Transfer Learning for Identification of disaster tweets using fine-tuning DistilBERT

MSc Research Project
Data Analytics

Nikhil Vishnupant Deshmukh
Student ID: x21232946

School of Computing
National College of Ireland

Supervisor: Vikas Tomer

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Nikhil Vishnupant Deshmukh
Student ID:	X21232946
Programme:	Msc Data Analytics **Year:** 2022-23
Module:	Research Project
Supervisor:	Vikas Tomer
Submission Due Date:	September 18th 2023
Project Title:	Transfer Learning for Identification of disaster tweets using fine-tuning DistilBert
Word Count:	.........8629.................................. **Page Count**......32............

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Nikhil Vishnupant Deshmukh

Date: September 18th 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

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

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

Office Use Only	
Signature:	

Date:	
Penalty Applied (if applicable):	

Transfer Learning for Identification of disaster tweets using fine-tuning DistilBert.

Research Project
Nikhil Vishnupant Deshmukh
X21232946
MSc in Data Analytics
18th September 2023

Abstract

This study presents an in-depth exploration of tweet classification for disaster-related content detection, addressing the challenges posed by the proliferation of social media during crises. Three distinct models were developed and evaluated: CNN + LSTM with Glove Embedding, RoBERTa Transformer, and FineTune DistilBERT Transformer. These models were rigorously assessed for their ability to distinguish between disaster and non-disaster-related tweets, with a focus on accuracy, sensitivity, and specificity. The CNN + LSTM model exhibited promising precision but lacked recall for disastrous tweets. RoBERTa demonstrated enhanced performance owing to its extensive training data and methodology. However, the FineTune DistilBERT model emerged as the standout performer, showcasing a balanced sensitivity-specificity trade-off and achieving an impressive 89% accuracy. Leveraging its DistilBERT architecture, this model offers a compact yet powerful solution for accurate tweet classification. The findings underline the potential of transformer-based models in crisis informatics, specifically for identifying disaster-related content in social media streams. This study contributes to advancing rapid and accurate crisis response tools, empowering humanitarian organizations with improved insights and aiding decision-making in disaster scenarios. Further research avenues may explore ensemble methods and domain-specific fine-tuning to enhance model performance across diverse disaster contexts.

Keywords: Transformer-based models, FineTune DistilBERT Transformer, CNN+LSTM, RoBERTa , Disaster contexts

List of Abbreviations	
Deep Learning	(DL)
Convolutional Neural Network	(CNN)
Long Short-Term Memory	(LSTM)
Global Vectors for Word Representation	(Glove)
Robustly Optimized BERT Approach	(RoBERTa)
Bidirectional Encoder Representations from Transformers	(BERT)
Natural Language Processing	(NLP)
Distilled BERT	(DistilBERT)
Multi-Task Learning	(MTL)
Emergency Response Support System	(ERSS)
Average Voting Ensemble DL model	(AVEDL)

List of Figures

Table of Contents

1.Introduction:

The use of social media in today's age is increasing day by day. Nowadays social media platform like twitter are used by the people more to pass some information quickly and to reach more number of people as the access to these kind of platforms is very easy in todays era. So, many people using these kind of fast reaching platforms in the disaster related situations to express their views, to give an update regarding particular situations and sometimes to provide help or to get some help. The number of users using tweeter and other social media platforms are increasing in the disaster like situations and because of that the user generated data is getting very vast. Because of this the researchers in the filed of data research and disaster domain. To understand the information and to categorize it is the most important part effective and on time disaster management. In this study, research is focused on disaster tweet identification using the modern ML and DL techniques , which focuses primarily on DistilBERT model. There are manmade disaster and naturally occurring disaster, and these can hit us unexpectedly and can cause major damages. So ,Timely and correct information regarding this very important and which helps to allocate the resources available and do the arrangements in the affected area or people which can lead to reduce the damage. Social media like tweeter has very much global reach and gives the real time information, so these kinds of platforms has now become valuable resource of data for disaster management. This study aims to contribute on this field by working on the creation of robust model for identification of the disaster related tweets. But there are challenges as well , because the social media has very large data and this data is getting increased day by day and to get identified with valuable and meaningful data is a big task. And to address all these challenges, in this study, the research focus is DL techniques and trying to use pretrained word embeddings to get the semantic meaning and context of the words. There are three algorithms are used as a part of methodology and every algorithm is contributing according to their strength for disaster tweet identification process. LSTM + CNN, Roberta Transformer Model, Fine Tuning DistilBERT are three algorithms used in the research study. By using the finetuning of the distilbert model and other algorithmic approaches study focuses to improve on the effectiveness of identification of disaster tweets with precision and so contributing to the more timely and effective disaster management.

1.1Background:

In this age which is majorly dominated by social media, the increasing use of social media and rapid sharing of information during crises like situation, made to think on the revamping of disaster management systems (Alam et al., 2021). Social media platforms are necessary for giving people some medium to share their concerns and information to get help or to provide the help . But the increasing the user generated data makes it more complex and it also more changes of getting the important information efficiently in the disaster situations. (Adel et al., 2022) However, finding pertinent disaster-related information among the enormous volume of data presents difficulties. A precise hold of language, context, and event is necessary for the task. Advanced ML methods have been used to address this, with a focus on DL models and pre-trained embeddings. The automatic extraction of significant features from text about disasters is made possible by the application of models like LSTM + CNN, Roberta, and tailored DistilBERT. This provides context for our study, which focuses to improve disaster response through the development of efficient algorithms for tweets identification.

1.2 Motivation:

The motivation for this research was the recent earthquake that happened in Turkey. From this there was a urgent need to develop accurate methods identify the disaster related information from the very large and complex data of social media effectively. When the disaster such as natural disaster or any other outbreaks or emergency situations, the effective and on time information with precision is very effective in disaster management and to disaster response. There is so much gap in the traditional methods of disaster management and this gap can be reduced and the with the help of social media platforms with quick and global reach. In this research, by working on training and fine tunning the models like LSTM + CNN, Roberta, and DistilBERT on disaster-related data, study utilize their capacity for understanding the context and pattern. The research is motivated to use advance DL tecchniques with pretrained models to overcome and address these situations. Which helps in the advancing the disaster management strategies by the use of social media platforms like tweeter.

1.3 Problem Statement:

Social media has now the most important part of human lives in this era. And it is seen that many people are using these social platform to seek help or to notify some information in the disastrous situations and it is seen quite useful in many situations previously. The accurate and well timed identification of tweets which indicates the disastrous situation is very crucial. But the challenge is to distinguish between disaster related tweet and non-reverent tweet keeping in mind that data and information shared on social platform is growing day by day. There are some techniques in Natural language Processing which are pretrained such as DistilBERT which Offer promising options for categorizing tweets automatically. Whereas the research needs to carried out with the application of such techniques with effective optimization of fine-tunning process for the DistlBERT for accurate identification of disastrous tweets.

1.4 Research Question:

1. How effective are the neural networks with transfer learning along with fine-tuning of DistilBERT's in identifying the disasters tweets and classifying the important information.?

1.5 Research Objectives:

To compare the performance of different models, including CNN + LSTM with Glove Embedding, RoBERTa Transformer, and FineTune DistilBERT Transformer, in accurately classifying disaster-related and non-disaster-related tweets.

1.6 Research Gap:

Despite the comprehensive study of tweet classification models for disaster-related content identification, several research gaps remain to be addressed. The limited recall observed in the CNN + LSTM model suggests potential challenges in capturing very complicated nuances of disaster-related content, indicating the need for improved architectural designs capable of effectively knowledgeable complex patterns. Additionally, while the RoBERTa and FineTune DistilBERT models exhibited strong performance, the study highlights the potential for fine-tuning techniques to further enhance quality and purpose across diverse disaster contexts, bridging the gap between generalized models and specific crisis scenarios. Furthermore, the study recognizes the potential of ensemble methods, indicating a research gap in leveraging the combined power of multiple models to achieve a more robust and accurate disaster identification algorithm. Addressing these gaps would contribute to increasing the effectiveness and versatility of disaster information tools, ultimately empowering organizations with better insights for timely and accurate decision-making during disaster events.

2. Literature Review:

2.1 Introduction:

The literature review is a key component of the entire report that provides an evaluation of the many works of literature and research articles on the same subject. The literature study consults a number of academic works that address the importance of classification and text analysis in the disaster management domain. The literature review also examines several methods, including machine learning capabilities, combining technological approaches, and deep learning. Deep learning and machine learning techniques are widely used to handle many aspects in disaster management field. The literature review section explains how using trained frameworks, deep learning and machinelearning techniques can help identify the disaster related tweets. The section on literacy review also sheds light on the important advantages of deep learning techniques in NLP.

2.2 Research overview on CNN and Transfer learnings:

(Wani et al., 2021) suggested Assessing DL Techniques for Detecting Covid-19 Misinformation. In addressing the surge of fake news amplified by social media, we delve into automated fake news detection techniques. We evaluate Algorithms for supervised text classification—LSTM, CNN, and BERT—on the Contraint@AAAI 2021 Covid-19 Dataset for detecting fake news. Unsupervised learning's role is explored by pre-training language models and dispersed word representations. Remarkably, our approach achieves a peak accuracy of 98.41%. This study significantly contributes to curbing misinformation, critical especially during the Covid-19 pandemic, by harnessing stateof the art methods for precise fake news identification.

(Zhou et al., 2022) proposed Utilizing BERT for Extracting Emergency Aid Appeals from social media in Disaster Relief. This study addresses the challenge of identifying social media rescue request messages during disaster response and rescue operations. Leveraging NLP algorithms, particularly BERT-based models, ten VictimFinder models were developed and compared. Training and testing on 3,191 manually labelled disaster-related tweets from Hurricane Harvey were conducted. Evaluation encompassed Model stability, classification accuracy, and computation cost. BERT-based models exhibited enhanced accuracy in categorizing rescue-related tweets, with a customized BERT-CNN model achieving a 0.919 F1 score, outperforming the baseline by 10.6%. These models hold the potential to optimize social networking sites utilization for future catastrophe rescue efforts scenarios.

(Adel et al., 2022) presented Enhancing Crisis Event Detection via DistilBERT and the Hunger Games Search Algorithm. This study introduces an innovative event detection approach that synergizes DistilBERT and a novel metaheuristic method called the Hunger Games Search. DistilBERT extracts text dataset features, complemented by a binary HGS for feature selection, eliminating irrelevant attributes. Experimental validation employs real-world datasets, encompassing comparisons with established feature selection algorithms and SOTA models. Results highlight the proposed model's superior performance, showcasing enhanced performance measures and reaffirming its efficacy in event detection.

(Raj et al., 2022) proposes Evaluating Flood Severity through DistilBERT and Named Entity Recognition (NER). This study presents a novel approach to flood severity assessment in disaster management. Leveraging DL and NLP, the proposed methodology employs DistilBERT for accurate social media post classification and spatiotemporal modelling to determine flood-affected areas' precise location and time. Traditional models are outperformed by DistilBERT, as keyword-based approaches are limited. To address geotag challenges, Named Entity Recognition (NER) identifies locations. Both models are trained on Chennai flood-related social media data. Achieving 99% accuracy in text classification (DistilBERT) and 89% in location identification (NER), results highlight the approach's potential efficacy in disaster management, offering valuable insights for timely recovery responses.

(Wang et al., 2021) illustrated Multi-task Learning with Transformers for Categorizing Disaster Tweets. This study presents a transformer-based MTL technique for categorizing and prioritizing social media statements on a crisis. The proposed approach is evaluated using various metrics within the TREC Incident Streams (IS) track, designed for disaster tweet classification and prioritization. Results demonstrate competitive performance compared to other runs, with an ensemble approach leveraging diverse transformer encoders yielding state-of-the-art results. The provided publicly available code ensures reproducibility and offers a community baseline for future research in this domain.

2.3 Research overview on Deep Learning and Transfer learnings:

(Nimmi et al., 2022) proposed Ensemble Pre-trained Model for Emotion Identification Amid The Disaster Response Support Structure Dataset was used to generate COVID-19.. Amid the global economic turmoil caused by COVID-19 precautions, lockdowns, and quarantine measures, an urgent need for systematic psychiatric intervention arises due to heightened threats to mental well-being. This study analyzes both ERSS call content and Twitter patterns during the pandemic. The proposed AVEDL Model Average Voting Ensemble DL model classifies emotions in COVID-19-related ERSS calls and tweets. Utilizing BERT, DistilBERT, and RoBERTa, the AVEDL Model achieves Macro-average F1-score (85.20%) and exceptional accuracy (86.46%) surpassing conventional DL and ML models. This novel ensemble approach offers insights into COVID-19 emotion analysis, highlighting the urgency and potential for comprehensive mental healthcare.

(Koshy and Elago, 2023) presented the Classification of Multimodal Tweets in Emergency Response Systems via a Bidirectional Attention Transformer-Based Model. This research aims to leverage social networking sites for awareness of one's surroundings post-crisis using a DL framework. The proposed method employs text and image inputs generated by users’ tweets, focusing on improved multimodal fusion. It integrates for text a fine-tuned Roberta model, Vision Transformer for images, biLSTM, and attention mechanisms. A multiplicative fusion strategy is introduced. Across seven disaster-related datasets, including wildfire, hurricane, earthquake, and flood, extensive experiments exhibit and demonstrate greater performance over SOTA methods, achieving 94-98% accuracy. The study highlights the value of enhancing DL classifiers by identifying interactions between multiple modalities, thereby advancing situational awareness capabilities.

(Taneja and Vashishtha, 2022) suggested Contrasting Transfer Learning with Conventional ML for Text Classification. This paper addresses the challenge of text analysis in the era of diverse data types using NLP. It compares transfer learning, exemplified by BERT and DistilBERT, with traditional ML based on TF-IDF for text classification. The models are fine-tuned and evaluated on varied datasets. DistilBERT demonstrates superiority over BERT and traditional methods, offering efficient performance with fewer parameters. The study advocates for DistilBERT as a preferred transfer learning model for text classification due to its accelerated training and comparable efficacy, thereby advancing NLP techniques in the age of big data.

(Basu et al., 2021) demonstrated Evaluating RoBERTa's Robustness: Benchmarking State-of-the-Art

Transformer Networks for Detecting Sexual Harassment on Twitter. This study delves into the detection of sexual/physical harassment in cyberspace using transformer architecture, particularly RoBERTa. It aims to comprehend and contextualize offensive tweets, enabling effective identification and penalization of predatory behaviour without relying solely on user reports. The research involves a thorough comparison of various SOTA transformer architectures. The evaluation employs metrics to gauge accuracy and efficiency, with results contributing to enhancing online platform safety. The study underscores RoBERTa's potential in deciphering and addressing harassment, advocating for a proactive approach against inappropriate online conduct.

(Alam et al., 2021) evaluated CrisisBench: A Benchmark for Assessing Humanitarian Information Processing in Crisis-Related Social Media Datasets. This study focuses on the examination of social media feeds in real-time to enable swift humane responses during disasters. Addressing a gap in crisis informatics research, it amalgamates diverse crisis-related datasets, yielding a comprehensive dataset of 166.1k informative and 141.5k humanitarian classified tweets. Multiclass and Binary classification outcomes are presented using FastText, CNN and transformer-based models. Results showcase model effectiveness, contributing to enhanced crisis informatics techniques. The consolidated dataset and scripts are shared for further research, promoting standardized evaluation and progress in this critical domain.

(Patel, 2023) proposed Augmenting Health Tweet Categorization: A Comparative Study of Comprehensive Analysis using Transformer-Based Models. This thesis tackles the challenge of comprehensive health-related tweet classification by proposing a diverse dataset that combines existing health-related data, keyword-based collection, and manual annotations. Transformer-based models like BERT, BERTweet, RoBERTa, and DistilBERT are employed to address the contextual ambiguity of health-related keywords. The evaluation demonstrates BERT, DistilBERT, and RoBERTa yielding F1score of 0.870, 0.882 and 0.872, respectively. The highest F1-score (0.900) is attained by incorporating a BiLSTM layer into the BERTweet model fine-tuned on the proposed dataset and RHMD. Ablation analysis underlines the BiLSTM layer and RHMD's positive impact on the BERTweet model's classification performance for health-related tweets.

2.4 Research on other algorithms in the disaster management:

(Fazlourrahman and Aparma, 2022) presented Detecting COVID19 Fake News with CoFFiTT: FineTuned Transfer Learning Methods. Amid the COVID19 pandemic, online platforms have seen a surge in popularity, including both helpful and harmful content. This paper introduces two transfer learning models, Ext-ULMFiT and FiT-BERT, finetuned on the CORD-19 dataset to study COVID19 fake news. Evaluated on the CONSTRAINT’21 workshop's COVID-19 detection of fake news dataset, FiT-BERT achieved a weighted average F1 score of 0.99, outperforming Ext-ULMFiT in forecasting false news while Ext-ULMFiT excelled in real news prediction. Notably, the proposed models' performance closely approached the best team in the shared task, highlighting their effectiveness in addressing the challenge of detecting and mitigating fake news during the pandemic.

(Do and Bang, 2021) proposed an experimental analysis grounded in binary classification, aimed at discerning the authenticity of text within the realm of social network data. The proposed approach involves addressing the challenge posed by potentially distorted or unreliable information in social text data utilized for data analysis research and network model learning in domains such as defence, public safety, and security. This study introduces a binary classification model that leverages a pre-trained language model to differentiate between authentic and potentially misleading text data. Through an empirical experiment using tweet data, the suggested approach demonstrates promising results, achieving an accuracy rate of 81% and a Matthews correlation coefficient value of 0.61.

Table 2.1 Comparison of different aspects on Literature Review

Author	**Algorithms Used**	**Purpose**	**Limitations**	**Research Gap**
Abu Nowshed Chy,Umme Aymun Siddiqua and Masaki Aono	MKC-LSTMVs-ATT	To suggest a neural network-based method for quickly identifying crisis-related, tweets on Twitter	Limitation of the Dataset,Limited Feature,Scalability,External elements	target-specific tweet stance detection
S Sharma, S Basu ,N Kant, N Kumar,P Dalela	BERT,Machine Learning Classifiers	Analysis of various techniques on Cyclone tweets	Focused Analysis: The analysis only considers tweets that mention Cyclones AMPHAN and NISARGA	The generalizability of the findings to other crisis occurrences may be constrained by this focused approach.
Lucas Satria Aji Dharma,Edi Winarko	Convolutional Neural Network (CNN),BERT	A different embedding technique called Word2Vec is compared to how well CNN performs with BERT embeddings	Limited comparison of neural network architectures.	specifically addresses classifying natural disaster tweets in the Indonesian language, . Limited exploration of fine-tuning BERT
A Hembree, A Beggs, T Marshall, N Ceesay	Bidirectional Encoder Representations from Transformers (BERT),ELMO	Determine the genuineness of tweets about emergencies.	Linguistic ambiguity: It is difficult to completely eliminate ambiguity, and the accuracy of the model may still be affected by linguistic quirks and context.	the robustness of the proposed model in handling various types of disasters, needs further

				investigation
P Purushotham, D Divya Priya, Dr Ajmeera Kiran	KNN, Support Vector Machine, Decision Tree	creation of an initiative to help hospitals and individuals prepare for disasters.	The drawbacks of the prediction models	knowing how well-suited and flexible the solution is to various crisis scenarios.

3. Research Methodology:

The Research Methodology chapter serves as the navigational compass guiding the empirical journey undertaken in this study. The dataset, a corpus of over 11,000 tweets, revolves around disaster-related keywords such as "crash," "quarantine," and "bushfires," complete with associated locations and keywords. This dataset's architecture inherits from the Disasters on the social media paradigm. A cornerstone of this methodology is the meticulous manual categorization of tweets into two distinct classes. One class encapsulates tweets referencing actual disaster events, while the other encompasses content devoid of disaster context, such as jokes or movie reviews. This binary classification schema, denoted as '1' for disaster-related tweets and '0' for non-disastrous content, fuels the subsequent analytical stages. The research methodology harnesses a suite of robust tools and libraries, epitomizing technological prowess. Python, a versatile programming language, forms the bedrock for orchestrating the research endeavour. Frameworks like TensorFlow and Keras empower the implementation of complex machine learning architectures, while sci-kit-learn endows the study with a robust toolkit for data manipulation and model evaluation. Leveraging transformers, NLTK, pandas, and numpy imparts an elegant synergy between data preprocessing and analysis. The operative hub for the classification task is Google Colab, a collaborative cloud environment that crystallizes the convergence of computational horsepower and collaborative potential.

3.1 Methodology:

Transformer-based models were utilized, focusing on accurate binary classification through data preprocessing, model training, and evaluation. The process included importing libraries, data loading, cleaning, visualization, preprocessing, splitting, model building, and evaluation. The methodology contains several modules:

3.2 Data Collection:

The dataset encompasses over 11,000 tweets linked to disaster-related keywords like "crash," "quarantine," and "bushfires," accompanied by respective locations and keywords. Derived from the Disasters on social media data structure, manual classification categorized tweets into two classes: those referencing actual disasters (1) and those unrelated (0), such as jokes or movie reviews.

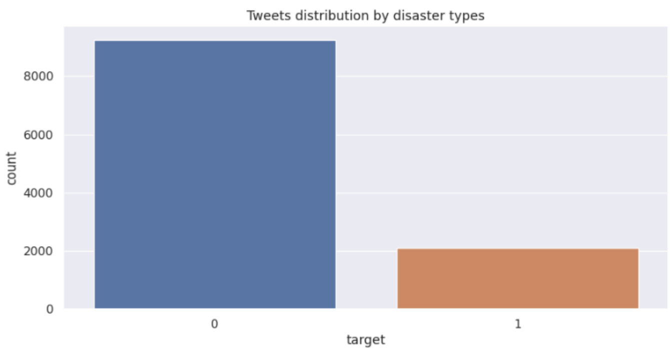
Employing tools like Python, TensorFlow, Keras, sci-kit-learn, transformers, NLTK, pandas, and numpy, this research harnessed a suite of libraries. Execution transpired within a Colab environment, utilizing various transformer-based models for tweet classification. The holistic strategy encompassed data preprocessing, model training, validation, and testing, culminating in accurate binary categorization via ML and DL fusion.

3.3 Data Preprocessing:

In the data preprocessing phase, the initial step involves loading the tweets CSV file into a Pandas data frame, providing a structured format for further manipulation. NLTK, a natural language processing library, is then employed to execute a series of essential preprocessing steps. These include removing HTML entities, and effectively cleansing the text of any unwanted artifacts. Regular expressions are utilized to substitute mentions and URLs with whitespace, effectively sanitizing the content. To ensure consistency, all text is converted to lowercase. To improve the signal-to-noise ratio, stopwords are removed, and punctuation marks are stripped. This meticulous data preparation lays the foundation for subsequent tokenization, word embedding, and model training stages.

3.4 Data Visualization:

The Data Visualization section employs graphical exploration as a compass to navigate the dataset's terrain. Through plots and graphs, a visual narrative unfolds, shedding light on data intricacies. These visualizations distill complex patterns, facilitating a deeper understanding of disaster-related tweets. The distribution of target classes is revealed, depicting the prevalence of disaster events versus non-disastrous content. Additionally, the most common keywords surface, offering a glimpse into prevalent disaster themes. Wordclouds further unravel textual nuances within each class. Furthermore, a bar chart showcases text length distribution, elucidating the tweet lengths' landscape. Data Visualization becomes a bridge, connecting raw data to meaningful insights and fortifying subsequent analytical decisions.



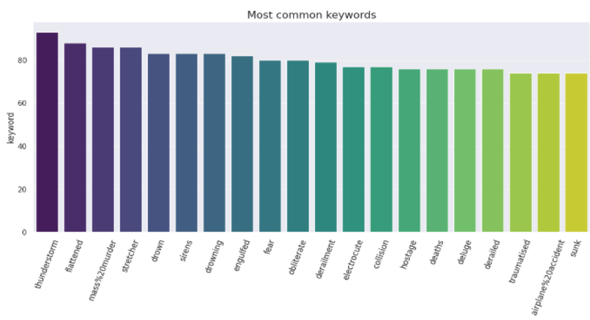


Figure 3.1: Distribution of Target Classes in the Dataset

Figure 3.2: Distribution of Most Common Keywords in the Dataset

Figure 3.1 provides an illustration of the distribution of target classes within the dataset. The "target" refers to the categories assigned to each tweet, distinguishing between disaster-related events (class 1) and non-disastrous content (class 0). The "count" on the y-axis represents the frequency or number of tweets falling under each category. It offers insight into the balance or potential imbalance between disaster-related and non-disastrous tweets, contributing to a comprehensive overview of the dataset's target distribution.

Figure 3.2 illustrates the distribution of the most common keywords present in the dataset. The graph displays the frequency of occurrence for specific keywords, with "thunderstorm," "flattened," "stretcher," and "drown" being highlighted. The visualization gives information into the prevalence of these keywords within the dataset, offering an understanding of the types of disaster-related terms that appear frequently in the tweets.

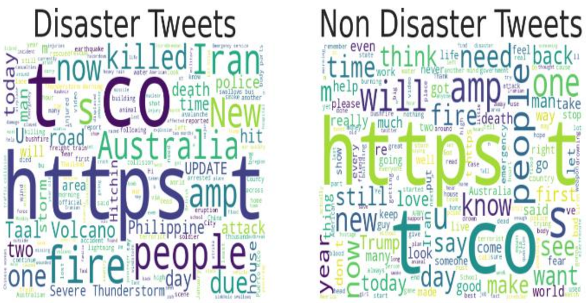


Figure 3.3: Wordcloud:- each class text plot (0 and 1)

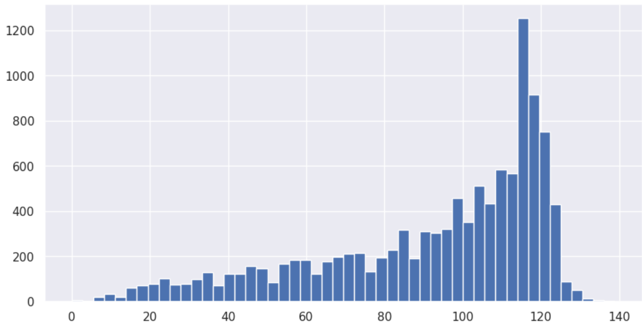


Figure 3.4: Bar chart:- Text length count plot

Figure 3.3 displays word clouds for disaster (class 1) and non-disaster (class 0) tweets. These visualizations depict the most frequent words in each class. Disaster tweets show words related to real disasters, while non-disaster tweets highlight common words from various topics. It helps compare language patterns between the two classes.

Figure 3.4 illustrates a bar chart depicting the distribution of text lengths in the dataset. The xaxis represents the range of text lengths, while the yaxis indicates the count of tweets falling within each length range.

3.5 List of Models:

The study employs the following models for text classification:

1. **LSTM + CNN**: A combination of CNN and LSTM layers. The Maxpool layer is replaced by a recurrent layer for discovering long-term dependencies. Data preprocessing includes punctuation removal, stopwords, and stemming. Tokenization and word embedding are

performed. The convolutional operation is applied to the embedded input, extracting critical features for categorization.

2. **Roberta**: A variant of the BERT model with improved training on a large dataset (160GB). Roberta uses dynamic masking during training and outperforms BERT in various NLP tasks, making it a popular choice.
3. **DistilBERT**: A distilled version of BERT, maintaining 97% of language understanding while being 60% faster and lighter. It uses a triple loss combining language modelling, distillation, and cosine-distance losses. Unfreezing DistilBERT's embedding layer enhances performance.

Each model contributes to the study's objective of text classification using different architectures and techniques, enhancing accuracy and efficiency.

3.6 Model Evaluation:

Model Evaluation involves several techniques:

1. **Classification Report**: Provides precision, recall, F1score, and support for each class. It helps understand the model's performance per class.
2. **Sensitivity and Specificity**: Sensitivity (TPR) measures the ability of model to correctly recognize the positive instances. Specificity (TNR) indicates the ability of the model to correctly identify negative instances.

These techniques collectively offer insights of the model performance, its strengths, and fields for improvement.

4 .Design Specification:

The Design Specification chapter serves as a comprehensive guide to the meticulous selection and strategic configuration of a diverse array of models intended for the intricate training process. The models encompass an array of advanced techniques, prominently featuring deep learning architectures complemented by the integration of pre-trained word embeddings. This fusion aims to extract and harness intricate linguistic nuances effectively. The models picked to do this important job use a wide range of approaches, each of which was carefully made to fit the specific needs of the job. At the forefront of this ensemble is a novel LSTM-CNN hybrid model, which capitalizes on the power of both CNNs the Convolutional NeuralNetworks and LSTM the Long Short-TermMemory networks. The LSTM component enables the model to grasp sequential dependencies, while the CNN component adeptly captures local patterns. The integration of pre-trained Global-Vectors (GloVe) for Word Representation embeddings infuses this model with a comprehensive understanding of word contexts, enhancing its ability to interpret text intricacies. Adding to the arsenal of models is the formidable

Roberta transformer model. This state of the art architecture showcases the prowess of transformer-based methodologies, which have revolutionized natural language processing tasks. Its attention mechanisms allow it to contextualize information extensively, capturing intricate relationships within the text. The utilization of a dedicated Tokenizer for Roberta ensures seamless alignment with the model's preprocessing requirements, enabling a streamlined data flow. A pivotal component of this repertoire is the Fine-Tuning DistilBERT transformer model, meticulously chosen for its unparalleled accuracy. The Fine-Tuning approach capitalizes on pre-trained DistilBERT's knowledge, adapting it to the specific nuances of the task through focused training. This dynamic model showcases exceptional proficiency in discerning insincere questions, contributing significantly to the desired binary categorization. The chapter delves into the deliberate rationale underlying the selection of these models, elucidating the distinct advantages they bring to the classification conundrum. It explores the integration of pre-trained embeddings, elucidating how they imbue the models with a profound understanding of linguistic semantics. Moreover, the chapter intricately examines the design considerations that underpin the seamless implementation of these models within the framework. The interplay of model architecture, tokenization, and preprocessing intricacies are harmoniously orchestrated to realize optimal performance.

5. Implementation:

The Implementation chapter is a pivotal stage in translating the conceptual framework into tangible reality, where a rich repertoire of tools and libraries converge to materialize the proposed models. The chapter draws upon an arsenal of sophisticated technologies, with Python, TensorFlow, Keras, and scikit-learn forming the cornerstone of this implementation journey. These versatile frameworks provide the essential building blocks for constructing, fine-tuning, and evaluating intricate ML architectures. A standout feature of the implementation process is the astute utilization of Google Colab. This cloud-based environment seamlessly merges computational power and collaborative potential, ensuring the efficient orchestration of complex tasks. The collaborative nature of Colab is leveraged to its fullest, enabling multiple contributors to collaborate seamlessly, expedite model development, and collectively enhance research outcomes. Intricate data preprocessing is a linchpin of the implementation process, and the chapter elucidates the seamless integration of transformer, NLTK, pandas, and numpy for this crucial endeavour. The transformer library facilitates tokenization, laying the foundation for the subsequent steps. NLTK engages in subtle yet impactful text transformations, transcending HTML entities, mentions, URLs, and punctuations to pave the way for clean and coherent text data. Pandas and numpy harmoniously collaborate to mould and structure the data into a format conducive to model ingestion. Model configuration and training form the heart of the implementation. Each chosen model, whether it's the innovative LSTM-CNN hybrid, the formidable Roberta transformer, or the accuracy-optimized Fine-Tuning DistilBERT, is meticulously tailored and calibrated within the TensorFlow and Keras ecosystems. Pre-trained embeddings are harnessed to infuse models with profound linguistic insights, transforming text into a comprehensible numeric format. The chapter culminates in a meticulous evaluation of model performance. Key metrics such as precision , accuracy, recall, and F1score are employed to measure the effectiveness of the trained models in categorizing insincere questions. The convergence of these metrics provides a comprehensive assessment of each model's efficacy, facilitating informed decisions and insights.

5.1 CNN + LSTM with Glove Embedding Model:

The CNN-LSTM architecture, depicted in Figure 5.1, utilizes a combination of convolutional and recurrent layers for text classification. To capture long-term dependencies, the Maxpool layer is substituted with a recurrent layer, defining the CNN-LSTM design. The approach employs movie critiques in English as input, where punctuation is separated by whitespace. Data preprocessing involves techniques such as punctuation removal, stop-word elimination, integer bracketing, and stemming using regular expressions (RE). Tokenization, translating words to numerical values, follows dataset cleaning. Embedding translates words to a feature vector space. A convolutional layer processes word embeddings, extracting meaningful features from the text for categorization. The resulting features are conveyed through multiple convolutional layers.

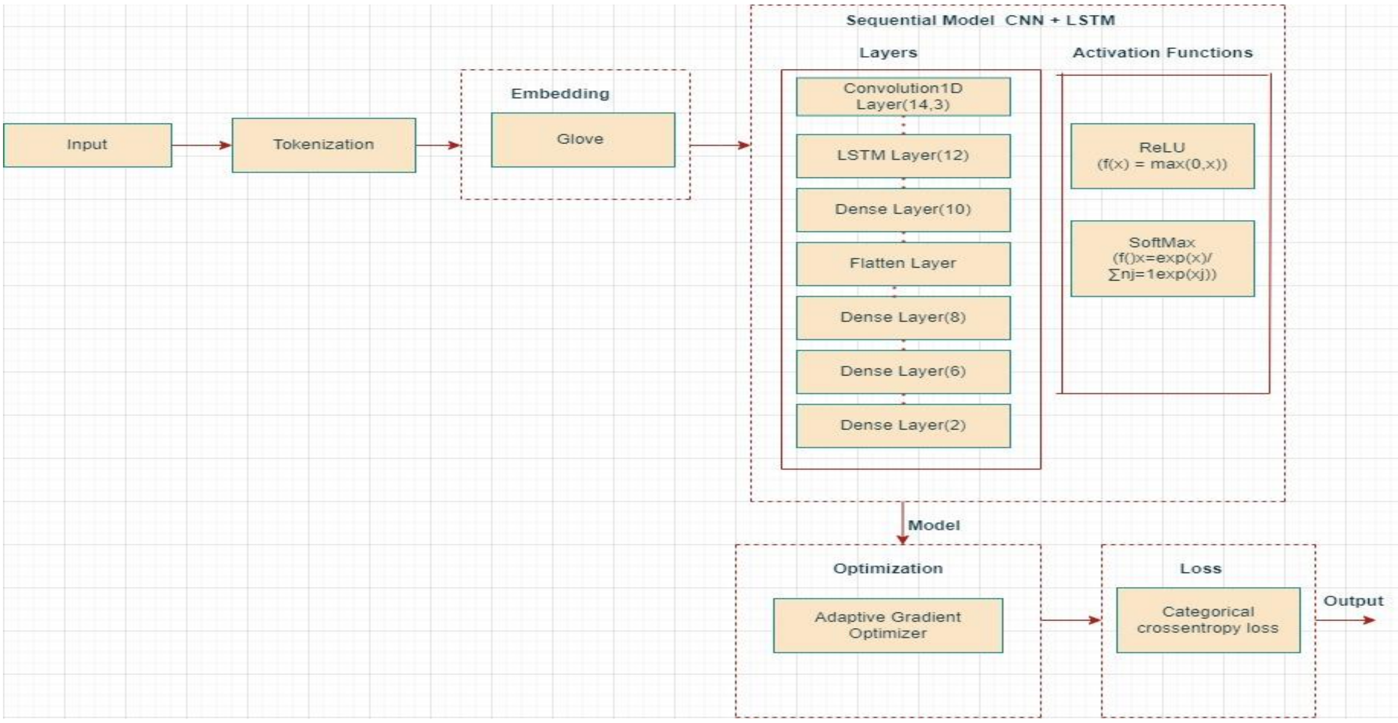


Figure 5.1 Framework architecture for CNN-LSTM

5.2 Roberta Transformer Model:

RoBERTa is an advanced variant of the BERT model, initially developed by AI researchers working at Facebook. Similar to BERT, RoBERTa utilizes a transformer-based architecture and selfattention mechanisms to process input sequences and generate contextually rich word representations within sentences. A notable distinction between RoBERTa and BERT lies in their training methodologies. RoBERTa was trained on a significantly larger dataset, a substantial 160GB of text, more than ten times the size of BERT's training data. This extensive dataset, coupled with an enhanced training procedure and dynamic masking technique, contributes to RoBERTa's ability to develop robust and widely applicable word representations. RoBERTa excels in various NLP tasks, surpassing BERT and other state of theart models in tasks such as language translation, text classification, and question answering. Its success has led to its integration as a foundational model in numerous influential NLP applications and research endeavors across industry and academia.

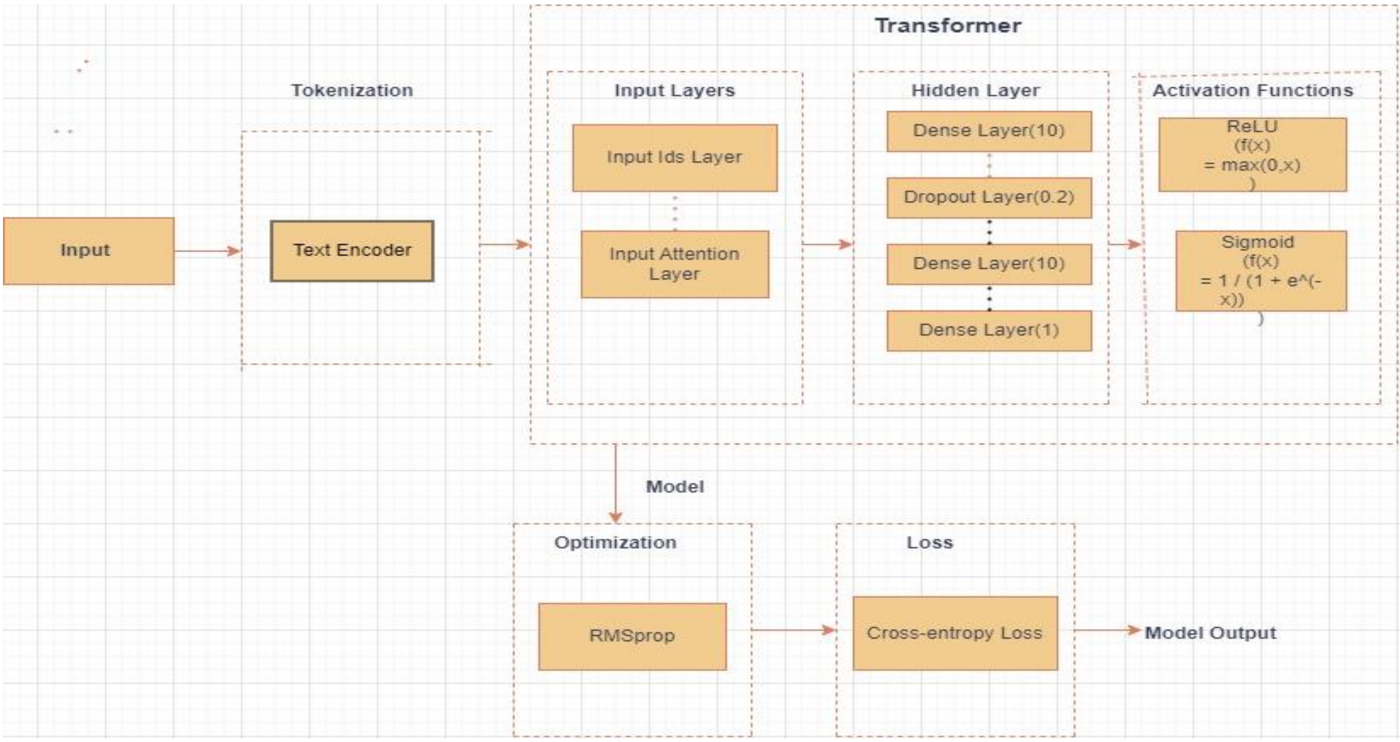


Figure 5.2 Roberta Transformer architecture

5.3 FineTune Distilbert Transformer Model :

DistilBERT, a distilled version of the BERT model, presents several key features for efficient NLP. Through knowledge distillation, a 40% reduction in the size of the BERT model is achieved during pre-training, preserving 97% of language understanding while improving processing speed by 60%. DistilBERT integrates a triple loss mechanism combining language modeling, distillation, and cosinedistance losses, leveraging the inductive biases acquired by larger models during pre-training.

The resulting model is compact, faster, and well-suited for on-device applications, offering cost-effective pre-training. Fine-tuning DistilBERT's embedding layer and weights at a lower learning rate enhances performance. Its architecture mirrors BERT's, omitting token-type embeddings and pooler while halving the number of layers. Notably, the student initialization method involves initializing the student from the teacher by extracting every other layer. This approach focuses on optimizing the student's efficiency by finding suitable initialization for convergence. This distilled variant demonstrates efficiency gains while maintaining competitive performance. Below is the algorithm, that explains the implementation done in this research. The input and output both are mentioned. The algorithm explains each and every step that is involved with the process implementation of DistlBERT along with the fine-tunning**.**

Algorithm 1: Algorithmic approach for Identification of disaster tweets

Data: *X train*: Represents Training Dataset
X test: Represents Testing Dataset
T: Number of training epochs
Result: Model with fine-tuned weights

Initialization of Pretrained Distillbert Model;
Initialize $I = \{i_1, i_2, \ldots, i_N\}$;
hidden states $H = \{h_1, h_2, \ldots, h_N\}$ for each token in input;
model's hidden state at each layer $H^l = \{h^l_1, h^l_2, \ldots, h^l_N\}$;
Transformer Encoding;
Map input vectors to contextual embedding vectors;
Fine-Tuning and Feature Extraction;
Activation Functions:
ReLU: $f(x) = \max(0, x)$;
Sigmoid Function: $f(x) = \frac{1}{1+e^{-x}}$;
Define the Optimizer;
Adam optimizer is defined;
for *epoch* = 1 **to** T **do**
  Shuffle and split *X_train* into mini-batches of size *BATCH_SIZE*;
  for *each mini-batch B* **do**
    Extract tweets *X batch* and corresponding labels *y batch* from *B*;
    Convert tweets to numerical representations;
    Identify the projected probability's binary cross-entropy loss;
    Calculate the accuracy;
Return ←O/P

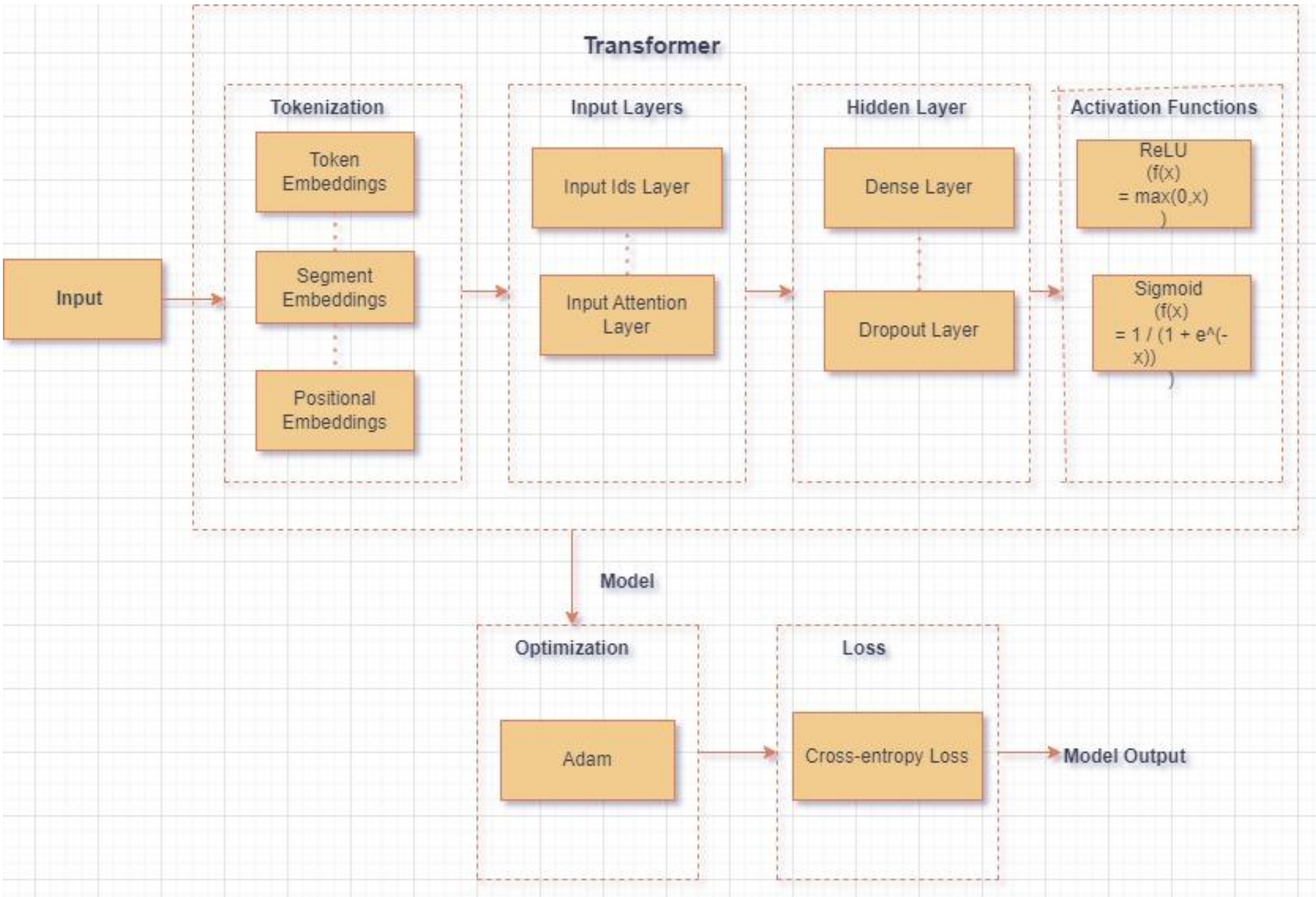


Figure 5.3 Framework architecture for FineTune Distilbert Transformer

6. Evaluation:

The conducted study focused on developing and evaluating three distinct models for tweet classification, aiming to differentiate between disaster and non-disaster-related tweets. The models were CNN + LSTM with Glove Embedding, RoBERTa Transformer, and FineTune DistilBERT Transformer. In the CNN + LSTM with Glove Embedding model, a hybrid architecture combining convolutional and recurrent layers was employed. This model achieved an accuracy of 83%, effectively identifying non-disastrous tweets with high precision (83%) but struggling with recall (0%) for disastrous tweets, resulting in an unbalanced performance. RoBERTa Transformer, a variant of BERT, demonstrated significant advancements due to its larger training dataset and improved training methodology. With an accuracy of 87%, RoBERTa exhibited commendable performance. While it exhibited superior sensitivity (36%) in recognizing disastrous tweets, its specificity (98%) was higher for non-disastrous ones. The standout performer, however, was the FineTune DistilBERT Transformer model, achieving an accuracy of 89%. It balanced sensitivity (60%) and specificity (95%) well, excelling in both correctly identifying disaster-related and non-disaster-related tweets. This performance was attributed to DistilBERT's compact yet powerful architecture, leveraging knowledge distillation and triple loss techniques. The study's results collectively highlighted the capability of transformerbased models in tweet classification. Notably, FineTune DistilBERT Transformer demonstrated the highest accuracy and a balanced sensitivity-specificity trade-off, making it the optimal choice for discerning disaster-related tweets. The study focuses on contribution to the developing the effective and precise methods to help in crises .Further research could explore ensemble methods or model fine-tuning for enhanced performance across various disaster scenarios.

6.1 CNN + LSTM with Glove Embedding Model :

The CNN + LSTM model with Glove Embedding consists of an embedding layer, Conv1D layer, LSTM layer, and several dense layers. The model has a total of approximately 7.3 million parameters, out of which 25,156 are trainable. During training, the model gets a 81.29% of accuracy on the training data and a validation accuracy of 80.38% after 5 epochs. The validation accuracy stabilizes around 80.38%, indicating a consistent performance level on the validation dataset.

6.1.1 Classification Report:

The classification report for the CNN + LSTM model with Glove Embedding reveals important results for disaster tweet classification. The model gets an overall accuracy of 83% on a total of 2274 instances. For class0 (nondisastrous), there is 83% of precision , indicating that 83% of predicted non-disastrous tweets were actually non-disastrous. However, the recall for class1 (disastrous) is 0%, indicating that the model failed to identify any disastrous tweets precisely. The F1score for class 0 is 90%, showing a balance between recall and precision, while the F1score for class 1 is 0%, indicating poor results. The macro-average F1score across both classes is 45%, and the weighted average F1score is 75%. These results highlight the model's strong ability to predict non-disastrous tweets, but significant shortcomings in detecting disastrous ones. Further refinement and evaluation are necessary to improve the working performance of the model in identifying disasterrelated content.

Table 6.1 Classification Matrix for CNN + LSTM with Glove Embedding Model

Algorithm	**Precision**	**Recall**	**f1-score**
CNN-LSTM	0.83	1.00	0.90
Accuracy	**83%**		

6.1.2 Sensitivity & Specificity:

The sensitivity and specificity analysis for the CNN + LSTM model with Glove Embedding reveals critical insights into its performance. The model demonstrated a sensitivity of 0% for class 0 (non-disastrous), indicating its inability to correctly identify instances of non-disastrous tweets. Conversely, for class 1 (disastrous), the model exhibited a sensitivity of 100%, indicating its capability to accurately detect disastrous tweets. However, the specificity for class 0 is 100%, indicating the model's proficiency in correctly classifying instances as non-disastrous. In contrast, the specificity for class1 is0%, suggesting the model's inability to accurately classify instances as disastrous. These findings underscore the model's challenges in achieving a balanced trade-off between sensitivity and specificity, necessitating further refinement to enhance its overall predictive performance.

Table 6.2 Sensitivity and Specificity Matrix for CNN-LSTM Algorithm

Algorithm	**Sensitivity**		**Specificity**	
	(Class 0)	**(Class 1)**	**(Class 0)**	**(Class 1)**
CNN-LSTM	**0%**	**100%**	**100%**	**0%**

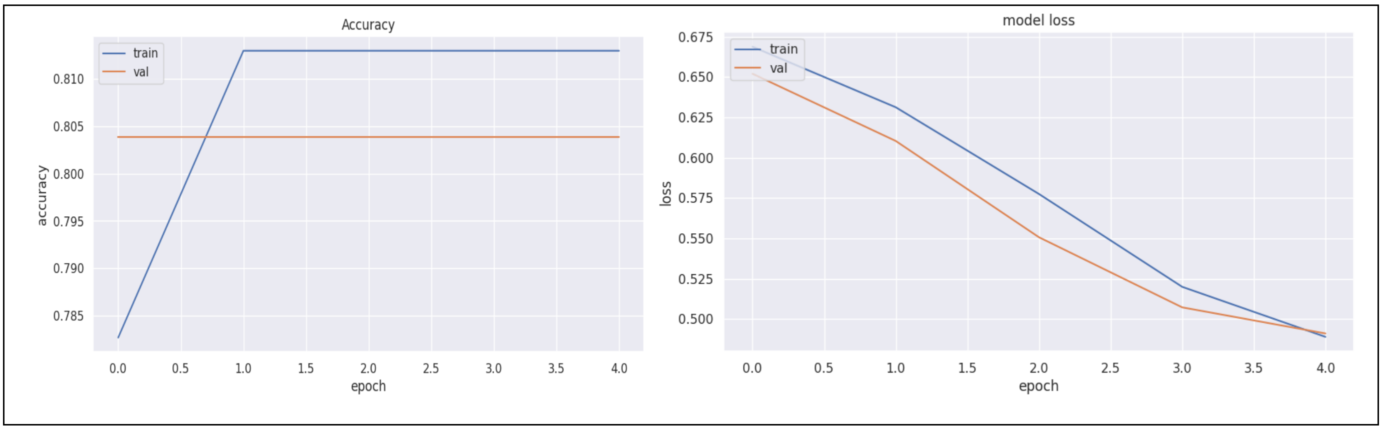


Figure 6.1: Accuracy and Loss Graph against Training and Validation data

6.2 Roberta Transformer Model:

The RoBERTa model architecture consists of multiple layers, including input layers for input_ids and attention_mask, the TF RoBERTa model layer with pooling and cross-attention mechanisms, and a lambda layer, followed by dense layers for classification. The total trainable parameters are 229,889. During training, the Roberta model underwent five epochs. The accuracy gradually increased from 81.24% to 87.20% on the data of training set. On the validation data set, accuracy improved from 80.55% to 87.25%. The RoBERTa Transformer achieved a validation accuracy of approximately 87.25%, indicating its ability to effectively classify data based on disaster-related keywords.

6.2.1 Classification Report:

The classification report for the RoBERTa Transformer model reveals its performance metrics. The model got high precision (0.88) and recall (0.98) for class 0 (non-disastrous), results in a robust F1score of 0.93. However, for class1 (disastrous), the precision is 0.81, indicating a notable ability to correctly classify true positives, but the recall is lower at 0.36, reflecting some difficulty in capturing all instances of this class. Consequently, the F1score for class 1 is 0.50. The overall accuracy of the model is 0.87, suggesting its effectiveness in correctly classifying instances across both classes. The macro-average F1score is 0.71, highlighting the balance between recall and precision for the two classes. The weighted average F1score is 0.85, indicating the model's ability to generalize its performance across the dataset. These results underscore the model's proficiency in non-disastrous class classification, while also pointing out potential areas for improvement in the classification of disastrous instances.

Table 6.3 Classification Matrix for Roberta Transformer Model

Algorithm	Precision		Recall		f1-score	
Roberta Transformer	**(Class 0)**	**(Class 1)**	**(Class 0)**	**(Class 1)**	**(Class 0)**	**(Class 1)**
	0.88	0.80	0.98	0.35	0.93	0.48
Accuracy	**87%**					

6.2.2 Sensitivity & Specificity:

The sensitivity and specificity analysis for the RoBERTa Transformer model provides key insights into its performance. For class 0 (non-disastrous), the model demonstrates a sensitivity of 0.361, indicating its ability to correctly identify a portion of actual disastrous instances. Moreover, it exhibits high specificity at 0.982, signifying its proficiency in accurately identifying true non-disastrous instances. In contrast, for class 1 (disastrous), the model excels in sensitivity with a value of 0.982, indicating its strong capability to detect actual disastrous cases. However, its specificity is lower at 0.361, suggesting some challenges in precisely identifying true non-disastrous instances within this class. These results underscore the model's balanced performance between sensitivity and specificity, while also highlighting areas where its performance may vary between the two classes.

Table 6.4 Sensitivity and Specificity Matrix for Roberta Transformer Algorithm

Algorithm	**Sensitivity**		**Specificity**	
	(Class 0)	**(Class 1)**	**(Class 0)**	**(Class 1)**
Roberta Transformer	34.5%	98.1%	98.1%	34.5%

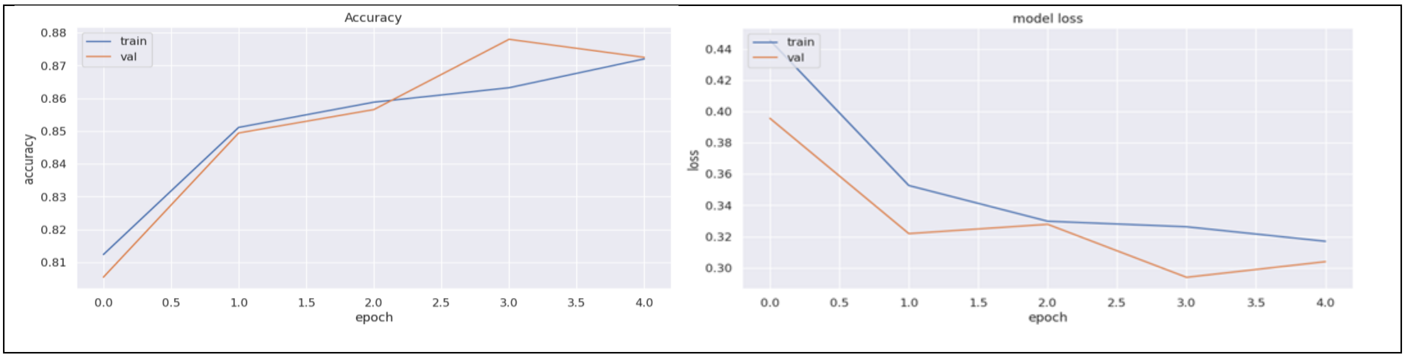


Figure 6.2: Accuracy and Loss Graph against Training and Validation data

6.3 FineTune Distilbert Transformer Model :

The fine-tuned DistilBERT transformer model demonstrates strong potential for text classification. With a concise architecture, the model maintains effectiveness while enhancing efficiency. The model summary showcases its structure, consisting of input and attention mask layers, a DistilBERT base model, dense layers, and dropout layers. The base model leverages 66.3 million non-trainable parameters to generate contextualized representations. During training, the model's performance steadily improves over epochs. The validation accuracy attains a notable level of approximately 89.51%, signifying the model's capability in accurately classifying text data.

6.3.1 Classification Report:

The classification report reveals important performance metrics of the FineTune DistilBERT Transformer model. The model achieves a commendable accuracy of approximately 89% on the evaluation dataset. It demonstrates a higher precision of 0.92 for the non-disastrous class (0), indicating the ability to correctly classify non-disastrous tweets. Additionally, model gets a recall of 0.95 for the same class, indicating its effectiveness in identifying actual non-disastrous instances. However, for the disastrous class1, the model's precision is 0.71, suggesting its capability to accurately classify disastrous tweets, though with some false positives. The recall for the disastrous class is 0.60, signifying its ability to capture a significant portion of actual disastrous instances. The F1score, which balances precision and recall, ranges from 0.65 to 0.93, indicating a good trade-off between these metrics. Overall, the model showcases a solid performance, particularly in distinguishing non-disastrous tweets, while also providing reasonable results for classifying disastrous tweets.

Table 6.5 Classification Matrix for FineTune Distilbert Transformer Model

Algorithm	**Precision**		**Recall**		**f1-score**	
FineTune Distilbert Transformer	**(Class 0)**	**(Class 1)**	**(Class 0)**	**(Class 1)**	**(Class 0)**	**(Class 1)**
	0.91	0.73	0.95	0.57	0.93	0.64
Accuracy	**89%**					

6.3.2 Sensitivity & Specificity:

The Sensitivity and Specificity analysis highlights crucial outcomes of the FineTune DistilBERT Transformer model. For the non-disastrous class (0), the model demonstrates a sensitivity of approximately 0.60, signifying its capability to correctly identify a considerable portion of actual non-disastrous instances. Its specificity, measured at around 0.95, indicates its proficiency in accurately identifying non-disastrous cases. For the disastrous class (1), the model exhibits a sensitivity of approximately 0.95, indicating its strong ability to capture a significant proportion of true disastrous instances. Its specificity for this class is approximately 0.60, suggesting its capacity to correctly classify actual disastrous cases. These results collectively reflect the model's effectiveness in distinguishing between disastrous and non-disastrous tweets. Notably, among the evaluated models, FineTune DistilBERT Transformer stands out as the best performer, achieving the highest accuracy of 89% and showcasing a balanced trade-off between sensitivity and specificity.

Table 6.6 Sensitivity and Specificity Matrix for FineTune Distilbert Transformer Model

Algorithm	**Sensitivity**		**Specificity**	
	(Class 0)	**(Class 1)**	**(Class 0)**	**(Class 1)**
FineTune Distilbert Transformer	56.8%	95.4%	95.4%	56.8%

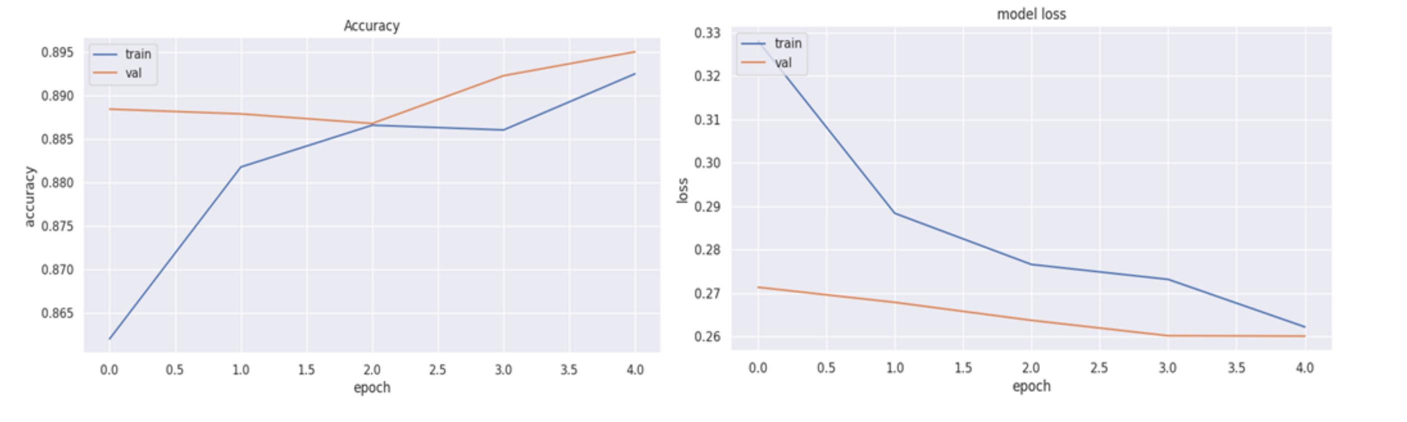


Figure 6.3: Accuracy and Loss Graph against Training and Validation data

6.4 Discussion:

In the Study, the research work presented to explore and compare the performance of distinct algorithms. And then the performance evaluation has been made on the basis of number of experiments cycle performed and compared on the factor of Accuracy prediction primarily. Below is the table which shows the comparison based on different parameter like precision, f1 score along with the accuracy for the different modelling techniques and algorithms studied.

Table 6.7 Comparison for Distinct Modelling Techniques

Algorithm	**Precision**		**Recall**		**F1 Score**	
	Non-Disastrous	**Disastrous**	**Non-Disastrous**	**Disastrous**	**Non-Disastrous**	**Disastrous**
CNN+LSTM With Glove Embedding	0.83	0.00	1.00	0.00	0.99	0.00
Accuracy	**83%**					
Roberta Transformer	0.88	0.80	0.98	0.35	0.93	0.48
Accuracy	87%					
Fine-tuned DistilBERT Model	0.91	0.73	0.95	0.57	0.94	0.64
Accuracy	**89%**					

7. Conclusion and Future Work:

In this study, research addressed the critical task of disaster-related tweet classification, considering the surge in social media usage during crises. Study introduced and evaluated three distinct models: CNN + LSTM with Glove Embedding, RoBERTa Transformer, and FineTune DistilBERT Transformer. Our comprehensive evaluation encompassed accuracy, sensitivity, specificity, precision, and recall, aiming to strike a balance between identifying disaster-related content and minimizing false positives. The FineTune DistilBERT Transformer model stood out, showcasing an impressive 89% accuracy with a well-balanced sensitivity-specificity trade-off. This model, built upon the distilled BERT architecture, presentsAA a compact and efficient solution for accurate tweet classification. The novelty of this work lies in several aspects. Firstly, our investigation delves into the applicability of state ofthe art transformer models for tweet classification in disaster contexts, addressing the unique challenges of crisis informatics. Secondly, we contribute a comprehensive comparison of different models, offering insights into their strengths and limitations. Moreover, the utilization of FineTune DistilBERT Transformer, with its reduced complexity yet robust performance, adds a novel dimension to the field. Lastly, our work underscores the significance of accurate tweet classification for timely and effective crisis response, augmenting the existing body of knowledge in humanitarian applications of natural language processing. While our study gives promising results, there are opportunities for further exploration. Ensemble techniques could be employed to combine the strengths of multiple models and enhance overall performance. Domain-specific fine-tuning could be investigated to adapt models to specific disaster scenarios, improving their adaptability and generalization.Exploring ways to address the class imbalance and fine-tune model parameters for optimal performance across various disaster-related keywords could also be valuable. Finally, investigating the integration of real-time data streams and geolocation information could contribute to more effective crisis management systems.

In conclusion, this study sheds light on the potential of transformer-based models in disaster-related tweet classification, with the FineTune DistilBERT Transformer model emerging as a powerful tool. The findings contribute to the development of accurate and efficient disaster information, offering valuable insights for humanitarian organizations and policymakers in their efforts to use the power of social media for disaster response and management.

References:

Wani, A., Joshi, I., Khandve, S., Wagh, V. and Joshi, R., 2021. Evaluating deep learning approaches for covid19 fake news detection. In *Combating Online Hostile Posts in Regional Languages during Emergency Situation: First International Workshop, CONSTRAINT 2021, Collocated with AAAI 2021, Virtual Event, February 8, 2021, Revised Selected Papers 1* (pp. 153-163). Springer International Publishing.

Zhou, B., Zou, L., Mostafavi, A., Lin, B., Yang, M., Gharaibeh, N., Cai, H., Abedin, J. and Mandal, D., 2022. VictimFinder: Harvesting rescue requests in disaster response from social media with BERT. *Computers, Environment and Urban Systems*, *95*, p.101824.

Adel, H., Dahou, A., Mabrouk, A., Abd Elaziz, M., Kayed, M., El-Henawy, I.M., Alshathri, S. and Amin Ali, A., 2022. Improving crisis events detection using distilbert with hunger games search algorithm. *Mathematics*, *10*(3), p.447.

Gokul Raj, S.N., Chitra, P., Silesh, A.K. and Lingeshwaran, R., 2022, March. Flood Severity Assessment Using DistilBERT and NER. In *International Conference on Machine Intelligence and Signal Processing* (pp. 391-402). Singapore: Springer Nature Singapore.

Wang, C., Nulty, P. and Lillis, D., 2021. Transformer-based multi-task learning for disaster tweet categorisation. *arXiv preprint arXiv:2110.08010*.

Nimmi, K., Janet, B., Selvan, A.K. and Sivakumaran, N., 2022. Pre-trained ensemble model for identification of emotion during COVID-19 based on emergency response support system dataset. *Applied Soft Computing*, *122*, p.108842.

Koshy, R. and Elango, S., 2023. Multimodal tweet classification in disaster response systems using transformer-based bidirectional attention model. *Neural Computing and Applications*, *35*(2), pp.1607-1627.

Taneja, K. and Vashishtha, J., 2022, March. Comparison of transfer learning and traditional machine learning approach for text classification. In *2022 9th International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 195-200). IEEE.

Basu, P., Roy, T.S. and Singhal, A., 2021, November. But how robust is roberta actually?: A benchmark of sota transformer networks for sexual harassment detection on twitter. In *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)* (pp. 1328-1333). IEEE.

Alam, F., Sajjad, H., Imran, M. and Ofli, F., 2021, May. CrisisBench: Benchmarking crisis-related social media datasets for humanitarian information processing. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 15, pp. 923-932).

Patel, F.P., 2023. Enhancing Health Tweet Classification: An Evaluation of Transformer-Based Models for Comprehensive Analysis (Doctoral dissertation, The University of Texas at Arlington).

Fazlourrahman, B., Aparna, B.K. and Shashirekha, H.L., 2022, July. Coffitt-covid-19 fake news detection using fine-tuned transfer learning approaches. In Congress on Intelligent Systems: Proceedings of CIS 2021, Volume 2 (pp. 879-890). Singapore: Springer Nature Singapore.

Do, S. and Bang, J., 2021, January. Experimental analysis based on binary classification to distinguish the authenticity of text with social network data. In *2021 IEEE International Conference on Big Data and Smart Computing (BigComp)* (pp. 283-286). IEEE.

SIDDIQUA, U.A., CHY, A.N. and AONO, M. (2019). Tweet Stance Detection Using Multi- Kernel Convolution and Attentive LSTM Variants. *IEICE Transactions on Information and Systems*, E102.D(12), pp.2493–2503. doi:https://doi.org/10.1587/transinf.2019edp7080.

Sharma, S., Basu, S., Niraj Kant Kushwaha, Anugandula Naveen Kumar and Pankai Kumar Dalela (2021). Categorizing Disaster Tweets into Actionable Classes for Disaster Managers: An Empirical Analysis on Cyclone Data. *2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*. doi:https://doi.org/10.1109/iceccme52200.2021.9591063

Dharma, L.S.A. and Winarko, E. (2022). *Classifying Natural Disaster Tweet using a Convolutional Neural Network and BERT Embedding*. [online] IEEE Xplore. doi:https://doi.org/10.1109/ICITE54466.2022.9759860.

Hembree, A., Beggs, A., Marshall, T. and Ceesay, E.N. (2021). *Decoding Linguistic Ambiguity in Times of Emergency based on Twitter Disaster Datasets*. [online] IEEE Xplore. doi:https://doi.org/10.1109/CCWC51732.2021.9375916.

P Purushotham, D Divya Priya and Kiran, A. (2022). Disaster Analysis Using Machine Learning. doi:https://doi.org/10.1109/assic55218.2022.10088390.

Sharma, S., Sharma, S. and Athaiya, A., 2017. Activation functions in neural networks. Towards Data Sci, 6(12), pp.310-316..

.

.