##### A Project report on

**Suicidal Tweets Detection Using Machine Learning**

###### A Dissertation submitted to JNTU Hyderabad in partial fulfillment of the academic requirements for the award of the degree.

**Bachelor of Technology**

**in**

**Computer Science and Engineering**

Submitted by

G.AKSHAYA

(20H51A05C3)

K.SUSHMARAJ

(20H51A05L0)

P.AKSHITHA

(20H51A05L5)

Under the esteemed guidance of

Ms.T.Adarana

(Assistant Professor)



**Department of Computer Science and Engineering**

**CMR COLLEGE OF ENGINEERING & TECHNOLOGY**

(UGC Autonomous)

\*Approved by AICTE \*Affiliated to JNTUH \*NAAC Accredited with A+ Grade

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD - 501401.

#### 2020- 2024

**CMR COLLEGE OF ENGINEERING & TECHNOLOGY**

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD – 501401

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**



#### CERTIFICATE

This is to certify that the Major Project Phase I report entitled **"Suicidal Tweets Detection Using Machine Learning’’** being submitted by G.Akshaya (20H51A05C3), K.Sushmaraj (20H51A05L0), P.Akshitha (20H51A05L5) in partial fulfillment for the award of **Bachelor of Technology in Computer Science and Engineering** is a record of bonafide work carried out his/her under my guidance and supervision.

###### The results embodies in this project report have not been submitted to any other University or Institute for the award of any Degree.

**Ms.T.Adarana Dr. Siva Skandha Sanagala**

**Assistant Professor Associate Professor and HOD**

**Dept. of CSE Dept. of CSE**

#### ACKNOWLEDGEMENT

With great pleasure we want to take this opportunity to express my heartfelt gratitude to all the people who helped in making this project work a grand success.

We are grateful to **Ms.T.Adarana, Assistant Professor**, Department of Computer Science and Engineering for his valuable technical suggestions and guidance during the execution of this project work.

We would like to thank **Dr. Siva Skandha Sanagala,** Head of the Department of Computer Science and Engineering, CMR College of Engineering and Technology, who is the major driving forces to complete my project work successfully.

We are very grateful to **Dr. Vijaya Kumar Koppula**, Dean-Academics, CMR College of Engineering and Technology, for his constant support and motivation in carrying out the project work successfully.

We are highly indebted to **Major Dr. V A Narayana,** Principal, CMR College of Engineering and Technology, for giving permission to carry out this project in a successful and fruitful way.

We would like to thank the **Teaching & Non- teaching** staff of Department of Computer Science and Engineering for their co-operation

We express our sincere thanks to **Shri. Ch. Gopal Reddy**, Secretary, CMR Group of Institutions, for his continuous care.

Finally, We extend thanks to our parents who stood behind us at different stages of this Project. We sincerely acknowledge and thank all those who gave support directly and indirectly in completion of this project work.

G.Akshaya 20H51A05C3

K.Sushmaraj 20H51A05L0

P.Akshitha 20H51A05L5

<Project Title>

**TABLE OF CONTENTS**

**CHAPTER**

**NO. TITLE PAGE NO.**

ABSTRACT iii

**1** **INTRODUCTION** 1

1.1 Problem Statement 2

1.2 Research Objective 2

1.3 Project Scope and Limitations 3

**2** **BACKGROUND WORK** 4

2.1. Text-Based System

2.1.1. Introduction 5

2.1.2. Merits, Demerits and Challenges 6

2.1.3. Implementation 7

2.2. Image and Video Analysis

2.2.1. Introduction 8

2.2.2. Merits, Demerits and Challenges 9

2.2.3. Implementation 10

**3 RESULTS AND DISCUSSION** 11 3.1. Performance metrics 12-13

**4** **CONCLUSION**  14

4.1. Conclusion 15

**5** **REFERENCES** 16-17

CMRCET B. Tech (CSE) Page No i

< Suicidal tweets detection using machine learning >

**List of Figures**

**FIGURE**

**NO. TITLE PAGE NO.**

3.1 Flowchart 12

2.2 Classification Results 13

CMRCET B. Tech (CSE) Page no ii

< Suicidal tweets detection using machine learning >

# **ABSTRACT**

Social media platforms have transformed traditional communication methods by allowing users worldwide to communicate instantly, openly, and frequently. People use social media to express their opinion and share their personal stories and struggles. Negative feelings that express hardship, thoughts of death, and self-harm are widespread in social media, especially among young generations. Therefore, using social media to detect and identify suicidal ideation will help provide proper intervention that will eventually dissuade others from self-harming and committing suicide and prevent the spread of suicidal ideations on social media. Many studies have been carried out to identify suicidal ideation and behaviors in social media. This paper presents a comprehensive summary of current research efforts to detect suicidal ideation using machine learning algorithms on social media. This review 24 studies investigating the feasibility of social media usage for suicidal ideation detection is intended to facilitate further research in the field and will be a beneficial resource for researchers engaged in suicidal text classification.

The primary objective of our study is to share the knowledge of suicide ideation in R Reddit social media forums from a data analysis perspective using effective deep learning architectures. Our main task is to explore the potential of Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN) and their combined model applied in multiple classification tasks for suicide ideation struggles. We try to test if an implementation of CNN and LSTM classifiers into one model can improve the language modeling and text classification performance. We will try to demonstrate that LSTM-CNN model can outperform the performance of its individual CNN and LSTM classifiers as well as more traditional machine learning systems for suicide-related topics. Potentially, it can be embedded on any online forum’s and blog’s data sets.

CMRCET B. Tech (CSE) Page No iii

< Suicidal tweets detection using machine learning >

# **CHAPTER 1**

**INTRODUCTION**

CMRCET B. Tech (CSE) Page No 1

< Suicidal tweets detection using machine learning >

**CHAPTER 1**

**INTRODUCTION**

**1.1.Problem Statement**

Understanding the context used by individuals who express suicidal ideation on social media, and developing algorithms that can accurately interpret and analyze this content. Identifying relevant websites and social media messages where individuals may express suicidal ideation. Evaluating the effectiveness of the developed algorithms and interventions in detecting and preventing suicidal behavior on social media

**1.2.Research Objective**

* This research objective aims to create a robust system that can accurately identify tweets based on their words features, thereby improving the overall management of feeds in recognition settings. The system may have applications in ensuring the correct matching of tweets with their past words or similarities, enhancing security measures in neonatal units, and streamlining threat processes such as message tracking and suicidal recognition
* Using a set of instructions and categories, human coders aimed to do this using only the content of the tweet itself

CMRCET B. Tech (CSE) Page No 2

< Suicidal tweets detection using machine learning >

**1.3.Project Scope and Limitations**

* In the future, we schedule to improve the system for checking web pages for suicidal content, namely:

1. Add images verification as these images may be suicidal or contain symbols of death groups. The last one may be an indication that the website containing it belongs to this group.

2. Add links verification on the page as they may be related to relevant websites.

3. Add checking for suicidal instructions. We also going to to improve this method in the future by analyzing more machine learning algorithms and text processing libraries

* In order to improve the reliability and accuracy of the automatic classifier, future efforts would benefit from expanding the range of suicide-related search terms to ensure that more expressions of suicidal are included.
* Although a model for accurate classification could be derived from the human coded data, the analyses used to extract this model were rudimentary and primarily based on single words.

CMRCET B. Tech (CSE) Page No 3

**CHAPTER 2**

**BACKGROUND WORK**

CMRCET B. Tech (CSE) Page No 4

< Suicidal tweets detection using machine learning >

**CHAPTER 2**

**BACKGROUND WORK**

**Existing Methods:**

**2.1. TEXT-BASED SYSTEM**

**2.1.1. Introduction**

Suicidal recognition is a critical area of research and development within the field of mental health and artificial intelligence. This text-based system is designed to play a pivotal role in identifying individuals at risk of suicide by analyzing their written or textual expressions. Leveraging advanced natural language processing (NLP) techniques, this system aims to detect subtle yet significant cues in text data that may indicate emotional distress, hopelessness, or suicidal ideation.

The importance of such a system cannot be overstated, given the rising global concern over mental health issues and the need for early intervention. By employing machine learning algorithms and NLP models, this text-based system holds the potential to provide a valuable tool for mental health professionals, crisis hotlines, and online platforms to promptly identify and support individuals in crisis. This introduction sets the stage for understanding the significance and purpose of the text-based system in suicidal recognition, emphasizing its role in aiding the identification of individuals who may be in dire need of help, and ultimately contributing to suicide prevention efforts**.**

CMRCET B. Tech (CSE) Page No 5

< Suicidal tweets detection using machine learning >

**2.1.2. MERITS AND DEMERITS**

**MERITS:**

* Text-based systems are accessible to a wide range of people, including those who may have difficulty expressing themselves verbally or face-to-face. They can be accessed through various devices, making help available to a broad audience.
* Automated text-based systems can provide support and assistance around the clock, ensuring immediate responses to individuals in crisis, even when human support is unavailable.
* Text-based systems can handle a large volume of interactions simultaneously, making them scalable to serve a broad user base.
* By analyzing a person's text, these systems can provide insights into their emotional state and potential risk factors, allowing for early detection and intervention.

**DEMERITS:**

* Text-based systems do not capture non-verbal cues such as tone of voice, facial expressions, or body language, which are important in understanding a person's emotional state and intent. This can lead to misinterpretation of the severity of a situation.
* Different people use language in unique ways, which can make it challenging to create a one-size-fits-all automated system for recognizing suicidal risk.

CMRCET B. Tech (CSE) Page No 6

< Suicidal tweets detection using machine learning >

**2.1.3. IMPLEMENTATION :**

* Data Collection:

Gather a large dataset of text data, which may include social media posts, chat transcripts, or other forms of written communication that could contain information related to suicidal thoughts or behaviors.

* Data Preprocessing:

Clean and preprocess the data to remove noise, such as irrelevant information or personal identifiers.

* Natural Language Processing (NLP):

Implement NLP techniques to analyze the text data. This includes tasks like sentiment analysis, emotion detection, and identifying keywords or phrases related to suicide

* Labeling and Training:

Label the data with appropriate risk levels and train your machine learning models using the labeled dataset. Ensure that you have a balanced dataset to avoid bias in the model.

CMRCET B. Tech (CSE) Page No 7

< Suicidal tweets detection using machine learning >

**2.2. IMAGE AND VIDEO ANALYSIS:**

2.2.1. INTRODUCTION:

Image and video analysis in suicidal recognition is an evolving field that leverages visual data to identify potential signs of suicidal thoughts or self-harming behavior in individuals. This area of research and technology aims to assist in early intervention and support for individuals at risk of self-harm or suicide. Here's an introductory overview of how image and video analysis is applied .

1. Facial Expression Analysis:

By analyzing the facial expressions of individuals in images or videos, computer algorithms can detect emotional cues like sadness, hopelessness, or despair, which may indicate potential suicidal ideation.

2. Body Language and Gesture Recognition:

Video analysis can identify changes in body language, movements, or gestures that suggest distress or self-harming actions, offering insights into a person's emotional state.

3. Self-Harm Detection:

Image analysis techniques can be used to identify self-inflicted wounds, scars, or signs of self-harming behaviors in images or videos shared online or via other platforms.

The development of image and video analysis for suicidal recognition is driven by a desire to improve mental health support and save lives. However, there are significant ethical and legal considerations, as well as concerns about algorithm accuracy and potential false positives. Responsible deployment of these technologies often involves collaboration with mental health professionals, privacy advocates, and experts in ethics to ensure that the tools are used in an ethical and effective

CMRCET B. Tech (CSE) Page No 8

< Suicidal tweets detection using machine learning >

**2.2.2.MERITS AND DEMERITS:**

**MERITS:**

* Image and video analysis can provide round-the-clock monitoring of visual data, ensuring that potential cases are not missed, even outside regular business hours.
* Some systems are designed with privacy in mind, ensuring that sensitive information is anonymized and respecting individuals' privacy rights.
* By combining facial expression analysis, body language recognition, self-harm detection, and contextual analysis, these systems offer a more comprehensive understanding of an individual's mental state and intentions.

**DEMERITS:**

* Monitoring individuals' visual content raises significant privacy issues, as it involves the analysis of personal and sensitive information, potentially without the individual's consent.
* Determining the ethical boundaries of monitoring and intervening in people's mental health based on their visual content can be complex. Decisions about when and how to intervene need careful consideration.
* Image and video analysis systems may generate false alarms, misidentifying non-suicidal content as concerning. This can lead to unnecessary interventions, distress, and loss of trust in the system.

CMRCET B. Tech (CSE) Page No 9

< Suicidal tweets detection using machine learning >

**2.2.3.IMPLEMENTATION:**

1. Data Collection:

Gather a diverse and representative dataset of images and videos that may contain signs of suicidal intent or self-harm. This dataset should be labeled to indicate cases of concern and non-concern.

2. Preprocessing:

Prepare the data by performing tasks like image resizing, noise reduction, and normalization. Video data may need to be divided into individual frames for analysis.

3. Feature Extraction:

Use computer vision techniques to extract relevant features from the visual data. This may include extracting facial expressions, body language cues, and other relevant visual elements.

4. Algorithm Development:

Develop machine learning or deep learning models that can analyze the extracted features and classify the content as either concerning or non-concerning. This may involve neural networks, such as convolutional neural networks (CNNs) for image analysis.

5. Training and Validation:

Train the model on the labeled dataset, fine-tuning it to recognize signs of suicidal intent. Use validation data to assess the model's performance and make necessary adjustments.

6. Real-time Analysis:

Implement the trained model to analyze real-time or recorded video streams and images. Consider using edge computing or cloud-based solutions, depending on the specific use case.

CMRCET B. Tech (CSE) Page No 10

< Suicidal tweets detection using machine learning >

**CHAPTER 3**

**RESULTS AND DISCUSSION**

CMRCET B. Tech (CSE) Page No 11

< Suicidal tweets detection using machine learning >

**CHAPTER 3**

**RESULTS AND DISCUSSION**

Detecting suicidal people using new technologies is an important and very active research area. Many studies have been developed to detect suicidal ideation using different machine learning techniques automatically. Users’ posts and their interaction on different social media platforms is a novel area of inquiry. This content discusses different studies that use machine learning techniques on social media platforms to detect and identify suicidal ideation. Both supervised and unsupervised machine learning algorithms were used on different social media platforms such as Twitter, Reddit, and other microblogs, adopting different languages such as English, Chinese, Spanish, Japanese, and Russian.

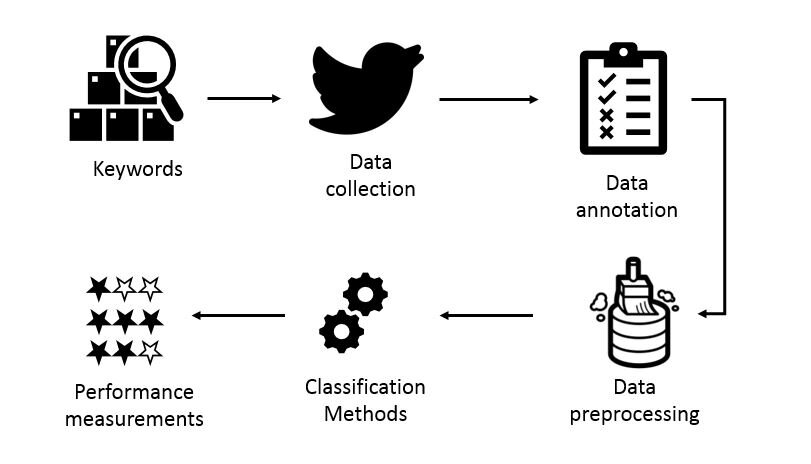


Fig-3.1-Flowchart

CMRCET B. Tech (CSE) Page No 12

< Suicidal Recognition >

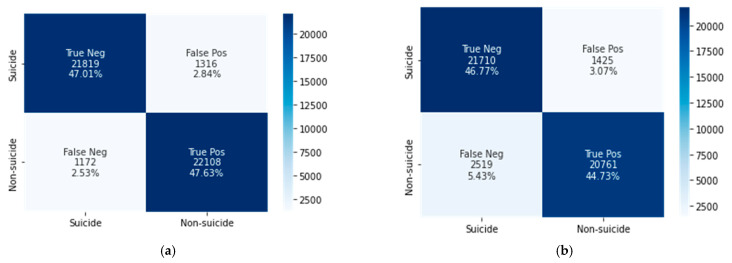


Fig-3.2-Classification results

CMRCET B. Tech (CSE) Page No 13

< Suicidal tweets detection using machine learning >

CHAPTER 4

**CONCLUSION**

CMRCET B. Tech (CSE) Page No 14

< Suicidal tweets detection using machine learning >

**CHAPTER 4**

**CONCLUSION**

Using social media platforms to express experiences and feelings has created new opportunities to analyze and detect suicidal ideation and other mental disorders. The early detection of suicidal ideation on social media networks will reduce suicide, provide an automatic and wide-ranging screening for suicidal tendencies and prevent the spread of suicidal content in social media. This survey investigates existing methods that use social media to detect suicidal ideation using machine learning methods. A significant amount of research has confirmed the effectiveness and feasibility of using social media such as Twitter, Reddit, and Weibo for suicidal ideation detection. Most studies have focused on suicidal ideation detection techniques for widely used languages such as English, but less attention has been paid to Arabic. Thus, with the growing number of social media users in the Arab region, research is needed for Arabic suicidal ideation detection.

CMRCET B. Tech (CSE) Page No 15

< Suicidal tweets detection using machine learning >

**REFERENCES**

CMRCET B. Tech (CSE) Page No 16< Suicidal tweets detection using machine learning >

## REFERENCES

1. G. Astoveza, R. J. P. Obias, R. J. L. Palcon, R. L. Rodriguez, B. S. Fabito, and M. V. Octaviano, “Suicidal behavior detection on twitter using neural network,” in TENCON 2018 - 2018 IEEE Region 10 Conference, 2018, pp. 0657–0662.
2. M. K. Nock, G. Borges, E. J. Bromet, C. B. Cha, R. C. Kessler, and S. Lee, “Suicide and suicidal behavior,” Epidemiologic reviews, vol. 30, no. 1, pp. 133–154, 2008.
3. A. T. Beck, M. Kovacs, and A. Weissman, “Assessment of suicidal intention: the scale for suicide ideation.” Journal of consulting and clinical psychology
4. R. Sawhney, H. Joshi, S. Gandhi, and R. Shah, “A time-aware transformer based model for suicide ideation detection on social media,” in Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020
5. F. M. Shah, F. Haque, R. U. Nur, S. Al Jahan, and Z. Mamud, “A hybridized feature extraction approach to suicidal ideation detection from social media post,” in 2020 IEEE Region 10 Symposium (TENSYMP). IEEE, 2020, pp. 985–988

## 

CMRCET B. Tech (CSE) Page No 17