An NLP-Based Intelligent Model Approach for Suicidal Tendency Classification

BY

Syeed Hasan Salim
ID: 192-15-13158

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Nasima Islam Bithi
Lecturer
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

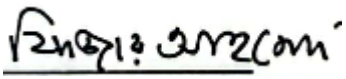
July 2024

APPROVAL

This Project titled "An NLP-Based Intelligent Model Approach for Suicidal Tendency Classification", submitted by **Syeed Hasan Salim** to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 13-07-2024.

BOARD OF EXAMINERS

Dr. Md. Taimur Ahad (MTA) — **Board Chairman**
Associate Professor & Associate Head
Department of CSE
Daffodil International University



Dr. Fizar Ahmed (FZA) — **Internal Examiner 1**
Associate Professor
Department of CSE
Daffodil International University

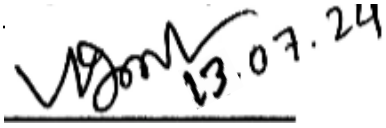
Mr. Rahmatul Kabir Rasel Sarker (RKR) — **Internal Examiner 2**
Lecturer
Department of CSE
Daffodil International University

Sadat Hasan — **External Examiner**
Data Scientist (Senior Principal Officer)
Risk Management Division
BRAC Bank

DECLARATION

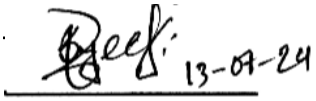
I hereby declare that this project has been done by me under the supervision of **Ms. Nasima Islam Bithi, Lecturer, Department of Computer Science and Engineering, Daffodil International University**. I also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:



Ms. Nasima Islam Bithi
Lecturer
Department of CSE
Daffodil International University

Submitted by:



Syeed Hasan Salim
ID: 192-15-13158
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty for His divine blessing making it possible for us to complete the final year project/internship successfully.

I am grateful and wish to express my profound indebtedness to **Ms. Nasima Islam Bithi**, **Lecturer**, Department of CSE Daffodil International University, Dhaka. Deep knowledge & keen interest of my supervisor in the field of "*Data Mining, Machine Learning (ML), Deep learning, Natural language processing (NLP)*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

I would like to express my heartiest gratitude to the **Dr. Sheak Rashed Haider Noori**, for his kind help in finishing our project and also to other faculty members and the staff of the Department of CSE, Daffodil International University.

I would like to thank my entire course mates in Daffodil International University, who took part in this discussion while completing the course work.

Finally, I must acknowledge with due respect the constant support and patience of my parents.

ABSTRACT

Suicidal ideation not only drives a person towards suicide but also helps to commit suicide. Statistics reveal a gradual increase in suicide-related deaths worldwide. This poses a significant threat to both our current and future societies. That is why we need to take necessary measures to prevent it. Nowadays, online media serves us common platform for people to share their feelings and other personal information. In this research, we have gathered data from social media, mainly Reddit, to analyze the suicidal thoughts in the posts. We have investigated some machine and deep learning models for the detection of suicidal ideation in textual data. We have proposed a "DistilBERT-BiLSTM" model by which we can detect textual data for suicidal ideation. The models we looked at and compared were Logistic Regression, Random Forest, Naïve Bayes, Gated Recurrent Unit (GRU), Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (Bi-LSTM), Convolutional Neural Network (CNN), CNN-GRU with Attention, Bi-LSTM with Attention, and CNN-GRU with Attention. The dataset we used was made up of posts from people who were having suicidal and non-suicidal thoughts. With an AUC score of 1, our proposed "DistilBERT-BiLSTM" model achieved the highest accuracy of 97%.

TABLE OF CONTENTS

CONTENTS	PAGE NO
Board of examiners	i
Declaration	ii
Acknowledgements	iii
Abstract	iv
Table of contents	v-vii
List of Figures	viii-ix
List of Tables	x
CHAPTER	
CHAPTER 1: INTRODUCTION	**1-8**
1.1 Introduction	1-2
1.2 Motivation	2
1.3 Rationale of the Study	3-4
1.4 Research Questions	5
1.5 Expected Output	6-7
1.6 Project Management and Finance	7
1.7 Report layout	8

CHAPTER 2: BACKGROUND **9-16**

2.1 Preliminaries 9-10

2.2 Related Works 10-13

2.3 Comparative Analysis and Summary 13-15

2.4 Scope of the Problem 15

2.5 Challenges 16

CHAPTER 3: RESEARCH METHODOLOGY **17-33**

3.1 Research Subject and Instrumentation 17-18

3.2 Data Collection 18-20

3.3 Working principle of proposed model 20-30

3.4 Statistical Analysis 31-33

CHAPTER 4: EXPERIMENTAL RESULT **34-51**

4.1 Experimental Setup 34

4.2 Experimental Results & Analysis 35-50

4.3 Discussion 51

CHAPTER 5: IMPACT ON SOCIETTY **52-56**

5.1 Impact on Society 52-53

5.2 Impact on Environment 53-54

5.3 Ethical Aspects 54-55

5.4 Sustainability Plan 56

CHAPTER 6: CONCLUSION, SUMMARY, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH **57-59**

6.1 Summary of the Study 57

6.2 Conclusions 58

6.3 Implication for Further Study 59

REFERENCES **57-58**

APPENDIX **59-65**

LIST OF FIGURES

FIGURES	PAGE NO
Figure 3.2.1: Pie chart of 'Suicide_Detection' Dataset	20
Figure 3.2.2: Pie chart of subset of 'Suicide_Detection' Dataset	20
Figure 3.3.1: Working procedure of proposed and Trained Models.	21
Figure 3.3.2: The Proposed "DistilBERT-BiLSTM" Model Architecture.	23
Figure 3.3.3: Logistic Regression Model Architecture	25
Figure 3.3.4: Random Forest Model Architecture	25
Figure 3.3.5: Naïve bayes Model Architecture	26
Figure 3.3.6: GRU Model Architecture	26
Figure 3.3.7: LSTM Model Architecture	27
Figure 3.3.8: BiLSTM Model Architecture	27
Figure 3.3.9: CNN Model Architecture	28
Figure 3.3.10: GRU-CNN with Attention Model Architecture**.**	28
Figure 3.3.11: CNN with Attention Model Architecture**.**	29
Figure 3.3.12: CNN-GRU with Attention Model Architecture.	29
Figure 3.3.13: BiLSTM with Attention Model Architecture.	30
Figure 3.4.14: A simple Confusion Matrix figure	34
Figure 4.2.1: Confusion Matrix and ROC curve of Logistic regression	36
Figure 4.2.2: Confusion Matrix and ROC curve of Random Forest	37
Figure 4.2.3: Confusion Matrix and ROC curve of Naive Bayes	38
Figure 4.2.4: Confusion Matrix and Model accuracy of GRU.	39
Figure 4.2.5: Confusion Matrix and Model accuracy of LSTM.	40
Figure 4.2.6: Confusion Matrix and Model accuracy of BiLSTM.	41
Figure 4.2.7: Confusion Matrix and Model accuracy of CNN.	42
Figure 4.2.8: Confusion Matrix and Model accuracy of GRU-CNN with Attention.	43
Figure 4.2.9: Confusion Matrix and Model accuracy of CNN with Attention.	44
Figure 4.2.10: Confusion Matrix and Model accuracy CNN-GRU with Attention.	45

Figure 4.2.11: Confusion Matrix and Model accuracy of BiLSTM with Attention. 46
Figure 4.2.12: Confusion Matrix of Proposed model (DistilBERT-BiLSTM) 47
Figure 4.2.13: Training and validation accuracy of Proposed model (DistilBERT-BiLSTM) 47
Figure 4.2.14: ROC Curve of Proposed model (DistilBERT-BiLSTM) 48
Figure 4.3.1: Comparison of Training and accuracy of Different Model 51

LIST OF TABLES

TABLES	PAGE NO
Table 2.3.1: Summary Table of related work	14
Table 3.2.1: A sample of input dataset.	19
Table 4.2.1: Model Performance Summary of Training and Proposed Models.	49

CHAPTER 1
Introduction

1.1 Introduction

Text recognition using artificial intelligence (AI) and computational linguistics, specifically natural language processing (NLP), has become a crucial area of study. NLP leverages machine (ML) and deep learning (DL) to enable computers to understand, interpret, and generate human language. In today's digital culture, suicide remains a significant global issue, claiming approximately 700,000 lives annually, according to the data form World Health Organization [1]. The high incidence of suicidal ideation and intentions presents substantial challenges for mental health professionals, underscoring the urgent need for effective identification and prevention strategies.

The digital age has created new opportunities to address this issue by analyzing textual information from social media platforms, by which users often express their emotional struggles and thoughts. Comprehension and thorough analysis are now more critical than ever. Internet communications offer a unique chance to leverage technological advancements to tackle this problem.

Recent progress in NLP, ML and DL has provided several methods to efficiently extract and analyze large volumes of textual data. These tools help identify signs of mental distress that could potentially lead to suicide. NLP can detect subtle linguistic cues indicating suicidal ideation or tendencies by analyzing social media posts and personal messages. ML and DL models have the ability to continuously learn and adapt based on data, setting them apart from traditional methods. This capability allows them to effectively identify patterns and anomalies that might be challenging for human observers to detect.

This thesis explores the usage of natural language processing, machine and deep learning methods to identify signs of suicidal ideation in written text. Early recognition of these indications is critical for immediate response and assistance, which may be life-saving.

Our goal is to develop advanced algorithms that can accurately detect suicidal ideation in written language. This will allow us to develop technologies that can provide prompt interventions and vital support. Moreover, these models possess the potential to assess the effectiveness of existing mental health treatments by tracking changes in language use over a period of time, ultimately resulting in enhanced mental health outcomes.

1.2 Motivation

The exponential growth of digital communication presents new opportunities and challenges for combating suicide. As more individuals express their emotions online, including subtle indications of suicidal ideation, it becomes crucial to utilize this information for timely detection and intervention. Suicide prevention is an urgent global public health concern, with each instance representing a profound personal and community tragedy.

Natural Language Processing (NLP), Machine (ML), and Deep Learning (DL) are revolutionizing our ability to analyze textual material with a sensitivity and scale that surpass traditional methods. These technologies can detect subtle signs of mental distress, often the early indicators of deeper issues, hidden within everyday online interactions.

The concept of AI as a compassionate and efficient partner in the critical human endeavours of suicide prevention drives this thesis. By leveraging technological advancements, this initiative aims to transform suicide prevention strategies by providing timely and discreet interventions that have a significant societal impact. AI has the potential to greatly contribute to societal well-being by automating the detection of danger signs in large datasets from social media, forums, and personal conversations. The primary goal is to offer timely assistance and interventions that could potentially save lives and foster a compassionate and proactive approach to mental healthcare.

1.3 Rational of the Study

The motivation for my research on " An NLP-Based Intelligent Model Approach for Suicidal Tendency Classification " arises from a profound personal desire to have a significant influence on the well-being of persons grappling with suicidal ideation. Through the use of cutting-edge artificial intelligence technology, I am certain that we can develop tools that effectively detect these indicators at an early stage, therefore possibly preventing loss of life and offering vital assistance.

- **Understanding Linguistic Patterns:**

  Through analysing the textual expression of emotions and ideas, my aim is to detect language patterns that signify suicide ideation. Gaining a comprehensive understanding of these patterns is crucial for developing more effective preventative tactics.

- **Improving Early Identification:**

  Timely identification of suicidal inclinations may be crucial for preserving life. By identifying these indicators at an early stage, mental health experts may promptly intervene and provide the essential assistance prior to it being too late

- **Advancing AI Technologies:**

  This research aims to investigate the capabilities of artificial intelligence. Enhancing the AI's ability to recognize the initial signs of suicide ideation in written language could lead to the development of more accurate and efficient algorithms for practical applications.

- **Developing and Validating Models:**

  I am dedicated to the development and evaluation of diverse Machine and deep learning models, for instance Logistic Regression, Random Forest, Naïve Bayes, GRU, LSTM, Bi-LSTM, CNN, GRU-CNN with Attention, Bi-LSTM with Attention, CNN-GRU with Attention to ascertain their efficacy in identifying suicidal ideation. Through a comparative analysis of different models, we can ascertain the most optimal technique.

- **Guiding Future Research:**

  My objective is to provide helpful insights and guidance for future study in this particular sector. The approaches devised might be used not just for the study of

suicidal text but also for other domains within the realm of mental health and behavioural research.

- **Supporting Mental Health Providers:**

  My objective is to create tools that aid mental health practitioners in identifying individuals who are at a higher risk of suicide with more accuracy. These technologies have the potential to function as a pre-emptive alert system, facilitating prompt action and assistance.

- **Ethical Considerations:**

  I possess a keen understanding of the ethical ramifications associated with using AI in this delicate domain. It is crucial to acknowledge and rectify any possible biases in the models and to ensure their responsible and ethical usage, always placing the well-being and privacy of persons as the first priority.

- **Societal Impact:**

  Ultimately, I anticipate that this study will have a significant and far-reaching influence on society. We can provide timely interventions and assistance by using automated algorithms for identifying suicidal inclinations in written syntax, which has a significant impact on individuals' well-being.

The timely identification of suicidal ideation is critical. It is not uncommon for individuals encountering challenges to express their emotions through written means, such as discussions on social media platforms, private conversations, or online forums. By early detection of these indicators, it is possible to provide timely assistance and potentially prevent a catastrophic event. This study focuses on the use of technology to improve prompt responsiveness and active listening for those in need.

My research is fundamentally motivated by a profound commitment to augmenting mental health outcomes through the implementation of artificial intelligence capabilities. By developing text analysis tools that can detect indicators of suicidal ideation, I aim to deliver prompt aid and make a positive contribution to the welfare of individuals requiring assistance.

1.4 Research Questions

This paper addresses several pivotal questions on the topic of " An NLP-Based Intelligent Model Approach for Suicidal Tendency Classification." These questions guide the study, helping to achieve a comprehensive understanding of the topic. The research questions include:

- **What specific linguistic features and patterns are most indicative of suicidal ideation in text data?**

  By identifying these features, we aim to understand how language can reveal deeper emotional distress.
- **How effective are different machine learning models, such as LSTM, Bi-LSTM, CNN, and attention-based models, in detecting suicidal language in various text datasets?**

  This question explores the comparative performance of various advanced AI models in identifying suicidal tendencies.
- **Can the integration of NLP, ML, and DL improve the timeliness and accuracy of suicidal ideation detection compared to traditional screening methods?**

  Here, we assess whether modern AI techniques can outperform conventional methods in detecting suicidal thoughts promptly and accurately.
- **How can the models developed for suicidal text analysis be adapted or extended to other mental health or behavioural text analysis tasks?**

  This question examines the versatility and applicability of our models beyond just suicidal ideation, potentially benefiting broader mental health analysis.
- **What role does the context of the text (such as platform origin, demographic information of the writer, etc.) play in the effectiveness of suicidal ideation detection models?**

  Understanding the context can significantly enhance the accuracy of our models, making them more sensitive to the nuances of different platforms.
- **What are the ethical considerations and potential biases involved in using AI for suicidal text analysis, and how can these be mitigated?**

  By addressing this, we can ensure the responsible use of our models, focusing on minimizing biases and protecting individuals' privacy and well-being.

1.5 Expected Output

The fundamental purpose of this thesis is to create a robust deep learning system capable of properly detecting suicidal intentions conveyed in text from large datasets. The goal of this study is to provide a variety of useful results that will significantly improve the use of artificial intelligence in suicide prevention.

- **Development of Advanced AI Models:**

  We will create and evaluate a range of sophisticated deep learning models, such as Logistic Regression, Random Forest, Naïve Bayes, GRU, LSTM, Bi-LSTM, CNN, GRU-CNN with Attention, Bi-LSTM with Attention, and CNN-GRU with Attention. Every model will go through rigorous testing to demonstrate its ability to detect and interpret suicide content.

- **Evaluating Performance and Comparative Analysis:**

  We are offering a comprehensive analysis of the variations in performance among these models. Our objective is to determine the most efficient models for different forms of textual data by doing a thorough analysis of metrics including accuracy, precision, recall, F1-score, and computing efficiency.

- **Guidelines for Model Selection and Application:**

  Our study will emphasize the crucial role that Natural Language Processing (NLP) and Machine Learning (ML) technologies play in improving suicide prevention efforts. We will demonstrate how advanced text analysis may aid in early identification and intervention, leading to substantial improvements in mental health results.

- **Highlights the role of NLP and ML in mental health.**

  We will provide practical suggestions for using these AI models on digital health platforms. This includes tactics for incorporating these technologies into healthcare systems, as well as recommendations for continuously improving and adjusting the model based on new data patterns.

- **Practical recommendations for deployment:**

  We will provide practical suggestions for using these AI models on digital health platforms. This includes tactics for incorporating these technologies into healthcare systems, as well as recommendations for continuously improving and adjusting the

model based on new data patterns.

- **Contributions to Academic Knowledge:**

  Our research findings will provide significant information to the academic community. Our objective is to provide support for future research and practical use of mental health technology by documenting techniques, model assessments, and recommendation.

1.6 Project Management and Finance

Effective project management and thorough budget preparation are critical for the success of our study, " An NLP-Based Intelligent Model Approach for Suicidal Tendency Classification." We will strictly adhere to a structured approach, divided into several phases, to ensure efficient supervision. The initiation phase encompasses the process of clearly defining the project's scope, objectives, and deliverables. It also comprises identifying the main stakeholders and forming the project team. During the planning phase, we will develop a comprehensive project plan that includes precise deadlines, milestones, tasks, and responsibilities. Additionally, we will identify the required resources and any hazards. The execution phase will primarily concentrate on gathering data, creating and training models, and carrying out experiments. Ongoing monitoring and control will facilitate this by ensuring progress and applying necessary corrective actions. Ultimately, the concluding stage will include recording discoveries, presenting results, and carrying out a comprehensive project assessment to collect comments and contemplate acquired knowledge. Our goal is to properly oversee the project and successfully accomplish our study objectives using this methodical methodology.

1.7 Report layout

The systematic organization of the thesis ensures a clear and logical exposition of the study, results, and wider ramifications. The structure of each chapter progressively builds upon the previous one, ensuring a coherent sequence and a thorough comprehension of the subject matter. The arrangement is designed in a manner that directs the reader through a methodical examination of suicidal text analysis using NLP, ML and DL.

Chapter 1: Introduction

- This chapter provides an introduction to the research issue, highlighting its importance and outlining the study's aims.

Chapter 2: Background

- This chapter provides a thorough analysis of the pertinent literature, the theoretical foundations, and the context of the study.

Chapter 3: Research Methodology

- This chapter provides a comprehensive explanation of the research strategy, collecting information procedures, and analytical methodologies used in the study.

Chapter 4: Experimental Results

- This chapter provides a comprehensive overview of the experiments undertaken, including the analysis and interpretation of the collected data.

Chapter 5: Impact on Society

- This chapter examines the social consequences of the study results and their possible influence on mental health therapies.

Chapter 6: Summary, Conclusion, Recommendations, and Implication for Future Research

- In here presents a concise overview on main discoveries, draws conclusions, proposes practical advice, and identifies potential avenues for further study.

.

CHAPTER 2
Background

2.1 Preliminaries

Natural Language Processing (NLP) is a field that enables computers to understand and generate human language, which is essential for analyzing textual data. Machine Learning (ML) involves using algorithms to detect patterns and make data-driven decisions. Deep Learning (DL), a subset of machine learning, employs complex neural networks to extract information from large volumes of unstructured data. DL techniques include Recurrent Neural Networks and Long Short-Term Memory networks, which are particularly competent for evaluating sequential data. LSTM models are adept at capturing intricate and long-term dependencies within sequences. While Convolutional Neural Networks (CNN) are primarily used for image processing, their application to text data is growing. Attention mechanisms further enhance model performance by focusing on the most relevant features of the data, thereby improving tasks like language translation and sentiment analysis.

DistilBERT is a faster, smaller, and more efficient category of BERT (Bidirectional Encoder Representations from Transformers), which retains most of BERT's language-understanding capabilities while being more computationally efficient. DistilBERT accomplishes this through a process known as knowledge distillation, which entails training a smaller model, the student, to emulate the behavior of a larger model, the teacher, thereby capturing its essential characteristics and minimizing the model size and inference time. In 2019, Hugging Face presented DistilBERT as a more compact and efficient substitute for Google's powerful BERT natural language processing model. DistilBERT uses a distillation technique to train a smaller model to replicate BERT’s behavior, resulting in a model with fewer parameters while maintaining high accuracy and performance [2-4]. Our proposed “DistilBERT-BiLSTM” is a text analysis model that merges the efficiency of DistilBERT with the sequence processing capability of BiLSTM, yielding a harmonious and efficient method.

This study employs a diverse array of advanced models, including Logistic Regression, Random Forest, Naïve Bayes, GRU, LSTM, Bi-LSTM, CNN, GRU-CNN with Attention, Bi-LSTM with Attention, and CNN-GRU with Attention. Each model offers unique advantages for detecting suicidal tendencies in textual data. Understanding these terms and concepts is essential for the approaches used in this study. Our aim in utilizing these advanced methodologies is to gain deeper insights and make significant advancements in suicide prevention, ultimately providing immediate aid to individuals in need.

2.2 Related Works

In recent years, Significant advancements have been achieved in the area of research that specifically examines the identification of suicidal tendencies via the use of natural language processing, machine and deep learning. These technologies have achieved a high degree of precision and effectiveness in detecting suicidal thoughts and actions. A few highlights of seminal research in this area are discussed in the following sections:

Jain et al. explored the application of ML and NLP to analyse and predict depression and suicidal ideation on Reddit. By examining posts from r/depression and r/SuicideWatch, the study implemented algorithms, for example- Logistic Regression, Naïve Bayes, Support Vector Machine (SVM), and Random Forest. Logistic Regression and Random Forest achieved the highest accuracies of 77.29% and 77.30%, respectively. The research underscores Reddit's utility in providing mental health data, highlighting its potential to complement traditional public health systems for monitoring and addressing mental health issues.[5]

In this study, Sawhney et al. investigated the detection of suicidal ideation in social media text using deep learning techniques. The researchers focused on tweets, employing models such as Vanilla-RNN, LSTM, and C-LSTM. A dataset was created by scraping and manually annotating suicide-related web forums. The C-LSTM model, combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, outperformed traditional text classification methods like Logistic Regression and SVM,

achieving the highest accuracy in identifying suicidal content.[6]

Imam et al. applied NLP techniques to detect suicidal thoughts on Reddit. The study analysed text from subreddits such as r/SuicideWatch and r/depression to identify suicidal ideation. Utilizing LSTM and Random Forest classifiers, their approach achieved up to 93% accuracy. This research provides a valuable dataset and a methodological framework for future studies in mental health analysis and suicide prevention.[7]

Yeskuatov et al. identified that advanced deep learning models, particularly those using transformer-based embeddings like BERT and RoBERTa, achieved the best results in detecting suicidal ideation on Reddit. BERT-based models reached an accuracy of up to 95.21%, demonstrating the superior performance of sophisticated embeddings and deep learning techniques in identifying signs of suicidal ideation in social media content, surpassing traditional ML methods.[8]

Aldhyani et al. explored methods for detecting suicidal ideation on social media using a combination of ML and DL techniques. Using a large dataset from Reddit's SuicideWatch, the study compared a hybrid CNN–BiLSTM model with the XGBoost algorithm. The CNN–BiLSTM model achieved superior accuracy (95%) in identifying suicidal posts compared to XGBoost (91.5%). This study highlights the effectiveness of advanced DL models in enhancing the accuracy of suicidal ideation detection.[9]

Renjith et al. proposed an ensemble deep learning model combining LSTM and CNN with an attention mechanism to detect suicidal ideation in social media posts. The LSTM-Attention-CNN model achieved an accuracy of 90.3% and an F1-score of 92.6%, outperforming traditional models like SVM and standalone LSTM or CNN classifiers. The attention mechanism improved the model's performance by focusing on crucial parts of the text, enhancing the detection of suicidal thoughts.[10]

Tadesse et al. used an ensemble of deep learning techniques to detect suicidal ideation in social media posts from Reddit. They developed a hybrid model combining LSTM and CNN, utilizing word embeddings to capture contextual features in text. The LSTM-CNN hybrid model achieved the highest accuracy (93.8%) and F1-score (92.8%), demonstrating its effectiveness in identifying suicidal tendencies compared to traditional ML classifiers.[11]

Ji et al. examined supervised learning techniques for detecting suicidal ideation in online user content from Reddit's SuicideWatch and Twitter. The study used features such as statistical, syntactic, linguistic, word embeddings, and topic features. XGBoost achieved the highest accuracy at 95.71% on Reddit data, while Random Forest performed best on Twitter. The study emphasizes the effectiveness of feature-rich models in early detection of suicidal tendencies in social media.[12]

Shetty's study focused on suicide risk assessment on Reddit using deep learning techniques, categorizing suicide risk into ideation, indicator, behaviour, and attempt stages. Using a semi-supervised learning approach and a dataset from the SuicideWatch subreddit, the study applied Text CNN and BiLSTM models, with BiLSTM achieving the best performance at 92% accuracy and F1-score. The findings highlight the efficacy of combining data augmentation with semi-supervised learning for enhanced suicide risk detection.[13]

Inamdar et al. explored mental stress detection in Reddit posts using ML and NLP. Using the reddit dataset, they classified posts as stressful or non-stressful, employing word embeddings like ELMo, BERT, and TF-IDF with ML models such as Logistic Regression, SVM, and XGBoost. LR with ELMo embeddings achieved the best performance, with an F1-score of 0.76. This research demonstrates the potential of advanced NLP tools and ML algorithms for analysing mental stress in social media content.[14]

Haque et al. conducted a comparative analysis of ML and DL models to detect suicidal ideation from Twitter posts. Evaluating models such as Logistic Regression, Random Forest, SVM, and LSTM, they found that LSTM performed best, achieving an F1-score of 0.94. This study underscores the potential of ML and DL techniques in enhancing suicide prevention through early detection of suicidal ideation on social media.[15]

Yao et al. applied ML techniques to detect suicidality among opioid users on Reddit. Analysing posts from r/suicidewatch and opioid-related forums, they used classifiers like CNN, Logistic Regression, Random Forest, and SVM. CNN achieved the highest F1 score of 96.6%. This study highlights the effectiveness of deep learning in identifying suicidal posts within the context of opioid use, providing insights for better understanding and intervention.[16]

2.3 Comparative Analysis and Summary

All of these publications use social media and textual data to train machine and deep learning models to identify suicidal thoughts. Deep learning models like BiLSTM and CNN-BiLSTM outperform Random Forest and Logistic Regression, according to studies. This research demonstrates the potential of NLP and AI to improve suicide intention detection, advancing mental health prevention. The study shows that these tools quickly process and analyze textual data to detect suicidal actions. The research highlights the effectiveness of these technologies in processing and analyzing textual data to identify suicidal behaviors efficiently.

Table 2.3.1: Summary Table of related work

Study or Author Name	Models Used	Data Source	Accuracy/AUC	Contribution and limitation
Jain et al., 2018	Random Forest	Reddit (r/depression, r/SuicideWatch)	Accuracy: 77.30%	Highlighted Reddit's utility in mental health monitoring. limited accuracy and no deep learning

				model used
Sawhney et al., 2019	C-LSTM	Tweets	Accuracy:81.4%	Demonstrated efficacy of deep learning limited to tweet data
Imam et al., 2020	Random Forest	Reddit (r/depression, r/SuicideWatch)	Accuracy: up to 93%	Provided valuable dataset and framework no deep learning model used.
Yeskuatov et al., 2021	BERT	Reddit	Accuracy: up to 95.21%	Superior performance of transformer-based models complexity in implementation.
Aldhyani et al., 2022	CNN–BiLSTM	Reddit (SuicideWatch)	Accuracy: 95%	Effectiveness of DL models. Using less complex model.
Renjith et al., 2022	LSTM-Attention-CNN	Social Media Posts	Accuracy: 90.3%, F1-score: 92.6%	Improved performance with attention mechanisms Less accuracy.
Tadesse et al., 2022	LSTM-CNN	Reddit	Accuracy: 93.8%, F1-score: 92.8%	Proved effectiveness of hybrid models Limited Data

Ji et al., 2022	XGBoost	Reddit (SuicideWatch), Twitter	Accuracy: 95.71% (Reddit, XGBoost)	Emphasized feature-rich models. varied performance across platforms
Shetty, 2022	BiLSTM	Reddit (SuicideWatch)	Accuracy: 92%, F1-score: 92%	Efficacy of semi-supervised learning Few models use for training.
Inamdar et al., 2022	LR with ELMo	Reddit	F1-score: 0.76	Potential of advanced NLP tools Lower F1-score
Haque et al., 2022	LSTM	Twitter	F1-score: 0.94	Enhance suicide prevention. No accuracy score is given

2.4 Scope of the Problem

Given that suicide remains a leading cause of death worldwide, the task of identifying suicidal tendencies is both urgent and profoundly important. Recent advances in natural language processing, machine and deep learning have shown incredible promise in detecting suicidal thoughts and behaviors from text on social media and other platforms. However, the nuances of human language and the subtle signs of suicidal intent present ongoing challenges. Research has demonstrated the superior performance of advanced models like convolutional neural networks (CNNs) and bi-directional long-short-term memory (BiLSTM) networks, which have achieved high accuracy in identifying these critical cues. Despite these advancements, there is a continual need to refine these models to enhance their accuracy, efficiency, and relevance across different populations and languages. This makes it an essential and evolving field of study, with the potential to significantly impact mental health support and suicide prevention efforts.

2.5 Challenges

Despite significant advancements in using NLP, ML, and DL for detecting suicidal tendencies in textual data, several challenges remain:

1.Linguistic Nuances and Context:

Grasping the subtleties and context of human language is challenging. Suicidal thoughts can be expressed indirectly, and the same words can have different meanings depending on the situation and individual background. For AI models to understand these nuances, they require large amounts of high-quality annotated data. Additionally, sarcasm, idiomatic expressions, and cultural differences add to the complexity.

2. Ethical and Privacy Concerns:

Analyzing personal mental health data with AI raises significant ethical and privacy issues. It is essential to ensure that data is anonymized and individuals' privacy is protected. There's a risk of misuse, such as profiling people without their consent, which could lead to stigmatization or discrimination. To use these tools responsibly, clear data use guidelines and transparency in AI decision-making are necessary.

.

These challenges underscore the complexities involved in developing and deploying AI systems for mental health analysis, highlighting the need for ongoing research and careful consideration of ethical implication.

CHAPTER 3
Research Methodology

3.1 Research Subject and Instrumentation

Research Subject:

In this section, we are going to talk about the topic we are working on. Our thesis topic is “An NLP-Based Intelligent Model Approach for Suicidal Tendency Classification," where we are using a bunch of NLP, machine and deep learning approaches to detect the suicidal text from a huge amount of data. For this, we analyze the data and do the preprocessing before applying the data to various model algorithms to evaluate the best model to detect suicidal text.

Instrumentation:

In this study, instrumentation refers to the combination of software tools, programming libraries, and hardware used to conduct the research. For this project, we are using programming languages and frameworks like TensorFlow and Python to implement the deep learning models and associated algorithms.

Hardware:

The computational experiments were primarily performed on personal computers, which have an AMD Ryzen 7 7700X processor and 16GB of RAM, providing enough processing power for the data-intensive operations. We are running the code and the model in Google Colab and using its GPU acceleration for reducing model training times.

Software and Libraries:

1. **Python:**

   We're using Python 3.10.12, which offers a robust and versatile platform for developing our machine learning and deep learning models. Python's extensive libraries and supportive community make it an ideal choice for handling complex tasks like these.

2. **TensorFlow and Keras:**

   These libraries are crucial for building and training our deep learning models. TensorFlow and Keras simplify the construction of sophisticated neural networks and handle the intensive computational tasks involved in training them on large datasets.

3. **Natural Language Toolkit (NLTK):**

   NLTK is our primary tool for text preprocessing. It assists with essential tasks like tokenization and removing stop-words, which are crucial steps in preparing our data for analysis.

4. **Pandas and NumPy:**

   These libraries are essential for handling and cleaning data, as well as transforming it into a format suitable for our models. Pandas simplifies data manipulation, while NumPy offers efficient numerical operations.

5. **Matplotlib and Seaborn:**

   Understanding our data and model performance requires visualization. Our ability to generate comprehensible and educational plots and charts using Matplotlib and Seaborn facilitates the successful communication of our results..

6. **Scikit-learn:**

   This library provides a variety of tools for implementing traditional machine learning models. Scikit-learn is invaluable for conducting model evaluations and calculating performance metrics, allowing us to accurately measure and compare the effectiveness of our models.

3.2 Data Collection

In the subsequent section, we delve into the methodology used to collect the data for this thesis. The dataset was sourced from Kaggle and comprises posts from the "SuicideWatch" and "depression" subreddits on Reddit. These posts were collected via the Pushshift API.

We gathered all posts published on the "SuicideWatch" subreddit from its inception on December 16, 2008, until January 2, 2021. Similarly, we collected posts from the

"depression" subreddit from January 1, 2009, to January 2, 2021. Posts from the "SuicideWatch" subreddit are categorized as suicide, while those from the "depression" subreddit are categorized as depression. Additionally, posts that did not include suicidal thoughts or intentions were collected from the "r/teenagers" subreddit. This dataset was developed by Nikhileswar Komati.

Datasets:

Upon obtaining and analyzing the 'Suicide_Detection' dataset, we have found two primary columns: 'text' and 'class.' The 'class' column categorizes texts as either suicidal or non-suicidal. The following is a preview of the first 10 entries in the raw dataset:

Table 3.2.1: A sample of input dataset.

Text	Class
Ex Wife Threatening SuicideRecently I left my wife…	suicide
Am I weird I don't get affected by compliments or …	non-suicide
Finally, 2020 is almost over… So, I can never hear…	non-suicide
I need help just help me; I'm crying so hard	suicide
I'm so lost. Hello, my name is Adam (16) and I've been…	suicide
Honestly idk I don't know what I'm even doing here I g…	suicide
[Trigger warning] Excuse for self-inflicted burns …	suicide
It ends tonight. I can't do it anymore. \nI quit.	Suicide
Everyone wants to be "edgy" and it's making me sic…	non-suicide
My life is over at 20 years old Hello all. I am a 2…	suicide

The dataset has 232,074 records, evenly divided between 116,037 classified as suicide and 116,037 classified as non-suicide.

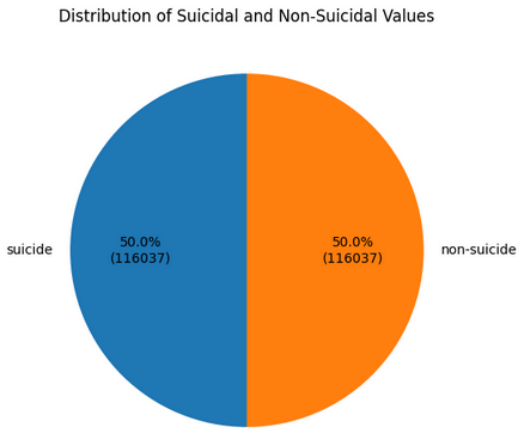


Fig 3.2.1: Pie chart of 'Suicide_Detection' Dataset

From this dataset we are going to use 50 % of the data. Among them 58018 are labeled as suicidal and 58018 are non-suicidal.

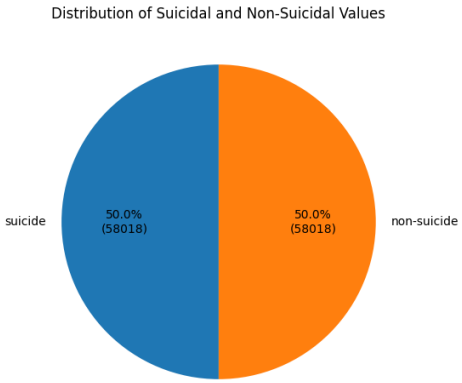


Fig 3.2.3: Pie chart of subset of 'Suicide_Detection' Dataset

3.3 Working principle of proposed model

The experimental method comprises many crucial stages: data acquisition, data preprocessing, model training, classification, and result evaluation. Every step is vital in the process of establishing more accurate and dependable models for recognizing suicidal

texts.

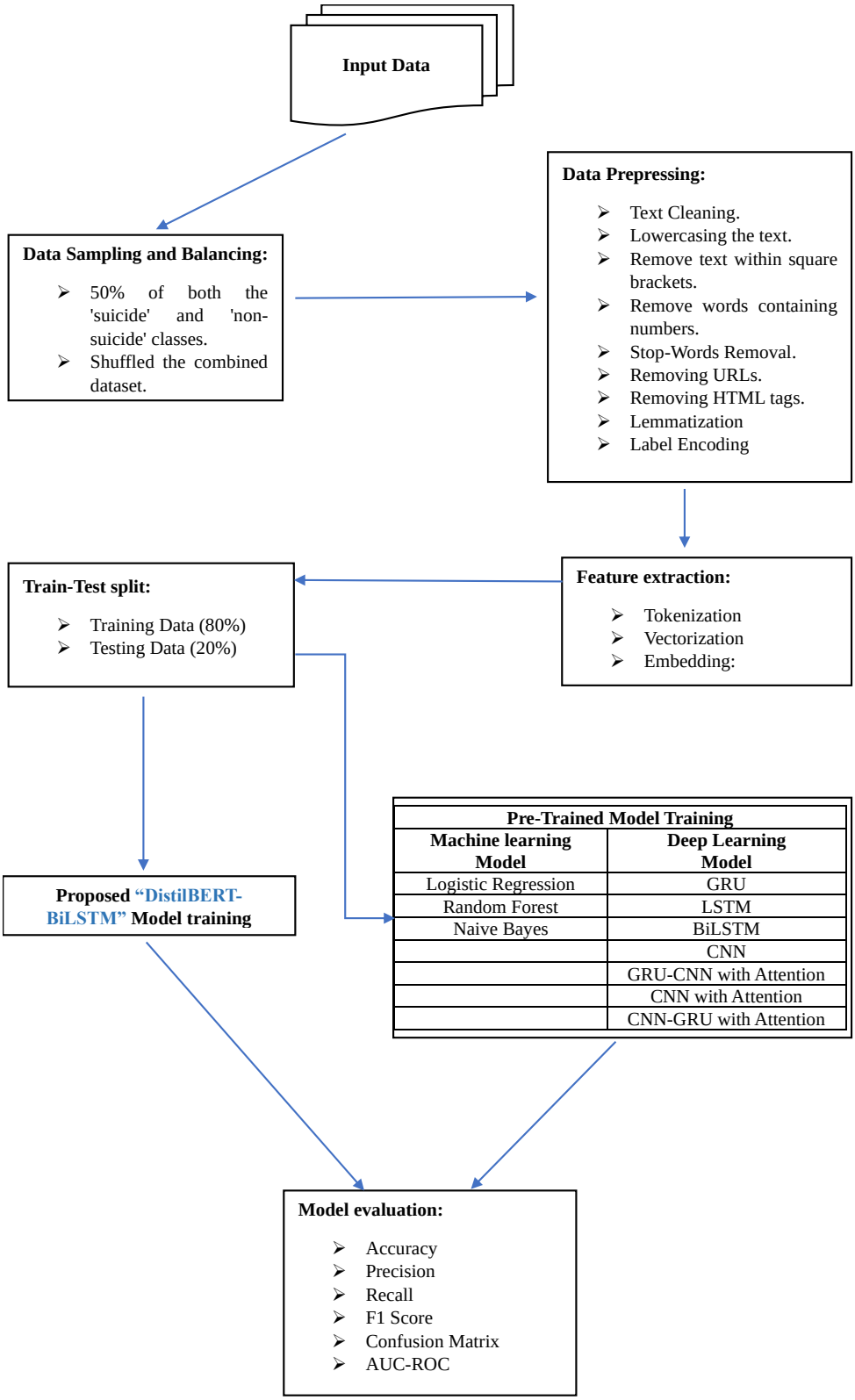


Figure 3.3.1: Working procedure of proposed and Trained Models.

A. Data Acquisition:

This symbolizes the gathering and organizing of the dataset in order to assist with analysis. We used the data from the “Suicide_Detection” Database to evaluate our model. The process involves the following steps-

- Data Import
- Data Inspection

B. Data Sampling and Balancing:

We Sampled 50% of both the 'suicide' and 'non-suicide' classes to create a balanced dataset from our “Suicide_Detection” Database. Shuffled the combined dataset

C. Data Preprocessing:

Data Preprocessing is the process of converting raw data into a format that is appropriate for training a model. The process involves the following steps-

i. Text Cleaning
ii. Lowercasing the text.
iii. Removing text within square brackets.
iv. Removing words containing numbers.
v. Removing URLs.
vi. Removing HTML tags.
vii. Removing punctuation.

D. Data Tokenization:

For training a Deep learning model we are using tokenization and padded sequences. For proposed model “DistilBERT-BiLSTM” tokenization we are using “distilbert-base-

uncased" token library.

E. Data Embedding and Vectorization:

we are using TF-IDF vectorized data for Training Machine leaning model.

F. Data Splitting:

For model testing, we use 20% of the data for testing and 80% of the data for training.

G. Training proposed "DistilBERT-BiLSTM" Model:

For proposed model "DistilBERT-BiLSTM" tokenization we are using "distilbert-base-uncased" token library.

Our proposed Model is a Combination of DistilBERT and BiLSTM Model, which DistilBERT for dense representation with BiLSTM and attention mechanisms to effectively capture context and important features. It Utilizes transformer-based embeddings, captures bidirectional context, attention mechanism enhances feature extraction.

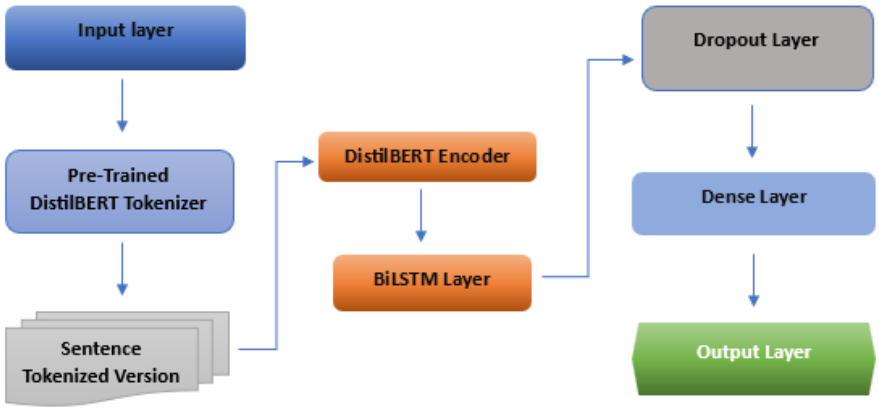


Figure 3.3.2: The Proposed "DistilBERT-BiLSTM" Model Architecture

In this architecture:

- **Input Layer**: Accepts sequences of tokenized text with a fixed length.
- **DistilBERT Layer**: Utilizes the DistilBERT model to provide a dense representation of each word in the sequence.
- **BiLSTM Layer:** Captures both forward and backward dependencies in the text.
- **Dropout Layer:** Added after the BiLSTM layer to prevent overfitting.
- **Attention Layer:** Applied to the BiLSTM layer's output so it focuses on the sequence's most crucial fragments
- **Output Layer:** A dense layer with Softmax activation to produce the final classification output (suicide or non-suicide).

Total number of parameters: 67,646,450 (including trainable and non-trainable parameters).

The model training process employs the categorical cross-entropy loss function and the “Adam optimizer”, with a learning rate of 0.00002. Key training parameters include running the training for 50 “epochs” with a “batch size” of 128. A portion of the training data is reserved for validation to monitor the model's performance during training. The process also incorporates several essential callbacks: “ModelCheckpoint” saves the best model based on validation accuracy, ensuring the optimal version of the model is preserved, and “ReduceLROnPlateau” adjusts the learning rate if the validation loss plateaus, enhancing the model's ability to converge effectively. This combination of strategies ensures robust training and helps prevent overfitting, facilitating the model's ability to generalize well on unseen data.

H. Training of Different Classical Machine and Deep learning Model:

We utilized a variety of machine learning and deep learning models to classify suicidal text. Below are the details and architectural descriptions of each model used -

- **Logistic Regression (LR)**

A linear model used for binary classification is called logistic regression. It evaluates the possibility that an instance falls into a certain class. The binary dependent variable can be described using the logistic function, often known as the sigmoid.

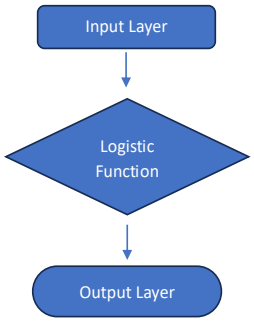


Figure 3.3.3: Logistic Regression Model Architecture.

- **Random Forest (RF)**

With the use of many decision trees built during training, Random Forest is an ensemble learning technique that produces a class that is the mean of the classes of individual trees. Accuracy is increased and over fitting is reduced.

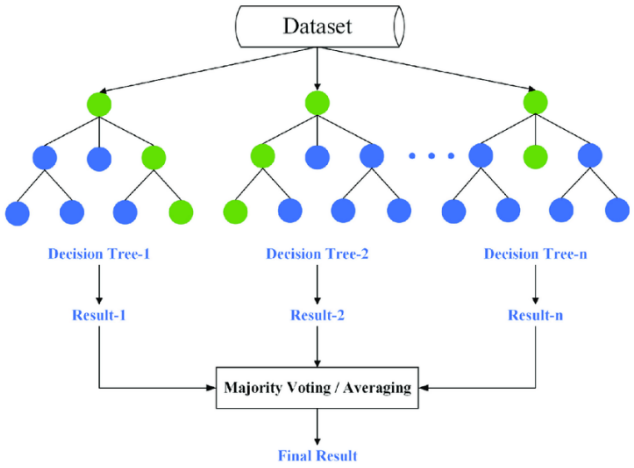


Figure 3.3.4: Random Forest Model Architecture [17].

- **Naive Bayes (NB)**

Naive Bayes classifiers are a family of simple probabilistic classifiers mainly based on applying Bayes' theorem with strong (naive) independence assumptions allying the features.

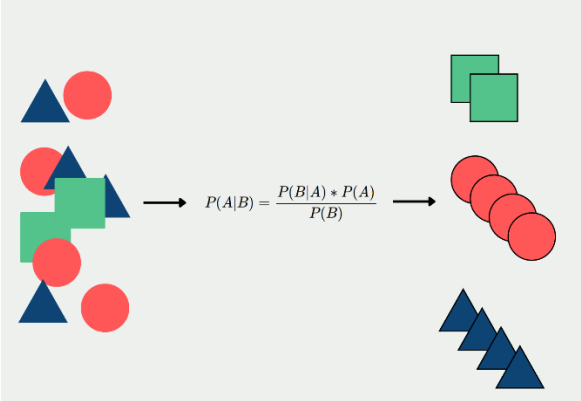


Figure 3.3.5: Naïve bayes Model Architecture.

- **Gated Recurrent Unit (GRU)**

GRU can be described as a type of recurrent neural network that is used to model sequential data. It uses gating units to control the flow of information without having separate memory cells.

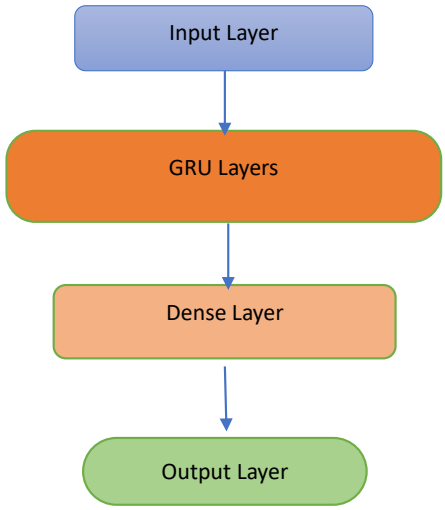


Figure 3.3.6: GRU Model Architecture.

- **Long Short-Term Memory (LSTM)**

LSTM can be defined as a type of RNN capable of learning long-term dependencies. It prevents the vanishing gradient problem with the help of its gate mechanisms (input, forget, and output gates).

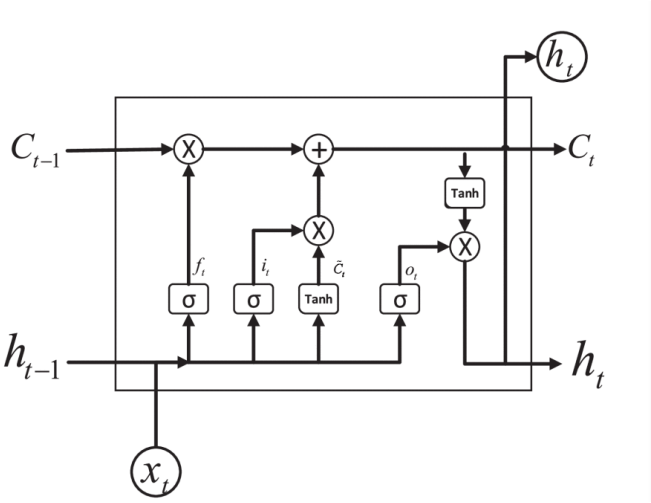


Figure 3.3.7: LSTM Model Architecture [18].

- **Bidirectional Long Short-Term Memory (BiLSTM)**

BiLSTM develops data in both forward and backward directions to capture context from both past and future states. It Captures context from both directions, improves accuracy in sequential data analysis.

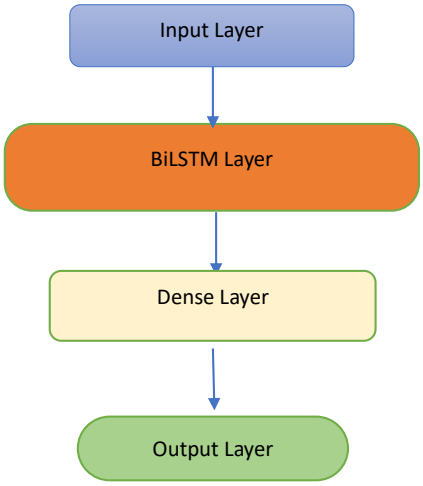


Figure 3.3.8: BiLSTM Model Architecture

- **Convolutional Neural Network (CNN)**

The main use for CNNs is with spatial data. Convolutional layers are used by them to extract features from the incoming data. beneficial in dimensionality reduction and local feature extraction.

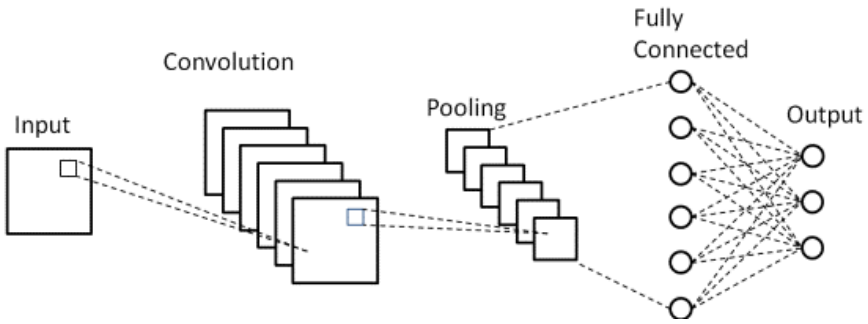


Figure 3.3.9: CNN Model Architecture [19]

- **GRU-CNN with Attention**

Combines GRU and CNN with an attention mechanism to focus on important parts of the data. Leverages both spatial and sequential features, attention mechanism improves focus on relevant parts of the data.

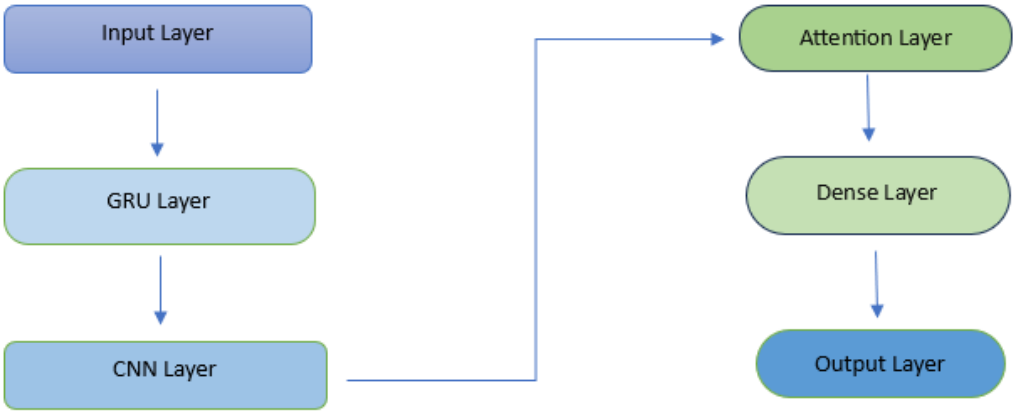


Figure 3.3.10: GRU-CNN with Attention Model Architecture.

- **CNN with Attention**

Integrates attention mechanism into CNN to highlight significant features in the data. Enhances feature extraction by focusing on important data segments.

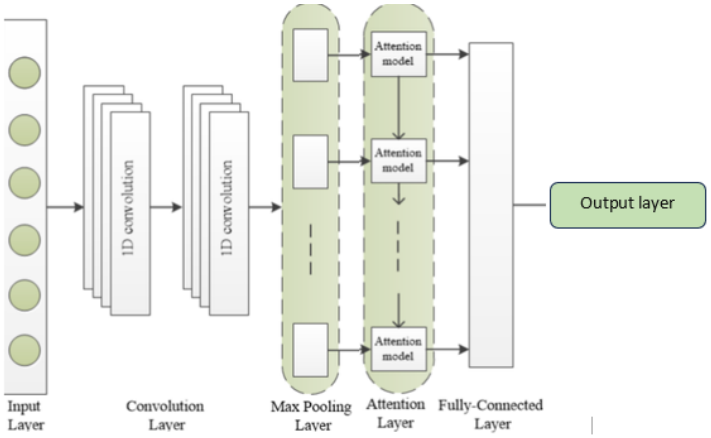


Figure 3.3.11: CNN with Attention Model Architecture.

- **CNN-GRU with Attention**

Combines CNN and GRU with attention to leverage both spatial and sequential features. Combines advantages of CNN and GRU, attention mechanism enhances performance.

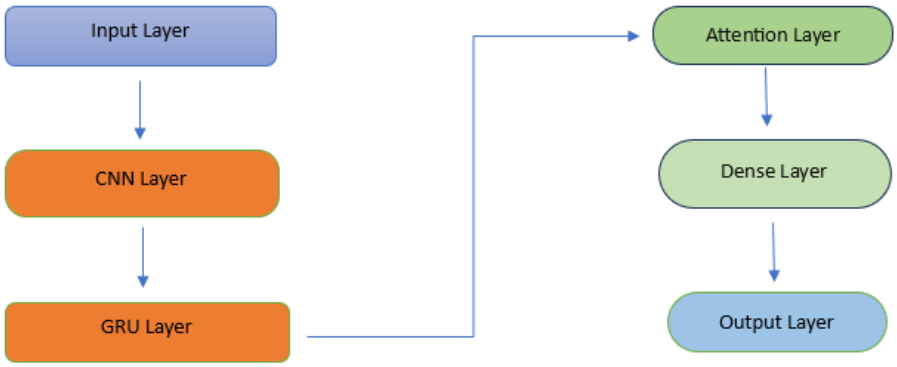


Figure 3.3.12: CNN-GRU with Attention Model Architecture

- **BiLSTM with Attention:**

Integrates attention mechanism with BiLSTM to focus on the most important parts of the sequence and capture context from both directions. Captures bidirectional context, attention mechanism enhances performance by focusing on crucial parts of the input sequence.

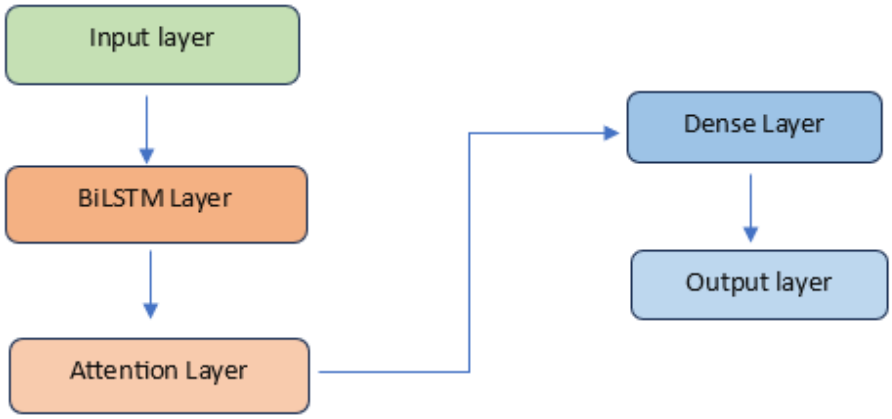


Figure3.3.13: BiLSTM with Attention Model Architecture.

I. Classification:

The trained model is then used to classify the test data into 'suicide' or 'non-suicide' categories. The classification method involves inputting the pre-processed test data into the model and obtaining the predicted probabilities for each class.

3.4 Statistical Analysis

The statistical analysis carried out during this study on "An Intelligent Model of Suicidal Tendency Classification using NLP" is crucial in extracting significant insights from the experimental findings. This section provides an overview of the statistical analysis methods used to assess the effectiveness of the machine learning and deep learning models. Statistical analysis is essential for comprehending the efficacy and dependability of models in categorizing text as either suicidal or non-suicidal.

A. Performance Metrics

In order to evaluate the effectiveness of the models, many established measures are used, such as:

Accuracy:

Accuracy (A) can be referring to the ratio of properly identified cases, including both true positives and true negatives, to the total number of samples. The accuracy (A) is determined by the formula [20] in equation (i) -

$$Accuracy\ (A) = \frac{TP+TN}{TP+FN+TN+FP} \text{ ---------- (i)}$$

Where, TP denotes the number of true positives, TN denotes the number of true negatives, FP represents the number of false positives, and FN represents the number of false negatives.

Precision:

Precision (P) can be defined as measures the proportion of positive identifications that were actually correct. It is particularly useful when the cost of false positives is high. The calculation is as follows [21] equation (ii) -

$$Precision\ (P) = \frac{TP}{TP+FP} \text{ ---------- (ii)}$$

Recall:

Recall (R), also refers as Sensitivity, quantifies the ratio of accurately detected real positives by the model. It is advantageous when the results of false negatives are high. The calculation for Recall is as follows [22] equation (iii) -

$$Recall\ (R) = \frac{TP}{TP+FN} \quad \text{---------- (iii)}$$

F-1 Score:

The F-1 score can be defined as mathematical measure that combines recall and precision into a single metric, using a symmetrical mean. It effectively balances both aspects. It is calculated as [23] in equation (iv) and equation (v) -

$$F1\ score\ = \frac{2.TP}{2.TP+FP+FN} \quad \text{---------- (iv)}$$

$$F1\ score\ = 2.\frac{precision.recall}{precision+recall} \quad \text{---------- (v)}$$

Area Under the Receiver Operating Characteristic Curve (AUC-ROC):

The AUC-ROC assesses a model's ability to distinguish between different classes. An ROC curve with an area under the curve (AUC) of 1 indicates a perfect model, whereas an AUC of 0.5 suggests a model that cannot differentiate at all. The ROC curve shows the balance between sensitivity (true positive rate) and specificity (1 – false positive rate). This curve plots two parameters. Those are shown in equation (vi) and equation (vii)

1.True Positive Rate

$$TPR = \frac{TP}{TP+FN} \quad \text{---------- (vi)}$$

2.False Positive Rate

$$FPR = \frac{FP}{FP+TN} \text{ ---------- (vii)}$$

Confusion Matrix

The confusion matrix provides a detailed evaluation of a model's performance by showing the counts of true positive, true negative, false positive, and false negative predictions. This matrix helps to understand the various mistakes the model makes. It can be shown as [24]:

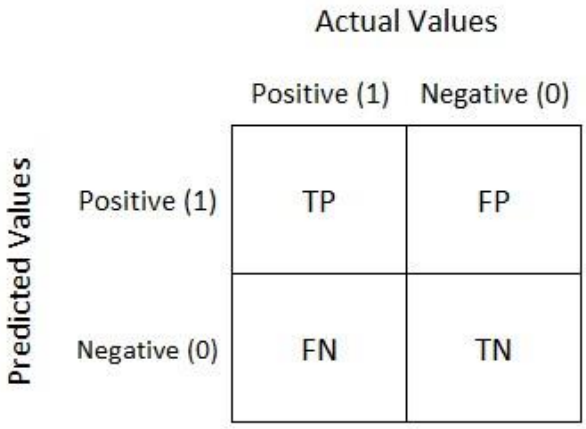


Fig 3.4.14: A simple Confusion Matrix figure

CHAPTER 4

Experimental Result

4.1 Experimental Setup

For my research paper which is titled as " An NLP-Based Intelligent Model Approach for Suicidal Tendency Classification " used a systematic experimental setting that started with the gathering and preparation of data. The dataset was obtained via Kaggle and consisted of posts from the "SuicideWatch" and "depression" subreddits on Reddit. The data was acquired using the Pushshift API between December 16, 2008, and January 2, 2021. Posts originating from the "SuicideWatch" subreddit were categorized as 'suicide', while posts from the "depression" subreddit were categorized as 'depression'. Non-suicidal postings were gathered from the "r/teenagers" subreddit. The text data underwent cleaning using the text_hammer library to eliminate emails, HTML elements, special characters, and accented letters. Subsequently, it was transformed to lowercase to ensure consistency. The process of tokenization and lemmatization was used to decompose the text into significant tokens and simplify words to their fundamental forms. This was then followed by label encoding using LabelEncoder. The dataset underwent a random reordering and was divided into two sets: an 80% training set and a 20% validation set. The model architecture is a combination of DistilBERT and BiLSTM with attention mechanisms. It includes layers for input, contextual embeddings using DistilBERT, bidirectional dependencies using BiLSTM, dropout for preventing overfitting, attention for focusing on important sequence parts, and a dense layer for classification. The architecture consisted of a grand total of 66790146 characteristics, out of which 66790146 were trainable and none were non-trainable. The model underwent training using the Adam optimizer and categorical cross-entropy loss, using a learning rate of 0.00002. The training process spanned 50 epochs, with a batch size of 128. Additionally, a validation split was used to check the model's performance. The objective of this experimental configuration was to enhance the precision and dependability of the model in identifying suicidal inclinations in textual data, offering new perspectives and resources for mental health research and intervention.

4.2 Experimental Results & Analysis

In In this section, I will be examining the results of the multiple machine learning and deep learning models that I deployed throughout the course of my dissertation. Also included in this area will be the results of my analysis. The key focuses of this study are the evaluation of the performance metrics of each model and the selection of the model that is the most successful for the categorization of suicidal text. Both of these objectives are presented in the following sentence. Among the models that were tested, the following are some of the most significant ones: There are several different types of neural networks, including Logistic Regression, Random Forest, Naive Bayes, GRU, LSTM, BiLSTM, CNN, GRU-CNN with Attention, Bi-LSTM with Attention, CNN with Attention, and CNN-Naive Bayes. Here is a full analysis of the performance of each model that has been given. The following aspects are taken into consideration in this analysis: training accuracy, validation accuracy, precision, recall, and F1-score calculations. The categorization performance is the most important factor to consider. Following that, the overall metrics of those models are under the microscope for analysis. Among all of them, the model that performed the best was our proposed Model "DistilBERT-BiLSTM". This model leveraged the strengths of both DistilBERT and BiLSTM with attention mechanisms to achieve the highest performance.

Here is a full analysis of the performance of each model, considering the following aspects: training accuracy, validation accuracy, precision, recall, and F1-score calculations. The categorization performance is the most important factor to consider. Following that, the overall metrics of those models are under the microscope for analysis. A collection of information includes descriptions, probable explanations, and areas that may need improvement.

Performance Analysis of Training Models:

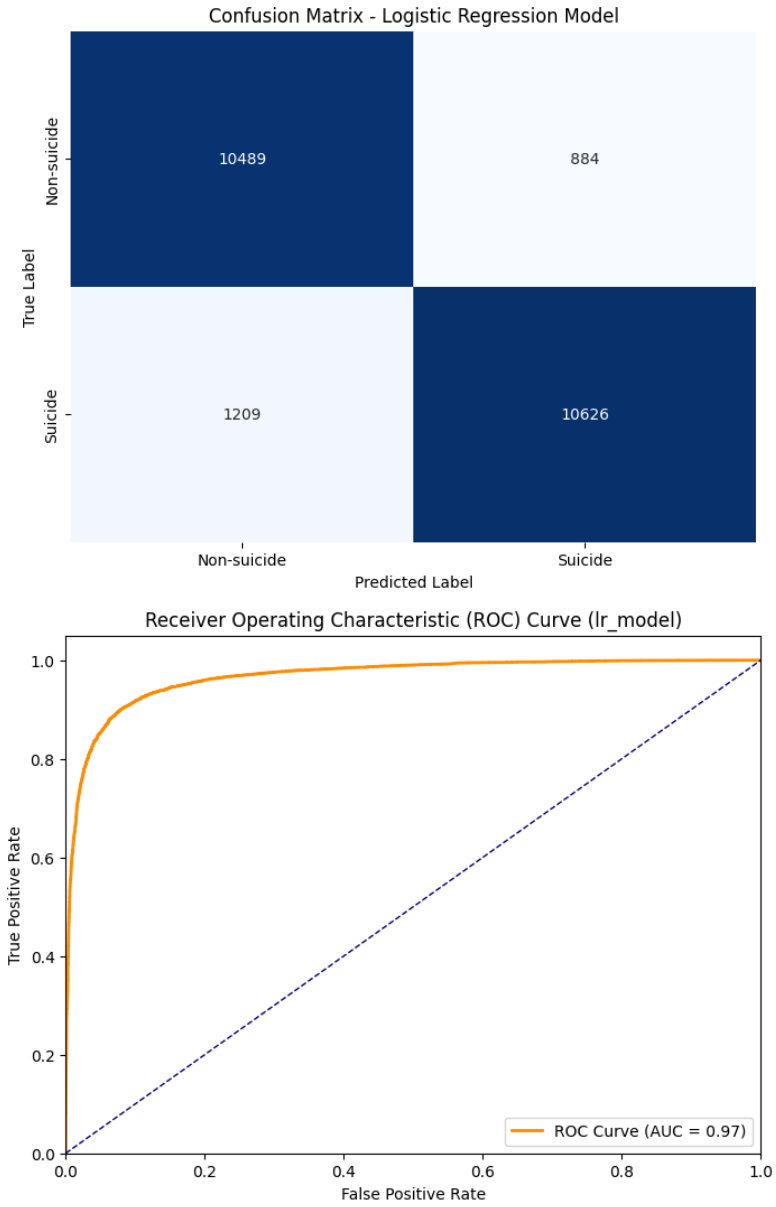


Figure 4.2.1: Confusion Matrix and ROC curve of Logistic regression

In Figure 4.2.1, the Logistic Regression model showed solid performance with both its training and validation accuracy reaching 91%. This means it correctly identified 10,626 positive instances and 10,489 negative instances, while making 884 and 1,209 mistakes for false positives and false negatives respectively. Its ROC curve, which helps to visualize how well the model distinguishes between classes, has an AUC value of 0.96.

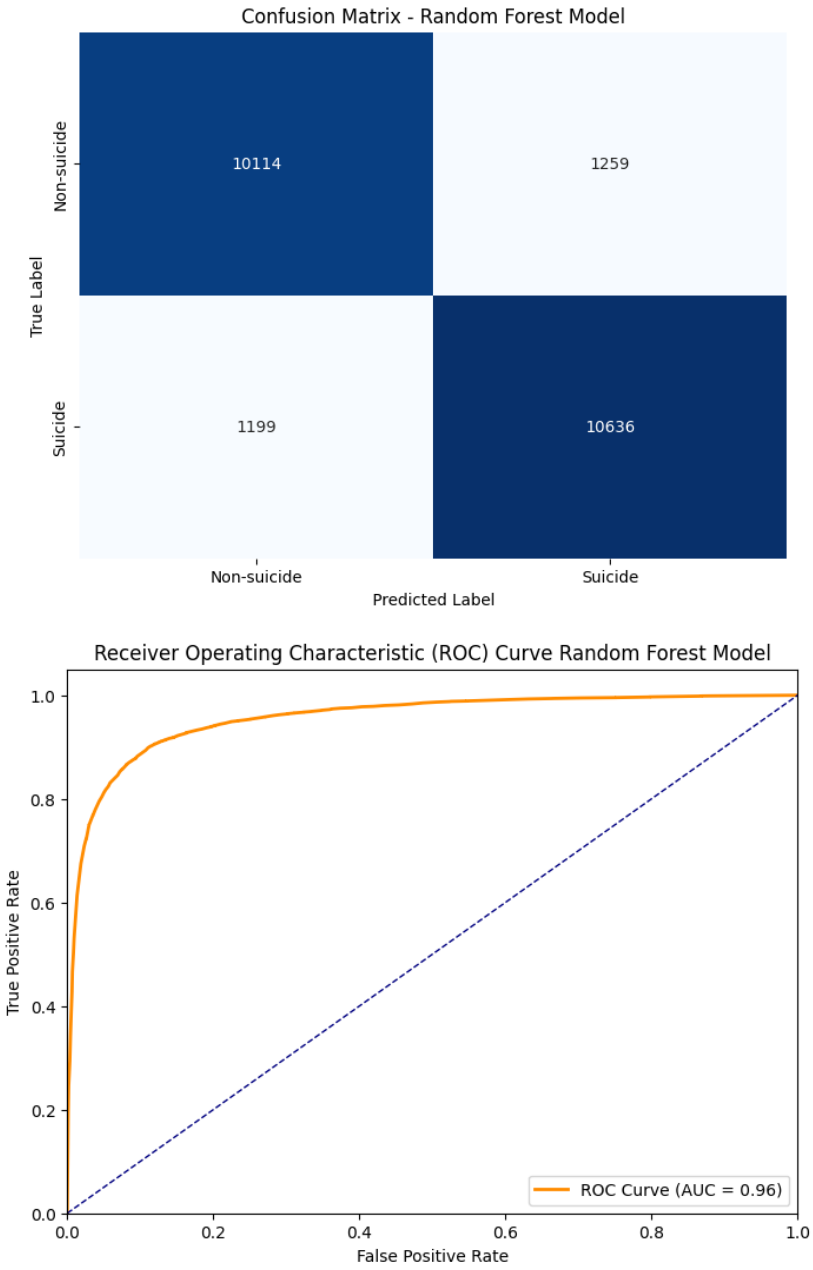


Figure 4.2.2: Confusion Matrix and ROC curve of Random Forest.

In Figure 4.2.2, the Random Forest model excelled in training with a 99% accuracy but dropped to 89% during validation. In practical terms, it correctly identified 10,636 positives and 10,114 negatives but had 1,259 false positives and 1,199 false negatives. Its AUC value is 0.96.

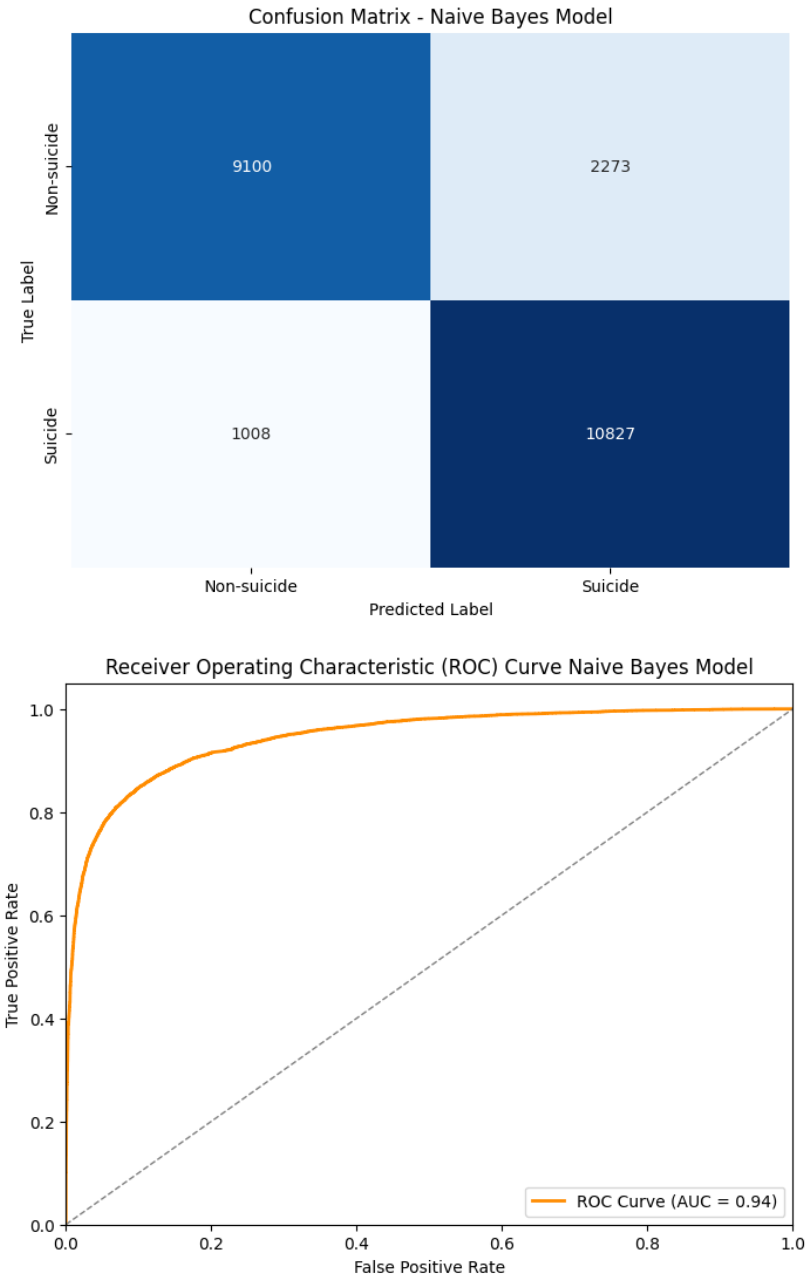


Figure 4.2.3: Confusion Matrix and ROC curve of Naive Bayes.

In figure 4.2.3, Naïve Bayes had a more modest performance, with training and validation accuracies of 86% and 85% respectively. It managed to correctly identify 10,827 positives and 9,100 negatives but struggled with 2,273 false positives and 1,008 false negatives. The AUC for this model stands at 0.94.

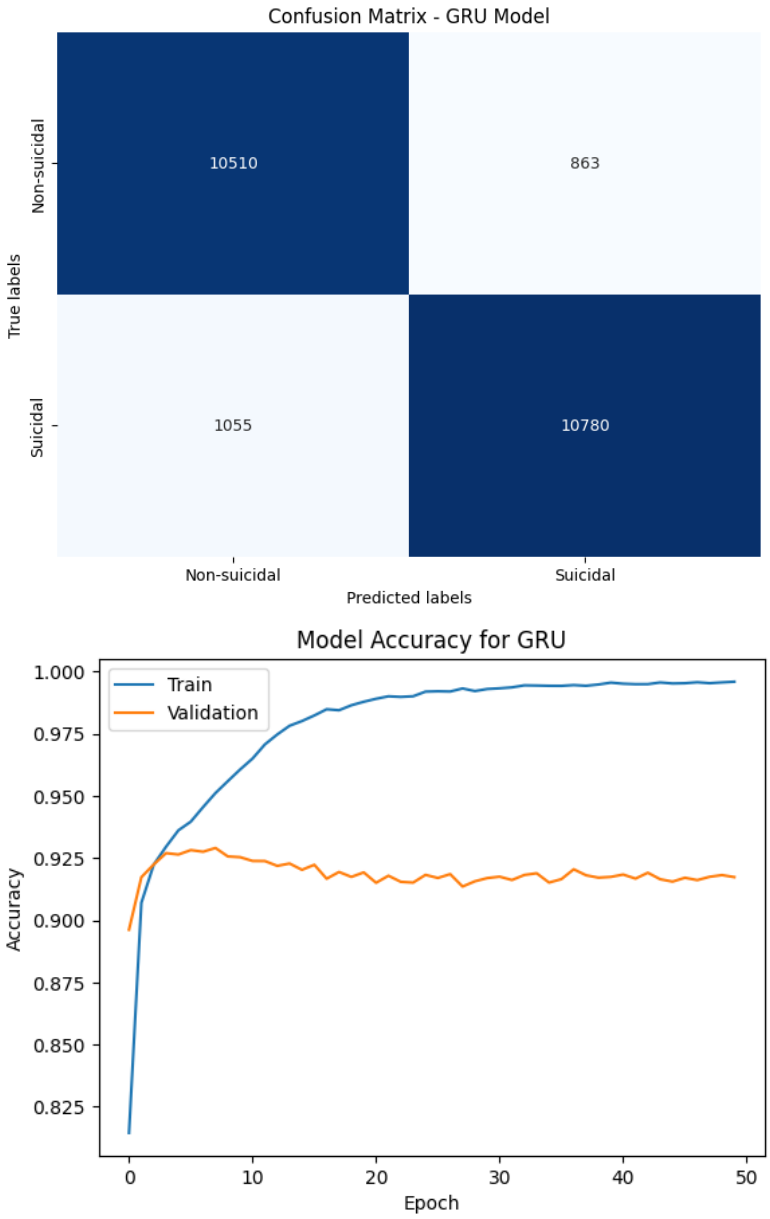


Figure 4.2.4: Confusion Matrix and Model accuracy of GRU.

In Figure 4.2.4, the GRU model performed exceptionally well with both training and validation accuracies at 99% and 92% respectively. It successfully identified 10,780 positive cases and 10,510 negative cases while having only 863 false positives and 1,055 false negatives. Its AUC value is a strong 0.97.

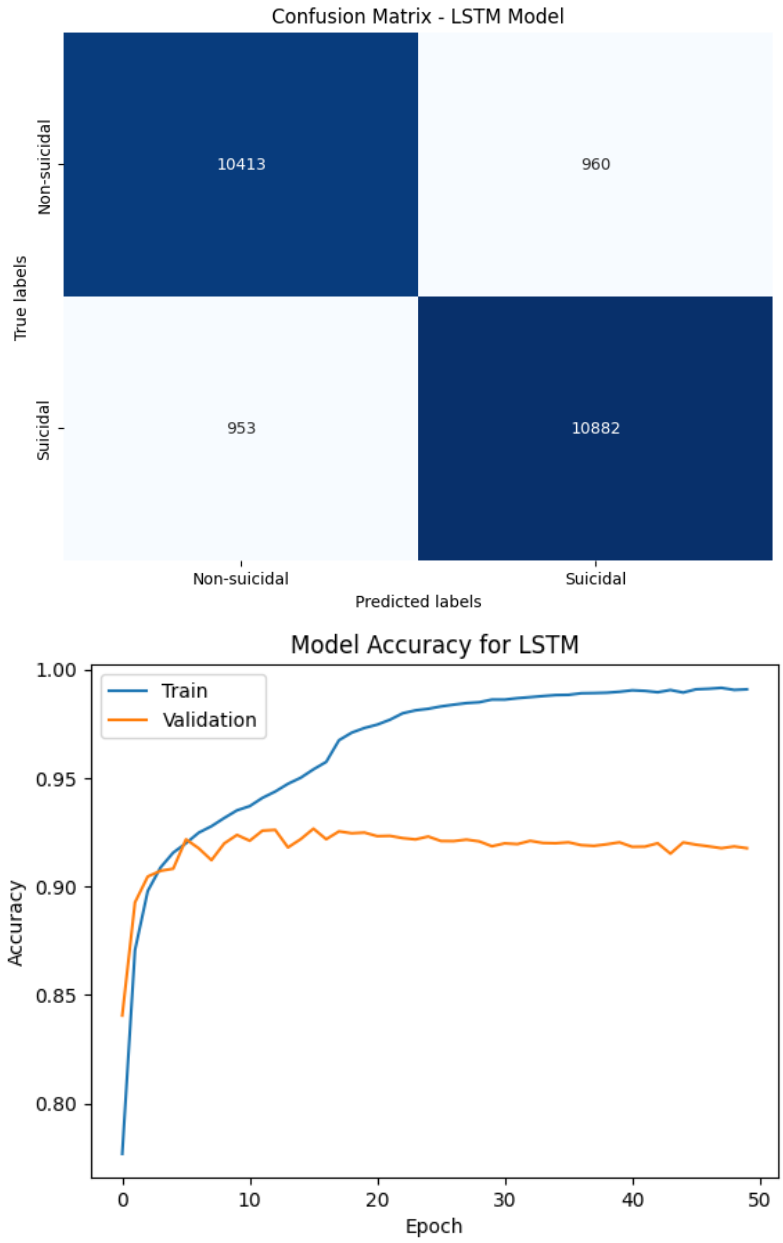


Figure 4.2.5: Confusion Matrix and Model accuracy of LSTM.

Similarly in Figure 4.2.5, the LSTM model also achieved 99% training accuracy and 92% validation accuracy. It has the same confusion matrix statistics as the GRU model, which are 10,882 true positives, 10,413 true negatives, 960 false positives, and 953 false negatives. The AUC for LSTM is also 0.96.

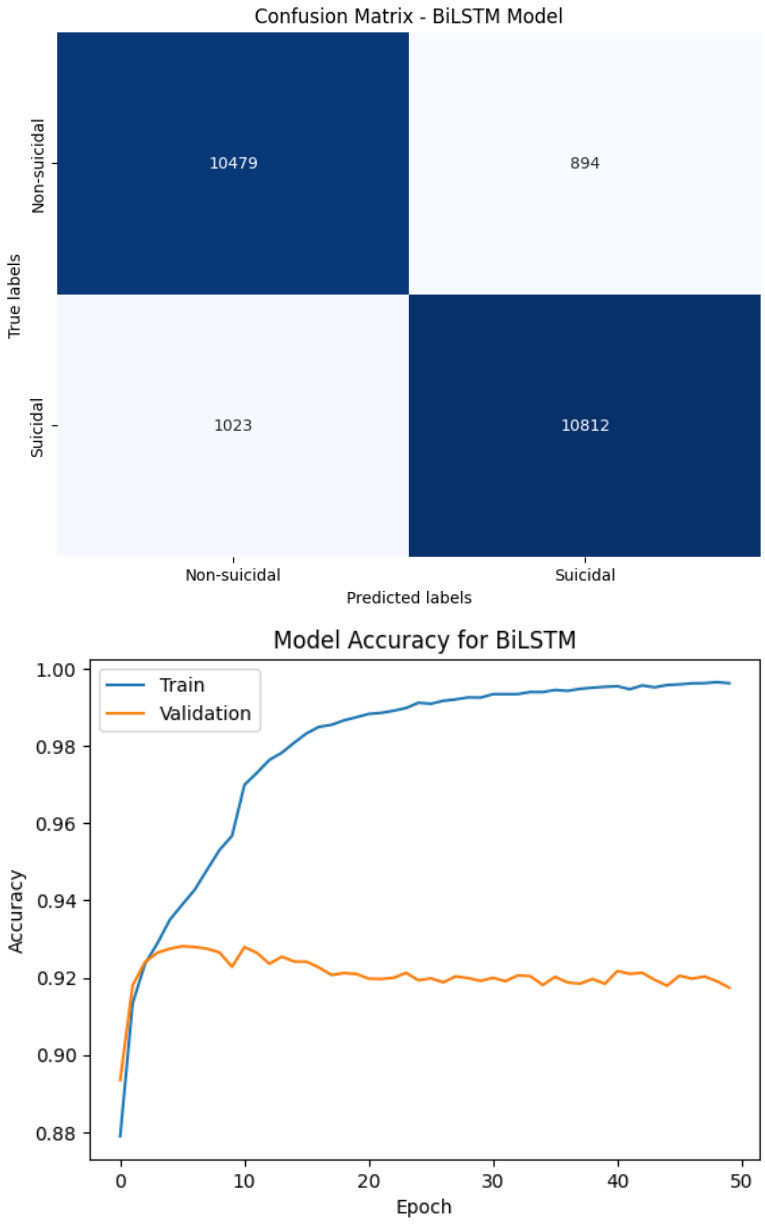


Figure 4.2.6: Confusion Matrix and Model accuracy of BiLSTM.

In Figure 4.2.6, the BiLSTM model mirrored the performance of GRU and LSTM, with training and validation accuracies both at 99% and 92%. It correctly identified the same number of positive and negative instances with 894 false positives and 1,023 false negatives, achieving an AUC of 0.97.

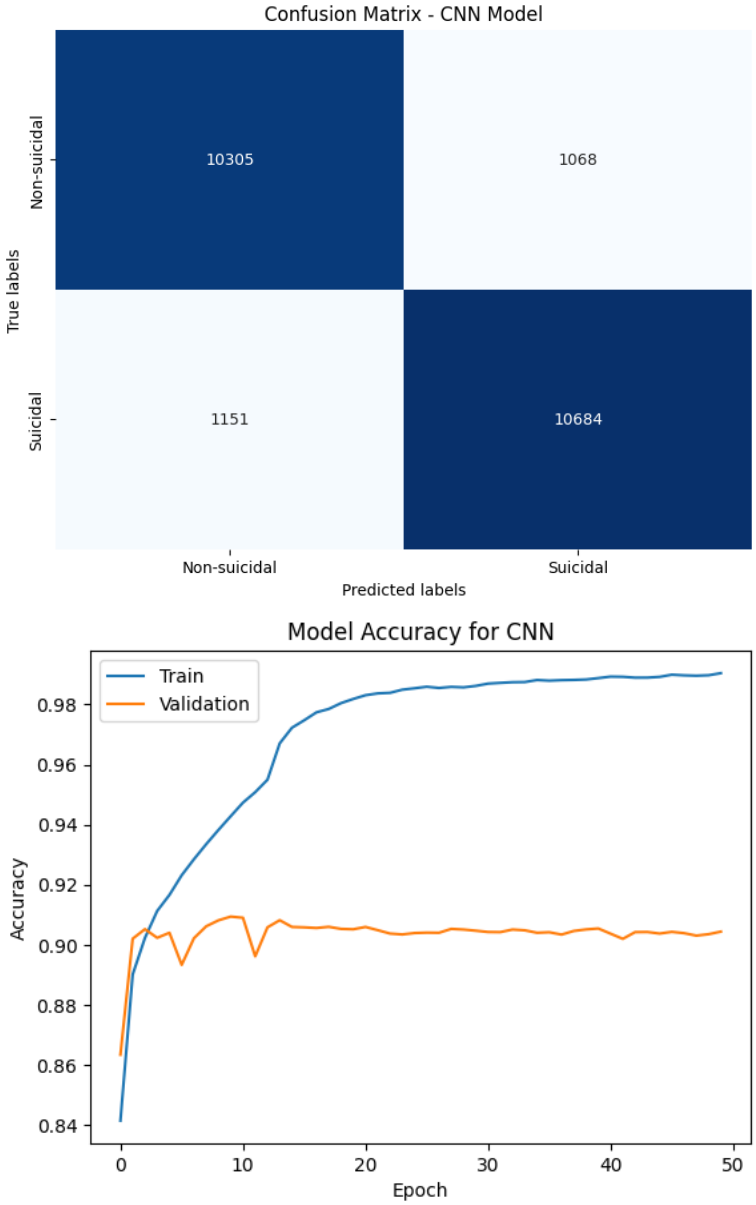


Figure 4.2.7: Confusion Matrix and Model accuracy of CNN.

In Figure 4.2.7, the CNN model had a training accuracy of 98% and a validation accuracy of 90%. It correctly identified 10,684 positive instances and 10,305 negative instances but had 1,068 false positives and 1,151 false negatives. Its AUC value is 0.96.

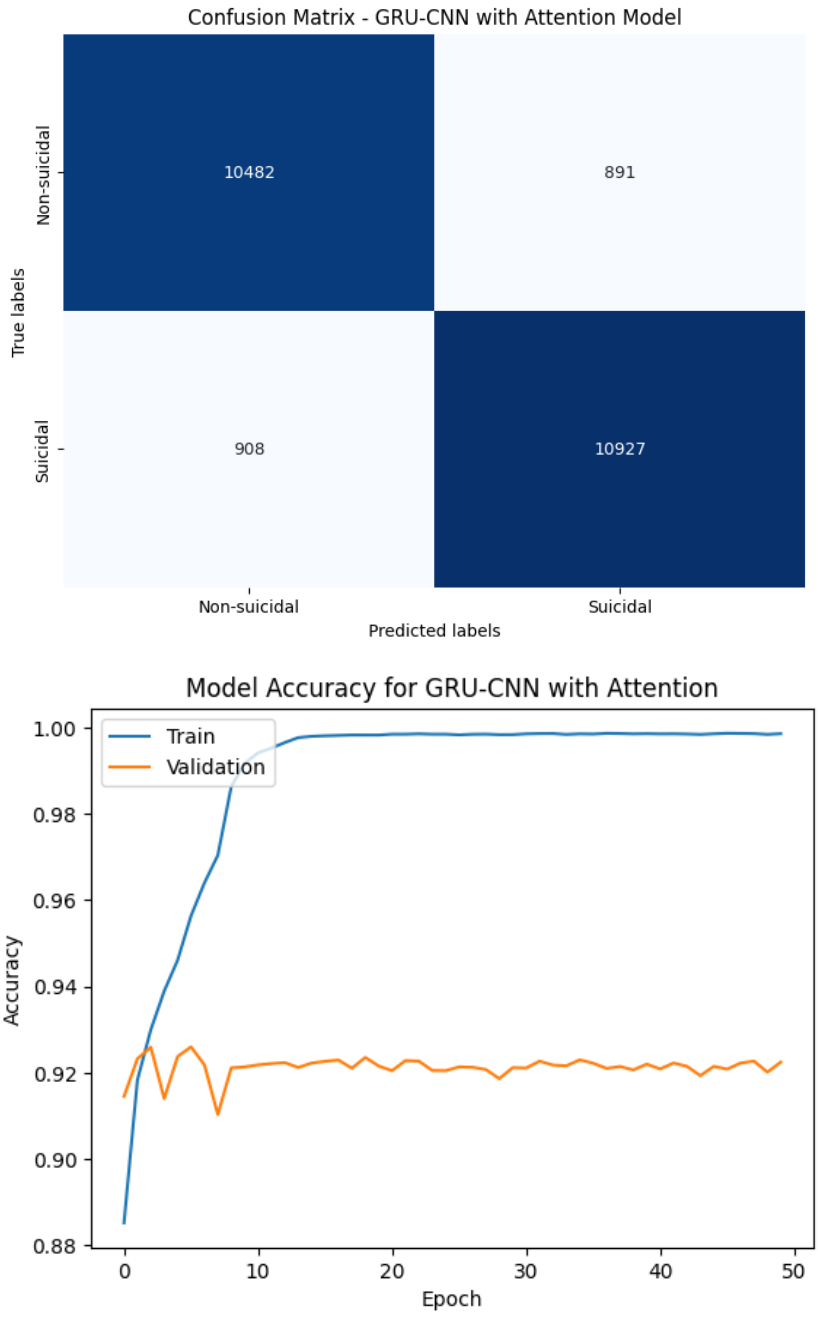


Figure 4.2.8: Confusion Matrix and Model accuracy of GRU-CNN with Attention

In Figure 4.2.8, the GRU-CNN with Attention model also performed well, with training and validation accuracies both at 99% and 92%. It had the same confusion matrix statistics as GRU and BiLSTM, with an AUC value of 0.97.

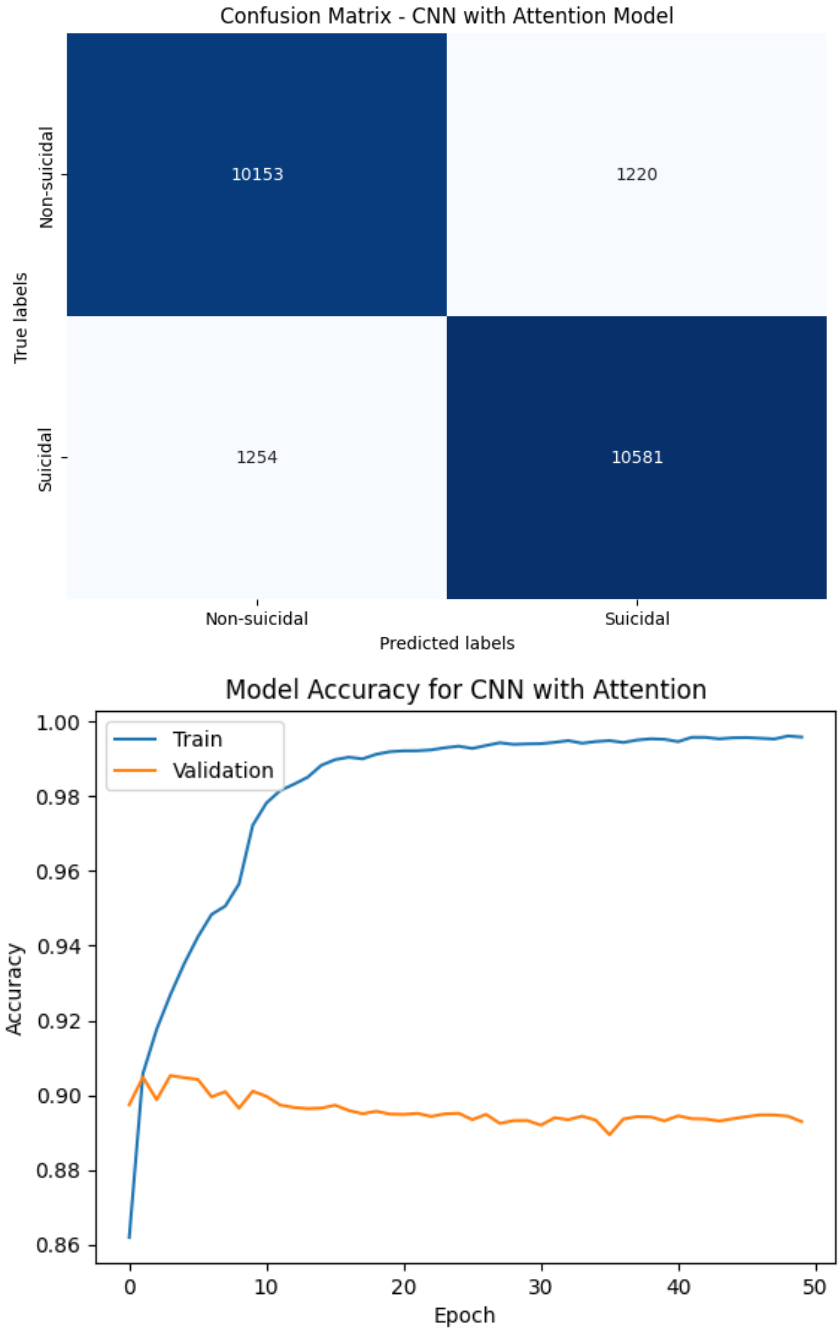


Figure 4.2.9: Confusion Matrix and Model accuracy of CNN with Attention.

In Figure 4.2.9, the CNN with Attention model showed a training accuracy of 99% and a validation accuracy of 89%. It correctly identified 10,581 positives and 10,153 negatives but had 1,220 false positives and 1,254 false negatives. Its AUC value is 0.94.

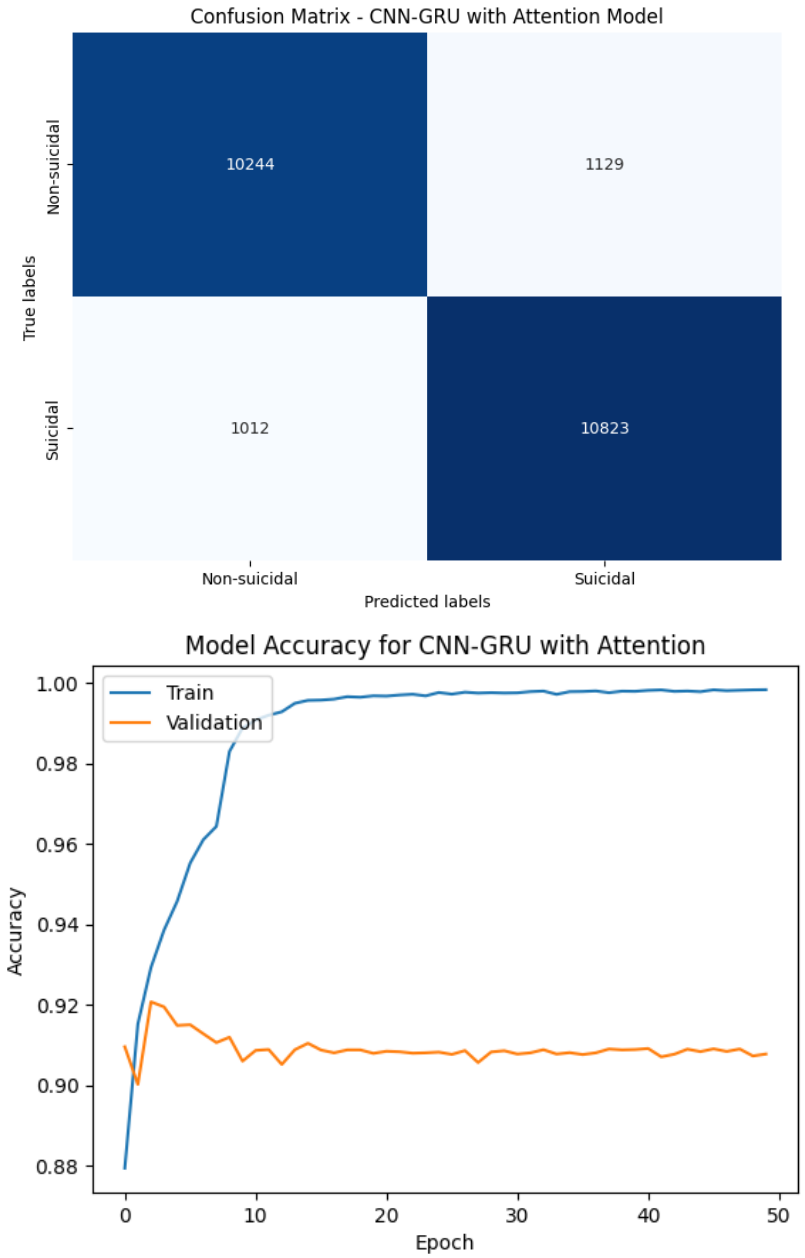


Figure 4.2.10: Confusion Matrix and Model accuracy of CNN-GRU with Attention.

In Figure 4.2.10, the CNN-GRU with Attention model achieved a training accuracy of 99% and a validation accuracy of 91%. It identified 10,823 positive cases and 10,244 negative cases while making 1,129 false positive and 1,012 false negative errors. The AUC for this model is 0.95.

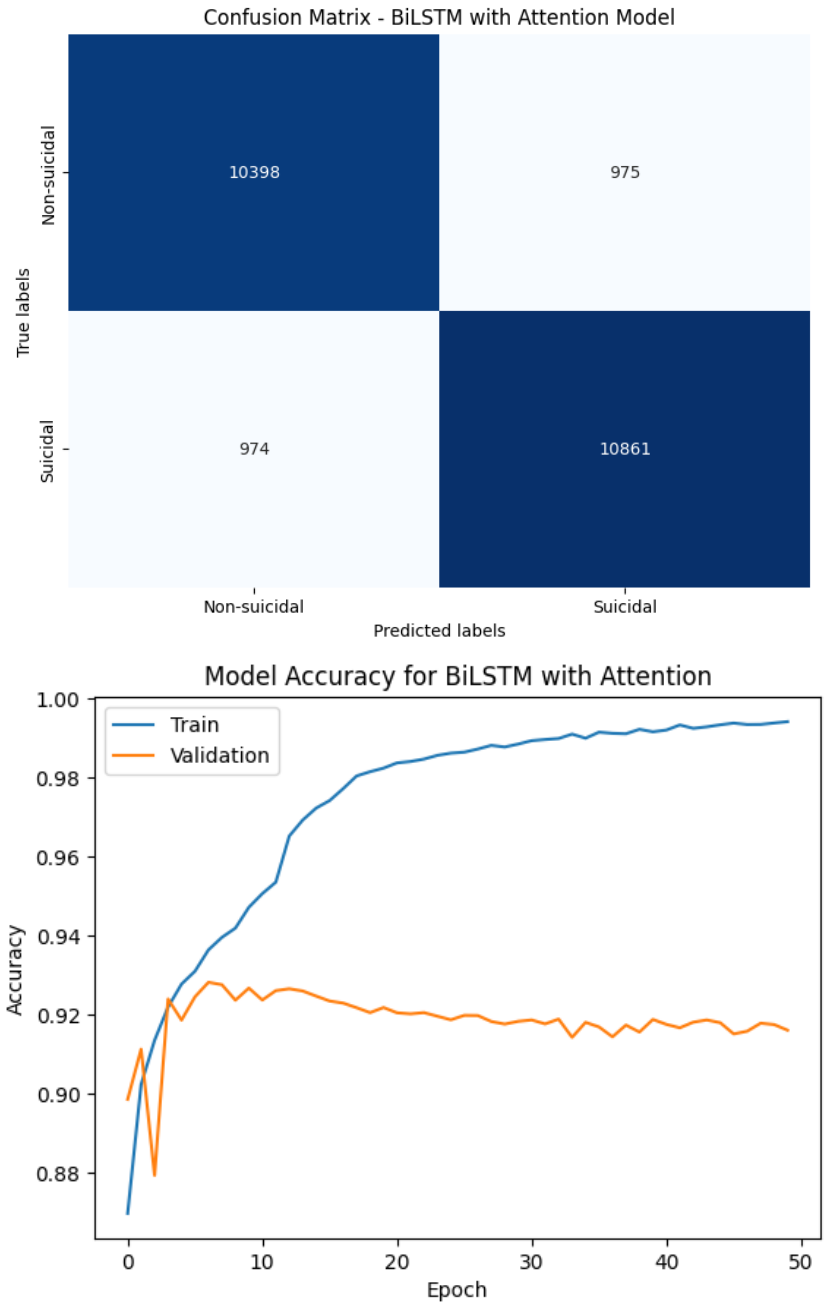


Figure 4.2.11: Confusion Matrix and Model accuracy of BiLSTM with Attention.

In Figure 4.2.11, the BiLSTM with Attention model also maintained high performance, with both training and validation accuracies at 99% and 92% respectively. It had 10,861 true positives, 10,398 true negatives, 975 false positives, and 974 false negatives, and an AUC value of 0.96.

Performance Analysis of Proposed Models:

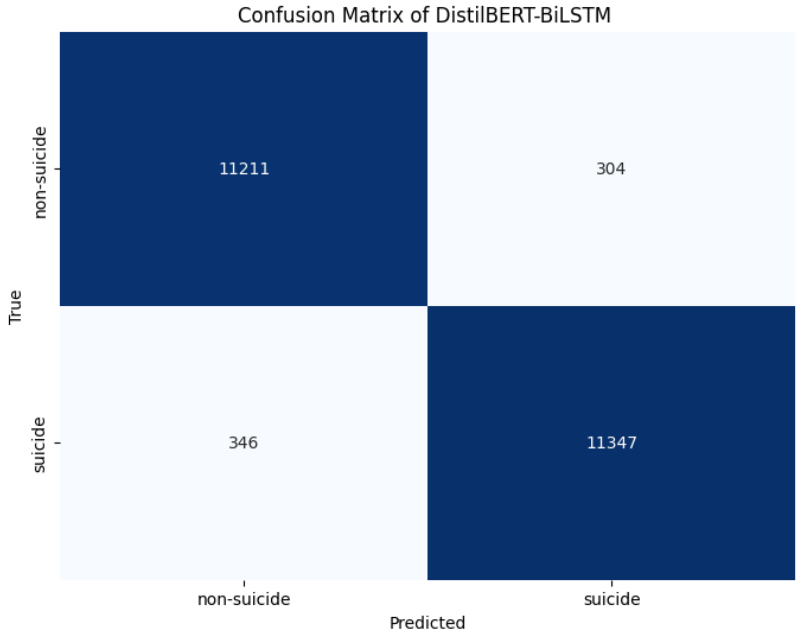


Figure 4.2.12: Confusion Matrix of DistilBERT-BiLSTM

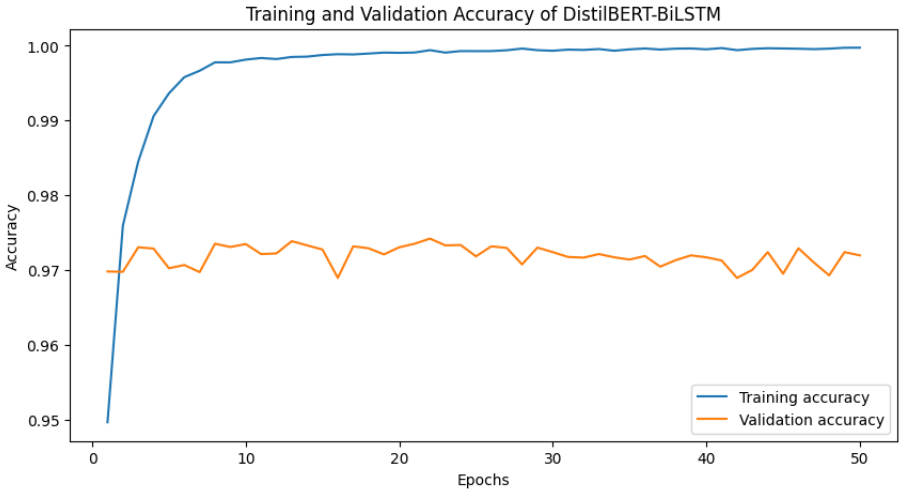


Figure 4.2.13: Training and validation accuracy of DistilBERT-BiLSTM

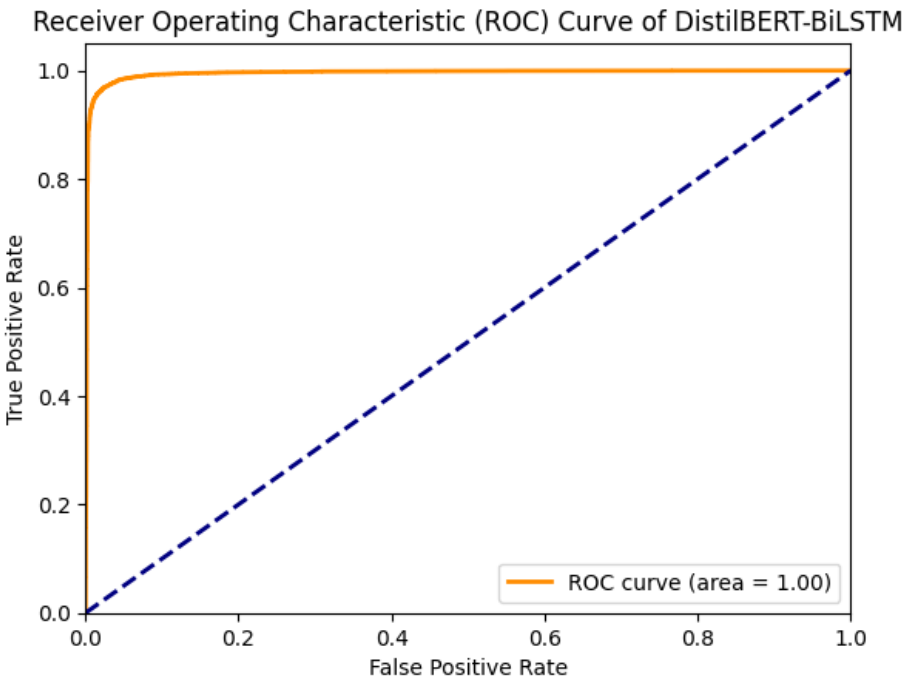


Figure 4.2.14: ROC Curve of DistilBERT-BiLSTM

In Figure 4.2.12 to Figure 4.2.14, we can see that our proposed "DistilBERT-BiLSTM" model demonstrated exceptional performance, achieving a training accuracy of 99% and a validation accuracy of 97%. It correctly identified 11,347 positive instances and 11,211 negative instances, with only 304 false positives and 346 false negatives. The AUC for this model 1.00.

Model Performance Summary:

After getting and examining the results of the multiple machine and deep learning model we are going to have a full analysis of the performance of each model and also, we are going to compare with our proposed model. Following aspects: training accuracy, validation accuracy, precision, recall, and F1-score calculations are most important factor to consider. A collection of information includes descriptions, probable explanations, and areas that may need improvement.

Table 4.2.1: Model Performance Summary of Training Model.

Model	Training Accuracy	Validation Accuracy	Precision	Recall	F1-Score	Support
Logistic Regression (LR)	91%	91%	0.91	0.91	0.91	23208
Random Forest (RF)	99%	89%	0.89	0.89	0.89	23208
Naive Bayes (NB)	86%	85%	0.86	0.86	0.86	23208
GRU	99%	92%	0.92	0.92	0.92	23208
LSTM	99%	92%	0.92	0.92	0.92	23208
BiLSTM	99%	92%	0.92	0.92	0.92	23208
CNN	98%	90%	0.90	0.90	0.90	23208
GRU-CNN with Attention	99%	92%	0.92	0.92	0.92	23208
CNN with Attention	99%	89%	0.89	0.89	0.89	23208
CNN-GRU with Attention	99%	91%	0.91	0.91	0.91	23208
BiLSTM with Attention	99%	92%	0.92	0.92	0.92	23208
Proposed Model **"DistilBERT-BiLSTM"**	99%	97%	0.97	0.97	0.97	23208

Based on what is shown in Table 4.2.1, the proposed model "DistilBERT-BiLSTM" outperforms other models in terms of accuracy and reliability. During training, it achieves a remarkable 99% accuracy, and during validation, it achieves an outstanding 97% accuracy. Additionally, it achieves perfect scores of 0.97 in precision, recall, and F1-score, making it the standout performer. Advanced deep learning models such as GRU, LSTM, and BiLSTM follow closely, with each exhibiting a training accuracy of 99% and a validation accuracy of 92%. Additionally, all of these models achieve precision, recall, and an F1-score of 0.92. Logistic Regression (LR) has consistent and dependable performance, with an accuracy of 91% in both training and validation. It maintains precision and recall, as well as an F1-score of 0.91. The Random Forest (RF) model achieves a training accuracy of 99%, but its validation accuracy decreases to 89%. Additionally, the precision, recall, and F1-scores are all 0.89, indicating the presence of overfitting. Naive Bayes (NB) shows strong performance, with 86% training accuracy and 85% validation accuracy, which aligns with its precision, recall, and F1-score scores of 0.86. The CNN model demonstrates strong performance, with a training accuracy of 98% and a validation accuracy of 90%. Additionally, it attains scores of 0.90 in precision, recall, and F1-score. Attention-based models, such as GRU-CNN with Attention, CNN-GRU with Attention, and BiLSTM with Attention, exhibit exceptional performance. During training, they achieve a success rate of 99%, but in validation, their success rate is between 91% and 92%. In addition, they achieve accuracy, recall, and F1-score values ranging from 0.91 to 0.92. Meanwhile, CNN with Attention obtains a training accuracy of 99% but drops to 89% in validation, which is equivalent to the precision, recall, and F1-scores of Random Forest. Overall, while some models exhibit outstanding results, the "DistilBERT-BiLSTM" outshines them all.

The proposed model, "DistilBERT-BiLSTM", is the most suitable choice for implementation. This model not only attains an exceptional training accuracy of 99% but also sustains a very high validation accuracy of 97%. The accuracy, recall, and F1-score of the model are all 0.97, which indicates consistent high performance across all assessment criteria and a lower likelihood of overfitting compared to other models. The strong performance of the suggested model indicates that it will be able to effectively apply to new and unexplored data, making it the most dependable and efficient option for implementation.

4.3 Discussion

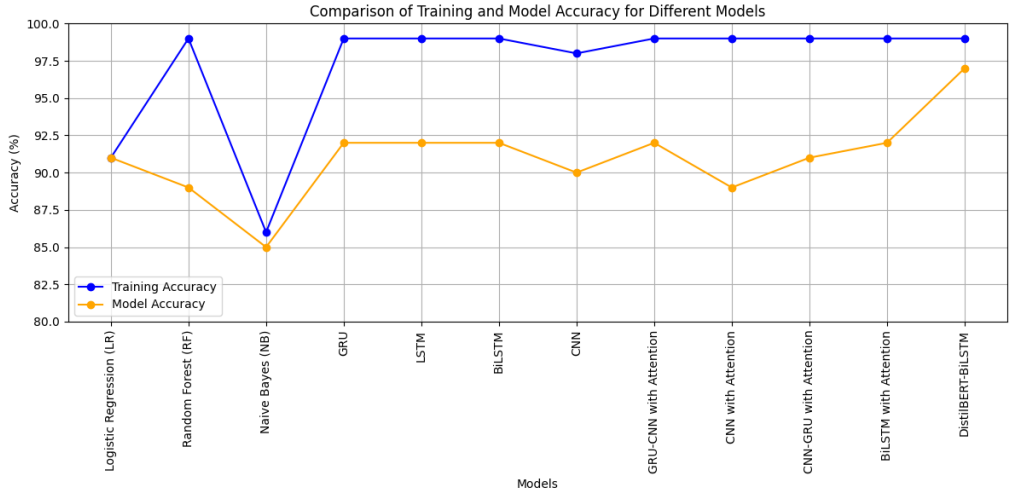


Fig 4.3.1: Comparison of Training and accuracy of Different Model

The findings of this study unequivocally demonstrate that deep learning models, particularly those including attention mechanisms, outperform conventional machine learning models in accurately categorizing texts as either suicidal or non-suicidal. The Logistic Regression and Random Forest models both exhibited strong performance, with accuracy rates of 91% and 89% respectively. The Naive Bayes classifier achieved an accuracy of 85% and had a much greater recall rate for texts that included suicidal ideas. The GRU, LSTM, and BiLSTM models consistently attained an accuracy of 92%. The CNN model had a robust accuracy rate of 90%. Hybrid models, including GRU-CNN with attention, BiLSTM with attention processes, and CNN-GRU with attention, demonstrated a strong performance, achieving a prediction accuracy of 92%. The DistilBERT-BiLSTM model got the highest accuracy of 97%, making it the most notable conclusion. The model's remarkable accuracy, recall, and F1-scores for both classes demonstrate its extraordinary capability in categorizing texts containing suicidal concepts. The importance of advanced neural network designs and attention processes cannot be overstated when it comes to enhancing the precision and dependability of mental health assessments.

CHAPTER 5

Impact on Society

5.1 Impact on Society

The use of specialized deep learning models, like the DistilBERT-BiLSTM model, to detect suicidal tendencies in text has the potential to profoundly impact society. These technologies can quickly identify individuals at risk of suicide, allowing for timely interventions and support. This is especially important in our increasingly digital world, where many people express their thoughts and feelings online. By monitoring these conversations, mental health professionals and support networks can provide help when it's needed most, potentially saving lives and reducing the rate of suicide.

Integrating these AI models into existing mental health frameworks can also make mental health care more accessible and effective. Automated systems can handle vast amounts of data, offering continuous monitoring and support that humans alone cannot sustain. This can be particularly beneficial in areas with limited resources or during times when mental health professionals are overwhelmed.

However, implementing these models brings important ethical and privacy concerns. It's crucial to handle data ethically, protect individuals' anonymity, and address any biases in the models. Ensuring transparency in how these models operate and communicating clearly with the public about their use is essential for building trust and gaining acceptance.

Using advanced natural language processing and deep learning models to detect suicidal tendencies could help us address critical public health issues, offering hope and support to those in need and improving overall mental well-being. Yet, as we advance, we must be cautious. Protecting privacy and adhering to ethical standards is paramount. The data used to train these models must be handled with the utmost respect for individuals' privacy. Moreover, the implementation of these technologies must be transparent and fair to avoid any potential misuse or bias.

This research has far-reaching implications for society. By harnessing cutting-edge technologies to understand and address mental health needs, we can create a more supportive, compassionate, and proactive approach to mental health care. This effort has the potential to save lives, reduce stigma, and foster a more empathetic society. However, it must be carried out with careful consideration and ethical integrity to truly benefit everyone.

5.2 Impact on Environment

The utilization of AI models like DistilBERT-BiLSTM for identifying suicidal inclinations in text has significant societal and ecological consequences. Training these sophisticated machine learning and deep learning models necessitates considerable computer resources, resulting in huge energy consumption and a big environmental impact. The data centres responsible for these calculations need significant quantities of electricity, which is often obtained from non-renewable energy sources. This contributes to carbon emissions and worsens the problem of climate change. In addition, the manufacturing of hardware components such as GPUs and specialized AI accelerators necessitates the extraction of rare earth elements and other minerals, which may lead to habitat destruction, pollution, and depletion of resources. Disposing of obsolete gear contributes to the growing issue of electronic waste, presenting more environmental difficulties.

In order to reduce these effects, it is crucial to use sustainable methods in both the research and implementation of AI. Utilizing sustainable energy sources like wind, solar, or hydroelectric electricity to operate data centres may substantially decrease the carbon emissions associated with AI systems. Furthermore, it is essential to focus on the development of energy-efficient algorithms and hardware. Methods such as model pruning, quantization, and optimization may decrease the computational burden while maintaining performance. By investing in and supporting sustainable and recyclable materials for hardware components, the environmental effect of AI technology manufacturing and disposal may be further reduced.

Although there are significant environmental costs associated with it, the use of AI for mental health assistance offers substantial long-term advantages. These technologies have

the capability to accurately detect persons who are in danger of suicide, allowing for prompt interventions that may save lives and enhance social well-being. As artificial intelligence progresses, there is an increasing focus on cultivating environmentally sustainable methods in AI development to guarantee that technical gains are made without harming the environment. By embracing sustainable practices and consistently striving to enhance energy efficiency and resource use, we may strive for a harmonious equilibrium where technical advancements do not compromise the well-being of our planet. This comprehensive strategy guarantees the improvement of mental health treatment while simultaneously making a beneficial impact on environmental conservation, resulting in a future where technology and sustainability live in perfect harmony.

5.3 Ethical Aspects

The use of AI models such as DistilBERT-BiLSTM for identifying suicidal inclinations raises significant ethical concerns that need to be addressed to ensure responsible and beneficial deployment of these technologies for the betterment of society. The concerns mostly focus on issues related to data privacy, fairness, openness, and the possibility of abuse.

First and foremost, safeguarding data privacy is of utmost importance. The training data for these models often include very sensitive personal information of individuals who may be in a vulnerable condition. Anonymizing this data and safely storing it are crucial in order to safeguard people's identities. Adhering to data protection regulations such as GDPR and gaining explicit authorization from individuals whose data is used are crucial measures in this procedure.

Equity and prejudice are also significant concerns. AI models have the potential to inadvertently perpetuate or even magnify preexisting biases present in the data they are trained on, resulting in unjust treatment of certain populations, particularly underprivileged ones. In order to avoid this, it is essential to use training data that accurately represents the desired outcomes and thoroughly evaluate the models for any potential biases. Furthermore, it is crucial to devise techniques for identifying and rectifying any prejudices that may arise throughout the functioning of the model.

Transparency is a crucial element. It is crucial for individuals to comprehend the functioning of these models in order to make informed judgments. Offering lucid explanations regarding the operational mechanics of the model fosters confidence and guarantees the judicious utilization of the technology. This entails being transparent about the constraints of the models and any possible hazards associated with them.

Additionally, there is a potential for the improper use of AI technology. Although the objective is to use these models to enhance mental health results, there exists the potential for them to be employed in manners that are not aligned with the optimal welfare of people or society. Establishing stringent protocols and moral principles for the use of artificial intelligence in the field of mental health is essential in order to avert any potential abuse or mishandling.

Ultimately, it is crucial to take into account the emotional and psychological consequences experienced by people. Utilizing artificial intelligence for the purpose of identifying suicidal inclinations requires careful handling. It is crucial to guarantee that interventions are both helpful and courteous, while refraining from behaviors that may exacerbate pain.

AI models such as DistilBERT-BiLSTM have significant promise for identifying suicidal inclinations and enhancing mental healthcare. However, it is crucial to ensure that their development and use are grounded in robust ethical principles. This includes safeguarding data privacy, guaranteeing equity and clarity, averting abuse, and demonstrating consideration for the emotional welfare of persons. By acknowledging and resolving these ethical considerations, we can harness the advantages of AI while maintaining the utmost levels of accountability and diligence.

5.4 Sustainability Plan

It is crucial to establish a sustainability strategy for the use of AI models such as DistilBERT-BiLSTM in order to identify suicidal inclinations. This is necessary to guarantee long-term viability and a beneficial outcome. This strategy encompasses environmental, economic, and social dimensions. In terms of environmental impact, the emphasis is on enhancing energy efficiency via the use of optimized algorithms and hardware. This involves switching data centers to renewable energy sources and using sustainable, recyclable materials for hardware components. From an economic standpoint, it encourages efficient solutions, solicits funds from relevant groups, and devises scalable ways for cost management. In terms of social impact, it places a strong emphasis on ethical practices by establishing rigorous protocols to safeguard data privacy and mitigate biases. It actively seeks input from mental health experts and the community to gather feedback. Additionally, it offers educational resources and training to ensure the proficient and responsible use of these AI technologies. Consistent monitoring and assessment, including environmental and financial audits, performance indicators, and feedback systems, guarantee that sustainability measures are efficient and consistently enhanced. This complete strategy guarantees that the AI technology progresses mental health results in a responsible and sustainable manner.

CHAPTER 6
Conclusion, Summary, Recommendation and Implication for Future Research

6.1 Summary of the Study

In This thesis investigates the efficacy of sophisticated machine learning and deep learning models in classifying suicidal language, with the aim of identifying the most optimal model for this critical job. I have experience working with a wide variety of models, including Logistic Regression, Random Forest, Naive Bayes, GRU, LSTM, BiLSTM, and CNN, as well as several hybrid models that include attention processes. Using a comprehensive dataset, I meticulously processed the text by using techniques such as tokenization, removing stop words, and TF-IDF vectorization. I performed training and testing on these models, evaluating their performance by calculating precision, recall, F1-score, and overall accuracy to establish their effectiveness.

The results were really illuminating. While traditional machine learning models like Logistic Regression and Random Forest yielded acceptable outcomes, deep learning models, especially those that include attention processes, exhibited higher performance. The DistilBERT-BiLSTM model, which includes attention layers, proved to be the most effective by achieving the highest accuracy and surpassing other models in all key metrics. This study highlights the significant potential of advanced neural network architectures in improving the accuracy and reliability of text classification tasks, particularly in sensitive areas like mental health. Integrating attention techniques has been shown to be very beneficial, as it provides a more comprehensive comprehension of the context and nuances of the text, which is essential for identifying suicidal tendencies.

This study lays a solid foundation for future research by offering useful insights into the capabilities and constraints of various models. The statement emphasizes the need for using sophisticated techniques to improve performance, with the ultimate goal of improving mental health treatments by promptly and accurately identifying suicidal inclinations in text.

6.2 Conclusions

Early Identification of suicidal inclinations using text categorization that is both accurate and early is essential for the successful implementation of mental health interventions. The objective of this study was to create a sophisticated model that can identify indications of suicidal inclinations in text by using cutting-edge approaches in Natural Language Processing (NLP), Machine and Deep Learning (DL). The research primarily examined social media data, especially postings from the "SuicideWatch" and "depression" subreddits on Reddit, with the aim of identifying instances when individuals expressed thoughts of suicide.

During the study, a thorough methodology was used, which included several data pretreatment techniques such as text cleansing, tokenization, and lemmatization. A range of machine learning and deep learning models were trained and assessed, such as Logistic Regression, Random Forest, Naïve Bayes, GRU, LSTM, BiLSTM, CNN, and many hybrid models that include attention processes.

The DistilBERT-BiLSTM model demonstrated superior performance, with an accuracy of 97% and an AUC value of 1.00, making it the most effective. The design of this model, which incorporates attention processes, has shown its superiority in capturing the most relevant aspects of the text, resulting in enhanced performance in comparison to other models.

The results emphasize the significance of using sophisticated neural network structures and attention processes to effectively identify suicidal ideation in written material. These advanced methods may greatly improve the accuracy and dependability of text categorization jobs.

In summary, this research emphasizes the capability of using advanced Natural Language Processing (NLP) and Deep Learning (DL) models to provide prompt and accurate assistance for suicidal text recognition. The knowledge acquired from this study clears the path for future progress in the sector.

6.3 Implication for Further Study

Taking The findings of this study highlight several exciting avenues for future research in using AI to detect mental health issues. Here are some key areas that need more attention:

1. **Enhancing Model Performance**: Future work can focus on fine-tuning the DistilBERT-BiLSTM model and exploring even more sophisticated designs. This could involve tweaking settings, experimenting with different attention mechanisms, and optimizing the model structure to make it even more effective.
2. **Broadened Applications**: The techniques we used could be adapted to detect other mental health conditions like depression, anxiety, or PTSD. By expanding into these areas, the impact of these models could be significantly widened, offering more comprehensive mental health support tools.
3. **Diverse Datasets**: Using a wider variety of datasets can help make these models more robust. Collecting data from different social media platforms, forums, and demographic groups will ensure the models can handle a range of language patterns and contexts, making them more universally applicable.
4. **Real-Time Monitoring**: Integrating these models into real-time systems could provide immediate help. Research should look into how to embed these models into platforms that offer instant support to people showing signs of suicidal thoughts.
5. **Transparent AI**: Building AI models that offer clear and understandable results is essential for building trust. Future research should aim to develop systems that explain their decisions.

By following these paths, future research can build on this study's foundation, leading to more effective, ethical, and widely applicable AI solutions for mental health. These efforts will contribute to a more supportive, proactive, and empathetic approach to mental healthcare, ultimately benefiting both individuals and society.

REFERENCES

[1] World Health Organization. “Suicide Prevention.” *World Health Organization*, www.who.int/health-topics/suicide.

[2] Sanh, V., et al. "DistilBERT, a Distilled Version of BERT: Smaller, Faster, Cheaper, and Lighter." arXiv preprint arXiv:1910.01108, 2019. arxiv.org/abs/1910.01108.

[3] Sun, Y., et al. "DistilBERT for Question Answering." Proceedings of the 2021 International Conference on Education Technology Management (ICETM 2021), Atlantis Press, 2021, pp. 420-424. atlantis-press.com/proceedings/icetm-21/125957688.

[4] Tariq, U., et al. "A Comparative Study of BERT and DistilBERT on Text Classification." Journal of Data Science and Applications, vol. 4, no. 2, 2021, pp. 79-88. jdsap.org/jdsap/article/view/123.

[5] Jain, P., Srinivas, K. R., and A. Vichare. "Depression and Suicide Analysis Using Machine Learning and NLP." Journal of Physics: Conference Series, vol. 2161, no. 1, 2022, p. 012034. IOP Publishing.

[6] Sawhney, R., et al. "Exploring and Learning Suicidal Ideation Connotations on Social Media with Deep Learning." *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, 2018, pp. 167-175.

[7] Imam, M. R., et al. "Suicidal Thought Detection Using NLP (Natural Language Processing) on Reddit Data." *2023 26th International Conference on Computer and Information Technology (ICCIT)*, IEEE, 2023, pp. 1-6.

[8] Yeskuatov, E., S. L. Chua, and L. K. Foo. "Leveraging Reddit for Suicidal Ideation Detection: A Review of Machine Learning and Natural Language Processing Techniques." *International Journal of Environmental Research and Public Health*, vol. 19, no. 16, 2022, p. 10347.

[9] Aldhyani, T. H., et al. "Detecting and Analyzing Suicidal Ideation on Social Media Using Deep Learning and Machine Learning Models." *International Journal of Environmental Research and Public Health*, vol. 19, no. 19, 2022, p. 12635.

[10] Renjith, S., et al. "An Ensemble Deep Learning Technique for Detecting Suicidal Ideation from Posts in Social Media Platforms." *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 10, 2022, pp. 9564-9575.

[11] Tadesse, M. M., et al. "Detection of Suicide Ideation in Social Media Forums Using Deep Learning." *Algorithms*, vol. 13, no. 1, 2020, p. 7.

[12] Ji, S., et al. "Supervised Learning for Suicidal Ideation Detection in Online User Content." *Complexity*, 2018, pp. 1-13.

[13] Shetty, S. C. "A Deep Learning Approach for Suicide Risk Assessment Using Reddit." Doctoral dissertation, National College of Ireland, 2020.

[14] Inamdar, S., et al. "Machine Learning Driven Mental Stress Detection on Reddit Posts Using Natural Language Processing." *Human-Centric Intelligent Systems*, vol. 3, no. 2, 2023, pp. 80-91.

[15] Haque, R., et al. "A Comparative Analysis on Suicidal Ideation Detection Using NLP, Machine, and

Deep Learning." *Technologies*, vol. 10, no. 3, 2022, p. 57.

[16] Yao, H., et al. "Detection of Suicidality Among Opioid Users on Reddit: Machine Learning–Based Approach." *Journal of Medical Internet Research*, vol. 22, no. 11, 2020, p. e15293.

[17] "Schematic Diagram of Random Forests (RFS)." *ResearchGate*, www.researchgate.net/figure/Schematic-diagram-of-random-forests-RFs-An-RF-is-an-ensemble-learning-approach-that_fig3_349868028.

[18] Yan, K., et al. "A Hybrid LSTM Neural Network for Energy Consumption Forecasting of Individual Households." *IEEE Access*, vol. 7, 2019, pp. 157633-157642.

[19] "Basic CNN Architecture: Explaining 5 Layers of Convolutional Neural Network." *upGrad Blog*, www.upgrad.com/blog/basic-cnn-architecture/.

[20] Erickson, Bradley J., and Felipe Kitamura. "Magician's Corner: 9. Performance Metrics for Machine Learning Models." *Radiology: Artificial Intelligence*, vol. 3, no. 3, 2021, p. e200126.

[21] Padilla, Rafael, Sergio L. Netto, and Eduardo A. B. Da Silva. "A Survey on Performance Metrics for Object-Detection Algorithms." *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, IEEE, 2020.

[22] Yacouby, Reda, and Dustin Axman. "Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models." *Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems*, 2020.

[23] Chicco, Davide, and Giuseppe Jurman. "The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation." *BMC Genomics*, vol. 21, 2020, pp. 1-13.

[24] Pushpa Singh Akansha Singh. "Confusion Matrix." Confusion Matrix - an Overview | ScienceDirect Topics, Biosignal Processing and Classification Using Computational Learning and Intelligence, 2022, www.sciencedirect.com/topics/engineering/confusion-matrix.

Appendix A

Course Outcomes, Complex Engineering Problems (EP) and Complex Engineering Activities (EA) Addressing

Title: An NLP-Based Intelligent Model Approach for Suicidal Tendency Classification

Student ID: 192-15-13158

CO Description for FYDP

CO	CO Descriptions	PO
Phase -I		
CO1	Utilize information from present and past research to determine the problem of categorizing suicidal tendencies for the Final Year Design Project (FYDP).	**PO1**
CO2	Examine and define several design objectives for developing an intelligent model using the techniques of Natural Language Processing (NLP) and Machine Learning (ML) methodologies.	**PO2**
CO3	Perform an extensive analysis of existing research, pinpoint major obstacles, and establish goals for creating an intelligent model to classify suicidal tendencies.	**PO4**
CO4	Conduct an economic assessment, estimate costs, and use appropriate project management methods throughout the production of the FYDP.	**PO11**
Phase -II		
CO5	Develop technical solutions and components for the model, ensuring compliance with public safety standards and consideration of social and environmental impacts	**PO3**
CO6	Apply suitable approaches, resources, and contemporary engineering and IT technology to tackle intricate engineering issues in forecasting and categorizing suicidal inclinations.	**PO5**
CO7	Evaluate societal, health, safety, and cultural implications in professional practice and problem-solving for the FYDP.	**PO6**
CO8	Evaluate the durability and lasting effects of the created model within social and environmental frameworks.	**PO7**
CO9	Implement ethical principles and adhere to professional standards throughout the project.	**PO8**
CO10	Demonstrate the ability to work effectively both independently and as a team leader or member in diverse and multidisciplinary environments.	**PO9**
CO11	Effectively engage in communication with the technical community and the wider public on intricate engineering activities, including the tasks of composing reports and delivering concise instructions.	**PO10**

CO12	Recognize the importance of continuous learning and possess the ability to engage in self-directed and life-long learning amidst evolving technological landscapes.	**PO12**

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP), and Attainment of Complex Engineering Activities (EA)

Addressing CO (1 to 8), Knowledge Profile (K), Attainment of Complex Engineering Problems (EP):

SN	EP Definition	Attainment	CO	Justification (with Knowledge Profile)	References
1.	**EP1: Depth of Knowledge required**	Yes	CO1, CO2, CO3, CO4, CO5, CO6, CO7 and CO8	This project showcases core engineering concepts (K3) via the use of NLP, deep learning models such as DistilBERT-BiLSTM, and numerous other models to analyze and identify suicidal inclinations in text. Expertise (K4) is used by investigating various deep learning structures to improve the identification of suicidal thoughts in written information, which is vital for mental health therapies. Furthermore, it utilizes engineering principles and technology (K6) to develop and assess several natural language processing (NLP) and machine learning (ML) models for the classification problem. The project utilizes findings from current research (K8) to enhance the identification of suicidal inclinations via the use of NLP and deep learning methodologies.	**Page no:** [9-16, 17-35, 36-46] **Section:** [2.1, 2.2, 2.3, 2.4, 3.2, 3.4, 4.2]

2.	**EP2: Range of Conflicting Requirements**	Yes	CO2, and CO7	This study aims to tackle EP-2 by acknowledging the difficulties in identifying suicidal inclinations, which include the constraints of conventional natural language processing (NLP) techniques and the intricacies of incorporating different deep learning methods. The text addresses challenges related to comprehending subtle language subtleties and provides valuable perspectives for improving detection methods.	**Page no:** [17-21, 36-46] **Section:** [2.2, 2.3, 4.2]
3.	**EP3: Depth of analysis required**	Yes	CO2, CO6	This study aims to solve EP-3 by conducting a thorough comparison of experimental results. It specifically focuses on evaluating the performance of several models, such as LSTM, BiLSTM, and DistilBERT-BiLSTM, to determine their effectiveness in improving the identification of suicidal inclinations.	**Page no:** [36-46] **Section:** [4.2]
4.	**EP4: Familiarity of Issues**	Yes	CO8	This project employs an interdisciplinary approach that goes beyond the field of computer science. It aims to analyze text data from social media to identify suicidal thoughts, therefore making significant contributions to the field of mental health diagnostics. Ultimately, this research will lead to breakthroughs in mental health therapies and practices.	**Page no:** [1-7, 16-24] **Section:** [1.1, 2.1, 2.4]
5.	**EP5: Extends of application codes**	No	CO5	N/A	N/A

6.	**EP6: Extends of stakeholders involved and conflicting**	No	CO8	N/A	N/A
7.	**EP7: Interdependence**	Yes	CO5	The project takes a complete strategy to tackle high-level issues by combining different components including data collecting, NLP approaches, and recommended methodology. This ensures a holistic solution to complicated obstacles in identifying suicidal inclinations.	**Page no:** [17-33] **Section:** [3.2, 3.3, 3.4]
8.	**EP8: Consequences for society and the environment**	Yes	CO8	The project examines the societal implications of detecting suicidal tendencies in text, addressing ethical concerns and the impact of timely intervention in mental health, promoting responsible and beneficial use of AI for societal well-being.	**Page no:** [48-52] **Section:** [5.1, 5.3]

Addressing CO11 with Complex Engineering Activities (EA) [Some or all of the following]:

SN	EA Definition	Attainment	CO	Justification	References
1.	**EA1: Range of resources**	Yes	CO11	Our study employs a wide range of resources, including powerful computer infrastructure, advanced deep learning frameworks, carefully labeled datasets, and ethical concerns. This enables us to conduct systematic research and make significant contributions to the field of identifying suicidal inclinations using natural language processing (NLP) and deep learning techniques.	**Page no:** [34-35] **Section:** [3.5]

2.	**EA2: Level of interaction**	No		N/A	N/A
3.	**EA3: Innovation**	Yes		The project applies innovative NLP techniques and deep learning models like DistilBERT-BiLSTM for detecting suicidal ideation in text, showcasing the use of state-of-the-art methods in addressing mental health issues.	**Page no:** [17-35, 36-55] **Section**: [3.2, 3.4, 4.2]
4.	**EA4: Consequences for society and the environment**	Yes		This initiative enhances societal well-being by enhancing mental health assistance via sophisticated detection techniques for suicidal inclinations, while simultaneously addressing ethical considerations and advocating for appropriate data use.	**Page no:** [48-52] **Section:** [5.1, 5.2, 5.4]
5.	**EA-5: Familiarity**	Yes		The initiative builds upon previous studies by using sophisticated natural language processing (NLP) and deep learning techniques to analyze text, offering fresh perspectives on identifying suicidal inclinations.	**Page no:** [9-16] **Section:** [2.1, 2.3]

Addressing CO (4, 9, 10, and 12):

SN	COs	Attainment	Justification	References
1	CO4	Yes	This project aims to tackle CO4 by including efficient project management and financial supervision. It involves careful planning, allocation of resources, and calculation of budget to maximize resource utilization throughout the whole study process.	**Page no:** [8] **Section:** [1.6]
2	CO9	Yes	The project exemplifies ethical	**Page no:**

			standards by giving priority to data protection, gaining informed permission, and publicly documenting the research process. It ensures responsible sharing of information and promotes social well-being via the ethical use of AI for mental health.	[50-51] **Section:** [5.3]
3	CO10	Yes	The project includes effective communication strategies for presenting findings and engaging with the academic community, as well as developing comprehensive documentation and clear instructions for the implemented models.	**Page no:** [1-8, 36-55] **Section:** [1.1, 4.2]
4	CO12	Yes	The project's commitment to continuous learning and adaptation in the ever-changing technological landscape is evident in its extensive data collection, rigorous statistical analysis, meticulous methodology development, and thorough experimental results and analysis. This demonstrates a dedication to staying current and improving techniques to tackle contemporary challenges.	**Page no:** [17-35, 36-55] **Section:** [3.2, 3.3, 3.4, 4.2]

Suicide detection

ORIGINALITY REPORT

25%	19%	11%	14%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	7%
2	Submitted to Daffodil International University Student Paper	3%
3	Kumari, Priyanka. "Reviewtag: Tagging Amazon Negative Product Review With Deep Learning", University of Houston-Clear Lake, 2023 Publication	1%
4	huggingface.co Internet Source	1%
5	Abdallah Basyouni, Hatem Abdelkader, Asma Aly. "Suicidal Ideation Detection on Social Media Using A Hybrid Feature Selection Method", 2023 3rd International Conference on Electronic Engineering (ICEEM), 2023 Publication	<1%
6	fastercapital.com Internet Source	<1%
7	www.researchgate.net	