Article:

Sarsam, SM, Al-Samarraie, H orcid.org/0000-0002-9861-8989, Alzahrani, AI et al. (2 more authors) (2024) Characterizing Suicide Ideation by Using Mental Disorder Features on Microblogs: A Machine Learning Perspective. International Journal of Mental Health and Addiction, 22 (4). pp. 1783-1796. ISSN 1557-1874

https://doi.org/10.1007/s11469-022-00958-z

Characterizing Suicide Ideation by Using Mental Disorder Features on Microblogs: A Machine Learning Perspective

Abstract

Despite the success of psychological and clinical methods, psychological studies revealed that the number of individuals exhibiting suicide ideation has highly increased in the recent decades. This study explored the potential of using certain sentimental features as a means for characterizing suicide. A total of 54,385 English-language tweets were collected and processed to extract suicide-related topics using the Latent Dirichlet Allocation (LDA) algorithm. Both suicidal polarity (positive, negative, and neutral) and emotions (anger, fear, sad-ness, and trust) were extracted via SentiStrength, time series, and NRC Affect Intensity Lex-icon methods. The results showed that suicidal tweets were less associated with trust, anger, and positive sentiments. In contrast, fear, sadness, and negative sentiments were highly associated with suicidal statements. The prediction results using this approach showed 97.64% accuracy in detecting suicide ideation. The obtained results from analyzing suicide-related tweets hold a promising future for characterizing suicide ideation worldwide.

Keywords: Suicide ideation; Sentiment analysis; Topic modelling; Twitter

1. Introduction

Suicide ideation has been identified in the existing literature as an individual's intention to end his or her life. Suicidal behaviors can be generally divided into three streams (suicide ideation, suicide planning, and attempted suicide) that are common among adolescents (Robinson et al., 2012). Suicide has also been identified as an individual's intention to develop a plan to kill oneself (Lewinsohn et al., 1996).

The literature revealed several factors contributing to the increase in suicide ideation/ behavior among people. Across ethnicities suicide has been found to be more present among boys than among girls, by an average ratio of 5:1 (Canetto & Lester, 1995). This ration has been widened in recent decades, especially in some ethnic minority groups. The widening gender gap is mostly due to many other psychological factors. Hopeless-ness has been identified as a key psychological construct for studying suicide (Weishaar & Beck, 1992). According to Beck et al.

(1989), individuals with hopelessness are known to experience depression that can be useful in the prediction of suicide act. Meanwhile, hopelessness resulted from other psychiatric disorders can predisposes individuals to suicidal behavior. Douglas (2015) reported the importance of understanding shared linguistic terminologies for suicidal behavior. The author emphasized that certain topics or terms can be useful in explaining individuals' feelings and suicide behavior.

Furthermore, the topics and terms used by people to express their suicide ideation can be linked to their context. According to Stack and Kposowa (2016), there are some cultural differences in the incidence of suicidal behavior. This includes the methods individuals utilize when committing suicide and the reasons for doing so. The literature, however, still assumes that certain cultural differences can be due to physiological differences between the members of the different cultures. Other explanations on why people commit suicide consist of psychological and social variables, such as the abuse of alcohol and the level of social integration and regulation.

It is believed that when competing cultures interact, members of minority culture/language groups are more prone to increased stress, thus increasing suicidality (Lester & Rogers, 2015). Despite numerous studies and efforts to better understand suicidal behaviors among people of different demographic backgrounds, the rates of suicide ideation and attempts have remained unchanged (Nock & Banaji, 2007). This can be attributed to the complex-ity in characterizing suicidal behaviors among a group of people. In addition, the current suicide identification approaches and methods used by healthcare organizations are mostly shaped by self-reports of suicidal thoughts and intentions. These classical methods can be less effective in identifying and monitoring changes in individual suicidal behaviors, especially in relation to certain events or situations (Sourirajan et al., 2020). Furthermore, sui-cide ideation and suicidal behavior may go unreported in places where healthcare services are limited. For instance, individuals who are developing plans to commit suicide are less likely to report depression and other mental health symptoms (Spokas et al., 2009). The literature has also showed that certain individuals may even lack an introspective awareness of the thoughts and feelings that drive suicidal behavior and thus lack the ability to inform others of these issues (Nock & Banaji, 2007).

Through the latest advancement in social media mining, researchers can benefit from information that is shared across social media networks to better understand people's behavioral and emotional changes in relation to certain events (Bide & Dhage, 2021). The potential of using sentiment analysis as a way to offer timely and effective predictions of online events has been widely addressed in the literature (Rahimi et al., 2019). However, the application of sentiment analysis in suicide detection has yet to reach its full potential (Sarsam, Al-Samarraie, Alzahrani et al., 2021a, 2021b). Based on these observations, this study aims to answer the following questions: (1) "What are the main topics useful for characterizing suicide ideation on Twitter?" and (2) "What are the types of emotions that can best describe suicide ideation on Twitter?". To answer these questions, this study investigates the potential of using topic modeling and time series analysis to characterize the main topics shared on Twitter in relation to suicide ideation and mental disorders. The study also investigates the predictive capability of certain emotions that are found in tweets in relation to suicide ideation. The findings from this study can offer mental health scientists and organizations a new and effective way to understand changes in people's suicidal behavior in a context-specific setting.

2. Suicide Ideation and Social Emotions

The process of identifying suicidal behaviors is recognized as an important public health problem that may result in unwanted psychological and economic burdens on a community (Okolie et al., 2017; Silverman, 2016). The detection and recognition of suicidal events have been reported to be understudied compared to other risk factors, such as gender, depression, or other psychopathological dimensions (Murni & Ibrahim, 2020). Several reasons may contribute to the existing difficulty in characterizing suicide ideation. The lack of advanced methods for screening early - stage suicidal behaviors play a significant role in recognizing suicide-related events (Ibrahim et al., 2019). In fact, the process of developing intelligent methods to characterize suicide ideation is a challenging task due to the complexity of linking shared information with certain suicidal events or situations. Also, the classical reporting methods of suicide ideation (questionnaires and clinical interviews) are not capable of capturing changes in suicide ideation across a population (Siau et al., 2017). Therefore, providing new ways to recognize suicidal behaviors in individuals has proven to aid the development of prevention strategies. The expression of individual sentiment (emotion) is shown to be an effective medium that can greatly contribute to the detection (prediction) process of suicide ideation. Previous studies such as Selby et al. (2009) revealed that people who struggle to regulate their emotions are more vulnerable to suicide attempts. Hence, individuals involved in different suicide attempts are likely to overcome their innate biological tendencies toward survival.

From the previous discussion, it can be observed that it is important to understand aspects related to individuals' emotions are critical contributors to humans' avoidance behaviors that lead a person to experience pain-related fear (Law et al., 2015). According to Abrutyn and Mueller (2014), social emotions can push certain groups of people to commit suicide. The authors explored the potential advantages of studying the relationship between emotions and cultural factors in an attempt to understand individual intentions and decisions to commit suicide. In line with that, social media websites can be viewed as a convenient medium for users to express their emotions, experience, and feelings about different personal matters (Sun et al., 2018; Wu et al., 2020). These emotions can be applied to provide an early screening mechanism for psychiatric disorders with the help of machine learning models. This assumption is supported by Sarsam, Al-Samarraie and Alzahrani et al. (2021a, 2021b) who studied the potential role of emotions in understanding mental disorders among people. The present study proposes a novel method for characterizing suicide ideation from microblogs with the help of data mining methods (sentiment analysis, topic modeling, and time series).

3. Method

The performed procedure consists of data collection, data pre-processing, cluster analysis, topic modeling, emotion extraction, and classification technique. The following sections explain the implemented process in detail.

3.1 Data Collection

A total of 54,385 English tweets were gathered within a time span of six months (September 1st, 2020, till April 30th, 2021). We collected the data using the Twitter free streaming Application Programming Interface (API) based on the recommendation of Sarsam et al. (2019). The desired tweets were obtained using a number of keywords: "suicide," "suicidal thoughts," and "contemplating suicide." It is also worth mentioning that two experts in mental health were invited in this study to help in assessing the relevancy of the collected tweets on suicidal ideation. Then, several data preparation steps were considered to enhance the quality and the reliability of tweets, which are extremely essential for the analysis stage.

3.2 Data Pre-processing

In this study, several pre-processing methods were used to obtain a solid knowledge out of it. In this context, the bag-of-words model was implemented where and the "Tokenization" technique was utilized to extract the tweets features (words) and build a dictionary of from these features. Then, all these features were converted to a lowercase form before applying the Stopwords list technique which was used to keep the necessary words in the dictionary and provide relevant information. Finally, the length of the tweets was normalized using the L2 norm (Sarsam et al., 2020). This helps to guarantee fair treatment to all the tweets by machine learning algorithms in the coming stages.

Figure 1 shows the distribution of the collected tweets on suicide worldwide. At the time of this search, both the USA and Afghanistan had the highest tweets on suicide topics. In addition, 62% of tweets were ported by male users and 38% by female users.

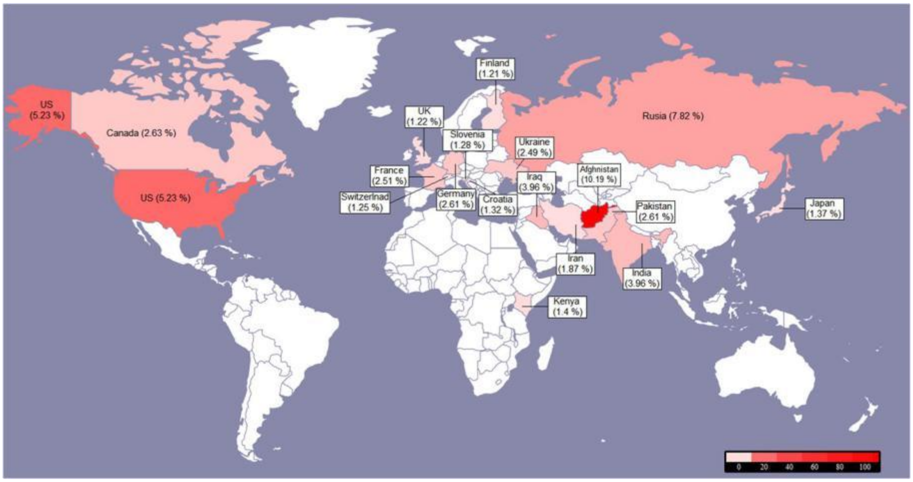

Figure 1: Breakdown of tweets on suicide across countries

3.3 Cluster Analysis

After the data pre-processing stage, the hierarchical clustering algorithm was used based on previous recommendations (Sarsam & Al-Samarraie, 2018a, 2018b). Hierarchical clustering was used to create a hierarchical decomposition of the dataset and find the hidden pattern in the data that share similar characteristics. The clustering result produced three groups/clusters. Finally, topic modeling was implemented on each cluster as a step to extract embedded themes in the tweets.

3.4 Topic Modeling

After extracting the three clusters from the collected tweets, the Latent Dirichlet Allocation (LDA) algorithm was applied using the LDAvis system (Jelodar et al., 2019; Sievert & Shirley, 2014). The LDAvis system was applied to each cluster in order to discover the latent topics in each tweet. LDA is unsupervised generative probabilistic method for modeling a corpus. It is the most popular topic modeling method. LDA allows presenting the documents as random mixtures over latent topics where a topic is explained by a distribution over data features (words). It presumes that each document can be represented as a probabilistic distribution over latent topics and that topic distribution in all documents share a common Dirichlet prior. Therefore, topic modeling was used as a step towards identifying suicide-related sentiment. To do so, three experts in the public health domain were asked to assess the content of each topic and provide proper label for its cluster (see the "Results of Topic Modeling" section) as step towards identifying suicide cluster. As a result, the experts suggested to label cluster 3 as "Suicide" category and both clusters 1 and 2 as "Non-suicide" category. Finally, sentiments associated with each category was extracted and examined to be used for the classification task; see the "Emotion Extraction" and "Classification task" sections.

3.5 Emotion Extraction

Users' sentimental features were identified and extracted from their textual data using NRC Affect Intensity Lexicon (Mohammad, 2017). This method is known for processing textual information in relation to users' mental conditions (Al-Samarraie et al., 2020). The NRC Affect Intensity Lexicon used in this study contained a list of English words that are linked with other associations. These words were adapted to classify tweets into four emotions: anger, fear, sadness, and trust using a scoring range from 0 (word conveys the lowest amount of emotion) to 1 (word conveys the highest amount of emotion). The extracted emotional features for each tweet were identified and labelled based a combination of words for a given lexicon. We applied the "SentiStrength" method in order to identify the polarity of the tweet using the Waikato Environment for Knowledge Analysis (WEKA) tool (Culpeper et al., 2018; Sarsam & Al-Samarraie, 2021; Thelwall, 2017). For each tweet, SentiStrength assigned scores ranging from " + 1" for "not positive" to " + 5" for "extremely positive" and " − 1" for "not negative" to " − 5" for "extremely negative." Based on these scores, we labeled the tweets with + 5 as "Positive" tweets, − 5 as "Negative" tweets, and − 1/ + 1 as "Neutral" tweets.

3.6 Classification Task

Four classification schemes were compared in this study to identify the best suicide classification (detection) method. These classifiers were *K*-nearest neighbor (IBk) (Sarsam, Al-Samarraie, & Alzahrani, 2021a), multinomial logistic regression (logistic) (Alafif et al., 2022), bagging (Al-Samarraie et al., 2016), and decision tree (Yaseen, 2021). These schemes were applied via WEKA tool (Al-Samarraie et al., 2016, 2018). In addition, stratified tenfold cross-validation method was applied to evaluate the overall prediction process based on the recommendation of Sarsam 2019. Several evaluation metrics were implemented to understand the prediction performance of the examined classifiers (see the "Classification Results" section). These metrics are accuracy, Kappa statistic, Root Mean Squared Error (RMSE), Receiver Operating Characteristic (ROC), and confusion matrix (Sarsam, Al-Samarraie, Alzahrani et al., 2021a, 2021b).

4. Results

4.1 Results of Topic Modeling

Figure 2 exhibits the result of the LDAvis tool for each cluster, where every circle rep-resents a specific topic from the collected tweets, while the size of a circle demonstrates the frequency of a topic. In addition, the distance between circles reflects the similarity between topics. From the figure, we can see that some topics are far apart (independent), whereas some topics are relatively close or even overlap (a high level of similarity).

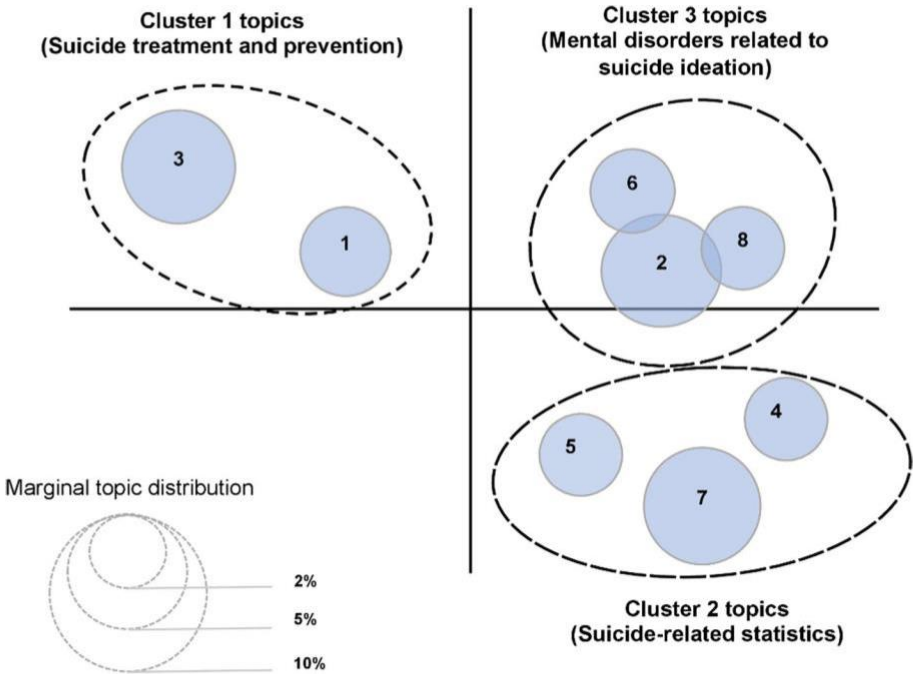

Figure 2: LDA results

The results of the topic modeling method implemented on the first cluster reveal that tweets in these groups are mostly concerned about providing advice on preventing suicide. Suicide treatment and prevention topics were discussed often, providing advice to help prevent suicides. For instance, many of these topics proposed engaging in psycho-social interventions, such as cognitive behavioral therapy (CBT). Additionally, some topics discussed accessing collaborative care as a step toward reducing suicidal thoughts. The second cluster, LDA results, shows that the tweets of this cluster focused on presenting statistical information about suicide, for example, that more than eight hundred thousand people die yearly because of suicide (i.e., one person every 40 seconds). Statistical information about the utilized suicide methods was also discussed in this cluster; for example, firearms were the most popular method utilized to commit suicide in the USA, where they annually contribute to more than 50% of all suicide deaths. Additional statistical information about suicide worldwide included, for example, that the suicide rate for men is twice as high as that for women. On the other hand, the LDA results on the third cluster show that these tweets were highly associated with mental disorder topics: (i) “depression,” (ii) “alcoholism,” and (iii) “schizophrenia disorders.” Regarding depression-related topics, the LDA results reveal that some of the people with depression who committed suicide had personal problems, such as the loss of a relationship or job.

On the other hand, suicide topics related to alcoholism indicated that individuals with suicidal thoughts often turn to alcohol, which increases suicidal thoughts when individuals drink alcohol before taking their lives. The topics also indicated that people use alcohol as a self-medication to treat, for example, anxiety or a personality disorder and to cope with a trauma. Finally, the topic modeling technique results related to schizophrenia disorders showed that schizophrenia patients are more likely to commit suicide if they have a family history of suicide or suffer from a long-term illness or chronic pain.

Once the topics were extracted from each cluster, we asked three experts in the public health domain to evaluate the content of each topic and provide proper labels for each cluster as a step toward identifying suicidal clusters. As a result, the experts suggested labeling Cluster 3 as the “Suicide” category and both Clusters 1 and 2 as the “Non-suicide” category. In light of this categorization, users’ sentiments associated with each category were extracted and examined to identify whether sentiments belonged to the suicide/non-suicide category during the classification task (see the “Results of the Emotion Extractions” and “Classification Results” sections).

4.2 Results of Emotion Extraction

To identify suicidal sentiments, suicidal polarity and emotions were extracted from the identified topics. SentiStrength was implemented to extract positive, negative, and neutral polarity types from each cluster. Then, we examined the temporal features of each topic, using the time series method over a period of 6 months (180 days) (see Figure 3). Finally, four types of emotions (anger, fear, sadness, and trust) were obtained via the NRC Affect Intensity Lexicon method. From this figure, it can be observed that each cluster has a specific type of sentiment that is dominant—the highest number of tweets carries a specific type of emotion in a particular cluster. We found that Cluster 1 and Cluster 3 contain positive and negative emotions, respectively, while neutral sentiments were found to be dominant in Cluster 2. In addition,

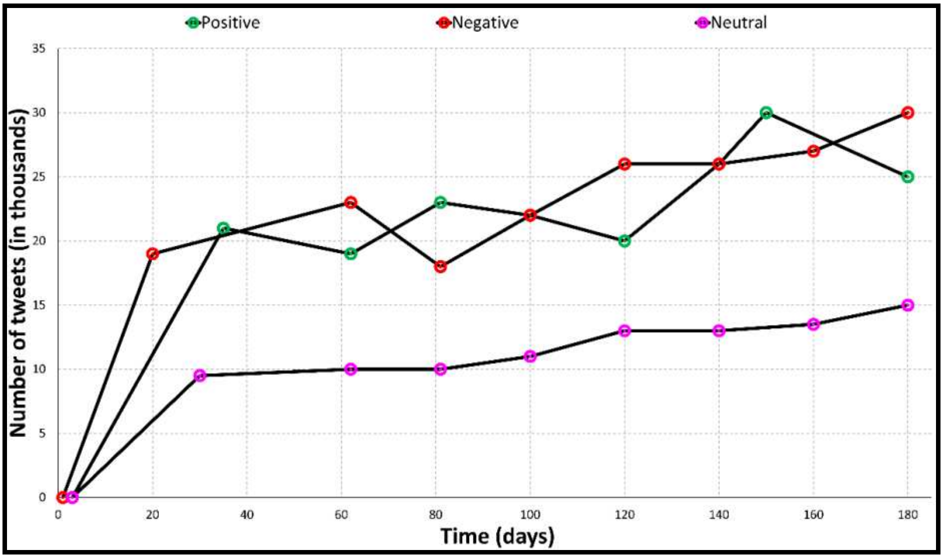

Figure 3: Time series results for users' sentiments in each cluster

the temporal features associated with both positive and negative sentiments are higher than those associated with neutral sentiments. Finally, suicide ideation-related tweets were assessed by three experts (the same experts involved in the topic modeling stage) to assess the relationship between suicide ideation as highlighted in the tweets and their relationship with polarity types. The experts concluded that the tweets that contain neutral sentiments have no relation with suicide ideation; thus, such sentiment was not recommended for identifying suicidal thoughts. Accordingly, we focused on analyzing only positive and negative sentiments via the NRC Affect Intensity Lexicon technique.

Four types of emotions were extracted, including anger, fear, sadness, and trust, via the NRC Affect Intensity Lexicon technique to understand suicidal and non-suicidal sentiments, and these results are summarized in Figure 4. From the figure, tweets from the suicidal group contained higher levels of fear (M = 1.67, SD = 0.14) and sadness (M = 1.98, SD = 0.20) than the fear (M = 0.11, SD = 0.08) and sadness (M = 0.07, SD = 0.03) in the tweets from the non-suicidal group. In contrast, non-suicidal tweets showed higher levels of anger (M = 0.68, SD = 0.18) and trust (M = 0.94, SD = 0.27) than suicidal tweets (anger (M = 0.10, SD = 0.02) and trust (M = 0.11, SD = 0.05)).

On the other hand, to assess the similarities and differences between tweets related (or not related) to suicidal and non-suicidal tweets, a *t* test was utilized, and the results (t = 39.22) showed that there was a significant difference (p< 0.05) between these two groups. As a result, non-suicidal tweets showed a high level of trust, anger, and positive sentiments. In contrast, fear, sadness, and negative sentiments were highly associated with suicidal statements.

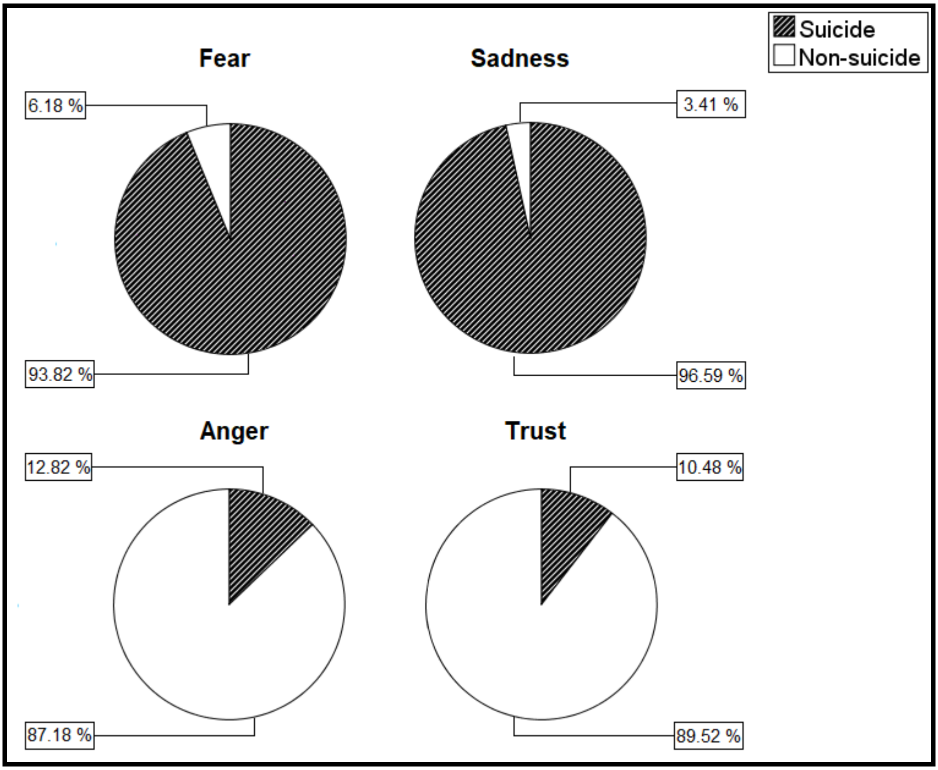

Figure 4: Emotion extractions

4.3 Classification Results

Our classification results are summarized in Table 1 and Figure 5, which reveal that the bagging classifier achieved the highest classification accuracy (97.64%), followed by the IBk (64.18%), J48 (56.82%), and logistic (49.43%) methods. Additionally, the classification results show that the bagging algorithm had the highest kappa statistic value (96%) compared to the IBk (68%), J48 (61%), and logistic (55%) methods. In contrast, the logistic classifier produced the highest RMSE value (91%), followed by the J48 (49%), IBk (35%), and bagging (3%) methods.

Table 1: Classification results

Learning algorithm	Accuracy (%)	Kappa statistic (%)	RMSE (%)
Bagging	97.64	96	3
IBk	64.18	68	35
J48	56.82	61	49
Logistic	49.43	55	91

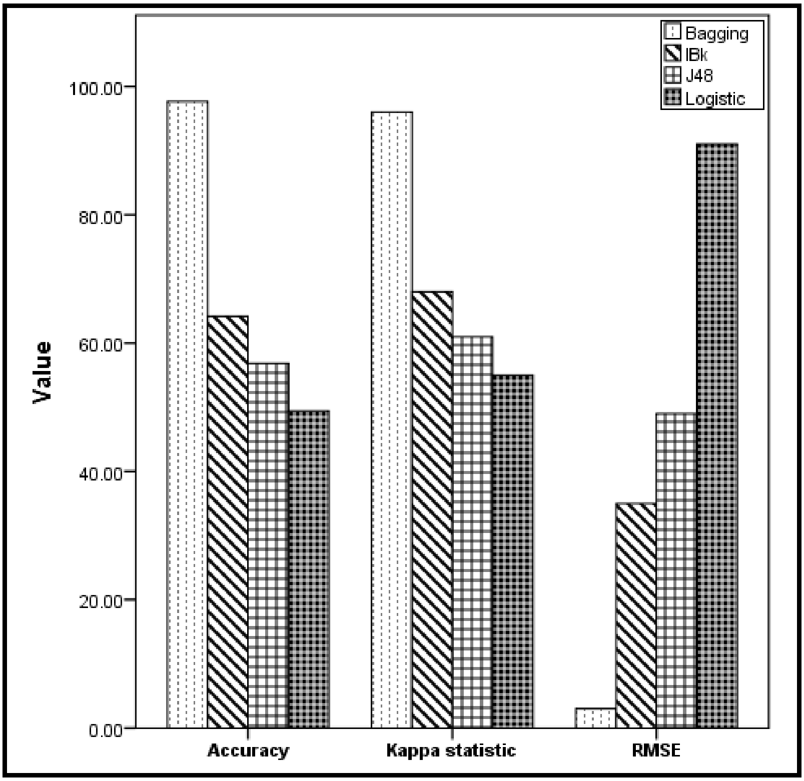

Figure 5: Evaluation metrics of the three algorithms

We further used the confusion matrix approach to evaluate the accuracy value of the four classifiers by assessing the relationship between the predicted and actual instances that are placed along the diagonal of the confusion matrix. The results (see Figure 6) reveal that the bagging classifier had the highest predictive capability between actual and predicted classes, i.e., 97.80% and 98% for the suicidal and non-suicidal categories, respectively. In conclusion, the bagging classifier had the best classification performance in detecting suicide ideation on the Twitter platform.

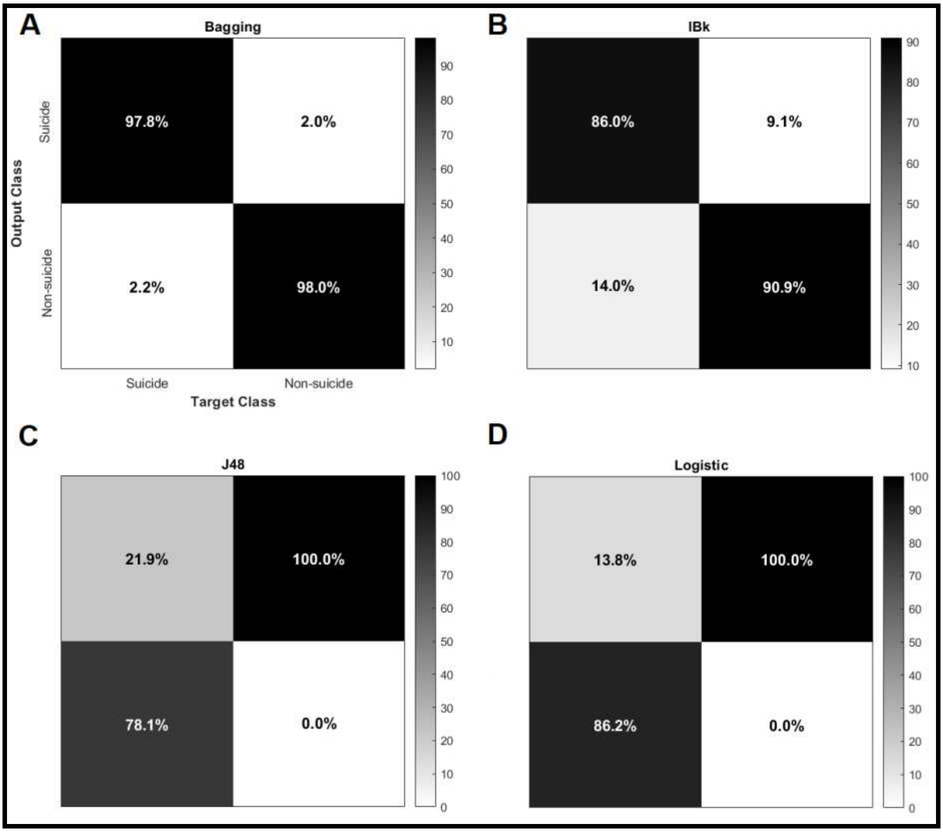

Figure 6: Confusion matrix results

5. Discussion

This study found that certain individual emotions embedded in Twitter messages have a great impact on the suicide classification process. The results showed that three main mental disorders, depression, alcoholism, and schizophrenia, are highly associated with suicidal behavior. In addition, the findings indicated that fear, sadness, and negative sentiments are highly linked to suicidal statements, whereas anger, trust, and positive sentiments are found to be related to non-suicidal content. The bagging classifier showed the highest prediction capability regarding suicidal statements compared to the other machine learning algorithms we examined. Furthermore, the LDA results indicate that mental disorder topics in tweets related to depression, alcoholism, and schizophrenia forcefully discuss suicide attempts. This could be attributed to the high probability of committing suicide among people with depression, alcoholism and schizophrenia; the risk of participating in suicidal and self-harming behaviors has been estimated to range between 5 and 8% for people with these mental disorders (Nordentoft et al., 2011). Depression is strongly related to both suicide ideation and attempts; adolescents with depression are six times more likely to commit suicide than non-depressed people (Chen et al., 2020). In addition to depression, suicide risk is correlated with high alcohol consumption, which might be related to the potential role of acute intoxication in increasing the risk of suicidal behavior, impairing judgment by reducing the inhibition of suicidal actions (Borges et al., 2017). In the existing literature, it can be observed that both depression and alcohol consumption are highly associated with suicide risk. According to Overholser et al. (2012), it is quite possible that the abnormal mental conditions in suicide victims are related to depressive and substance abuse disorders. This could explain our LDA results, which indicates the occurrence of depression and alcoholism in our data. The topic modeling results show that several Twitter messages that dis-cussed suicide attempts identified schizophrenia as one of the important factors influencing suicide deaths and attempts. The suicide risk among patients with schizophrenia can be related to delusions and hallucinations as well as disorganized thoughts. However, it has been found that suicide risk is considered to be higher among patients with schizophrenia during their first year of illness compared with those with long-term schizophrenia (Ven-triglio et al., 2016). The reason behind this could be due to the loss of trust among schizophrenic patients together with other negative emotions such as fear and lack of social confidence. Our sentiment analysis results showed that sadness is the most frequent type of emotion expressed in the relevant tweets (Pestian et al., 2012). This finding confirms the results of previous studies (e.g., Sarsam, Al-Samarraie, Alzahrani et al., 2021a, 2021b), which have found that fearful and negative sentiments are associated with suicidal behavior. This might be the reason for the strong relationship between fear and anxiety emotions found among suicide victims.

The obtained results from analyzing suicide-related tweets hold a number of implications for research and policy. Specifically, traditional methods for identifying suicide ideation among a specific population, which typically compare responses from the general public's suicide ideation scores across groups, may not be timely enough to capture suicide attempts and related behaviors. This might explain why the characterizations of suicidality through annotating and detecting emotions on Twitter can offer an alternative means for obtaining risk-related information.

In addition, identifying and using users' emotions can improve sentimental data analytics effectively and efficiently, thus enabling policy makers to better understand and plan for future interventions. This study also greatly contributes to suicide prevention by aiding current clinical decision support systems that implement a temporal risk profile for suicidal behaviors. Finally, the proposed mechanism can be applied to predict other psychiatric disorders on microblogs via the analysis of the user sentiments that are embedded in their online posts.

6. Limitations and Future Work

Despite the effectiveness of the proposed approach, it has some limitations. Tweets in English were collected and analyzed since it is the most popular communication language in the world. Additionally, specific emotion-related features were extracted in the current work, so in the future, other emotions could be explored, and their relationship with suicide ideation could be examined further. In this work, specific mental aspects were examined, but additional aspects could also be examined.

Finally, suicide was used here as an example of a mental health disorder due to its dangerous impact on people and society. Hence, future studies could apply the proposed mechanism to predict other psychiatric disorders by using the content of social media platforms.

7. Conclusion

We proposed a novel approach for suicide detection via users' sentiments existed in Twitter messages. We analyzed several types of users' polarity and emotions embedded in their tweets via SentiStrength, time series, and NRC Affect Intensity Lexicon methods, respectively. Then, bagging classifier was used to predict suicide-related content. Result showed that suicide messages contained low-level trust, anger, and positive sentiments. In contrast, fear, sadness, and negative sentiments were highly associated with suicide statements. Besides, our result indicated that bagging classifier (accuracy of 97.64%) has the best classification result in detecting suicide ideation. This study shows the significant role of individual emotions in predicting suicidal behavior on microblogs. The proposed method allows predicting suicide ideation from online platforms as a step toward boosting the decision-making process in characteristics mental disorders from online texts.

References

Abrutyn, S., & Mueller, A. S. (2014). Are suicidal behaviors contagious in adolescence? Using longitudinal data to examine suicide suggestion. *American Sociological Review*, *79*(2), 211-227.

Al-Samarraie, H., Sarsam, S. M., Alzahrani, A. I., & Alalwan, N. (2018). Personality and individual differences: the potential of using preferences for visual stimuli to predict the Big Five traits. *Cognition, Technology & Work*, *20*(3), 337-349.

Al-Samarraie, H., Sarsam, S. M., Alzahrani, A. I., & Alalwan, N. (2019). Reading text with and without diacritics alters brain activation: The case of Arabic. *Current Psychology*, 1-10.

Al-Samarraie, H., Sarsam, S. M., & Guesgen, H. (2016). Predicting user preferences of environment design: A perceptual mechanism of user interface customisation. *Behaviour & Information Technology*, *35*(8), 644-653.

Alafif, T., Alotaibi, R., Albassam, A., & Almudhayyani, A. (2021). On the prediction of isolation, release, and decease states for COVID-19 patients: A case study in South Korea. *ISA Transactions*.

Beck, A. T., Brown, G., & Steer, R. A. (1989). Prediction of eventual suicide in psychiatric inpatients by clinical ratings of hopelessness. *Journal of consulting and clinical psychology*, *57*(2), 309.

Bide, P., & Dhage, S. (2021). Similar Event Detection and Event Topic Mining in Social Network Platform. 2021 6th International Conference for Convergence in Technology (I2CT),

Borges, G., Bagge, C., Cherpitel, C. J., Conner, K., Orozco, R., & Rossow, I. (2017). A meta-analysis of acute alcohol use and the risk of suicide attempt. *Psychological medicine*, *47*(5), 949.

Canetto, S. S., & Lester, D. (1995). Gender and the primary prevention of suicide mortality. *Suicide and Life-Threatening Behavior*, *25*(1), 58-69.

Chen, R., Liu, J., Cao, X., Duan, S., Wen, S., Zhang, S., Xu, J., Lin, L., Xue, Z., & Lu, J. (2020). The relationship between mobile phone use and suicide-related behaviors among adolescents: the mediating role of depression and interpersonal problems. *Journal of affective disorders*, *269*, 101-107.

Culpeper, J., Findlay, A., Cortese, B., & Thelwall, M. (2018). Measuring emotional temperatures in Shakespeare's drama. *English Text Construction*, *11*(1), 10-37.

Douglas, J. D. (2015). *Social meanings of suicide* (Vol. 1242). Princeton University Press.

Ibrahim, N., Che Din, N., Amit, N., Ghazali, S. E., & Mohd Safien, A. (2019). Development and validation of Yatt Suicide Attitude Scale (YSAS) in Malaysia. *PloS one*, *14*(2), e0209971.

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, *78*(11), 15169-15211.

Law, K. C., Khazem, L. R., & Anestis, M. D. (2015). The role of emotion dysregulation in suicide as considered through the ideation to action framework. *Current Opinion in Psychology*, *3*, 30-35.

Lester, D. (2013). Suicide and culture. *Understanding suicide: A global issue*, 209-232.

Lewinsohn, P. M., Rohde, P., & Seeley, J. R. (1996). Adolescent suicidal ideation and attempts: prevalence, risk factors, and clinical implications. *Clinical Psychology: Science and Practice*, *3*(1), 25.

Mohammad, S. M. (2017). Word affect intensities. *arXiv preprint arXiv:1704.08798*.

Murni, N. I., & Ibrahim, M. A. (2020). Moderation Effect of Impulsivity in The Relationship Between Depression and Suicide Ideation Among Late Adolescent in Malaysia. *Sciences*, *10*(10), 110-119.

Nock, M. K., & Banaji, M. R. (2007). Prediction of suicide ideation and attempts among adolescents using a brief performance-based test. *Journal of consulting and clinical psychology*, *75*(5), 707.

Nordentoft, M., Mortensen, P. B., & Pedersen, C. B. (2011). Absolute risk of suicide after first hospital contact in mental disorder. *Archives of general psychiatry*, *68*(10), 1058-1064.

Okolie, C., Dennis, M., Thomas, E. S., & John, A. (2017). A systematic review of interventions to prevent suicidal behaviors and reduce suicidal ideation in older people. *International psychogeriatrics*, *29*(11), 1801-1824.

Overholser, J. C., Braden, A., & Dieter, L. (2012). Understanding suicide risk: Identification of high-risk groups during high-risk times. *Journal of clinical psychology*, *68*(3), 349-361.

Pestian, J. P., Matykiewicz, P., & Linn-Gust, M. (2012). What's in a note: construction of a suicide note corpus. *Biomedical informatics insights*, *5*, BII. S10213.

Rahimi, M., Naghi Zadeh Kakhki, E., Winter, S., & Stevenson, M. (2019). The Effectiveness of Sentiment Analysis for Detecting Fine-grained Service Quality.

Robinson, J., Yuen, H. P., Gook, S., Hughes, A., Cosgrave, E., Killackey, E., Baker, K., Jorm, A., McGorry, P., & Yung, A. (2012). Can receipt of a regular postcard reduce suicide-related behaviour in young help seekers? A randomized controlled trial. *Early intervention in psychiatry*, *6*(2), 145-152.

Sarsam, S. M. (2019). Reinforcing the decision-making process in chemometrics: Feature selection and algorithm optimization. Proceedings of the 2019 8th international conference on software and computer applications,

Sarsam, S. M., & Al-Samarraie, H. (2018a). A first look at the effectiveness of personality dimensions in promoting users' satisfaction with the system. *Sage Open*, *8*(2), 2158244018769125.

Sarsam, S. M., & Al-Samarraie, H. (2018b). Towards incorporating personality into the design of an interface: a method for facilitating users' interaction with the display. *User Modeling and User-Adapted Interaction*, *28*(1), 75-96.

Sarsam, S. M., & Al-Samarraie, H. (2021). Early-stage detection of eye diseases on microblogs: glaucoma recognition. *International Journal of Information Technology*, 1-10.

Sarsam, S. M., Al-Samarraie, H., & Al-Sadi, A. (2020). Disease discovery-based emotion lexicon: a heuristic approach to characterise sicknesses in microblogs. *Network Modeling Analysis in Health Informatics and Bioinformatics*, *9*(1), 1-10.

Sarsam, S. M., Al-Samarraie, H., & Alzahrani, A. I. (2021). Influence of personality traits on users' viewing behaviour. *Journal of Information Science*, 0165551521998051.

Sarsam, S. M., Al-Samarraie, H., Alzahrani, A. I., Alnumay, W., & Smith, A. P. (2021). A lexicon-based approach to detecting suicide-related messages on Twitter. *Biomedical Signal Processing and Control*, *65*, 102355.

Sarsam, S. M., Al-Samarraie, H., & Omar, B. (2019). Geo-spatial-based emotions: A mechanism for event detection in microblogs. Proceedings of the 2019 8th international conference on software and computer applications,

Selby, E. A., Anestis, M. D., Bender, T. W., & Joiner Jr, T. E. (2009). An exploration of the emotional cascade model in borderline personality disorder. *Journal of abnormal psychology*, *118*(2), 375.

Siau, C. S., Wee, L.-H., Ibrahim, N., Visvalingam, U., & Wahab, S. (2017). Cross-cultural adaptation and validation of the attitudes toward suicide questionnaire among healthcare personnel in Malaysia. *Inquiry: The Journal of Health Care Organization, Provision, and Financing*, *54*, 0046958017707295.

Sievert, C., & Shirley, K. (2014). LDAvis: A method for visualizing and interpreting topics. Proceedings of the workshop on interactive language learning, visualization, and interfaces,

Silverman, M. M. (2016). Challenges to defining and classifying suicide and suicidal behaviors. *The international handbook of suicide prevention*, *11*.

Sourirajan, V., Belouali, A., Dutton, M. A., Reinhard, M., & Pathak, J. (2020). A machine learning approach to detect suicidal ideation in us veterans based on acoustic and linguistic features of speech. *arXiv preprint arXiv:2009.09069*.

Spokas, M., Wenzel, A., Stirman, S. W., Brown, G. K., & Beck, A. T. (2009). Suicide risk factors and mediators between childhood sexual abuse and suicide ideation among male and female suicide attempters. *Journal of Traumatic Stress*, *22*(5), 467-470.

Stack, S., & Kposowa, A. J. (2016). Culture and suicide acceptability: A cross-national, multilevel analysis. *The Sociological Quarterly*, *57*(2), 282-303.

Sun, X., Zhang, C., Li, G., Sun, D., Ren, F., Zomaya, A., & Ranjan, R. (2018). Detecting users' anomalous emotion using social media for business intelligence. *Journal of Computational Science*, *25*, 193-200.

Thelwall, M. (2017). The Heart and soul of the web? Sentiment strength detection in the social web with SentiStrength. In *Cyberemotions* (pp. 119-134). Springer.

Ventriglio, A., Gentile, A., Bonfitto, I., Stella, E., Mari, M., Steardo, L., & Bellomo, A. (2016). Suicide in the early stage of schizophrenia. *Frontiers in psychiatry*, *7*, 116.

Weishaar, M. E., & Beck, A. T. (1992). Hopelessness and suicide. *International Review of Psychiatry*, *4*(2), 177-184.

Wu, P., Li, X., Shen, S., & He, D. (2020). Social media opinion summarization using emotion cognition and convolutional neural networks. *International Journal of Information Management*, *51*, 101978.

Yaseen, A. (2021). Data Classification Using Decision Trees J48 Algorithm for Text Mining of Business Data. *LGURJCSIT*, *5*(2), 71-79.