DETECTING SUICIDAL TENDENCY USING MACHINE LEARNING

Shaik Juhiya Afreen, Padala Rohini Aparna
Chaitanya Bharathi Institute of Technology

Abstract: Suicide is a critical issue in modern society. There are several factors that can lead a person to commit a suicide, for instance, stress, depression, failure, disappointment, pessimism, unemployment, among others. A recent study by the Center for Disease Control (CDC) of the United States of America estimated that the economic toll of suicide on society is immense as well. Suicides and suicide attempts cost the nation almost $70 billion per year in lifetime medical and work-loss costs alone. Detecting and preventing suicide attempts therefore becomes crucial for the authorities of the land and should be addressed in order to save people's lives and preserve the social fabric of the community.

     Nowadays social media has become a way for people to express themselves and thus may be used to convey suicidal tendencies. Due to the anonymity of online media and social networks, people tend to express their feelings and sufferings in various online communities' Traditional modes of prevention include interactions between likely suicidal individuals and an expert, a therapist or a social worker. But this mode of operations is susceptible to create some delays in the diagnosis of the patient, which could lead to a fatality.

     In order to prevent suicides more effectively, the ideation must be detected as early as possible. This can be done by analyzing users' posts for suicidal related content. The key objective being to present an automatic recognition of suicidal posts using Machine learning techniques. We focus on the online communities of Reddit and Twitter. The project will consist of developing a model of classification of various social media posts into classes that determine whether the user has suicidal tendencies or not.

Index Terms: Suicides, Suicidal Detection, Machine Learning, Earlier detection, Tweets.

I. INTRODUCTION

Suicide is a serious issue in modern society. There are several factors that can lead a person to commit a suicide, for instance, stress, depression, failure, disappointment, pessimism, unemployment, among others. A recent study by the Center for Disease Control (CDC) of the United States of America estimated that the economic toll of suicide on society is immense as well. Suicides and attempts to suicide cost the nation approximately $70 billion per year in lifetime medical and work-loss costs alone.

Detecting and preventing suicide attempts therefore becomes crucial for the authorities of the land and should be addressed in order to save people's lives and preserve the social fabric of the community.

Nowadays social media has become a way for people to express themselves isnd thus may be used to convey suicidal tendencies. Due to the addiction of social networks, people express their feelings and sufferings in

various online communities, modes of prevention include interaction between likely suicidal individuals and experts, therapists or social workers. But this mode of operation is susceptible to create some delays in the diagnosis of the patient, which could lead to a fatality.

In order to prevent suicides more effectively, the ideation must be detected as early as possible. This can be done by analyzing users' posts for suicidal related content. The key objective being to present an automatic recognition of suicidal posts using Machine learning techniques. We focus on the online Twitter users. The project will consist of developing a model of classification of various social media posts into classes that determine whether the user has suicidal tendencies or not. The goal of text-based suicide classification is to determine whether candidates have suicide ideations or not through their posts.

Traditional modes of prevention rely on clinical methods, including self-reports and face-to-face interviews that's interactions between likely suicidal individuals and an expert, a therapist, or a social worker. The main objectives of this project are an earlier detection of suicidal ideation, an automatic recognition of suicidal content in social networking posts like twitter, raddit.

Suicide is quite possibly the most recognized reasons for death on the news around the world. There are a few elements and factors that can lead an individual to submit this suicide, for instance, stress, confidence, sadness, among others. We propose a new methodology with a machine learning using twitter dataset; such information mirrors the young adult populace with self-destructive inclination in many. We at last propose a technique that decides the suicidal detection in posts.

Self-destruction has consistently had its situation among the best 10 reasons for death everywhere in the world. It was assessed by the World Health Organization (WHO) that consistently roughly 1,000,000 individuals ended it all, which brings the death rate to 16 individuals for each 100,000 or one demise each 40 seconds. It is anticipated that the self-destruction rate is going to ascend to one at regular intervals. More youthful ages have now supplanted older guys as the gathering at most elevated danger in many nations. Psychological well-being messes, (for example, discouragement and substance misuse) are credited with more than 90% of all instances of suicide. Numerous nations perceive the need and positive effect of Suicide Prevention Strategies and are attempting to guarantee they are set up. One such anticipation system is early detection of suicidal ideation or melancholy among people.

The scope of this project is an earlier Detection Of Suicidal Ideation (DSI). The key objective is to present an automatic recognition of suicidal posts using Machine learning techniques. The goal of text-based suicide classification is to determine whether candidates, through their posts, have suicidal ideations. Machine learning methods and NLP also will be applied in this field.

We have used Twitter as our dataset for this project. Twitter provided us with four keys - consumer_key, consumer_secret, access_token, access_secret. The dataset was split into 80% training and 20% testing. The data was provided in text format and we have performed preprocessing to remove unwanted data like emojis, signs (!/@$:*&^%) and stopwords and then we have stored it in csv format.

II. LITERATURE REVIEW

1) Suicidal Tendency Detection:

This paper was written by the Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain and it was published by IEEE in the year 2019.

They have used the Convolutional Neural Network (CNN) in this project to get an accuracy of 77%. The drawback of this approach was that they have used image based predictive models For detecting depression and suicide.

2)Performance Evaluation of Different Machine Learning Techniques using Twitter Data for Identification of Suicidal Intent:

This paper was written by Anirudh Ramachandran, Akshara Gadwe, Dishank Poddar, Saurabh Satavalekar and Sunita Sahu and was published by IEEE in 2020 at the International Conference on Electronics and Sustainable Communication Systems (ICESC).

Research and Evaluation based on online behavior have been conducted repeatedly. Using machine learning, this online trail of data that a person leaves behind can be used to gain insights on the behavior and psychological status.

In this paper, different machine learning techniques have been used, studied and gauged their effectiveness for suicidal tendency detection to prove that Machine Learning Algorithms like Logistic Regression can correctly identify residing Suicidal Tendency of a Twitter user.

They have used algorithms such as Linear Regression, Logistic Regression, Naive Bayes , Random Forest, GBDT, XGBoost, MLFFNN on the twitter dataset to achieve an accuracy of 76.3%. The only drawback of this approach is that it uses logistic regression which is a long process for detection of suicides

3) DEPRESSION AND SUICIDAL TENDENCY IDENTIFICATION:

This paper was written by Seung Young Ryu ,Hyeongrae Lee,Dong-Kyun Lee AND Kyeongwoon Park Department of Mental Health Research ,National Center for Mental Health, Seoul, Republic of Korea and was published on 30th December 2019.

This approach uses the Random Forest algorithm to get an accuracy of 85%.

4) Simulation of Suicide Tendency by Using Machine Learning:

This paper was written by Hugo D. Calderon-Vilca, William I. Wun-Rafael and Roberto Miranda-Loarte and was published by IEEE on 16th October 2017 at the 36th International Conference of the Chilean Computer Science Society (SCCC).

They have proposed a simulation with a systematically generated dataset. Such data reflect the adolescent population with suicidal tendencies in Peru. They have evaluated three algorithms of supervised machine learning as a result of the algorithm C4.5 which is based on the trees to classify in a better way the suicidal tendency of adolescents. They finally propose a desktop tool that determines the suicidal tendency level of the adolescent.

The algorithms used are C4.5, JRip algorithm and Naive Bayes algorithm and have achieved an accuracy of 90.7%. This approach falls short as it was implemented only on adolescents.

5) Detection of Suicidal Ideation on Social Media: Multimodal, Relational, and Behavioral Analysis:

This paper was written by Diana Ramírez-Cifuentes , Ana Freire , Ricardo Baeza-Yates , Joaquim Puntí , Pilar Medina-Bravo , Diego Alejandro Velazquez, Josep Maria Gonfau, Jordi Gonzàlez and was published in the year 2020 in the Journal of Medical Internet Research.

This paper aimed to describe an approach for the suicide risk assessment of Spanish-speaking users on social media. We aimed to explore behavioral, relational, and multimodal data extracted from multiple social platforms and develop machine learning models to detect users at risk.

They characterized users based on their writings, posting patterns, relations with other users, and images posted. They also evaluated statistical and deep learning approaches to handle multimodal data for the detection of users with signs of suicidal ideation.To evaluate the performance of the models, they distinguished 2 control groups: users who make use of suicide-related vocabulary (focused control group) and generic random users (generic control group).

The algorithms used were random forest, multilayer perceptron, logistic regression, and support vector machines as classifiers which gave an accuracy of 82%. But this is just an observational study. Results can be improved by enhancing the contribution of the textual and relational features

III. PROPOSED METHODOLOGY

Methodology is to implement a machine learning classifier to improve a performance of a language modeling and text classification for detecting suicide ideation in social media.

The algorithms that are used are:

1. **Support Vector Machine Algorithm**

   Support Vector Machine or SVM is can be used for both classification and regression.The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. The training set will be fitted to the SVM classifier. To create the SVM classifier, we will import SVC class from Sklearn.svm library.

2. **Naive Bayes Algorithm**

   Naive Bayes classifiers are a collection of classification algorithms based   on Bayes' Theorem. Bayes' Theorem finds the probability of an event occurring given the probability of another event that has already occurred. Bayes' theorem is stated mathematically as the following equation:

   $$P(A|B) \ = \ [P(B|A)\,P(A)]/P(B)$$

   where A and B are events and P(B) ? 0.

   In our project, we have used Multinomial Naive Bayes Classifier. This is the event model typically used for document classification. It takes into account the frequencies of terms and uses conditional probability to classify the text data into different classes.

3. Decision Tree Algorithm

Decision tree is the most powerful and popular tool for classification and prediction. A decision tree is a flowchart like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label.

The method proposed in this project first identifies the suicidal keywords, and then we use these keywords to extract tweets from Twitter using Twitter Streaming Application Programming Interface (API).After that we pre-process the data or extract features from the text document.Then SVM and Decision Tree with three types of weight optimizers are imposed on the dataset and at last we determine the model efficiency on the basis of accuracy score, precision, recall and F1 scores.Tweets are retrieved from twitter data and intensity of suicidal tendency is calculated based on the weightage given to the word. Among the two types of machine learning approach we followed supervised learning for this proposed method. A flowchart that represents the proposed model is as follows

FlowChart:
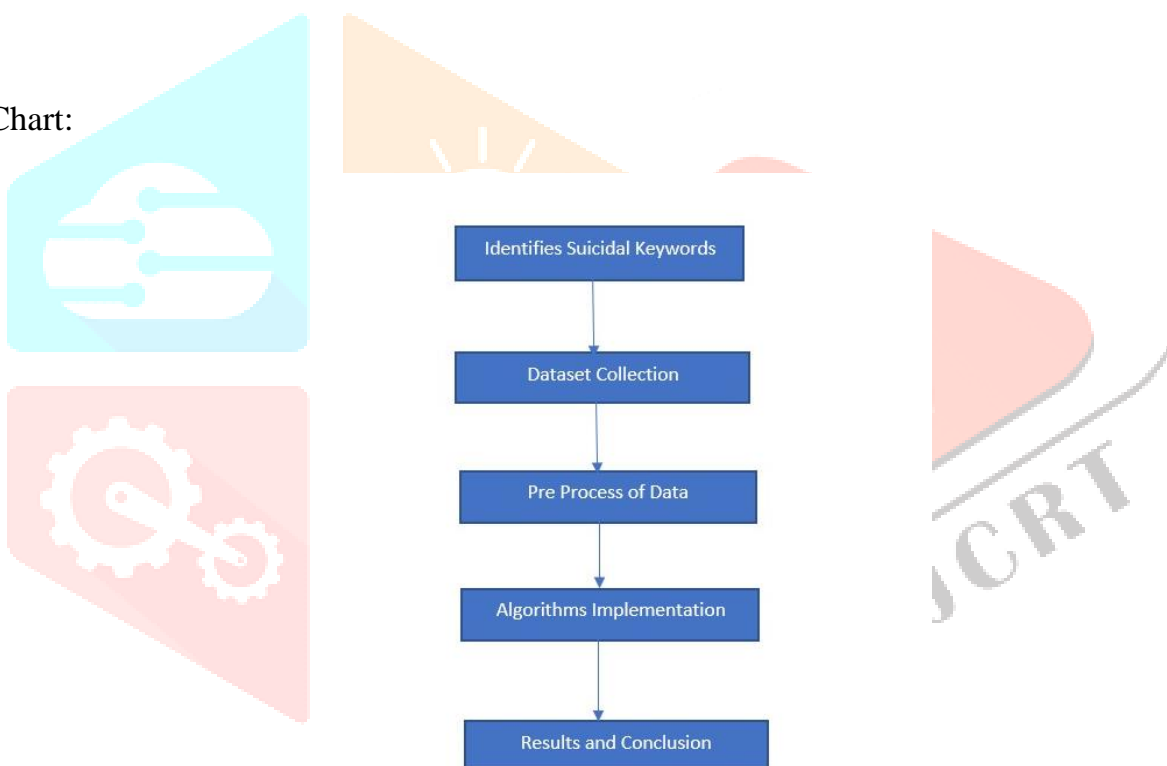
This proposed method can be divided into five modules:

1. Identify Suicidal Keywords

For collecting tweets from Twitter we need certain keywords that contain suicidal intent. A research that collected texts from several sources and analyzed these texts to identify keywords that are usually within a suicidal note was very helpful in this regard. Then keywords were further examined by two experts of suicide identification field and finally there was a list of 62 keywords that happen to exist in suicidal notes . We used these keywords to extract tweets from Twitter.

2. Search And Extract Tweets

In this part of the proposed system we used Twitter Streaming API to extract real time tweets containing those specific keywords from Twitter. A tweet object contains many portions among them we extracted portions that are necessary for us such as user id and tweet text. After extracting these portions we need to process these data as these contain several unnecessary constituents such as special characters, emoticons. We used regular expressions to remove these unnecessary portions.

3. Feature Extraction

extraction is transforming arbitrary data such as text and images into numerical features that can be used for machine learning. For example, "the", "a", "is" such types of words remain present in text documents in large numbers and carry less importance. Even in some cases it may hamper model efficiency. Using Tf-idf we get rid of these words. Doing so we extracted features from the tweets.

4. Introducing SVM,Naive Bayes And Decision Tree For Evaluation

We have used Naive Bayes Algorithm, decision Tree algorithm and Support Vector Machine algorithm for evaluation. We have imported each algorithm from sklearn.

Multinomial Naive Bayes is a specialised version of Naive bayes which takes into account the frequencies of terms and uses conditional probability to classify the text data into different classes. It is mainly designed for classification with discrete features.

5. Model Efficiency Evaluation

Way of evaluating model efficiency is the accuracy score. For this we split the dataset for training and testing. After training on the dataset we let the model predict the outcome for some cases. From there we calculate the percentage of accuracy. We also implement k fold cross validation for model evaluation as sometimes only training and test split is not sufficient for determining the efficiency of the model. K fold cross validation provides train/test indices to split data in train/test sets.Split dataset into k consecutive folds. Each fold is then used once as a validation while the k-1 remaining folds form the training set. For our case the value of k is 10. For further evaluation we introduced precision, recall and f1 scores. Precision is the value derived from dividing true positive by addition true positive and false positive. On the other hand recall is the value that we get by diving true positive by addition of true positive and false negative and The F1 score can be interpreted as a weighted harmonic mean of the precision and recall.For clarification, a true positive is an outcome where the model correctly predicts the positive class.Similarly, a true negative is an outcome where the model correctly predicts the negative class and a false positive is an outcome where the model incorrectly predicts the positive class.

Model Diagram:
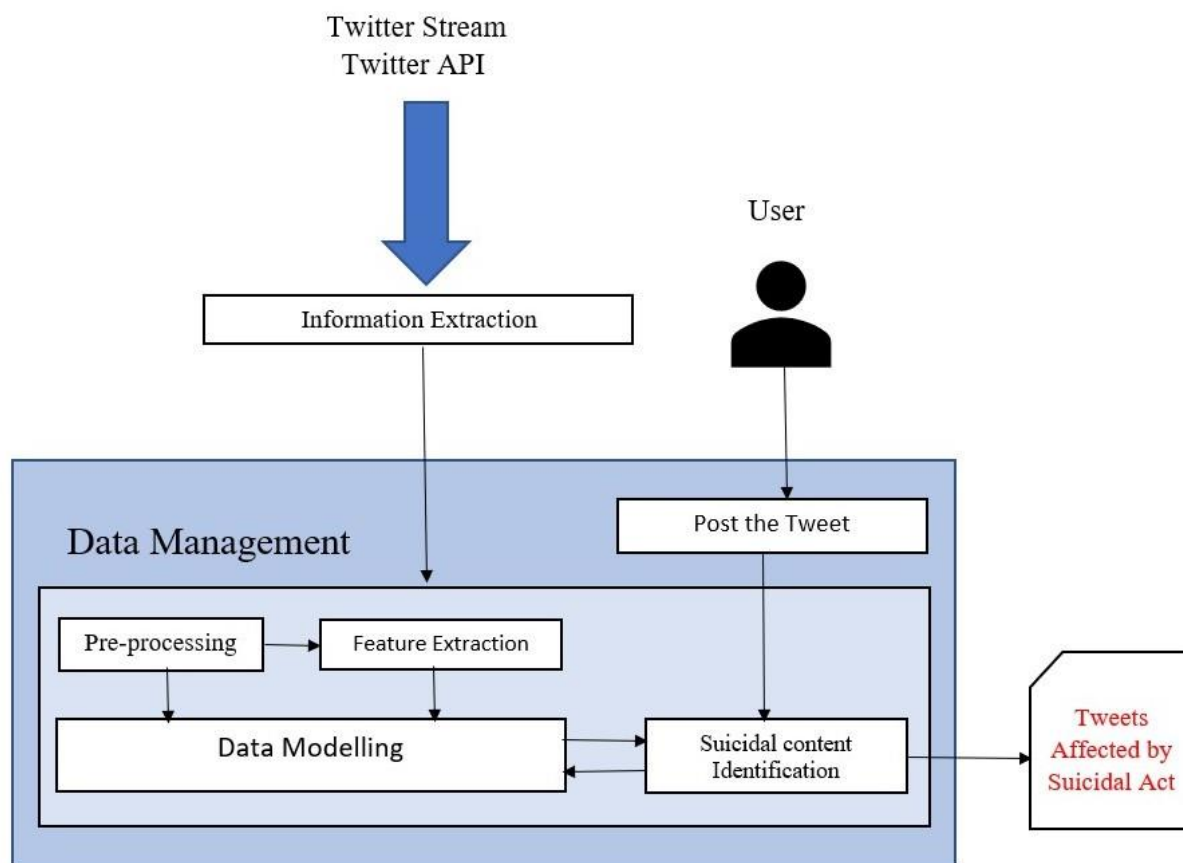
IV. RESULTS

Our model was assessed on the basis of five variables: accuracy ,10-fold cross validation, recall, f1 score, and precision.

Results of classification results using features based on tweets.

When we use only features based on tweets, below table shows the accuracy, precision and recall, and Completion Speed of our model in recognizing suicidal posts on tweeter. A 10-fold cross validation was used in the experiments. Table shows that the Decision Tree classifier achieves the best accuracy of 97.89%, precision of 97.89%, recall 97.38% and Naïve Bayes classifier complete a speed of 5.35 second which is the fastest speed among the remaining algorithms.

Classification results using features based on tweets.

Algorithms Implemented	Accuracy	precision	recall	Completion Speed (unit in second)
Support vector Machine	**92.89%**	**93.53%**	**91.41%**	**1105.52**
Decision Tree	**97.89%**	**97.89%**	**97.38%**	**18.70**
Naïve Bayes	**93.29%**	**89.23%**	**89.04%**	**5.35**

Fig:1 Confusion Matrix for Décision Tree Classifier   Fig:2 Confusion Matrix Naive Bayes Classifier
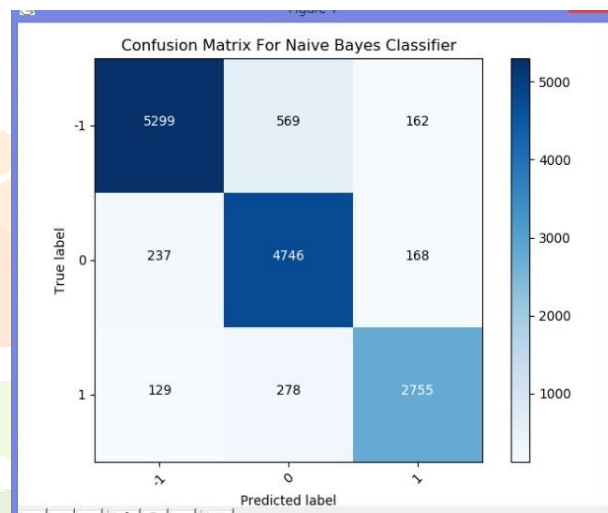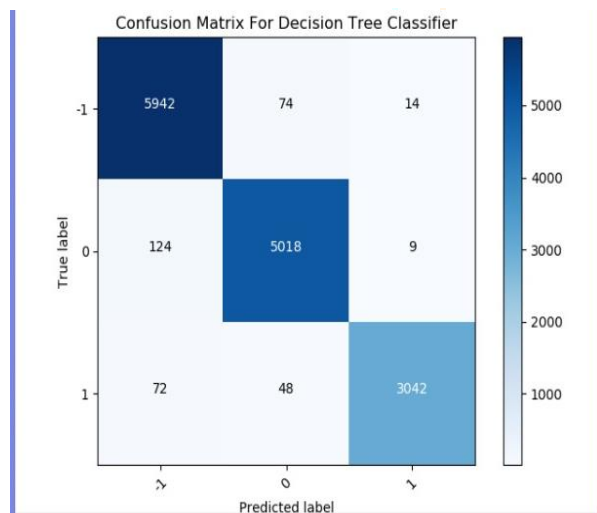

Fig :3 Confusion Matrix For Support Vector Machine Classifier   Fig :4 Performance Comparison
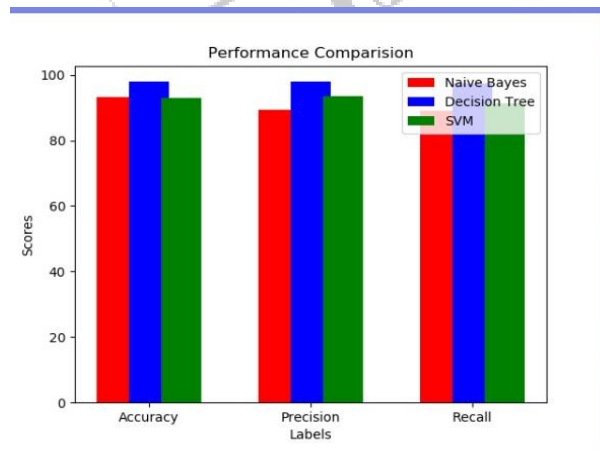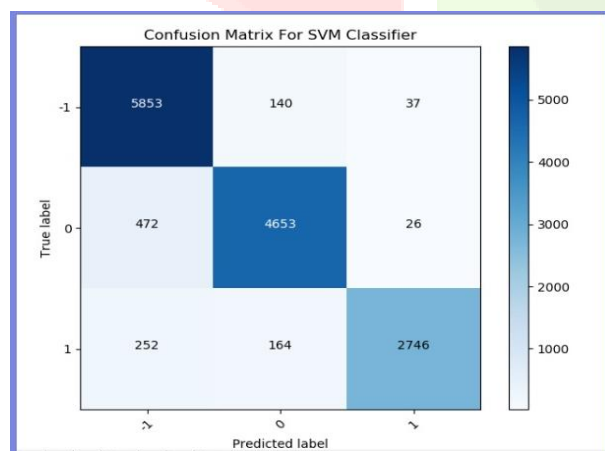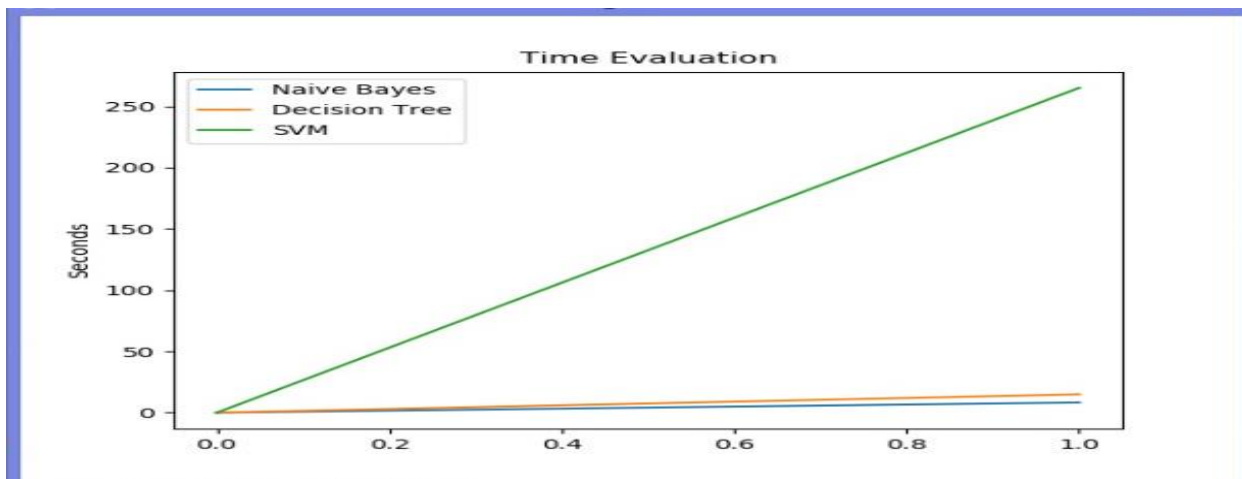

Fig:5 Time Evaluation

Fig 1,2 and 3 are confusion matrix of the three implemented algorithms decision tree, Naïve bayes and support vector machine. A confusion matrix is a description of classification problem prediction output. The number of accurate and inaccurate predictions is totaled and subdivided by class using count values. The trick to the uncertainty matrix is that when it makes predictions, it is perplexed. The three figures 1,2 and 3 shows an N x N matrix which is used to evaluate the efficiency of a classification model, where N is the number of target groups. The matrix compares the real goal values to the machine learning model's predictions. This provides us with a comprehensive picture of how well our classification model is doing and the types of errors it makes.

Fig 4, demonstrates the performance comparison between decision tree, naïve bayes and support vector machine based on accuracy, precision, and recall of our model in identifying suicidal posts on Twitter. In the experiments, a 10-fold cross validation was used. Fig 4 reveals that the Decision Tree classifier has the highest accuracy (97.89%), precision (97.89%), recall (97.38%), where Support vector machine gave us accuracy (92.89%), precision (93.53%), recall (91.41%) and Nave Bayes obtained an accuracy (93.29%), precision (89.23%), recall (89.04%) classifier has the fastest speed (5.3).Fig 5 shows the completion speed of the three algorithms per second, support vector machine complete a speed of 1105.52 seconds, Decision Tree complete 18.70 seconds and Nave Bayes complete a speed of 5.35 seconds.

From the results obtained by both figures 4 and 5 we can conclude that Decision Tree classifier has the highest accuracy (97.89%), precision (97.89%), recall (97.38%), and Nave Bayes classifier has the fastest speed (5.3).

```
Retrved Tweet  ☺ frickin sad to see whats happening in this world... js wnna adopt all those kids☹☹☹
Retrved Tweet  RT @TDZ66: @DFisman If the government valued the mental health of children then there would be smaller classes, investment in ventila
Retrved Tweet  Leave Charles alone•••
Retrved Tweet  Something sinister in the tone
Told me my secret must be known:
Word I was in the house alone
Somehow must have got… https://t.co/yT0o90Kf9a
Retrved Tweet  Sedat Peker: (Süleyman Soylu'ya) Boynuna tasma takıp seni sokaklarda gezdireceğim #sadıksoylu
weightage of  sad  is  10
total suicidal weightage of  ☺ frickin sad to see whats happening in this world... js wnna adopt all those kids☹☹☹  is  10
total suicidal weightage of  RT @TDZ66: @DFisman If the government valued the mental health of children then there would be smaller classes, investme
weightage of  alone  is  20
total suicidal weightage of  Leave Charles alone•••  is  20
weightage of  alone  is  20
total suicidal weightage of  Something sinister in the tone
Told me my secret must be known:
Word I was in the house alone
Somehow must have got… https://t.co/yT0o90Kf9a  is  20
weightage of  sad  is  10
total suicidal weightage of  Sedat Peker: (Süleyman Soylu'ya) Boynuna tasma takıp seni sokaklarda gezdireceğim #sadıksoylu  is  10

Process finished with exit code 0
```

Intensity of Tweets

V.CONCLUSION AND FUTURE SCOPE

Machine Learning applied to detect suicide intention and depression amongst individual is effective as traditional approaches are hindered by factors like face-to-face conversation and shyness to express themselves. In the future I would like to delve more into context analysis via the retweet history and given links to external sites. Even though ML seems to be a better method, yet it entails some shortcomings, like eventually having to use human intervention to approve the predictions by the ML model, it's limited to detecting suicidal tendencies and depression and contextual analysis as the data collected doesn't come with prior background explanation.

The amount of text keeps increasing with the popularization of social networking services. And suicide detection and prevention remain a crucial task in our modern society. It is therefore essential to develop new methods to detect online texts containing suicidal ideation in the hope that suicide can be prevented. In the project an automatic recognition of suicidal posts is presented using machine learning techniques

As a future scope we can implement some more algorithms to improve the efficiency and quality of work and we can implement the same project on videos using image processing techniques

VI.BIBLIOGRAPHY

[1] "Suicide:one person dies every 40 seconds", World Health Organization 9 September 2019.

[2] Sher, Leo, and René S. Kahn. "Suicide in schizophrenia: an educational overview." Medicina 55.7 (2019): 361.

[3] Ji, S., Yu, C. P., Fung, S. F., Pan, S., & Long, G. (2018). Supervised learning for suicidal ideation detection in online user content. Complexity, 2018.

[4] Vijaykumar, Lakshmi. "Suicide and its prevention: The urgent need in India." Indian journal of psychiatry 49.2 (2007): 81.

[5] Ryu, S., Lee, H., Lee, D. K., & Park, K. (2018). Use of a machine learning algorithm to predict individuals with suicide ideation in the general population. Psychiatry investigation, 15(11), 1030.

[6] O'dea, B., Wan, S., Batterham, P. J., Calear, A. L., Paris, C., & Christensen, H. (2015). Detecting suicidality on Twitter. Internet Interventions, 2(2), 183-188.

[7] Jashinsky, J., Burton, S. H., Hanson, C. L., West, J., Giraud-Carrier, C., Barnes, M. D., & Argyle, T. (2014). Tracking suicide risk factors through Twitter in the US. Crisis.

[8] Coppersmith, G., Leary, R., Crutchley, P., & Fine, A. (2018). Natural language processing of social media as screening for suicide risk. Biomedical informatics insights, 10, 1178222618792860.

[9] Coppersmith, G., Harman, C., & Dredze, M. (2014, May). Measuring post traumatic stress disorder on Twitter. In Eighth international AAAI conference on weblogs and social media.

[10] Sueki, Hajime. "The association of suicide-related Twitter use with suicidal behaviour: a cross-sectional study of young internet users in Japan." Journal of affective disorders 170 (2015): 155-160.

[11] Sawhney, R., Manchanda, P., Singh, R., & Aggarwal, S. (2018, July). A computational approach to feature extraction for identification of suicidal ideation in tweets. In Proceedings of ACL 2018, Student Research Workshop (pp. 91-98).

[12] Vioulès, M. J., Moulahi, B., Azé, J., & Bringay, S. (2018). Detection of suicide-related posts in Twitter data streams. IBM Journal of Research and Development, 62(1), 7-1.

[13] W. Wang, L. Chen, M. Tan, S. Wang, A. P. Sheth, "Discovering finegrained sentiment in suicide notes", Biomedical informatics insights, vol. 5, pp. 137, 2012.

[14] S. Chattopadhyay, "A Study on Suicidal Risk Analysis," 2007 9th International Conference on e-Health Networking, Application and Services, Taipei, 2007, pp. 74-78.