Algorithm	Configuration Parameters
Support Vector Machine (SVM)	C = 1.0, kernel='poly', shinking = True, probability = True, coef0 = 0.0, degree = 3, decision_function_shape='ovo', max_iter=-1, Verbose = 10, cache-sixe = 200, class_wight = None tol = 0.0001, gamma='auto deprecated', random_state='110'.
Random Forest (RF)	n_estimators = 300, bootstrap = True, njobs = 6 random_state = 42, verbose = True, oob score = False, Criterion='entropy', warm state = False max_depth = 30, maximum_features='sqrt', minimum samples split = 5, class weight = None, minimum samples leaf = 1, minimum weight fraction leaf = 0.0, minimum impurity decrease = 0.0, maximum leaf nodes = None, minimum impurity split = None.
Extreme boosting Gradient (XGB)	base_score = 0.5, 'reg_alpha = 1.2', col_sample_bytree = 1, 'reg_lambda = 1.3' booster='gbtree', n_jobs = 6, nestimators = 250, objective='multi:softprob', verbose = 1, validate = False, gamma = 0, subsample = 0.8.

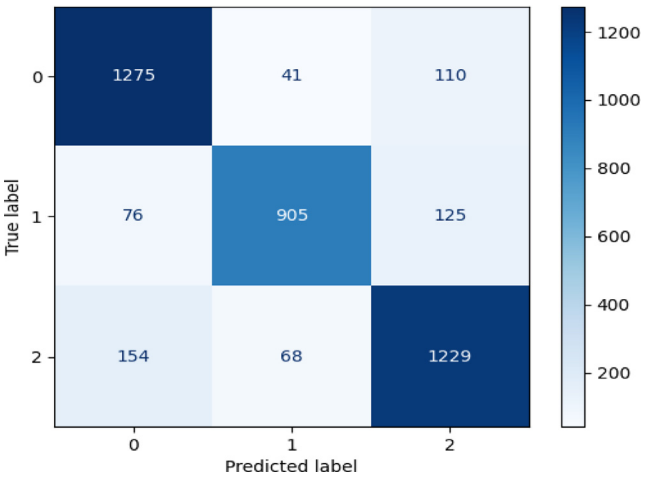

Fig. 5. Confusion Matrix generated through SVM_EFASRI.

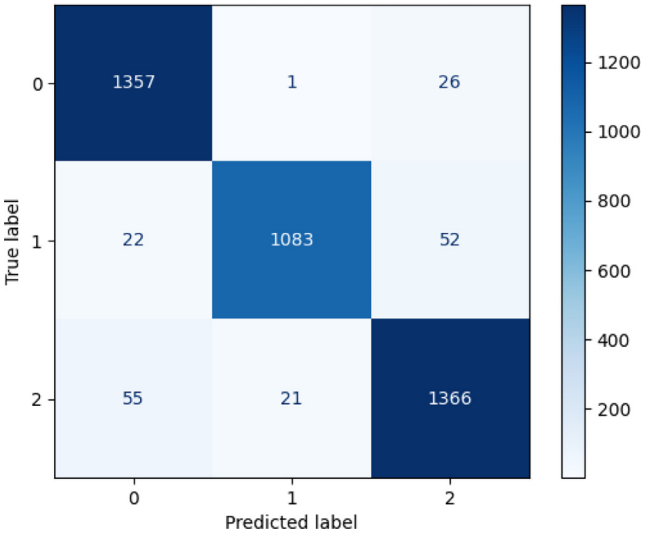

Fig. 6. Confusion Matrix generated through RF_EFASRI.

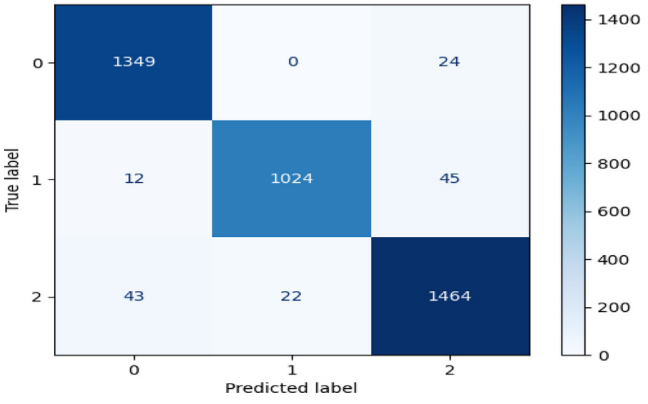

Fig. 7. Confusion Matrix generated through XGB_EFASRI.

Table 5
Classification report of various algorithms based upon our proposed approach (EFASRI)

Classifier	Metrics	No Risk	Moderate Risk	High Risk
SVM-EFASRI	**Accuracy**		**0.8558**	
	Precision	0.8540	0.8905	0.8430
	Recall	0.8945	0.8243	0.8507
	F-measure	0.8702	0.8523	0.8205
RF-EFASRI	**Accuracy**		**0.9555**	
	Precision	0.9523	0.9820	0.9523
	Recall	0.9843	0.9432	0.9534
	F-measure	0.9638	0.9643	0.9535
XGB-EFASRI	**Accuracy**		**0.9633**	
	Precision	0.9621	0.9821	0.9523
	Recall	0.9834	0.9535	0.9635
	F-measure	0.9700	0.9623	0.9643

rectly predicted 55 as No Risk and 21 as Moderate Risk. Fig. 7 depicts the confusion matrix generated by applying XGBEFASRI. Out of 1373 instances that truly belong to the No Risk class, the model correctly predicted 1349 and incorrectly predicted 0 as Moderate Risk and 24 as High Risk. Out of 1081 instances that truly belong to the Moderate Risk class, the model correctly predicted 1024 and incorrectly predicted 12 as No Risk and 45 as High risk. Out of 1529 instances that truly belong to the High-Risk class, the model correctly predicted 1464 and incorrectly predicted 43 as No Risk and 22 as Moderate Risk. The majority of the elements in the left diagonal of the confusion matrices indicates that the our model is performing well in predicting the correct class for most instances. Table 5 shows the detailed comparison of the results generated through application of our proposed feature extraction. mechanism and implementation of various machine learning algorithms. We used the various standard metrics like Precision, Recall and F-measure to evaluate the performance of the machine learning algorithms using our proposed approach that provides the better picture of the model's performance. As our problem statement was about the disease prediction where the high cost is associated with the false negative values, the high recall indicates that the model is able to correctly identify a higher proportion of true positive cases while minimizing false negatives. This is an important outcome in disease prediction, where correctly identifying positive cases is critical to ensure timely treatment and minimize negative consequences.

Table 6 shows the variation in performance accuracy of the XGB, the best performing algorithm on inclusion of different features sets. The First column depicts the combination of feature sets and last four columns reflects the accuracy in terms of various evaluation metrics. It is observed that there is an increase in precision, Recall and F measure when additional features are added and thus validates its use for increase in overall accuracy.

5. Comparison with the previous research

Limited research has been conducted on the classification of social media data related to suicidal ideation. In this study, we compared our proposed approach with the most recent and relevant studies [30,38] regarding the identification of suicidal ideation. Our findings revealed that none of the previous studies achieved a level of accuracy comparable to our work. Furthermore, the prior research employed simplistic feature extraction techniques. In contrast, our approach achieved higher precision, recall and overall accuracy due to the hybrid feature engineering approach for extracting the most relevant features. Table7 comparison with the recent works.

6. Conclusion

In light of the increased prevalence of Social Networking Sites (SNS) and the associated social stigma, individuals have become more comfortable sharing their personal feelings on these platforms. In this article, we presented a text classification approach for detecting suicidal ideation on SNS, utilizing a hybrid feature engineering mechanism. The extracted features were then supplied to three machine learning algorithms, resulting in a maximum achieved accuracy of 96.33%, demonstrating successful replication of human accuracy. The high precision and recall values obtained using the proposed feature extraction approach indicate that it could play a vital role in developing a reliable model with stable precision and recall. The findings of this research article offer valuable insights for psychologists, psychiatrists, and patients, shedding light on the detection of suicidal ideation in a novel and effective manner. Some of the Future directions of the work are listed as under:

1. Prediction models developed for binary classification and multiclass classification can be deployed as a product that can generate the intervention messages. People can chat anonymously with mental health experts/therapists without leaving their homes and worrying about social stigma. When the user will give feedback about how intervention messages helped them, it can be channelized to improve the model.
2. The work can be extended to photos and videos related to suicidal ideation by including image and video processing.

Table 6
Variation in performance accuracy of the XGB

Features	Precision	Recall	F- Measure	Accuracy
Statistical features (SF) +TFIDF	0.8850	0.8630	0.8730	**0.8850**
SF + TFIDF + latent semantic indexing features	0.9013	0.9215	0.9330	**0.9300**
SF + TFIDF_latent semantic indexing features + Average Risk Similarity (ARS) feature	0.9210	0.9315	0.9450	**0.9415**
SF + TFIDF + latent semantic indexing features + Average Risk Similarity (ARS) feature + Topic model features	0.9523	0.9535	0.9643	**0.9633**

Table 7
Comparison of our proposed work with the previous research

Study	Features Used	Accuracy Achieved
[30]	TFIDF, BOW	**0.9292**
[38]	Word2vec	**0.9500**
Proposed Work (XGBEFASRI)	Statistical features, TFIDF, Latent Semantic Indexing features, Average risk similarity features based upon cosine similarity and topic modelling features	**0.9633**

3. Retrospective longitudinal analysis of those users needs to be analyzed who died due to suicide that will help understand suicidal behavior in a better way and further improve the model.
4. Questionnaires in consultation with mental health experts can be used to extract the self-reported diagnosis (SRD) data that will be supplied to the machine model to make it more accurate.
5. The algorithm could be developed that can detect repeated name-calling and abuse directed at a particular user even though the person does not complain about it or express his/her thoughts in implicit ways.
6. Further fine-grained classification of suicidal at-risk individuals needs to be performed on the basis of emotions like anger, disgust, etc. The classifier detecting the same will help redirect the different kinds of suicidal individuals to particular resources for intervention.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Suicide, https://www.who.int/health-topics/suicide#tab=tab_1, (Accessed on 09/11/2022).
[2] J. Bilsen, Suicide and youth: Risk factors, https://www.frontiersin.org/articles/10.3389/fpsyt.2018.00540/full, (Accessed on 09/11/2022) (10 2018).
[3] Värnik P. Suicide in the world. Int J Environ Res Public Health 2012;9(3):760–71. doi: https://doi.org/10.3390/ijerph9030760. https://www.mdpi.com/1660-4601/9/3/760.
[4] Keith Hawton KVH, The international handbook of suicide and attempted suicide — wiley, https://www.wiley.com/en-in/The+International+Handbook+of+Suicide+and+Attempted+Suicide-p-9780470849590, (Accessed on 09/11/2022) (07 2002).
[5] Risk factors, protective factors, and warning signs — afsp, https://afsp.org/risk-factors-protective -factors-and- warning-signs/, (Accessed on 09/11/2022).
[6] Shioiri T, Nishimura A, Akazawa K, Abe R, Nushida H, Ueno Y, Kojika-Maruyama M, Someya T. Incidence of note-leaving remains constant despite increasing suicide rates. Psychiatry Clin Neurosci 2005;59(2):226–8. doi: https://doi.org/10.1111/j.1440-1819.2005.01364.x. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1440-1819.2005.01364.x.
[7] Foster T. Suicide note themes and suicide prevention. Int J Psychiatry Med 2003;33(4):323–31. doi: https://doi.org/10.2190/T210-E2V5-A5M0-QLJU. pMID: 15152783. arXiv:https://doi.org/10.2190/T210-E2V5-A5M0-QLJU.
[8] Wang N, Luo F, Shivtare Y, Badal VD, Subbalakshmi K, Chandramouli R, Lee E, Learning models for suicide prediction from social media posts, arXiv preprint arXiv:2105.03315.
[9] Aladağ AE, Muderrisoglu S, Akbas NB, Zahmacioglu O, Bingol HO. Detecting suicidal ideation on forums: proof-of-concept study. J Med Internet Res 2018;20(6):e9840.
[10] Fu K-W, Cheng Q, Wong PWC, Yip PSF. Responses to a self-presented suicide attempt in social media: A social network analysis. Crisis 2013. doi: https://doi.org/10.1027/0227-5910/a000221.
[11] Jashinsky J, Burton SH, Hanson CL, West J, Giraud-Carrier C, Barnes MD, Argyle T. Tracking suicide risk factors through twitter in the us. Crisis 2014;35(1):51–9. doi: https://doi.org/10.1027/0227-5910/a000234. pMID: 24121153. arXiv:https://doi.org/10.1027/0227-5910/a000234.
[12] Ahuja AK, Biesaga K, Sudak DM, Draper J, Womble A, Suicide on Facebook, Journal of Psychiatric Practice 20 (2). URL: https://journals.lww.com/practicalpsychiatry/Fulltext/2014/03000/Suicide_on_Facebook. 8.aspx.
[13] O'Dea B, Larsen ME, Batterham PJ, Calear AL, Christensen H. A linguistic analysis of suicide-related Twitter posts. Crisis 2017. doi: https://doi.org/10.1027/0227-5910/a000443.
[14] Khanday AMUD, Khan QR, Rabani ST. Identifying propaganda from online social networks during COVID-19 using machine learning techniques. Int J Inform Technol 2021;13(1):115–22. doi: https://doi.org/10.1007/s41870-020-00550-5.
[15] Khanday AMUD, Rabani ST, Khan QR, Rouf N, Mohiud Din M, Machine learning based approaches for detecting COVID-19 using clinical text data, Int J Inform Technol. doi:10.1007/s41870-020-00495-9. URL: https://doi.org/10.1007/s41870-020-00495-9
[16] Rabani ST, Khan QR, Khanday AMUD. Detection of suicidal ideation on twitter using machine learning & ensemble approaches. Baghdad Sci J 2020;17(4):1328. doi: https://doi.org/10.21123/bsj.2020.17.4.1328. URL: https://bsj.uobaghdad.edu.iq/index.php/BSJ/article/view/5245.
[17] Choudhury MD, Kiciman E, The language of social support in social media and its effect on suicidal ideation risk, 2017. URL: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15662/14792.
[18] Ji S, Yu CP, Fung S-F, Pan S, Long G. Supervised Learning for Suicidal Ideation Detection in Online User Content. Complexity 2018;2018:6157249. doi: https://doi.org/10.1155/2018/6157249.
[19] Sikander D, Arvaneh M, Amico F, Healy G, Ward T, Kearney D, Mohedano E, Fagan J, Yek J, Smeaton AF, Brophy J. Predicting risk of suicide using resting state heart rate. In: 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA). p. 1–4. doi: https://doi.org/10.1109/APSIPA.2016.7820833.
[20] Moreno MA, Jelenchick LA, Egan KG, Cox E, Young H, Gannon KE, Becker T. Feeling bad on facebook: depression disclosures by college students on a social networking site. Depression Anxiety 2011;28(6):447–55. doi: https://doi.org/10.1002/da.20805. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/da.20805 URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/da.20805.
[21] Shen JH, Rudzicz F, Detecting anxiety through reddit. In: Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology-From Linguistic Signal to Clinical Reality, 2017, pp. 58–65.
[22] Cash SJ, Thelwall M, Peck SN, Ferrell JZ, Bridge JA. Adolescent suicide statements on myspace. Cyberpsychology, Behavior, Social Networking 2013;16(3):166–74. doi: https://doi.org/10.1089/cyber.2012.0098. pMID: 23374167. arXiv:https://doi.org/10.1089/cyber.2012.0098.
[23] Harris KM, McLean JP, Sheffield J. Suicidal and online: How do online behaviors inform us of this high-risk population? Death Studies 2014;38(6):387–94. doi: https://doi.org/10.1080/07481187.2013.768313. pMID: 24666145. arXiv:https://doi.org/10.1080/07481187.2013.768313.
[24] Homan S, Gabi M, Klee N, Bachmann S, Moser A-M, Michel S, Bertram A-M, Maatz A, Seiler G, Stark E, et al. Linguistic features of suicidal thoughts and behaviors: A systematic review. Clinical Psychol Rev 2022;95:102161.
[25] Desmet B, Hoste V. Online suicide prevention through optimised text classification. Inf Sci 2018;439:61–78.
[26] Chiang W-C, Cheng P-H, Su M-J, Chen H-S, Wu S-W, Lin J-K. Socio-health with personal mental health records: suicidal-tendency observation system on facebook for taiwanese adolescents and young adults. In: 2011 IEEE 13th International Conference on e-Health Networking, Applications and Services. IEEE; 2011. p. 46–51.
[27] O'dea B, Wan S, Batterham PJ, Calear AL, Paris C, Christensen H. Detecting suicidality on twitter. Internet Interventions 2015;2(2):183–8.
[28] De Choudhury M, Kiciman E, Dredze M, Coppersmith G, Kumar M, Discovering shifts to suicidal ideation from mental health content in social media. In: Proceedings of the 2016 CHI conference on human factors in computing systems, 2016, pp. 2098–2110.
[29] Vioules MJ, Moulahi B, Azé J, Bringay S. Detection of suicide-related posts in twitter data streams. IBM J Res Dev 2018;62(1). 7–1.
[30] Chadha A, Kaushik B. Performance evaluation of learning models for identification of suicidal thoughts. Computer J 2022;65(1):139–54.
[31] Abboute A, Boudjeriou Y, Entringer G, Azé J, Bringay S, Poncelet P, Mining twitter for suicide prevention. In: Natural Language Processing and Information Systems: 19th International Conference on Applications of Natural Language to Information Systems, NLDB 2014, Montpellier, France, June 18–20, 2014. Proceedings 19, Springer, 2014, pp. 250–253.
[32] Okhapkina E, Okhapkin V, Kazarin O. Adaptation of information retrieval methods for identifying of destructive informational influence in social networks. In: 2017 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA). IEEE; 2017. p. 87–92.
[33] Cheng Q, Li TM, Kwok C-L, Zhu T, Yip PS. Assessing suicide risk and emotional distress in chinese social media: a text mining and machine learning study. J Med Internet Res 2017;19(7):e243.
[34] Shing H-C, Nair S, Zirikly A, Friedenberg M, Daumé III H, Resnik P, 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.
[35] Mbarek A, Jamoussi S, Charfi A, Hamadou AB. Suicidal profiles detection in twitter. In: WEBIST. p. 289–96.
[36] Mbarek A, Jamoussi S, Hamadou AB. An across online social networks profile building approach: Application to suicidal ideation detection. Future Generation Computer Syst 2022;133:171–83.
[37] Burnap P, Colombo G, Amery R, Hodorog A, Scourfield J. Multi-class machine classification of suicide-related communication on twitter. Soc Netw Media 2017;2:32–44.
[38] Chadha A, Kaushik B. A survey on prediction of suicidal ideation using machine and ensemble learning. Computer J 2021;64(11):1617–32.
[39] Aldhyani TH, Alsubari SN, Alshebami AS, Alkahtani H, Ahmed ZA. Detecting and analyzing suicidal ideation on social media using deep learning and machine learning models. Int J Environ Res Public Health 2022;19(19):12635.
[40] Kowsari K, Jafari meimandi K., Heidarysafa M, Mendu S, Barnes L, Brown D, Text Classification Algorithms: A Survey. Information 10 (4).

[41] Gasparetto A, Marcuzzo M, Zangari A, Albarelli A. A survey on text classification algorithms: From text to predictions. Information 2022;13(2):83.
[42] Huang X, Zhang L, Chiu D, Liu T, Li X, Zhu T. Detecting suicidal ideation in chinese microblogs with psychological lexicons,. In: 2014 IEEE 11th Intl Conf on Ubiquitous Intelligence and Computing and 2014 IEEE 11th Intl Conf on Autonomic and Trusted Computing and 2014 IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops. IEEE; 2014. p. 844–9.
[43] Anand N, Goyal D, Kumar T, Analyzing and preprocessing the twitter data for opinion mining. In: Proceedings of International Conference on Recent Advancement on Computer and Communication, Springer, 2018, pp. 213–221.
[44] Gholamy A, Kreinovich V, Kosheleva O, Why 70/30 or 80/20 relation between training and testing sets: a pedagogical explanation.