

Original Article

An Approach to Detect Suicidal Bengali Posts from Social Media Using Machine Learning Algorithms

Shaikat Chandra Paul¹, Busrat Jahan¹, Abdullah Al Mamun², Md Jakir Hossen³

¹Department of Computer Science and Engineering, Feni University, Feni, Bangladesh.

²School of Information and Communication, Griffith University, Brisbane, Australia.

³Faculty of Engineering and Technology, Multimedia University, Melaka, Malaysia.

³Corresponding Author : jakir.hossen@mmu.edu.my

Received: 26 October 2023

Revised: 01 February 2024

Accepted: 19 April 2024

Published: 26 May 2024

Abstract - In this modern era, suicide is one of the critical issues. According to the WHO, more than seven million people die due to suicide every year. Suicide is also the second cause of unnatural death for persons between the ages of 15 and 29. Youth in nations like Bangladesh struggle with schoolwork, employment, relationships, drug use, and family issues, all of which are significant or minor contributors on the road to depression. In Bangladesh, people are uncomfortable discussing this ailment openly and frequently mistake this problem as madness. Many at-risk persons use social platforms to talk about their issues or get knowledge on related topics. This study aims to prevent suicide by identifying suicidal posts on social media. We collected suicidal-related data from Kaggle. (We use nine algorithms for three features). The prediction model achieved good performance. Stochastic Gradient Descent is the best model with the highest accuracy for unigram features, 87.23%. For bigram features, the Multinomial Naive Bayes is the best model with the highest accuracy, 88.69%. The best model with the highest accuracy for trigram features, 86.13%, is Stochastic Gradient Descent. This research demonstrates the chance that a machine-learning strategy can reduce the risk of suicide. Hopefully, this model will serve as a guide for lowering potential suicide risk in the future. The study concludes with a summary of several practical concerns that may be considered to improve model performance.

Keywords - NLP, N-grams, Machine Learning Approaches, Social Media, Bengali Post.

1. Introduction

Our beloved world is doing great with the help of modern technology and skilful people. We are still trying to increase the productivity of modern technology because modern technology has made our lives much easier. Because of this, we are becoming busier and busier, and most of the time, we are not thinking about our mental health. By ignoring that situation, we are taking more pressure both physically and mentally, and that causes us a significant problem called depression. Sometimes, this depression leads one to commit suicide. Suicide is death caused by injuring oneself with the intent to die. More than seven million people die due to suicide every year [1]. Suicide is the fourth leading cause of death among 15–19-year-olds. Most of the people who commit suicide are mentally disturbed (Depression and alcohol). In recent years, a common behaviour has been shown in many of them. They are posting something on social media before committing suicide. Many reputed media already cover this trend [2][3][4]. The young generation is more active on social media platforms such as social networking sites (e.g., Facebook, Twitter, Google+), Video Sites (e.g. YouTube), Blogging Websites (e.g., Blogspot),

and electronic bulletin boards, as well as text messaging, e-mail, and video chat [7]. People are constantly using social media as a way of communication. Among them, Facebook is the biggest social networking media having more than 2.7 billion monthly active users [8]. In Bangladesh, about 90% of social networking media users are using Facebook. The general purpose of using these social media is to communicate with friends and family through sharing posts or messaging with them. It is shown that a Facebook user normally has 338 friends and spends about 1 hour or 60 minutes every day [9]. According to the WHO, every year, almost 1 million youth commit suicide [1]. If we dig further, we will see the actual number of suicides is much higher than the published report because, in some situations, the case is not reported due to several factors like social-cultural, political issues and economic motives [12][13]. It has now become a major problem in this modern era. It is declared by who is behind a successful suicide; there are at least 20 suicide attempts [14]. Many people post various ideas, processes and plans on social media to commit suicide [15]. This trend is now becoming a tough challenge for the authority [16]. In recent years, the tendency to commit



suicide has increased rapidly, especially during the time of the COVID-19 pandemic, because of fear and uncertainty. In most cases, these people are suffering from different mental causes. The total number of this type of patient is about 15 million [17]. From the above things, we can assume that the tendency to commit suicide will increase in the coming days. People will be more open to social media before committing suicide. Some studies show that there are many suicidal posts found on social media, but our authority seems not to have taken any remarkable steps to handle this type of situation.

Even if the world is changing and everyone accepts novel ideas, mental illness is still not well two received in many civilizations. It appears to be taboo, particularly in many Asian cultures, which is why those suffering from it are less outspoken about their issues, conceal their disease, and attempt to get better on their own. As a result, their sickness may get worse, which might impair both their physical and mental health and cause them to experience extreme depression, which will, in turn, make individuals more likely to engage in suicidal thoughts and behaviours. Suicide is now a serious issue for our society, the country, and the world. We can even define it as a curse for people, society, and our country. When people commit suicide, many people bear the loss. Many families have collapsed due to their main earning member committing suicide. A two-stage study was conducted to honour suicidal ideation and found that the average life expectancy of suicidal ideation was 5 per cent in young adults. Each year, many meritorious students lose their lives as a result of suicide. The rate of suicide is increasing day by day, and we need proper steps to stop it. Sometimes, psychiatrists help people to deal with depression and suicidal thoughts. However, most people do not want to talk about their problems in person. Some of them indirectly post on different social media about their problem.

The main objectives of the study are:

- To predict suicidal posts.
- To Classify sentiment into-
- Positive.
- Suicidal Post.
- To Find the accuracy of various algorithms for the dataset.
- To Compare the different models to find out which one got more accuracy.

2. Literature Review

The use of machine learning techniques may be used to create automatic processes. Data gathering and goal setting are the initial steps in creating a machine learning model, which involves appropriately characterizing the problem. Machine learning is in the modelling phase. By 2022, the death toll might reach 1.5 million, according to estimates. The frequency rapidly increases, and you must halt properly. Psychologists frequently assist patients in overcoming

depression, as well as any suicidal thinking or thoughts. Too many authors who submit work fail to discuss their flaws. However, their suicidal ideas change their conduct, and the majority of them end their lives. Many parents are ignorant of the health issues affecting their kids. Some work has been done on this topic. Some of them are discussed below. This research would help to predict whether their status is suicidal or not. By compiling resources related to suicide, some researchers want to alleviate the dearth of terminological materials on the subject. This study also suggests looking at Weka as a data mining tool for a more thorough investigation. They concluded their investigation using machine learning algorithms that can extract useful information from the Twitter data that Twitter has gathered. The outcome of cross-validation of classifier assessments for tweets assuming suicide risk was naïve bias 87.50% [19].

According to this research, suicide attempts are classed as schizophrenia. Based on a machine learning strategy, they use sociocultural and medical aspects. They presented a cross-sectional estimate based on a sample of 345 individuals who have been diagnosed with schizophrenia spectrum illnesses. Using the CSSRS and the BSS, they were able to distinguish between suicide attempters and non-attempters. To train the models, they made use of four classification methods. The Support vector machine provided the best accuracy after training with their dataset, at 67% [20].

Some researchers use machine learning techniques to develop prediction models for the high suicide risk among Korean teenagers. The goal of this study was to use machine learning approaches to enhance the prediction model for identifying Korean teenagers who are at high risk of suicide. They made use of a dataset from the Korean Youth Risk Behavior Web-based Survey that was nationally representative. With the use of five machine learning methods, they were able to attain the best accuracy for the XGB model, which is 79.0% [21].

One study paper attempted to use a machine learning algorithm to create a model that might predict people who had suicidal thoughts. The prediction model performed better, with an AUC of 0.85% [22]. Based on machine learning, a study forecasts the future risk of suicidal thoughts. They collected information from Twitter. On their gathered dataset, they trained several neural networks toward psychological factors connected to Suicide. Using Neural Network results, a random forest model was trained to predict a binary state of suicidal ideation. An AUC of 0.88% was attained in this study [23]. A study article suggested that by using population-based longitudinal data, machine learning may be used to predict future suicidal behaviour. The random forest algorithm was shown to be the best algorithm for predicting suicidal thoughts in this study. The most accurate method for predicting suicide attempts is gradient boosting [24].

Critical risk variables for suicide attempts and mortality were evaluated in a study. They found that individuals who engaged in suicide conduct had higher results for each of the five symptoms ($p < 0.05$) [18]. To determine the diagnostic effectiveness of suicide prediction models by predicting future suicide and suicide attempts, one study paper conducted a systematic review and simulation. Reviewing

7306 abstracts, 17 group studies, and 64 different prediction models led to a higher global classification accuracy [10]. Based on BPNN, a study developed a prediction model for a suicide attempt. Following implementation, the research attained the following metrics: specificity (93.9%), sensitivity (67.6%), negative predictive value (84.1%), and positive value (86.0%) [6].

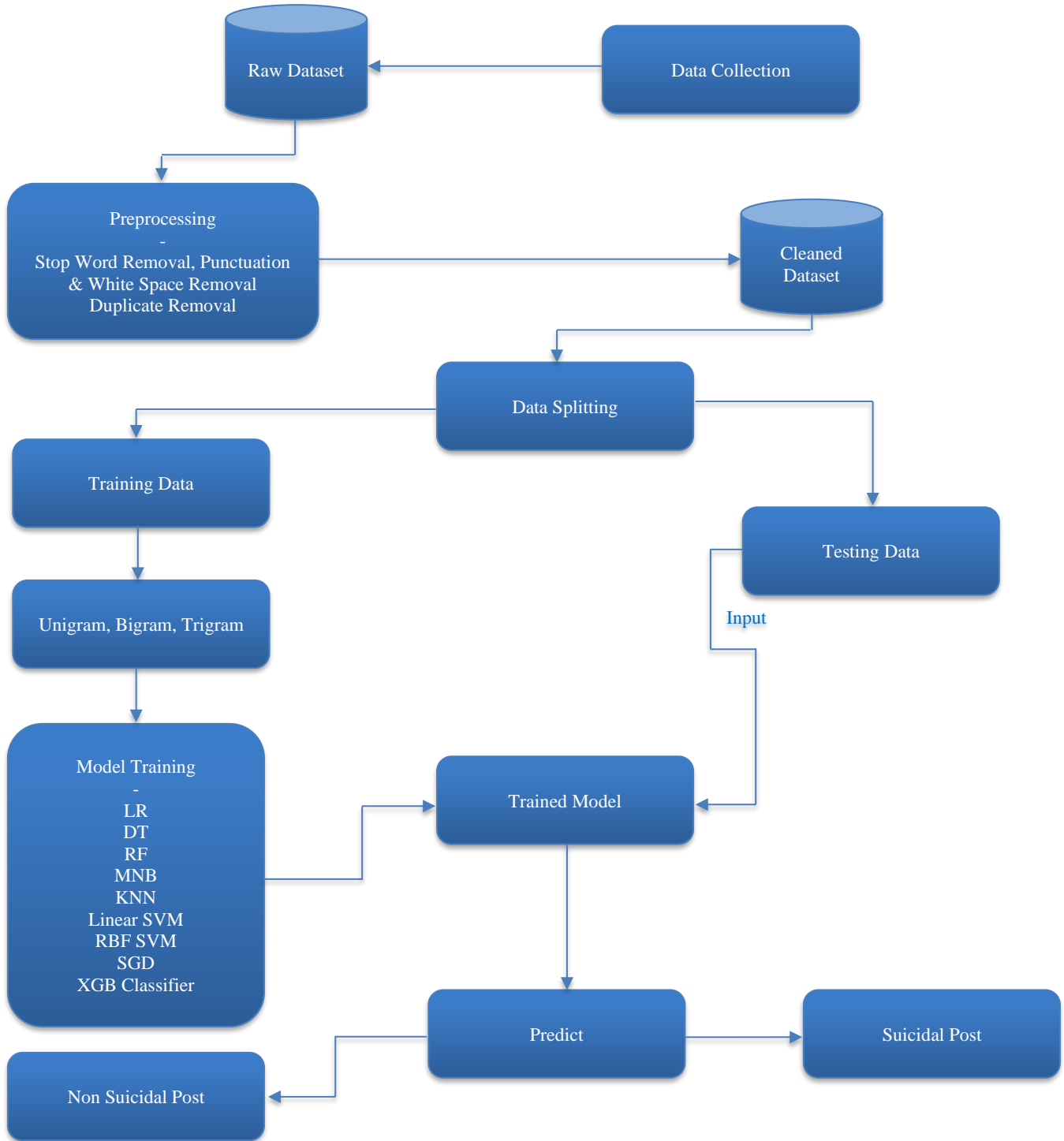


Fig. 1 Process flow diagram of this system

Table 1. Dataset sample

Index	Data	Label
0	আমার নিজের পরিবার আমরা সাথে এমন ধোকা বাজি করবে কখনো ভাবি নি। ভালো থাকুন আপনারা। সুখে থাকুন। সব ছেড়ে দিলাম আপনাদের জন্য।	suicidal_post
1	পিজাটা চমৎকার ছিল। আমি এবং আমার বন্ধুদের এটি পছন্দ হয়েছে।	positive
2	আজ আমার বোন,কাল আপনার বোন, পরে আপনার মেয়ে,তার আরেকজন এর পালা, এই ভাবে চলতে থাকলে, আমাদের বোনেরা,নেয়েরা লেখা করতে পারবেনা,নিরাপদে,রাস্তায় চলা ফিরা করতে পারবেনা, আসুন সবাই মিলে প্রতিবাদ করি।এর দায়িত্ব শুধু কুমিল্লাহ র ছাত্র- ছাত্রী ভাই,বোন দেব না। আমার, আপনার, সকলের। আমি মনে করি বাংলাদেশের সকল শিক্ষা প্রতিষ্ঠানের ছাত্র,ছাত্রী,শিক্ষক,কর্মচারী,অভিবাবক,সচেতন নাগরিক মাত্রই প্রতিবাদ করা উচিত।	positive
3	তনু হত্যাকারীদের অবিলম্বে গ্রেফতার করে দৃষ্টান্তমূলক শাস্তি কমনা করছি যাতে করে আর কেউ এমন নেঙ্কারজনক ঘটনা ঘটাতে আর সাহস করতে না পারে, তাই এই ব্যাপারে মাননীয় প্রধানমন্ত্রীর দৃষ্টি কামনা করছি।	positive
4	কেন আমায় ধোঁকা দিলে? কি দোষ করেছি আমি? কেন ছলনা করলে আমার সাথে ? আমিতো শুধু চেয়েছিলাম তোমার সাথে ভালো থাকতে। এটা কি আমার অপরাধ ছিল?	suicidal_post
6	আমার বেঁচে থাকার ইচ্ছে নেই। এত কষ্ট আর সহ্য হয় না। আমার মৃত্যুর জন্য কেউ ধায়ী নয়।	suicidal_post

Machine Learning is the process through which experience is transformed into knowledge or skill. In machine learning techniques, training data are inputs that reflect the experience and are used to train and learn the algorithm. The outcome of the experiment resembles knowledge or expertise that was learnt with the aid of the training examples [5]. The best sets of the function $h: X \rightarrow Y$ [11] should be aimed for in mathematics.

The array of objects in the domain that we wish to label (X). Vector features can represent the domain set. For instance, mango colour and softness may be used as vector features to decide whether a mango is tasty. This is the outcome that an algorithm produces for the specified task. For instance, a machine must be able to determine if a mango tastes well. So let Y be 0, 1. where 1 stands for delicious and 0 for not tasty. Label sets may also not be binary.

$S = ((x_1, y_1) \dots (x_m, y_m))$ is the training Data. The model is trained using the provided set of input and associated output. For instance, if x_1 is a mango, y_1 indicates whether or not that fruit is tasty. The same applies to x_2, y_2 , and so on till x_m, y_m .

The learner's output is h , a function known as the predictor and the optimal collection of functions that can reasonably predict the outcome. Finding the optimal collection of predictors is done by $h: X \rightarrow Y$, where X is the set of potential inputs, and Y is the set of potential outcomes for those inputs. Success is measured by the rate of incorrectly forecasting the outcome. The error rate in the mango example fails to accurately predict if a mango is tasty or not [5].

The goals of this paper are to develop a suicidal post-prediction model and to compare classifier algorithms in order to select the best one. To get the highest performance for this predicting model, we have applied supervised learning classifier approaches in our research.

First, we have prepared a dataset with all the information we need for the study. A solid dataset is necessary for any machine learning model to build an efficient automated system. A better dataset was necessary for every machine learning model in order to create an accurate automated system. We have utilized a variety of classification techniques to build this model.

3. Proposed System Models

In this phase, our proposed methodology is shown in Fig. 1. We follow the steps shown in Fig. 1 to complete this work.

3.1. Data Collection

The collection of data is the most important for this work. We took the dataset from Kaggle. The dataset is made up of genuine data that are collected from several social media networking sites.

3.2. Data Description

Table 2 shows the summary of the dataset. Here, the dataset contains 3657 posts, where the number of suicidal posts is 1467 and the number of positive posts is 2190.

Table 2 . The Summary of the dataset

Total Data	Suicidal Post	Positive
3657	1467	2190

3.3. Data Preprocessing

The adjustments made to our data prior to feeding it to the algorithm are referred to as pre-processing. Data pre-processing is a method for transforming unclean data into clean data sets.

In other words, when data are received from various sources, they are gathered in raw form, which makes analysis impossible. Data pre-processing is crucial since it gets the data ready in the most useful way for the ensuing in-depth analysis. Emoji, punctuation, white space, commas, and other symbols are included in the raw dataset and are not ideal for machine learning to process.

Pre-processing the data set can be done using a variety of methods. Pre-processing plays a crucial part in machine learning data analysis since it can increase the accuracy of the results when done properly before the analysis. In our research, we processed our data set for analysis using some main pre-processing processes. These include removing numbers, whitespace, stop words, a single character, and a new line.

3.3.1. Punctuations and Special Character Remove

When using text in Natural Language Processing, cleaning the text or pre-processing is necessary (NLP). We need to clean this sort of noisy text data before supplying it to the machine learning model since real-life human-written text often contains words with incorrect spellings, short words, special symbols, emojis, etc. Special characters like [!"#\$%&'()*+,-./:;>?@[_`] are eliminated from our data since they do not have any significant value for our sentiment analysis work.

3.3.2. Whitespace Removal

The online platform does not have any set guidelines for how a user should compose a post or remark. As a result, when writing, individuals frequently pay little attention. The user frequently uses extraneous whitespace in their post content, for example, “আমি আগেই জানতাম, আমরাই নিতবো” in this post has too much uncommon whitespace, making it challenging to interpret the context for our model. Our analysis methods do not benefit from leading or trailing spaces. Therefore, removing it improves the accuracy of our dataset for our algorithm. We eliminated all the extra white space from posts using a built-in Python method.

3.3.3. Stop-word Removal

Any human language has a variety of stop words. By getting rid of these terms, we can make our text more focused on the key information by eliminating the low-level information. In other words, we may state that the model we train for our task does not exhibit any detrimental effects as a result of the removal of such phrases. The most frequent words in a language are stop words, such as “আপার” ,

“আপনি” , “আবার” , “আমরা” , “আমাকে” , “আমাজের” , “আমার” , “আমি” , “আর” and “আরও” etc. In literature, these terms are typically eliminated because they are less important. Stop words should be eliminated since they increase the dataset size and lengthen training time because there are fewer tokens to train on.

4. Results Analysis

According to our methodology, this distinguishes a post from all the other posts that are genuinely suicidal. Next, we will analyze each method’s accuracy and f-measure to determine which algorithm will best fit our model.

Accuracy in classification issues refers to the overall number of predictions made by the system out of all possible guesses (A_γ).

$$A_\gamma = \frac{TP + TN}{TP + FP + FN + TN} \times 100\%$$

Where TP stands for “True Positive,” TN for “true negative,” FP for “False Positive,” and FN for “false negative.” The denominator includes all of the algorithm’s predictions, both true and false, whereas the numerator only includes the accurate predictions (true positive and true negative).

Precision is the measuring metric that counts both genuine positive and false positive, positive patterns that are accurately predicted from all positive predictions (P_γ).

$$P_\gamma = \frac{TP}{TP + FP} \times 100\%$$

The number 1 indicates the greatest accuracy score, while the value of 0 indicates the poorest precision score. Recall is the parameter that estimates the percentage of positive patterns that are accurately categorized (R_γ).

$$R_\gamma = \frac{TP}{TP + FN} \times 100\%$$

When the recall is 1, we can claim that the result is the most accurate; yet, when the recall is 0, we can say that the result is the least accurate.

By definition, the F1-score is the harmonic mean of precision and recall. It combines precision and recalls into a single number using the following formula:

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \times 100\%$$

Notice that the F1-score takes both precisions and recall into account, which means it accounts for FPs and FNs. The higher the precision and recall, the higher the F1 score. F1-score ranges between 0 and 1. The closer it is to 1, the better the model.

Table 3. Performance table of all algorithms for unigram feature

Algorithms	Accuracy	Precision	Recall	F1 Score
LR	82.48	76.74	94.29	84.62
DT	82.12	80.13	86.43	83.16
RF	83.21	79.75	90.00	84.56
MNB	84.67	92.24	76.43	83.59
KNN	75.18	91.86	56.43	69.91
Linear SVM	79.20	72.43	95.71	82.46
RBF SVM	79.20	72.43	95.71	82.46
SGD	87.23	84.31	92.14	88.05
XGBClassifier	82.85	80.39	87.86	83.96

Table 4. Performance table of all algorithms for bigram feature

Algorithms	Accuracy	Precision	Recall	F1 Score
LR	85.40	80.86	93.57	86.75
DT	82.12	80.54	85.71	83.04
RF	85.77	84.35	88.57	86.41
MNB	88.69	90.37	87.14	88.73
KNN	79.56	82.81	75.71	79.10
Linear SVM	82.85	76.88	95.00	84.98
RBF SVM	84.31	79.39	93.57	85.90
SGD	88.32	86.49	91.43	88.89
XGBClassifier	81.02	77.85	87.86	82.55

Table 5. Performance table of all algorithms for trigram feature

Algorithms	Accuracy	Precision	Recall	F1 Score
LR	79.20	72.19	96.43	82.57
DT	81.39	79.87	85.00	82.35
RF	83.21	78.66	92.14	84.87
MNB	81.75	93.27	69.29	79.51
KNN	76.64	94.19	57.86	71.68
Linear SVM	67.15	61.06	98.57	75.41
RBF SVM	78.10	71.51	95.00	81.60
SGD	86.13	82.28	92.86	87.25
XGBClassifier	81.75	78.48	88.57	83.22

Table 3 shows the performance of All Algorithms for Unigram Feature. Though the accuracy of all algorithms is fine, we can see that the Stochastic Gradient Descent is the best, with a higher accuracy of 87.23%. The results are represented in percentage format using various metrics that help to compare the performance results with other models.

Table 4 shows the performance of different algorithms for Bigram Feature. Though the accuracy of all algorithms is fine, we can see the Multinomial Naive Bayes is the best, with a higher accuracy of 88.69%. The results are represented in percentage format using various metrics that help to compare the performance results with other models.

Table 5 shows the performance of different algorithms for Trigram Feature. Though the accuracy of all algorithms is fine, we can see that the Stochastic Gradient Descent is the best, with a higher accuracy of 86.13%. The results are

represented in percentage format using various metrics that help to compare the performance results with other models.

5. Conclusion

Depression and suicidal thoughts are the most common mental illnesses or conditions, and this large area of inquiry has many consequences for medicine and psychology. Due to different socio-demographic factors, young individuals, especially those who live in lower-middle-income nations, are more susceptible to contracting this disease. Depression has roots that are deeply embedded in lifestyle, culture, routines, and behaviour. Analyzing suicidal ideas among Bangladeshi residents has not received significant attention. However, given that it is a significant social and medical issue that is only worsening, it must be dealt with immediately. The bulk of recent increases in suicide rates across the nation, according to newspaper publications, was brought on by depression.

Data processing might be done using a variety of sophisticated techniques, and the model could be displayed beautifully by varying how algorithms are used. By using more diverse algorithms, it is feasible to improve accuracy. Given that so many people today commit suicide for various reasons, the issue of posting on social media before committing suicide has become crucial. This upward trend in suicide must become a decreasing trend.

Because of this, we have created a model that can determine whether or not a person is considering suicide. To determine the best method for identifying a person's suicidal thoughts, a comparison examination of several classifier algorithms is conducted. This thesis applied the supervised classifier technique to predict a person's suicidal post. Out of 9 classifier techniques, we found the "Multinomial Naive

Bayes" strategy to be the most dependable and effective one, with an accuracy level of 88%. This model will aid in reducing the effects of people's continued growth and their evolving ability to live with the ecosystem in the interest of a sustainable world.

Acknowledgements

Thanks to Multimedia University.

Author Contributions

Both Saikat Chandra Paul and Busrat Jahan have finished all of the tests, as well as the design and drafting of the article. Busrat Jahan and Abdullah Al Mamun have acted as research supervisors for this project. For the sake of this study, model tuning and paper editing were carried out by Abdullah Al Mamun and Md Jakir Hossen, respectively.

References

- [1] Suicide, World Health Organization, 2023. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/suicide>
- [2] Karen McVeigh, Facebook 'Friends' did not Act on Suicide Note, The Guardian, 2011. [Online]. Available: <https://www.theguardian.com/technology/2011/jan/05/facebook-suicide-simone-back>
- [3] Andreas M. Kaplan, and Michael Haenlein, "Users of the World, Unite! The Challenges and Opportunities of Social Media," *Business Horizons*, vol. 53, no. 1, pp. 59-68, 2010. [CrossRef] [Google Scholar] [Publisher Link]
- [4] Thomas D. Ruder et al., "Suicide Announcement on Facebook," *Crisis*, vol. 32, no. 5, 2011. [CrossRef] [Google Scholar] [Publisher Link]
- [5] Shai Shalev-Shwartz, and Shai Ben-David, *Understanding Machine Learning: From Theory to Algorithms*, Cambridge University Press, pp. 1-397, 2014. [Google Scholar] [Publisher Link]
- [6] Juncheng Lyu, and Jie Zhang, "BP Neural Network Prediction Model for Suicide Attempt among Chinese Rural Residents," *Journal of Affective Disorders*, vol. 246, pp. 465-473, 2019. [CrossRef] [Google Scholar] [Publisher Link]
- [7] Dave Chaffey, Global Social Media Statistics Research Summary, Smartinsights, 2024, [Online]. Available: <https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/>
- [8] Number of Monthly Active Facebook Users Worldwide as of 4th Quarter 2023, Statista. [Online]. Available: <https://www.statista.com/statistics/264810/number-of-monthly-active-facebookusers-worldwide/>
- [9] Social Media Stats Bangladesh, Statcounter, 2024. [Online]. Available: <https://gs.statcounter.com/social-media-stats/all/bangladesh>
- [10] Bradley E. Belsher, Derek J. Smolenski, and Larry D. Pruitt, "Prediction Models for Suicide Attempts and Deaths A Systematic Review and Simulation," *JAMA Psychiatry*, vol. 76, no. 6, pp. 642-651, 2019. [CrossRef] [Google Scholar] [Publisher Link]
- [11] Pierre Lison, "An Introduction to Machine Learning," University of Oslo, pp. 1-35, 2012. [Google Scholar] [Publisher Link]
- [12] George C. Patton et al., "Global Patterns of Mortality in Young People: A Systematic Analysis of Population Health Data," *The Lancet*, vol. 374, no. 9693, pp. 881-892, 2009. [CrossRef] [Google Scholar] [Publisher Link]
- [13] Soron Tanjir Rashid, "Decriminalizing Suicide in Bangladesh," *Asian Journal of Psychiatry*, vol. 39, pp. 91-92, 2019. [CrossRef] [Google Scholar] [Publisher Link]
- [14] Mental Health and Substance Use, World Health Organization. [Online]. Available: <https://www.who.int/teams/mental-health-and-substance-use/data-research/suicide-data>
- [15] Sean D. Young, and Renee Garrett, "Ethical Issues in Addressing Social Media Posts About Suicidal Intentions During an Online Study Among Youth: Case Study," *JMIR Mental Health*, vol. 5, no. 2, pp. 1-5, 2018. [CrossRef] [Google Scholar] [Publisher Link]
- [16] David D. Luxton, Jennifer D. June, and Jonathan M. Fairall, "Social Media and Suicide: A Public Health Perspective," *American Journal of Public Health*, vol. 102, no. S2, pp. S195-S200, 2012. [CrossRef] [Google Scholar] [Publisher Link]
- [17] Rafia Tasnim et al., "Suicidal Ideation among Bangladeshi University Students Early during the COVID-19 Pandemic: Prevalence Estimates and Correlates," *Children and Youth Services Review*, vol. 119, pp. 1-8, 2020. [CrossRef] [Google Scholar] [Publisher Link]
- [18] Elizabeth D. Ballard et al., "Acute Risk Factors for Suicide Attempts and Death: Prospective Findings From the STEP-BD Study," *Bipolar Disorders*, vol. 18, no. 4, pp. 363-372, 2016. [CrossRef] [Google Scholar] [Publisher Link]
- [19] Marouane Birjali, Abderrahim Beni-Hssane, and Mohammed Erritali, "Machine Learning and Semantic Sentiment Analysis Based Algorithms for Suicide Sentiment Prediction in Social Networks," *Procedia Computer Science*, vol. 113, pp. 65-72, 2017. [CrossRef] [Google Scholar] [Publisher Link]
- [20] Nuwan C. Hettige et al., "Classification of Suicide Attempters in Schizophrenia Using Sociocultural and Clinical Features: A Machine Learning Approach," *General Hospital Psychiatry*, vol. 47, pp. 20-28, 2017. [CrossRef] [Google Scholar] [Publisher Link]

- [21] Jun Su Jung et al., “Prediction Models for High Risk of Suicide in Korean Adolescents Using Machine Learning Techniques,” *PLoS One*, vol. 14, no. 6, pp. 1-12, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [22] Seunghyong Ryu et al., “Use of a Machine Learning Algorithm to Predict Individuals with Suicide Ideation in the General Population,” *Psychiatry Investigation*, vol. 15, no. 11, pp. 1030-1036, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [23] Arunima Roy et al., “A Machine Learning Approach Predicts Future Risk to Suicidal Ideation from Social Media Data,” *NPJ Digital Medicine*, vol. 3, no. 78, pp. 1-12, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [24] Kasper van Mens et al., “Predicting Future Suicidal Behaviour in Young Adults, with Different Machine Learning Techniques: A Population-Based Longitudinal Study,” *Journal of Affective Disorders*, vol. 271, pp. 169-177, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]