# ANALYZING THE ACCURACY RATE FOR SUICIDAL TWEET DETECTION USING SEQUENTIAL MINIMAL OPTIMIZATION OVER NAIVE BAYES

**Ashvika Venkatesan[1], S. Magesh Kumar[2*]**

**Article History: Received:** 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

**Abstract**

**Aim:** The objective of the work is to detect and determine tweets that indicate suicide ideation using SMO over Naive Bayes. To achieve accuracy, a novel SMOClassifier function was used. **Materials and Methods:** Accuracy and Loss are performed with SUICIDAL_DATA dataset from the Github library. The total sample size is 242. The two groups considered were Sequential Minimal Optimization and Naive Bayes. **Result:** The accuracy of SMO is 93.5% and loss is 6.5%, which appears to be better than Naive Bayes whose accuracy is 81.7% and loss is 18.3%. Finally, SMO appears significantly better than the Naive Bayes algorithm. The two algorithms, Sequential Minimal Optimization (SMO) and Naive Bayes (NB) with independent sample T-Test value achieved is $p=0.662$ ($p>0.05$), statistically insignificant.. **Conclusion:** Detecting suicidal tweets significantly seems to be better in Sequential Minimal Optimization (SMO) than Naive Bayes (NB).

**Keywords**: Machine Learning, Suicide Ideation, Sequential Minimal Optimization, Naive Bayes, Novel SMOClassifier, Twitter, Social Media.

[1]Research Scholar, Department of Computer Science Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India. Pincode: 602105.

[2*]Department of Computer Science Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India. Pincode: 602105.

Email: [1]venkatesana18@saveetha.com

## 1. Introduction

In this modern age, about half of the world's population dwells on the internet. People are engaged very much in a virtual world such as social media like Twitter. This platform is used extensively to share their feelings, sentiments and even cryptic messages (Astoveza et al. 2018). The importance of this report is that it investigates online social content for early detection of suicidal ideation. With the growing issue of suicide Ideation at hand and as social media sites grow in popularity, this platform is increasingly being recognized for identifying those at a danger of suicide (Fodeh et al. 2019). The applications such as Suicidal Tweet Detection machine learning model helps in identifying and providing professional mental health care to those in need. This system can be used by Non-governmental organizations (NGOs) or psychologists to monitor suicidal persons that have a past history of suicide attempts (Luxton 2015). Another application of the suicidal tweet detection model is that it can be used for the study and research of suicidal behaviour in people (Biernesser et al. 2021).

Suicidal tweets detection is carried out by researchers and there are 8 related research articles on it in IEEE digital Xplore and 8 articles are published in Research Gate. Social networking sites have gained the attention of the research community that tries to understand, among other, their structure and user interconnection, as well as interactions among users. People tend to express their feelings and tell about their daily life activities through Twitter (Li et al. 2012). 86% of users of social media are young users in the age group 18-29. These young people are more prone to express their emotions on social media (Himelboim 2017). It has been evident in many suicide ideation cases that the suicide victims leave behind messages on social media (Côté et al. 2021). They have expressed their feelings of hopelessness, talking about their intentions, or having no reasons to live on Twitter before ending their life (Fahey, Matsubayashi, and Ueda 2018). Previous studies have shown that the use of certain language patterns on Twitter and their frequencies in subjects' tweets are indicative of personal traits such as loneliness. But most of the time, their posts are not taken seriously or are unattended and lead to death (Safa, Bayat, and Moghtader 2021). With the high rate of people dying from suicide, the speed of information disseminated through Twitter should be an advantage to make good use of it to detect the signs of suicide in a faster way (Brownlie et al. 2021).

Our institution is keen on working on latest research trends and has extensive knowledge and research experience which resulted in quality publications (Rinesh et al. 2022; Sundararaman et al. 2022; Mohanavel et al. 2022; Ram et al. 2022; Dinesh Kumar et al. 2022; Vijayalakshmi et al. 2022; Sudhan et al. 2022; Kumar et al. 2022; Sathish et al. 2022; Mahesh et al. 2022; Yaashikaa et al. 2022). The research gap identified is that the Naive Bayes algorithm has less accuracy. Detection of suicidal tweets is shown at a low percentage of accuracy while analysing using Naive Bayes algorithm. The classification of suicidal tweets is a challenging task since the amount of tweets keep growing everyday with more people entering the world of Twitter due to its booming popularization. It is also a very taxing task to understand and decode the slang/accent used in the tweets. The model is aimed at being capable of predicting individuals at risk as well as the odds of risk of suicidal thoughts within a given time frame. The approach involves training ML techniques to measure existing patterns in tweets that are predictive of suicide risk but do not yet explicitly express suicidal thoughts (Belfort, Mezzacappa, and Ginnis 2012). The aim of research work is to improve accuracy of detecting suicidal tweets, and to reduce loss of data while training and testing dataset. The novel SMOClassifier is used to achieve accuracy.

## 2. Materials and Methods

The study setting of the proposed work is done at the Data Analytics lab at Saveetha School of Engineering. Two groups were identified for the study setting: group one SMO algorithm and group two NB algorithm. The sample size is taken as 10 for each group. The computation is performed using G-power as 80% with a confidence interval at 95% and alpha value is 0.05 and beta value is 0.2 (Fodeh et al. 2019).

The dataset named 'SUICIDAL_DATA' is downloaded from the public domain Github library. In our experiments here we used the suicidal_data.csv dataset. Detailed description of the features/attributes in the dataset can be found below in Table 1. The dataset consists of 242 instances and contains tweets and labels. The dataset was splitted into two parts namely the training part and testing part. 70% of the data was used for training and the remaining 30% was used for testing. The algorithm was implemented by evaluating the train and test datasets. For exhibiting this research work, a Jupyter notebook is used along with a laptop with Intel i3 processor, 4GB RAM, 64 bit Microsoft Windows 11 Pro Operating system and other required specifications.

**Sequential Minimal Optimization (SMO) Algorithm**

The sequential minimal optimization algorithm (SMO) has been shown to be an effective method for training support vector machines (SVMs) on classification tasks defined on sparse data sets. SMO differs from most SVM algorithms in that it does not require a quadratic programming solver. Instead of previous SVM learning algorithms that use numerical quadratic programming (QP) as an inner loop, SMO uses an analytic QP step (Pentrakan, Yang, and Wong 2021). The amount of memory required for SMO is linear in the training set size, which allows SMO to handle very large training sets. Because matrix computation is avoided, SMO scales somewhere between linear and quadratic in the training set size for various test problems, while the standard chunking SVM algorithm scales somewhere between linear and cubic in the training set size. SMO's computation time is dominated by SVM evaluation, hence SMO is fastest for linear SVMs and sparse data sets. On realworld sparse data sets, SMO can be more than 1000 times faster than the chunking algorithm (Andina and Pham 2007). Table 2 illustrates the pseudocode for the SMO algorithm. Table 4 represents accuracy of Suicidal Tweets Detection using SMO algorithm.

**Naive Bayes (NB) Algorithm**

The Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. It is a probabilistic classifier, which means it predicts on the basis of the probability of an object (Zhu et al. 2020). Some popular examples of Naïve Bayes Algorithm are spam filtration, Sentimental analysis, and classifying articles. It is called Naïve because it assumes that the occurrence of a certain feature is independent of the occurrence of other features (Bayram, Minai, and Pestian 2018). The formula for Bayes' theorem is given by Equation 1.

$$P(A|B)=P(B|A).P(A)/P(B) \quad (1)$$

Where, equation (1)

P(A|B) is Posterior probability: Probability of hypothesis A on the observed event B. P(B|A) is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true. P(A) is Prior Probability: Probability of hypothesis before observing the evidence. P(B) is Marginal Probability: Probability of Evidence.

Table 3 illustrates the pseudocode for Naive Bayes algorithm. Table 5 represents accuracy of Suicidal Tweets Detection using Naive Bayes algorithm.

**Statistical Analysis**

For statistical implementation, the software used here is IBM SPSS V22.0. Statistical Package for Social Sciences (SPSS) is used for calculating the statistical calculations such as mean, standard deviation, and also to plot the graphs etc. The independent variables are the tweet and its label. The dependent variable is 'accuracy'. The dataset is prepared using 10 as sample size for each group and accuracy is given as the testing variable.

## 3. Results

The experimental results are carried out on the Novel SMOClassifier and Naive Bayes algorithm where the performance is measured based on accuracy. The accuracy of the Sequential Minimal Optimization Algorithm is 93.5% and the Naive Bayes algorithm is 81.7%.

Table 7 represents the independent sample test that has been performed on the novel Sequential Minimal Optimization model and Naive Bayes algorithm for calculating the equal variance assumed and equal variance not assumed and it also shows mean difference, standard error differences with a confidence level of 95%. Independent Samples T-test shows significance value achieved is $p=0.662$ ($p>0.05$), which shows that two groups are statistically insignificant.

Figure 1 shows the simple bar graph of accuracy by a group of the novel Sequential Minimal Optimization model and Naive Bayes algorithm. It is observed that a novel Sequential Minimal Optimization model algorithm has a higher significance when compared to the Naive Bayes algorithm. The error bars are shown in the graph and error rate is less for Sequential Minimal Optimization algorithm compared to Naive Bayes.

For Sequential Minimal Optimization (SMO) and Naive Bayes (NB), compared the two algorithms with their accuracy rate. For both proposed and existing algorithms 10 iterations were taken for each iteration the predicted accuracy was noted for analyzing accuracy. The results of statistical packages of social sciences (IBM-SPSS v22) are used for data analysis. With values obtained from the iterations, an Independent Sample T-Test was performed. Group statistics values and Independent Sample T-test Result of proposed and existing algorithms are shown in Table 6 & Table 7 where t-test equality is calculated. Confidence interval of the difference as lower and upper values range.

The bar graph is plotted by selected mean accuracy on Y-axis and the Group on X-axis. From the

graph, it is clear that SMO has significantly higher accuracy than NB as shown in Fig. 1. The error bars are shown in the graph and the error rate is less for SMO compared to NB.

## 4. Discussions

In this study, the Sequential Minimal Optimization algorithm has better suicidal tweets detection accuracy than Naive Bayes algorithm (p=0.000, Independent sample t-test). The accuracy is improved and loss is reduced for Sequential Minimal Optimization (Accuracy = 93.5%, Loss = 6.5%) than Naive Bayes (accuracy = 81.7%, Loss = 18.3%).

Various machine learning algorithms are used to detect suicidal tweets. Classification model was proposed to solve the difficulties encountered in detection of tweets expressing suicide ideation using Sequential Minimal Optimization and Naive Bayes (Walsh, Ribeiro, and Franklin 2017); (Kaski and Östergård 2006). SMO has a higher accuracy of 93.5% obtained for classification models in order to classify the tweet (De Choudhury et al. 2016); (Bach, Lanckriet, and Jordan 2004). Our machine learning algorithm achieved balance between precision and recall values due to the application of feature selection techniques for extracting relevant features (Luo et al. 2020).

The limitation in this study is that the amount of tweets keeps growing everyday with more people entering the world of twitter due to its booming popularization. Hence it is difficult to keep up with the huge volume of data. It is also a very challenging task to understand and decode the slang/accent used in the tweets since it is a global platform used by people from all around the world.

## 5. Conclusion

In this research work, the results indicate that our proposed Sequential Minimal Optimization (SMO) based model with novel SMOClassifier function can be used to detect tweets indicating suicide ideation with improved accuracy of 93.5% and Naive Bayes whose accuracy is 81.7%

**Declsaration**

**Conflict of Interests**
The authors do not have any conflict of interest associated with this manuscript.

**Authors contributions**
Author AV involved in data collection, data analysis, manuscript, writing. Author SMK involved in conceptualization, data validation, and critical review of manuscript.

**Acknowledgement**
The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

**Funding**
We thank the following organizations for providing financial support that enabled us to complete the study.

1. Insysness Pvt. Ltd.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

## 6. References

Andina, Diego, and Duc Truong Pham. 2007. *Computational Intelligence: For Engineering and Manufacturing*. Springer Science & Business Media.

Astoveza, Ghelmar, Randolph Jay P. Obias, Roi Jed L. Palcon, Ramon L. Rodriguez, Bernie S. Fabito, and Manolito V. Octaviano. 2018. "Suicidal Behavior Detection on Twitter Using Neural Network." *TENCON 2018 - 2018 IEEE Region 10 Conference*. https://doi.org/10.1109/tencon.2018.8650162.

Bach, Francis R., Gert René Georges Lanckriet, and Michael Irwin Jordan. 2004. *Fast Kernel Learning Using Sequential Minimal Optimization*.

Bayram, Ulya, Ali A. Minai, and John Pestian. 2018. "A Lexical Network Approach for Identifying Suicidal Ideation in Clinical Interview Transcripts." *Unifying Themes in Complex Systems IX*. https://doi.org/10.1007/978-3-319-96661-8_17.

Belfort, Erin L., Enrico Mezzacappa, and Katherine Ginnis. 2012. "Similarities and Differences Among Adolescents Who Communicate Suicidality to Others via Electronic Versus Other Means: A Pilot Study." *Adolescent Psychiatrye*. https://doi.org/10.2174/2210676611202030258.

Biernesser, Candice, Jamie Zelazny, David Brent, Todd Bear, Christina Mair, and Jeanette Trauth. 2021. "Automated Monitoring of Suicidal Adolescents' Digital Media Use: Qualitative Study Exploring Acceptability Within Clinical Care." *JMIR Mental Health* 8

(9): e26031.

Brownlie, Julie, Justin Chun-Ting Ho, Nikki Dunne, Nichole Fernández, and Tim Squirrell. 2021. "Troubling Content: Guiding Discussion of Death by Suicide on Social Media." *Sociology of Health & Illness* 43 (3): 607–23.

Côté, David, Marissa Williams, Rabia Zaheer, Thomas Niederkrotenthaler, Ayal Schaffer, and Mark Sinyor. 2021. "Suicide-Related Twitter Content in Response to a National Mental Health Awareness Campaign and the Association between the Campaign and Suicide Rates in Ontario." *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie* 66 (5): 460–67.

De Choudhury, Munmun, Emre Kiciman, Mark Dredze, Glen Coppersmith, and Mrinal Kumar. 2016. "Discovering Shifts to Suicidal Ideation from Mental Health Content in Social Media." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI Conference* 2016 (May): 2098–2110.

Dinesh Kumar, M., V. Godvin Sharmila, Gopalakrishnan Kumar, Jeong-Hoon Park, Siham Yousuf Al-Qaradawi, and J. Rajesh Banu. 2022. "Surfactant Induced Microwave Disintegration for Enhanced Biohydrogen Production from Macroalgae Biomass: Thermodynamics and Energetics." *Bioresource Technology* 350 (April): 126904.

Fahey, Robert A., Tetsuya Matsubayashi, and Michiko Ueda. 2018. "Tracking the Werther Effect on Social Media: Emotional Responses to Prominent Suicide Deaths on Twitter and Subsequent Increases in Suicide." *Social Science & Medicine*. https://doi.org/10.1016/j.socscimed.2018.10.004.

Fodeh, Samah, Taihua Li, Kevin Menczynski, Tedd Burgette, Andrew Harris, Georgeta Ilita, Satyan Rao, Jonathan Gemmell, and Daniela Raicu. 2019. "Using Machine Learning Algorithms to Detect Suicide Risk Factors on Twitter." *2019 International Conference on Data Mining Workshops (ICDMW)*. https://doi.org/10.1109/icdmw.2019.00137.

Himelboim, Itai. 2017. "Social Network Analysis (Social Media)." *The International Encyclopedia of Communication Research Methods.* https://doi.org/10.1002/9781118901731.iecrm0236.

Kaski, Petteri, and Patric R. J. Östergård. 2006. *Classification Algorithms for Codes and Designs*. Springer Science & Business Media.

Kumar, J. Aravind, J. Aravind Kumar, S. Sathish, T. Krithiga, T. R. Praveenkumar, S. Lokesh, D. Prabu, A. Annam Renita, P. Prakash, and M. Rajasimman. 2022. "A Comprehensive Review on Bio-Hydrogen Production from Brewery Industrial Wastewater and Its Treatment Methodologies." *Fuel*. https://doi.org/10.1016/j.fuel.2022.123594.

Li, Rui, Kin Hou Lei, Ravi Khadiwala, and Kevin Chen-Chuan Chang. 2012. "TEDAS: A Twitter-Based Event Detection and Analysis System." *2012 IEEE 28th International Conference on Data Engineering*. https://doi.org/10.1109/icde.2012.125.

Luo, Jianhong, Jingcheng Du, Cui Tao, Hua Xu, and Yaoyun Zhang. 2020. "Exploring Temporal Suicidal Behavior Patterns on Social Media: Insight from Twitter Analytics." *Health Informatics Journal* 26 (2): 738–52.

Luxton, David D. 2015. *Artificial Intelligence in Behavioral and Mental Health Care*. Academic Press.

Mahesh, Narayanan, Srinivasan Balakumar, Uthaman Danya, Shanmugasundaram Shyamalagowri, Palanisamy Suresh Babu, Jeyaseelan Aravind, Murugesan Kamaraj, and Muthusamy Govarthanan. 2022. "A Review on Mitigation of Emerging Contaminants in an Aqueous Environment Using Microbial Bio-Machines as Sustainable Tools: Progress and Limitations." *Journal of Water Process Engineering*. https://doi.org/10.1016/j.jwpe.2022.102712.

Mohanavel, Vinayagam, K. Ravi Kumar, T. Sathish, Palanivel Velmurugan, Alagar Karthick, M. Ravichandran, Saleh Alfarraj, Hesham S. Almoallim, Shanmugam Sureshkumar, and J. Isaac JoshuaRamesh Lalvani. 2022. "Investigation on Inorganic Salts K2TiF6 and KBF4 to Develop Nanoparticles Based TiB2 Reinforcement Aluminium Composites." *Bioinorganic Chemistry and Applications* 2022 (January): 8559402.

Pentrakan, Amarawan, Cheng-Chia Yang, and Wing-Keung Wong. 2021. "How Well Does a Sequential Minimal Optimization Model Perform in Predicting Medicine Prices for Procurement System?" *International Journal of Environmental Research and Public Health* 18 (11). https://doi.org/10.3390/ijerph18115523.

Ram, G. Dinesh, G. Dinesh Ram, S. Praveen Kumar, T. Yuvaraj, Thanikanti Sudhakar Babu, and Karthik Balasubramanian. 2022. "Simulation and Investigation of MEMS Bilayer Solar Energy Harvester for Smart Wireless Sensor Applications." *Sustainable*

*Energy Technologies and Assessments*. https://doi.org/10.1016/j.seta.2022.102102.

Rinesh, S., K. Maheswari, B. Arthi, P. Sherubha, A. Vijay, S. Sridhar, T. Rajendran, and Yosef Asrat Waji. 2022. "Investigations on Brain Tumor Classification Using Hybrid Machine Learning Algorithms." *Journal of Healthcare Engineering* 2022 (February): 2761847.

Safa, Ramin, Peyman Bayat, and Leila Moghtader. 2021. "Automatic Detection of Depression Symptoms in Twitter Using Multimodal Analysis." *The Journal of Supercomputing*, September, 1–36.

Sathish, T., V. Mohanavel, M. Arunkumar, K. Rajan, Manzoore Elahi M. Soudagar, M. A. Mujtaba, Saleh H. Salmen, Sami Al Obaid, H. Fayaz, and S. Sivakumar. 2022. "Utilization of Azadirachta Indica Biodiesel, Ethanol and Diesel Blends for Diesel Engine Applications with Engine Emission Profile." *Fuel*. https://doi.org/10.1016/j.fuel.2022.123798.

Sudhan, M. B., M. Sinthuja, S. Pravinth Raja, J. Amutharaj, G. Charlyn Pushpa Latha, S. Sheeba Rachel, T. Anitha, T. Rajendran, and Yosef Asrat Waji. 2022. "Segmentation and Classification of Glaucoma Using U-Net with Deep Learning Model." *Journal of Healthcare Engineering* 2022 (February): 1601354.

Sundararaman, Sathish, J. Aravind Kumar, Prabu Deivasigamani, and Yuvarajan Devarajan. 2022. "Emerging Pharma Residue Contaminants: Occurrence, Monitoring, Risk and Fate Assessment – A Challenge to Water Resource Management." *Science of The Total Environment*. https://doi.org/10.1016/j.scitotenv.2022.153897.

Vijayalakshmi, V. J., Prakash Arumugam, A. Ananthi Christy, and R. Brindha. 2022. "Simultaneous Allocation of EV Charging Stations and Renewable Energy Sources: An Elite RERNN-m2MPA Approach." *International Journal of Energy Research*. https://doi.org/10.1002/er.7780.

Walsh, Colin G., Jessica D. Ribeiro, and Joseph C. Franklin. 2017. "Predicting Risk of Suicide Attempts Over Time Through Machine Learning." *Clinical Psychological Science*. https://doi.org/10.1177/2167702617691560.

Yaashikaa, P. R., P. Senthil Kumar, S. Jeevanantham, and R. Saravanan. 2022. "A Review on Bioremediation Approach for Heavy Metal Detoxification and Accumulation in Plants." *Environmental Pollution* 301 (May): 119035.

Zhu, H., X. Xia, J. Yao, H. Fan, Q. Wang, and Q. Gao. 2020. "Comparisons of Different Classification Algorithms While Using Text Mining to Screen Psychiatric Inpatients with Suicidal Behaviors." *Journal of Psychiatric Research* 124 (May): 123–30.

Preethi, P. S., Hariharan, N. M., Vickram, S., Manian, R., Manikandan, S., Subbaiya, R., ... & Awasthi, M. K. (2022). Advances in bioremediation of emerging contaminants from industrial wastewater by oxidoreductase enzymes. Bioresource Technology, 127444.

**Tables And Figures**

Table 1. Sample Dataset Containing Tweet and Label

| S.No | Tweet | Label |
|---|---|---|
| 1 | my life is meaningless 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 | 1 |
| 2 | muttering i wanna die to myself daily for a few months now i feel worthless shes my soulmate i can't live in this horrible world without her i am so lonely i wish i could just turn off the part of my brain that feels | 1 |
| 3 | work slave i really feel like my only purpose in life is to make a higher man money parents forcing me through college and i have too much on | 1 |

| | | |
|---|---|---|
| | my plate i owe a lot of money i know this is the easy way out but i am really tired all of these issues are on top of dealing with tensions in america as well i want to rest | |
| 4 | i did something on the 2 of october i overdosed i just felt so alone and horrible i was in hospital for two days now when i walk down the hallways of my school they always look at me weird and say i should take more pills and i hate it i have no one i have this voice in my head now and it won't go away and i can't be myself anymore thanks for reading | 1 |
| 5 | i feel like no one cares i just want to die maybe then i ' d feel less lonely | 1 |
| 6 | i am great and wonderful i am worth it except not enough to be anyone's first choice everyone tells me how wonderful i am but not enough to be loved like i love others i put aside everything for people but i am too crazy to hold a job too nothing to be really loved i am not entitled and i don't even have the right to die on my own terms and i am an asshole for being angry about it for being upset that i am there when other people treat me like shit and cant be bothered when i'm hurt | 1 |
| 7 | i ll be dead just you wait and see my last words before my death for whoever is interested i am sorry but you're better off without me you'll learn to live without me it won't be difficult now i shall die | 1 |
| 8 | health anxiety prompting some bad thoughts in my head i have been struggling for 2 months now with some health issues as a 26 year old male my pessimistic nature just makes me think about the worst my hands and feet are currently tingling and burning i just keep picturing myself on a wheelchair being a burden to my family girlfriend and so on suicide thoughts come to my mind as i prefer to put a sudden end to everything instead of deteriorating myself day after day losing motor and cognitive functions life is already hard as it is and now my health is failing for the first time | 1 |
| 9 | everything is okay but nothing feels okay i ve always been a bit unhappy as a kid too i think although i can t remember much of my childhood i dont want to kill myself but sometimes that thought just comes creeping and it scares me a little a few weeks ago a problem came up it was a financial problem quite fixable but i just couldn t handle it i tied myself a noose and everything i was gonna do it i was all alone in the house with my dog so there was really no one that would be able to stop me i didnt do anything but i felt like i could have done it completely on impulse over a fixable problem leaving behind everything i love and my hopeful future i feel it now too creeping up on me everything should be fine but i cant help the feeling that i should just do it like everything would be easier for everyone if they would just realize how little they need me | 1 |
| 10 | ptsd and alcohol i had some extremely horrible violent stuff happen to me a few years ago i was 21 nowi am26 i forgot about repressed it or whatever for several years something unrelated one day made me | 1 |

| | | |
|---|---|---|
| | remember everything it all came flooding back into my mind and it was like i was reliving it all and felt like i was having a never ending panic attack for about 4 days this buried trauma explained a lot about why my alcohol pornography cigarette usages were all insanely high to the point they fucked up my life and relationships with people close to me in signifigant waysi amafraid to talk about what happened to anyone even people i trust like my family or a potential therapist due to extreme irrational paranoia about the people involved finding out and hurting me again and sometimesi amjust completely consumed by negative horrible thoughts and cant escape them i tried getting a sliding scale therapist a couple years ago and even t and gave up on the idea of therapy i dont know what to do i dont want to give up | |

Table 2. Pseudocode for Sequential Minimal Optimization Algorithm

| |
|---|
| **// I : Input dataset records** |
| 1. Import the required packages. |
| 2. Convert the string values in the dataset to numerical values. |
| 3. Assign the data to X_train, y_train, X_test and y_test variables. |
| 4. Using train_test_split() function, pass the training and testing variables and give test_size and the random_state as parameters. |
| 5. Import the SMOClassifier from sklearn library. |
| 6. Using SMOClassifier, predict the output of the testing data. |
| 7. Calculate the accuracy. |
| **OUTPUT**<br>**//Accuracy** |

Table 3. Pseudocode for Naive Bayes Algorithm

| |
|---|
| **// I : Input dataset records** |
| 1. Import the required packages. |
| 2. Convert the string values in the dataset to numerical values. |
| 3. Assign the data to X_train, y_train, X_test and y_test variables. |
| 4. Using train_test_split() function, pass the training and testing variables and give test_size and the random_state as parameters. |
| 5. Import the NBClassifier from sklearn library. |
| 6. Using NBClassifier, predict the output of the testing data. |
| 7. Calculate the accuracy. |
| **OUTPUT**<br>**//Accuracy** |

Table 4. Accuracy of Suicidal Tweet Detection using Sequential Minimal Optimization Algorithm (Accuracy = 93.5)

| Test | Accuracy | Loss |
|---|---|---|
| Test 1 | 93.5 | 6.5 |
| Test 2 | 92.3 | 7.7 |
| Test 3 | 91.6 | 8.4 |
| Test 4 | 90.9 | 9.1 |
| Test 5 | 89.1 | 10.9 |
| Test 6 | 88.4 | 11.6 |
| Test 7 | 87.7 | 12.3 |
| Test 8 | 86.5 | 13.5 |
| Test 9 | 85.8 | 14.2 |
| Test 10 | 84.2 | 15.8 |

Table 5. Accuracy of Suicidal Tweet Detection using Naive Bayes Algorithm (Accuracy = 81.7)

| Test | Accuracy | Loss |
|---|---|---|
| Test 1 | 81.7 | 18.3 |
| Test 2 | 80.8 | 19.2 |
| Test 3 | 79.4 | 20.6 |
| Test 4 | 77.5 | 22.5 |
| Test 5 | 76.3 | 23.7 |
| Test 6 | 75.8 | 24.2 |
| Test 7 | 74.1 | 25.9 |
| Test 8 | 73.6 | 26.4 |
| Test 9 | 72.2 | 27.8 |
| Test 10 | 71.9 | 28.1 |

Table 6. Group Statistics Results represented for Accuracy and Loss for Sequential Minimal Optimization and Naive Bayes algorithm

| Algorithm | N | Mean | Std. Deviation | Std.Error Mean |
|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| Accuracy | SMO | 10 | 89.0000 | 3.03864 | 0.96090 |
| | NB | 10 | 76.3300 | 3.48012 | 1.10051 |
| Loss | SMO | 10 | 11.0000 | 3.03864 | 0.96090 |
| | NB | 10 | 23.6700 | 3.48012 | 1.10051 |

Table 7. Independent Samples T-test shows significance value achieved is p=0.662 (p>0.05), which shows that two groups are statistically insignificant.

| | **Levene's test for equality of variances** | | **t-Test for Equality of Means** | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **F** | **Sig.** | **t** | **df** | **Sig. 2-tailed** | **Mean Difference** | **Std.Error Difference** | **95% confidence interval of the difference** | |
| | | | | | | | | **Lower** | **Upper** |
| **Accuracy Equal variance Assumed** | 0.198 | 0.662 | 8.672 | 18 | 0.000 | 12.67000 | 1.46098 | 9.60060 | 15.73940 |
| **Equal variance not Assumed** | | | 8.672 | 17.679 | 0.000 | 12.67000 | 1.46098 | 9.59660 | 15.74340 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Loss** | | | | | | | | | |
| **Equal variance Assumed** | 0.198 | 0.662 | -8.672 | 18 | 0.000 | -12.67000 | 1.46098 | -15.73940 | -9.60060 |
| **Equal variance not Assumed** | | | -8.672 | 17.679 | 0.000 | -12.67000 | 1.46098 | -15.74340 | -9.59660 |

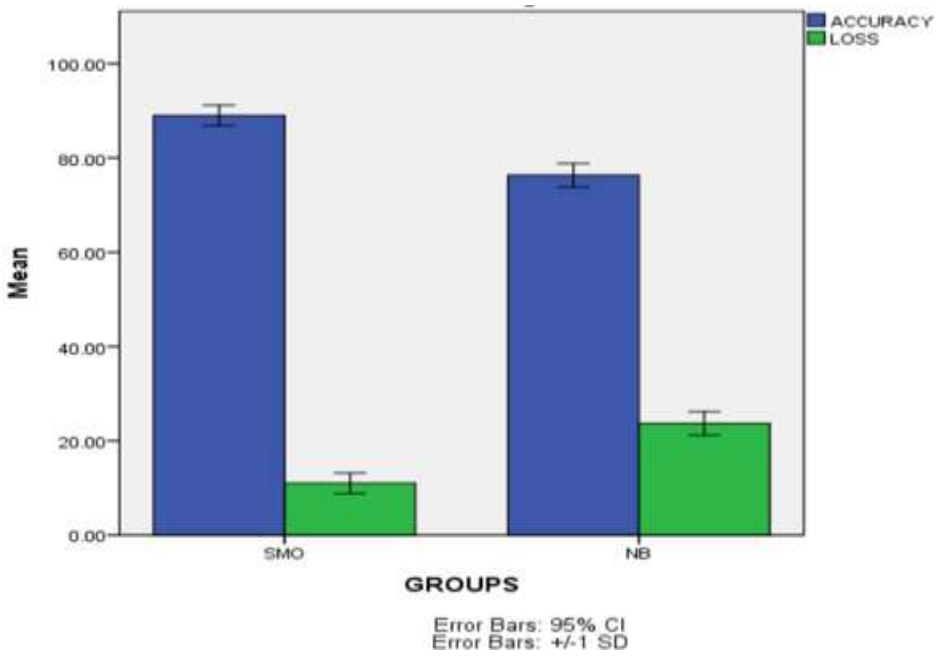


Fig. 1. Comparison of Sequential Minimal Optimization algorithm and Naive Bayes Algorithm in terms of mean accuracy. Mean accuracy of Sequential Minimal Optimization is better than Naive Bayes and standard deviation of Sequential Minimal Optimization is slightly better than Naive Bayes. X Axis: Sequential Minimal Optimization vs Naive Bayes. Y Axis : Mean Accuracy of detection = +/- 1 SD with Confidence Interval of 95%.