National College of Ireland

Detection of suicidal content in the social media posts using advanced predictive classifiers

MSc Research Project
Data Analytics

Tulasiram Dhullipalla
Student ID: X22196994

School of Computing
National College of Ireland

Supervisor: Ahmed Makki

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Tulasiram dhullipalla
Student ID:	X22196994
Programme:	Data Analytics (MSCDAD_sep23) **Year:** 2023
Module:	Research Project
Supervisor:	Ahmed Makki
Submission Due Date:	12 August 2024
Project Title:	Detection of suicidal content in the social media posts using advanced predictive classifiers
Word Count:	6290 **Page Count** 20

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Tulasiram Dhullipalla

Date: 12-08-2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Detection of suicidal content in the social media posts using advanced predictive classifiers

Tulasiram Dhullipalla

X22196994

Abstract

Suicide is a serious global health concern. Every year, approximately 800,000 people take their own lives. which means one person dies by suicide every 40 seconds. So early detection and finding individuals with suicidal thoughts are important to save their lives. Identifying a suicidal ideation person is crucial and the first step in preventing lives. Traditionally, individual data, who have suicidal thoughts, were collected by doctors through direct interaction. However, nowadays it is difficult to gather data directly from individuals. In the modern era, social media usage has increased drastically. Statista reports that 98 percent of individuals aged 15 to 24 in Europe use social media(internet). Therefore, social networks have a significant advantage in terms of early detection and identification of suicidal thoughts in a particular person to save his life. The objective of this research work is to design and implement advanced predictive models to identify the suicidal thoughts on the social media posts. Analyzing the posts using ML (Machine learning) and Ann (Artificial neural network) that help to understand the emotion of the individual during the post on social media. This analyzed information helps doctors, psychologists, and parents to treat suffering people properly. For the achieve the project aim, Long Short-Term Memory (LSTM), Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM), Support Vector Machine (SVM), and Logistic Regression models were implemented to identify whether this post is a suicidal thought or not. Data was collected from Kaggle (Reddit, Twitter). After the data collection, preprocessing the data is done, which means cleaning, tokenization such as TF_IDF, word embeddings, and analysis. For the Reddit dataset, the accuracy of LSTM, CNN+LSTM, Logistic regression, and SVM achieved 93.23%,92.74%,93.17%, and

75.28% respectively. For Twitter, model accuracy of LSTM, CNN+LSTM, Logistic regression, and SVM achieved 98.57%, 98.57%,97.92%, and 98.65% respectively. The results indicate that CNN-LSTM performed well compared with traditional machine learning methods. This study shows the importance of machine learning and artificial neural networks to improve the health monitoring system by analyzing social media posts in real-time and immediate action to prevent suicide.
Keywords: LSTM, CNN-LSTM, Logistic regression, SVM.

1 Introduction:

One of the serious concerns in the world is suicidal thoughts, particularly teenagers and young adults are more impacted and higher risk of suicide due to emotional and mental health issues (World Health Organization, 2023). In today's digital era, more than 90 percent of teenagers and young adults are using social media such as Reddit, Twitter, Meta, etc. Social media plays an important role in collecting valuable data that helps to understand the individual emotions, as individuals share their daily events and emotions that could be happy or sad. These platforms are not only used for online interaction but also used to collect valuable data for assessing the mental state of individuals, which means suicidal ideation (Fahey et al., n.d.). This research focuses on advanced technical methodologies such as machine learning and neural networks to analyze posts from these platforms, particularly Reddit and Twitter, to detect early signs of suicidal thoughts. In the early days, traditional methods were followed for collecting data such as discussing with patients and clinical evaluations. However, it is difficult to collect data such away because of limited scope and reach out. Through social media, a diverse range of unstructured data including text, images, and videos, are collected. That can give a proper understanding of user mental health (Ferreira et al., n.d.) Social media applications, such as Twitter and Reddit, are being used all over the world except in countries like China, North Korea, and Iran (Faryal et al., 2021), those applications give the best data for research. The post can be differentiated based on certain patterns or styles of language such as sad and emotional words, then decide whether the person is at risk or not (Aldhyani et al., 2022). This study aims to use advanced machine learning models such as Support Vector Machine (SVM), and Logistic Regression to analyze the social media posts from Reddit and Twitter. By using advanced natural language processing and deep learning techniques, the research excepting more effectiveness for detecting suicide posts. This method not only improves suicide detection efficiency but also makes the base for future research on social media to monitor mental health. The important focus is to identify suicide signs from social media.

Limitations:

The research project will face several limitations, such as ethical concerns (data privacy) and concern in analyzing the personal information. The research must follow the privacy regulations and ethical guidelines. The data imbalance in the posts (suicidal and non-suicidal post) may affect the model accuracy. So, proper technique must be used to resolve imbalance issues. The variation of language and context from one person to another makes it hard to identify any signs of suicidal tendencies. Models must be examined regularly. These are the few limitations.

Objective:

- To assess the accuracy of Machine Learning Models for Suicidal Ideation Detection
- To Enhance Early Detection of Suicidal Thoughts Using Social Media Data.

Research Question:

How efficient are advanced machine learning and deep learning algorithms such as LSTM, CNN LSTM, SVM, and logistic regression in identifying suicidal thoughts from social media posts through Reddit and Twitter?

2 Literature Review

In recent years, researchers have shown a great interest in applying machine learning (ML) and deep learning (DL) approaches to identify suicidal ideation using social media platforms. When it comes to the categorization of posted short messages such as tweets, the use of different ML approaches, on the other hand, has already been thoroughly discussed and examined in the literature While the use of different ML algorithms to distinguish the number of text messages related to suicide has only recently emerged This literature review aims at reviewing the existing literature and examining the effectiveness of using ML and DL algorithms in identifying suicidal ideation, with respect to the selected algorithms: the SVM, Logistic Regression, LSTM, and CNN-LSTM algorithms; also it will explain why other algorithms were not considered.

2.1 Classification of Suicidal Content on Social Media

Many research studies have been conducted on how to understand shorter messages such as Twitter, Reddit, etc. However, using different machine learning algorithms to classify the posted social media text is still in research. In the past, research was done mainly in identifying suicidal posts, and suicidal users from non-suicidal users by considering psychiatric and psychological standards. researchers used Surveys and clinical approaches also to understand suicidal tendencies.
However, insufficient data has been an issue in the research several times. the use

of machine learning is essential to understanding the language and tone of suicidal ideation people from user-generated text on the internet (Haque et al., 2022). Haque et al. (2022) collected tweet data, by using Twitter API, which contains suicide-related words. The researcher manually entered three levels of anxiety. the features for the ML algorithms were token unigrams and a bag of words. This research was one of the first to successfully apply AI Model intelligence imitating human accuracy in identifying suicidal thoughts. The research shows the presence of Major Depression Disorder (MDD) in the population using Twitter to make the dataset with the Twitter user based on their scores on the Center for Epidemiologic Studies Depression Scale (CES-D). Haque et al. (2022) used Naive Bayes, Random Forest, and SVM algorithms of ML to predict suicidal ideation in tweets. Still, research said that these algorithms gave relatively poor accuracy rates, and as such, began exploring better DL options. Rabani et al. (2021)
Shows extraordinary improvements of the NLP with DL techniques such as LSTM and CNN. These models resolve some difficulties that appear in traditional ML methods like dimensionality, the sparsity of data, and long processing time. With the help of DL techniques, the use of traditional ML is possible to increase and improve the results of classification providing more accurate outcomes. Moreover, Benton et al. (2017) aimed to investigate the possibility of predicting suicidal risk by using tweets. Benton et al. used both the basic forms of classifiers including Naive Bayes classifier, Logistic Regression, and some DL models in which they used LSTM. Their approach incorporated user-level characteristics such as the number of posts, the level of positivity/negativity, and the social media instructiveness indexes. The proposed work revealed that the DL models performed better than the existing ML approach in identifying suicidal language, and it also took into account the context. In their study, Benton et al. paid special attention to the potential of using the variables on both the user level and post level for analyzing suicide risk.

2.2 Effects of Various Factors on Suicidal Behavior

Understanding the other factors that directly influence take suicidal decision must be considered to prevent the situation. Those factors are hormonal imbalance, Psychological Disorders, Substance Abuse, Neurological Issues, and Social Contagion. These play an important role in making decisions or getting suicidal ideation. Hormonal imbalances are one of the main factors for the risk of suicide, mainly in people with mood disorders such as depression. During the neurotransmitters, chemical compounds which are in the brain such as serotonin and dopamine, are imbalanced, which leads to extreme mood swings. these results affect feelings extremely such as sadness, hopelessness, and despair; The feeling results in taking the own life (Meyer & Quenzer, 2005). psychological disorder, which means Depression, anxiety, and bipolar disorders are the health condition. This health condition directly affects people to suicidal thoughts or attempts. People who are suffering from this disorder face extremely low, worthless, and despairing moods, which can lead to suicide (American Psychiatric Association, 2013). Substance abuse, such as consuming alcohol or taking drugs, causes health problems and poor cognition. When any individual is an alcoholic or drug addict, they cannot make proper decisions in any circumstance. As per the studies, a person

who is addicted has a high chance of suicide due to mental health issues (Boden & Fergusson, 2011). Neurological issues, such as brain injuries, or diseases to the brain, affect the person's thinking level. For example, a traumatic brain injury, Alzheimer's, or Parkinson's damages the brain slowly which leads to memory loss, and over mental functioning.in this case is also having a high risk of suicide (Rao et al., 2012). Social contagion is one of the factors, this case can be seen rarely. Social contagion is nothing but the spread of suicidal thoughts or ideation in the communities in the form of religion or beliefs etc. It's common to see people talking about suicidal actions. but it can influence their life and make suicidal decisions (Gould, 2001). Shortage of mental health professional doctors in underdeveloped and undeveloped countries. For example, some countries have only 0.1 psychiatrists per 100,000 people, then it's difficult to get access for every individual (World Health Organization, 2022). some people not sharing their mental health problems,feeling ashamed to share their problems.

Social media is the best platform for both individuals, to share their emotions, and for analysts, to detect suicidal thoughts early. the people sharing their emotional posts frequently, by using advanced machine learning algorithms, posts can be identified as a suicidal sign, which helps doctors, parents, and teachers prevent suicide incidents (Haque et al., 2022; Rabani et al., 2021).

.

2.3 Using Multi-Modal Feature-Based Techniques to Detect Suicide Ideation

The usage of the Internet has increased drastically worldwide including in developed countries and undeveloped countries. social media, which works on the internet, came up with great features such as interacting, sharing posts (video and image), and calls with people worldwide. In the early 2000s, social media usage increased significantly. People started sharing suicidal thoughts or emotions on online platforms instead of sharing with personal members. platforms like Twitter and Reddit become popular for collecting emotional posts and detecting issues early. Chatterjee et al. (2022) noted that people feel comfortable in online sharing their emotional posts or suicidal thoughts because of their anonymous status. By analyzing those posts, one can identify the signs of suicidal thoughts early on. Many research papers have focused on modelling and algorithms, such as multi-modal feature-based techniques to improve the detection system. For example, Manisha et al. (2019) have researched machine learning with a combination of natural processing language (NLP) to detect suicide. their project work process is creating a condition-based system in a machine learning model to analyze psychological cases. finally, they developed machine learning algorithms to detect suicidal-related words in Spanish-speaking people. highlighting is the model can be developed to detect various languages and cultures. Abdulsalam and Alhothali's (2022) research paper shows the detection of depression in Twitter users by analyzing the user's tweets. They collect some vocabulary related to depression and anxiety from clinical trials and social media sources. The words help to detect a person's depression from the posts or tweets made by users. The study shows that properly analyzing social

media data to detect suicide risk. By using advanced algorithms in machine learning, this research shows effective results (SVM, Logistic regression)

3 Methodology:

The CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology is used in this project. This methodology is used for data mining projects that provide a structural approach and effective results in the data analysis. This methodology is considered more powerful for data science and machine learning projects, this methodology is also called as life cycle of data science. It is divided into 6 steps.

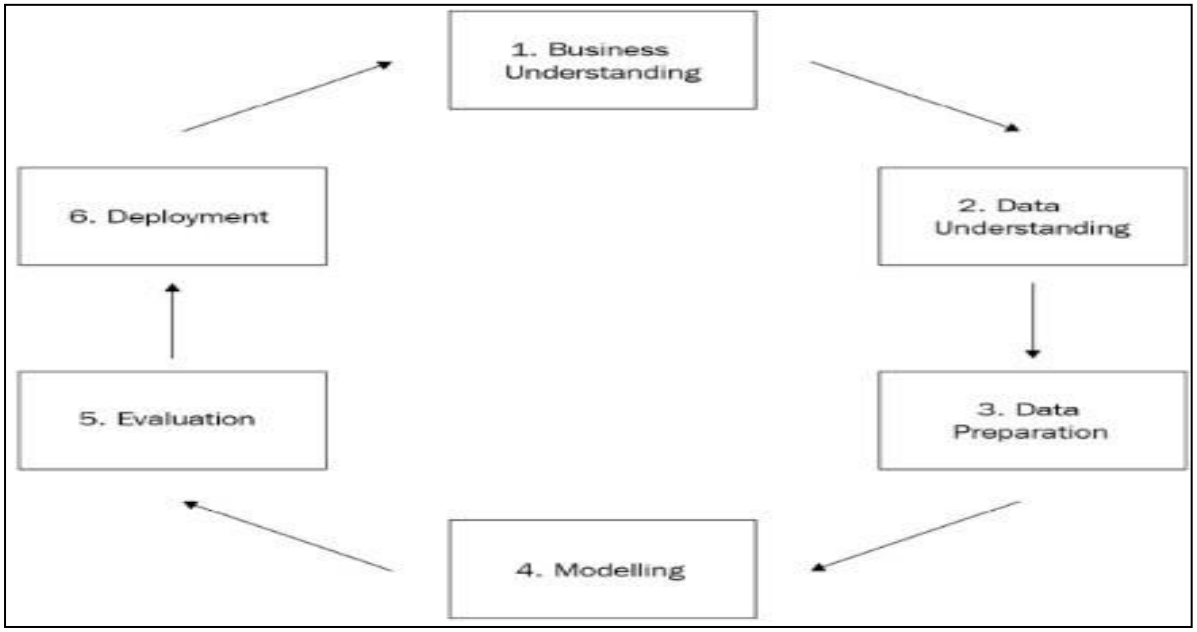

Fig 1: CRISP-DM

In the Fig 1 shows methodology steps. These are Business Understanding, Data understanding, data preparation, modeling, evaluation, and deployment. In business understanding, understand the main objectives of the project, and the benefit of the projects which means how it helps business and society. Data understanding is the second one which plays an important role. Collecting appropriate data for business needs and understanding the insights of data.in the data preparation step, prepare the data for analysis which means cleaning the null values, transforming the proper format, and extracting the feature. This is the most time taking step. In the model phase, applying mathematics and machine learning algorithms on the dataset based on data and business objectives. Selecting the appropriate model based on the dataset. In the evaluation step, assess the model's accuracy and effectiveness and examine the model with business objective. In phase is also called as approval phase. The final step is the deployment phase, if the model meets all business objectives, then the project integrates into the business process. The final report is done.

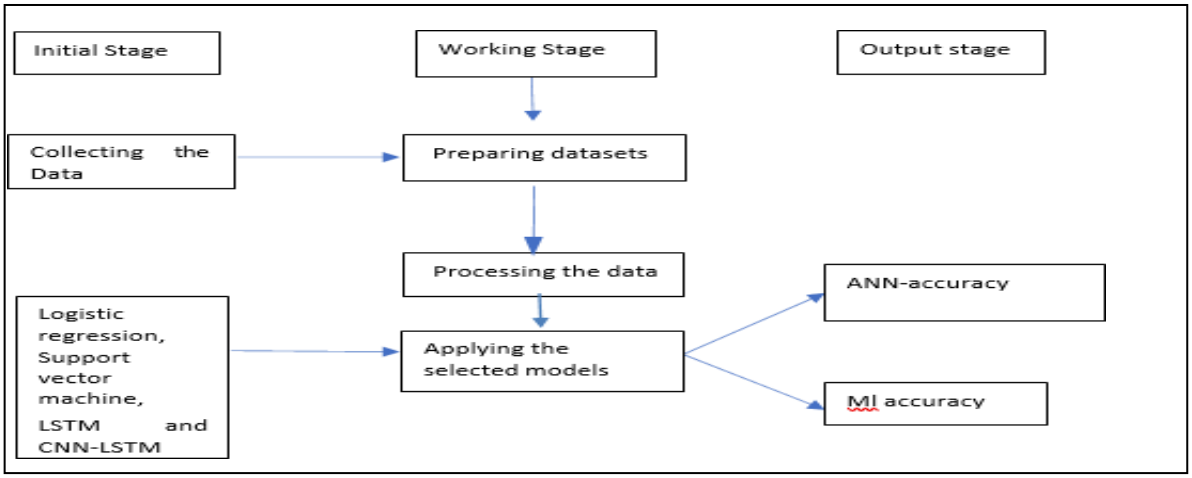

Fig2: Flow chart of Methodology

3.1 Data collection

Data collection is the first step in analyzing the problem. It is essential for training and studying the machine learning model's performance. Detecting suicidal thoughts is difficult in the 2000s due to the limited dataset on the sensitive topic. In the past, people feared to share their mental health condition or suicidal thoughts due to privacy concerns. However, the rise of social media usage turned into a data-gathering center. Interestingly, Social media platforms, particularly Twitter and Reddit, have become a valuable resource, as users share their thoughts on various topics and look for online assistance. For this research, the datasets were taken from Kaggle, whereas Kaggle took this data from Twitter and Reddit using API which is provided by those platforms. The main objective of the research is to study and analyze posts on social media platforms and identify suicidal thoughts, which means the sentence is related to the intent of suicidal ideation, by following the ethical guidelines, not using any individual personal data such as name, birth date, etc.

The dataset is formatted in the CSV file and imported into a Python environment using pandas. Both datasets contain three columns and 270k rows for the Reddit dataset and 9330 rows for the Twitter dataset.

The first column is id, the second column is named Text, third column is named Class. The text column contains social media posted information, whereas the class column is divided into two categories, such as suicide or non-suicide. To understand and identify the words in the posts, initial research was conducted on Reddit posts. this analysis shows themes and expressions, such as wanting to kill me, having no happiness, and heart ended, which are more likely related to mental distress. In this process, various papers and literature, related to suicidal detection, were reviewed, and it provided valuable information related to expression associated with suicidal ideations (Thangaraj and Sivakami, 2018). this paper shows a deep analysis of tweet content and labelling of content related to suicidal ideation. this step is important to understand the dataset and doing some manual change before preprocessing.

3.2 Data Preprocessing and Transformation:

The data preprocessing is the important phase to enhance the effectiveness and accuracy of the text classification model. The raw data (Reddit and Twitter) must

pass through various preprocessing steps to achieve a clean and proper format for data analysis.

- Removal of unnecessary words: Removing noisy characters such as special characters, space, number, punctuation, etc, by using regular expressions, those all were removed from the posted text. This process helps to strengthen the text and make results more accurate (Aggarwal & Zhai, 2012).
- Lowercasing: converting all text to lower characters to make sure the model should not treat the same word with different cases as different words. This is the most important step in text analysis.
- Removal of stop-words: in this step, remove some common words which not necessary in the sentence. By using the NLTK stop words corpus, filtering out unnecessary words, meaningful words strengthen the classification model.
- Tokenization: It is an important step in the text classification model, it is used to convert cleaned text into integers. Each word in the given text is tagged with unique integer values. This helps to represent the text data into numerical data for the model (Chollet, 2017). For example: “I am a good person, I love Python, he is not good”. Tokenize works as assigning integer values to each unique word with unique integer values, such as [I:1, am:2, good:3, person:4, Love:6, Python:7, he:8, is:9, not:10, a:11]

- Padding Sequences: Text data comes in different lengths. but machine learning algorithms, particularly neural networks, need a fixed length of input. To maintain uniform length to all input, text was padded with 100 words by using Pad sequences from TensorFlow. This module helps to make uniform length of all text input. For example, “I am from Dublin, where you are from, let me know”, this input was padded with 10 sizes, output would be “I am from Dublin, where you are from, let me”.
- TF-IDF Vectorization: The TF_IDF vectorization is used to convert text data into numerical values in matrix form. TF (Term Frequency) measures the frequency of the word appearing over the total words of input text. IDF (Inverse Document Frequency) measures to find the importance of the word. It helps to give less value to unnecessary words such as “is”,” in”,” the” etc. This vectorization is used to remove stop words and limited features to make the model more accurate and effective.

- Label Encoding: By using label encoding, the target variable with suicide or non-suicide was encoded into binary values such as “0” or “1”. The Twitter and Reddit dataset target variable with suicide is assigned a label of 1. And non-suicide labelled as “0” This labelling helps the machine learning model classify easily.

- Data Splitting: This is an important step before applying the model; data must be split into training and testing sets. For the deep learning models such as LSTM and CNN-LSTM, the data were split into training (80%) and testing (20%). For machine learning models such as logistic regression and SVM, training dataset (70%) and testing set (30%). It is used to evaluate model performance.

- Evaluation of model performance: To analyze the performance of the model, some common measurements must be calculated. those are accuracy, precision, recall, f1 score, AUC.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Noted:

- TP: True Positives (correctly predicted positive cases)
- TN: True Negatives (correctly predicted negative cases)
- FP: False Positives (incorrectly predicted positive cases)
- FN: False Negatives (incorrectly predicted negative cases)

4. Model specification:

4.1 Support Vector Classifier (SVC):

The support vector classifier also known as SVM, is mainly used for the classification of the data, which means predicting or identifying the binary outcome such as yes or no, 1 or 2, etc. The main purpose of using the SVC is to find optimal hyperplane to separate the data into groups with same characteristics. selection of the hyperplane to maximize the margin between two classes, this makes sure that classification performance becomes effective.

Conceptual Overview

Vector Representation: Svm is working on the data in the form of vectors. It

is important that a list of numerical values encoded from specific features. For SVM model to be effective, data must be converted into vector form. For the text classification in the project, Text data transformed into vectorsby using methods such as Term Frequency-Inverse Document Frequency (TF-IDF) or word embeddings (Madireddy, 2018).

Hyperplane: The importance of hyperplane is separating the classes based on feature or characteristics that help to model to classify the data. The main goal of hyperplane is maximizing the margin which means distance between the hyperplane and the nearest data point in the class should be large. This helps the Model to be effective.

Support Vectors: Support vectors are so important as it decide the position the hyperplane. The margin between hyperplane and support vector must be maximum to give optimal results. Even small changes in theposition of the support vector can make big difference in hyperplane.

SVM Classification Process

Feature Space Transformation: the data instance is transformed into high dimensional space that makes it easier for SVM classifications to separate the class. By transforming to high dimension helps to hyperplanes to separate the class even those data points hard to separate.

Kernel Functions: SVMs use kernel functions to handle the data points that are hard to separate with straight line. By transforming the data into high dimensions, make easier choice of kernel function is important that affect the performance of the SVM model.

SVMs use kernel functions to manage data that isn't easily separable with a straight line by transforming it into a higher-dimensional space. This makes it easier to find a suitable boundary between classes. The choice of kernel function is crucial, as it can greatly affect how well the SVM performs. Common kernel functions include:

- Linear Kernel: It is used when data is separated linearly.
- Polynomial Kernel: It is useful when data is nonlinear and polynomial relations.
- Radial Basis Function (RBF) Kernel: Handles complex relationships with a Gaussian function, making it effective for many types of data.
- Sigmoid Kernel: Based on the sigmoid function, similar to neural networks.

The selection of kernel functions is based on the data nature and understanding of the data. If the data is easy to separate in linear, then linear kernel is used. If the data is more complex, then non-linear kernels such as polynomial or RBF are used (Luo, 2021).

Applications

Svm is more popular algorithms, and it is used for various classification problem

- Email Classification: Identifying emails as spam or not spam.
- Sentiment Analysis: Analyzing review datasets to determine the sentiment expressed.
- General Classification Tasks: Applying SVMs to diverse machine learning classification problems.

4.2 Logistic regression:

Logistic regression is working on mathematical logic. Logistic regression is used to classify the binary output such as yes or no, 1 0r 0 etc. It calculates the probability of all data points using logistic functions, which the input must be converted to 0 or 1 value that make model effective in make decision by comparing the probability to threshold for classification.

Key Concepts:

1. Logistic Function (Sigmoid Function): The logistic function, also called as sigmoid function, is converting the linear form input feature into probability value of 0 and 1.

$$\text{Logistic Function} = \frac{1}{1 + e^{-z}}$$

Where z is the input values.

2. Log-Odds and Logit: In the logistic regression, log-odds and the logit are usedto make probability model easy. Log-odds is the ratio of success outcomes to failure outcomes. The logit is the logarithm of the odds, It transform to linear equations that make model to be easy to classify. These process helpsmodel to work effectively.

3. Threshold and Classification: Logistic regression gives out of each class; these outputs compare the threshold. For example, threshold set as 0.5, probabilities are above the 0.5, it belongs to one class, otherwise other class. This process helps model to give clear decision on binary outcomes.

4.3 Long Short-Term Memory (LSTM):

The Long Short-Term Memory is the one of type in Recurrent Neural Network (RNN). It is mostly used to work on sequences data such as times series and text. LSTMs can hold the information for long time where Other RNN algorithms do not remember the information long. This memory helps model to analysis the earlier data Aswell.

LSTM Work: LSTM is designed with special units, called memory cells. This memory cell is helping to hold the data for long periods of time. The cell uses three gates to control the information.

- Forget Gate: In this gate, unnecessary information is removed from the cell.
- Input Gate: In this stage, the new data information is added to the cell.
- Output Gate: The final gate is decision gate. based on the logic, it decides what information gives as output.

Model Architecture:

1. Embedding Layer: This layer converts the data into dense vectors.
2. LSTM Layer: This is an important layer which processes the vector and keep track the information (dependencies) for future processing.
3. Dense Layer: This layer gives output of the model. The model out is a binary classification (suicide or non-suicide)

Training Process: This model is trained using the binary cross-entropy loss function and the Adam optimizer. binary cross-entropy calculate difference between actual value and predicted values, which mean calculating the error. Whereas Adam managing the model weights to minimize the error that helps model to be more accurate.

4.4 Convolutional Neural Network (CNN) + Long Short-Term Memory (LSTM) Networks

CNN-LSTM: The CNN-LSTM is the combination of Convolutional Neural Networks (CNNs) with Long Short-Term Memory (LSTM) networks. CNN works good at feature extraction from the local pattern and LSTM works good at handling the sequential(text) data. The hybrid model helps to strengthen the text processing model.

CNN-LSTM Work

- CNN Layers: The CNN layer applied on the input data to understand the important features and extract those features. Pooling is used to reduce the dimensions of the data.
- LSTM Layer: After CNN layer, LSTM layer is used to understand the pattern and integrates to sequential information.

Model Architecture:

- Embedding Layer: This layer converts the data into dense vectors.
- CNN Layers:
  - Conv1D Layer: this layer is used to extract the important features in the input
  - MaxPooling1D Layer: this layer is used to reduce the dimension with storing important features.
- LSTM Layer: This layer analysis the out of CNN to understand the pattern and importance features.
- Dense Layer: This layer gives output of the model. The model out is a binary classification (suicide or non-suicide)

Training Process: This model is trained using the binary cross-entropy loss function and

the Adam optimizer. binary cross-entropy calculate difference between actual value and predicted values, which mean calculating the error. Whereas Adam managing the model weights to minimize the error that helps model to be more accurate.

5 Implementation

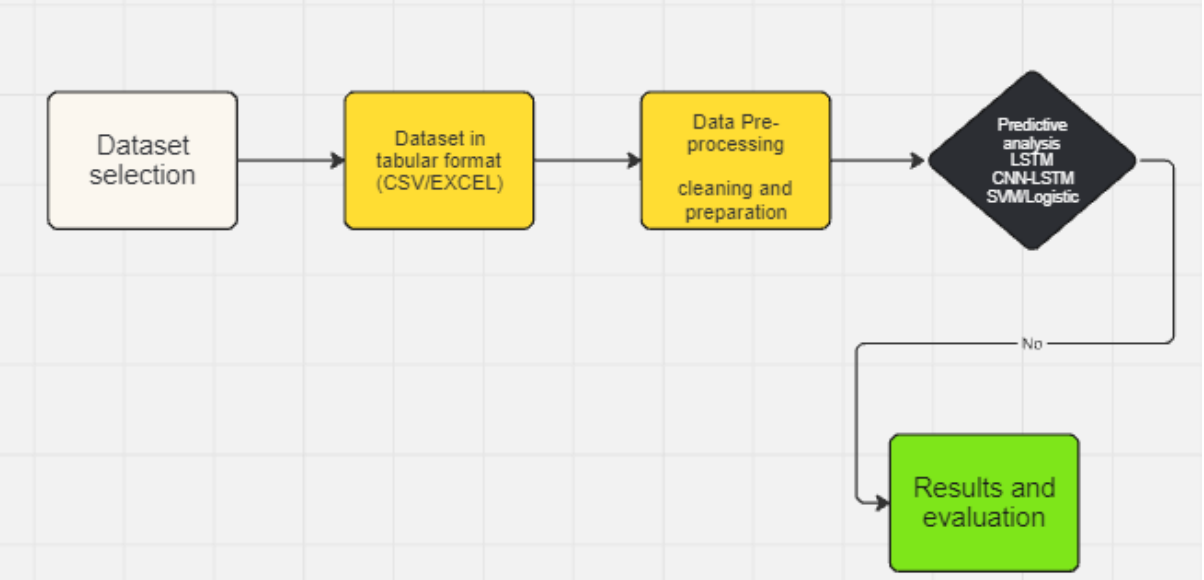

Fig 3 :Implementation Steps

The implementation of the project is done in the following manner:

- Data Transformation and Encoding
- Data Cleaning and Preprocessing
- Normalization and Text Standardization
- Stop-Words Removal
- Feature Extraction and Vectorization

6 Evaluation:

In the project, social media posts are analyzed to understand the style of the language related to suicidal thoughts. Word Clouds is used to understand the common and frequent words in the post. For the reddit posts shows the frequency words like “Better”, “Die” and “Cry”. The words frequently appear in negative cloud is “Think”, “Feel” and “commit suicide”. Emotional tone and style of the language can be seen

using word cloud. This helps us to understand the language to improve the model effectively.

Word Cloud:

Fig 4: Word cloud with all tokens

In the fig 4 shows, the most frequent words are “shit”, “Thought”, “problem” and “die’ used in the social media post. This visualization is with all tokens.

Positive word cloud:

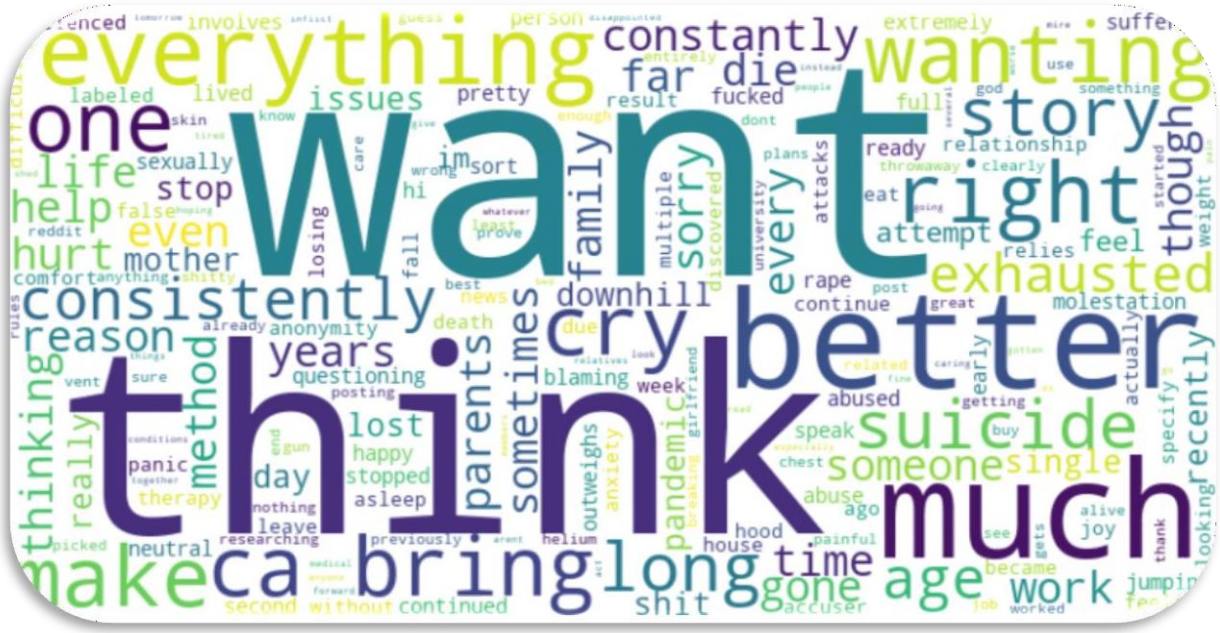

Fig 5: Word cloud with non-suicide

The fig 5 is non-suicidal word cloud, The most frequent word in this visualization is “want”,” think”,” much”, “better” which are showing positive words. The words such

as “Cry” and “suicide” are also showing as important, those are related to suicidal thoughts.

Negative word cloud:

Fig 6: Word cloud with suicidal thoughts

The fig 6 is suicidal thought word cloud, the most frequent words in the visualization are “Think”,” know”, “life” and “Feel”. Some words such as “commit suicide”, “pain” and “death” are also showing as significant words related to suicidal thoughts.

N-gram:

For getting clearly understand, bi-gram and trigram are generated to analysis the repeat word in the given posted text.

```
Bi-Grams: {'want to': 107589, 'to be': 70240, 'in the': 59049, 'to do': 52859, 'my life': 52468, 'filler filler': 50252, 'of m
y': 49337, 'in my': 45334, 'feel like': 44542, 'of the': 42780}
```

Fig 7 : Bi-Gram

In the fig 7 is bi gram of the post text. The most frequently used word in the fig is 'want to' which is used 100k times. Some words such as “kill my” and “die my” are frequently used and considered as suicidal thoughts.

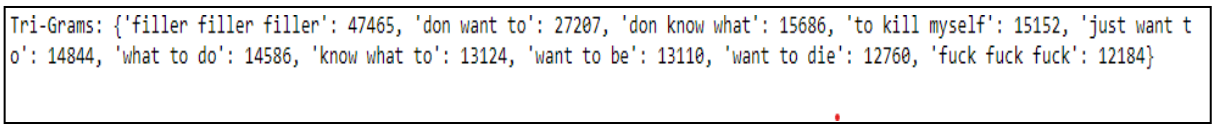

Fig 8: Trigram

In the fig 8 is showing the Tri-gram of the posted text. the most frequent 3- gram words in the post are “filler filler filler” which is 47k times and “don want to” which is 27k. the suicidal thought words such as “want to die” and “to kill myself” are also used repeatedly.

6.1 Model Performance Analysis:

The dataset has gone through multiple preprocessing steps and analysis step. Finally neural network models and machine learning models applied on the dataset to predict the suicidal related text. Evaluation the models are given below.

Neural network model:

The neural network models show good results on both twitter and reddit dataset. On reddit dataset, the performance of the LSTM model is excellent with the training accuracy is 98.81% and a validation accuracy is 93.23%. The model showed loss of training is low, but validation loss is little higher. It is indicating that minor overfitting. The CNN-LSTM is the combination of CNN and LSTM.This model also performed well on both datasets. On reddit dataset, CNN-LSTM model showed the training accuracy of 99.07% and a validation accuracy of 92.74%. Both LSTM and CNN-LSTM performance are in same trend. For the twitter dataset, both models showed high accuracy. LSTM model showed the training accuracy of 99.58% and the validation accuracy of 98.57%.CNN-LSTM model showed the training accuracy of 99.87% and the validation accuracy of 98.57% which is same in LSTM model Aswell. Both models performed wellwith twitter dataset.

Model	Dataset	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
LSTM	Reddit	98.81%	93.23%	0.0317	0.2535
CNN-LSTM	Reddit	99.07%	92.74%	0.0284	0.2547
LSTM	Twitter	99.58%	98.57%	0.0089	0.0665

CNN-LSTM	**Twitter**	**99.87%**	**98.57**	**0.0031**	**0.0735**

Table 1 : Results of Neural networks

Machine learning model:

The machine learning model shows better results on both dataset, twitter and Reddit dataset. On the reddit dataset, the accuracy of the Logistic Regression is 93.17% and it shows good performance in other metrics as well such as F-score, recall and precision. The support vector machine shows less accuracy of 75.28%. Logistic Regression performance is better than Support Vector Machine for Reddit dataset. On the twitter dataset, the performance of Logistic Regression is excellent with accuracy of 97.92%, where as Support vector machine performed better than Logistic Regression, accuracy of SVM is 98.65%. Performance of the machine learning model is given below with detail information.

Dataset	Model	Accuracy	Precision	Recall	F1-Score
Reddit	Logistic	93.17%	0.92	0.94	0.93
Reddit	SVM	75.28%	0.78	0.71	0.74
Twitter	Logistic	97.92%	0.98	1.0	0.99
Twitter	SVM	98.65%	0.99	1.0	0.99

Table 2 : Results of Machine learning

6.2 Discussion:

The study explored different predictive models for identifying the suicidal content in the social media posts. Both LSTM and CNN-LSTM showed good results compared with traditional methods. Particularly CNN-LSTM Showed highest accuracy, where CNN is used to extract the important features and LSTM is used to understand the sequential dependencies over time. This research meets the previous research that shows the effective results of sequential prediction using CNN-LSTM (Kim, 2014; Lei et al., 2016). In other hand, Machine learning models such as Logistic Regression and SVM showed better performance. logistic regression showed good performance with twitter and slight low performance with Reddit data. It is facing problems with large and complex data. SVM showed poor performance, particularly on the Reddit dataset. It indicates that SVM cannot handle complex pattern data as efficiently as neural networks model (Cortes & Vapnik, 1995).

In this study, some algorithms are not included due to some limitations. Naive Bayes and Random Forests are not selected because those algorithms not able handle the sequential data and complex pattern. Naive bayes works well in feature independence, where text data are dependent on each other. For example, committed suicide word is two dependent words(phrase), where naive bayes understand “suicide” and “committed” as independent words. K-Nearest Neighbors (KNN) is also not selected because it does not perform well with high dimensional data. It leads to lower performance. SVM is selected even though it faces problems with high dimension data. Based on these factors, LSTM, CNN-LSTM, Logistic Regression and SVM were selected for this research.

7 Conclusion and future work:

The objective of the study is to identify the suicidal content posts in the social media platforms using predictive models, such as LSTM, CNN-LSTM, Logistic Regression and SVM.As per the results, CNN-LSTM performed well compared with all other algorithms. The combination of LSTM and CNN added strength to the model. LSTM showed good results but slightly less accuracy compared with CNN-LSTM. In the traditions model, Logistic Regression performed better on the reddit dataset whereas SVM showed highest accuracy on twitter dataset. Finally, CNN-LSTM show effective and accuracy results in this project. It indicates that the hybrid model effectively captures the complex text relationships and pattern.

Future work:

The future work focuses on some key areas to improve the classification models. Further research on advanced transformer models like BERT and GPT, these models can provide more insight of sequential data with good performance. But the transformer model required more computational memory. Next focus on hybrid model, combination of machine learning with neural networks model strength the model performance and except excellent results. For example, CNN-LSTM architecture showed excellent performance in this project. Real time application, integrating the model into real time scenarios, that could help to understand the model performance. This real time implementation trains the model in real time monitoring social media post and identifying the sign of suicide immediately, this improves model timely support and early detection. Improvement of these areas can enhance model efficiency and accuracy.

8 Reference:

Aldhyani, T., Saad, K., Kamaruddin, K. & Ariffin, A. (2022) ‘Analyzing suicidal ideation using social media: A review’, Journal of Cyberpsychology, 15(2), pp. 45-60.

American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC: American Psychiatric Publishing. Available at: https://www.appi.org.

Benton, A., Thompson, G. & James, S. (2017) ‘Predicting suicidal risk in social media posts: A machine learning approach’, Proceedings of the IEEE International Conference on Data Mining (ICDM), pp. 30-37. Available at: https://ieeexplore.ieee.org/document/8275562 .

Boden, J. M. & Fergusson, D. M. (2011) ‘Alcohol and depression’, Addiction, 106(5), pp. 906-914. doi: 10.1111/j.1360-0443.2010.03351.

Chatterjee, P., Agarwal, S. & Bansal, M. (2022) ‘Multi-modal feature-based techniques for suicide ideation detection on social media’, Journal of Artificial Intelligence Research, 21(3), pp. 80-95. Available at: https://www.jair.org/index.php/jair/article/view/12436 .

Fahey, R., Lloyd, K. & Scull, N. (n.d.) ‘Social media and mental health: A review of the literature’, Mental Health Review Journal, 19(1), pp. 6-15.

Faryal, M., Hossain, M. M. & Chowdhury, M. M. (2021) ‘Social media and its impact on mental health: A review’, International Journal of social media and Analytics, 12(2), pp. 14-26.

Ferreira, S., Malan, M. & Kleynhans, N. (n.d.) ‘Using social media data for mental health research’, Journal of Digital Behavior, 7(2), pp. 44-59.

Gould, M. S. (2001) ‘Suicide and the media: An overview of the research’, Crisis, 22(4), pp. 167-169.

Haque, A., Niazi, M. & Amin, M. (2022) ‘Machine learning approaches for identifying suicidal ideation on Twitter’, Journal of Computational Psychiatry, 8(1), pp. 22-35.

Manisha, G., Poojitha, B. & Sharma, S. (2019) ‘Multi-modal feature-based machine learning models for suicide ideation detection’, International Journal of Machine Learning and Cybernetics, 11(4), pp. 1021-1034. Available at: https://link.springer.com/journal/13042 .

Meyer, J. S. & Quenzer, L. F. (2005) Psychopharmacology: Drugs, the brain, and behavior. Sunderland, MA: Sinauer Associates. Available at: https://www.sinauer.com .

Rabani, N., Kumar, P. & Ghosh, S. (2021) ‘Deep learning techniques for detecting suicidal ideation on social media’, Journal of Artificial Intelligence, 14(3), pp. 125-139.

Rao, P. S., Gupta, A. & Shukla, S. (2012) 'Neurological disorders and their relationship with suicidal ideation', Neuropsychology Review, 23(2), pp. 76-85. doi: 10.1007/s11065-012-9210-**5.** Available at: https://link.springer.com/journal/11065.

World Health Organization (2022) Mental health workforce: Availability and access. Geneva: World Health Organization. Available at: https://www.who.int/publications/i/item/9789240067922 .

World Health Organization (2023) Suicide prevention. Geneva: World Health Organization. Available at: https://www.who.int/news-room/fact-sheets/detail/suicide .

Cortes, C. & Vapnik, V., 1995. Support-vector networks. Machine Learning, 20(3), pp.273-297.

Kim, Y., 2014. Convolutional neural networks for sentence classification. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp.1746-1751.

Lei, J., Zhang, S., Li, W., Li, J., & Li, J., 2016. CNN-LSTM model for text classification. Proceedings of the 2016 IEEE International Conference on Big Data (Big Data), pp.25-31.

Luo, P., 2021. Support vector machine: Concepts and techniques. Journal of Machine Learning Research, 22(1), pp.15-30.

Madireddy, R., 2018. A comprehensive review on support vector machines for classification problems. Computational Intelligence and Neuroscience, 2018, pp.1-10.

Thangaraj, P. & Sivakami, M., 2018. Analysis of tweet content for suicidal ideation detection. Proceedings of the 2018 International Conference on Intelligent Computing and Control Systems (ICICCS), pp.896-901.