Disaster Tweets NLP: EDA & BERT With Transformers

A PROJECT REPORT

Submitted by

20MIA1031 - SANJAY.M
20MIA1117 - PILLARAM MANOJ
20MIA1162 - GUNA SHANKAR.S

School of Computer Science and Engineering
Vellore Institute of Technology
Vandalur - Kelambakkam Road, Chennai - 600 127

Winter semester 2022-23

S.NO	CONTENTS
1.	CERTIFICATE
2.	ACKNOWLEDGEMENT
3.	ABSTRACT
	INTRODUCTION
3	MOTIVATION
4.	LITERATURE REVIEW
5.	CHALLENGES
6.	SOFTWARE REQUIREMENTS System Support
7.	System Design
8.	METHODOLOGY
9.	RESULTS & DISCUSSIONS
10.	CONCLUSION
11.	FUTURE WORK
12.	REFERENCES

CERTIFICATE

This is to certify that the report entitled Disaster Tweets NLP: EDA and BERT With Transformers is prepared and submitted by Pillaram Manoj(20MIA1117), Sanjay.M (20MIA1031) and Guna Shankar(20MIA1162) to Vellore Institute of Technology, Chennai, in partial fulfillment of the requirement for the award of the degree of Master of Technology in Business Analytics(5 year Integrated Programme) and as part of SWE1017 –Natural language processing Project is a bona-fide record carried out under my guidance. The project fulfills the requirements as per the regulations of this University and in my opinion meets the necessary standards for submission.

Guide/Supervisor

Name: Dr.krithiga R
Date: 13-04-2023

HoD

Name: Dr.Sivabalakrishnan
Date: 13-04-2023

ACKNOWLEDGEMENT

We would like to express our heartfelt gratitude and appreciation to Prof.Krithiga R, Faculty of the School of Computer Science and Engineering, for providing us with the opportunity to work on this wonderful project titled Disaster Tweets NLP: EDA And BERT With Transformers as well as for providing us with proper guidance and suggestions to help us develop it.This aided us in conducting extensive research and learning about many new topics. Second, we would like to thank our friends and groupmates for their assistance in finalizing our project within the limited time frame. We'd also like to thank our family for their unwavering support throughout this difficult period.

Abstract:

Natural language processing (NLP) is gaining a lot of interest recently because of the diverse range of applications it may be used for, including disaster management. During emergencies, social media sites like Twitter are frequently utilised to disseminate current information. As a result, more businesses are considering automating Twitter tracking. we outline a natural language processing method for identifying tweets about disasters. In this study, we describe a method for categorising tweets about disasters using deep learning models like BERT (Bidirectional Encoder Representations from Transformers) and Exploratory Data Analysis (EDA). In order to learn more about the data, including data distribution, class imbalance, and tweet length analysis, we do EDA on a disaster twitter dataset that is available to the general public. On the basis of our catastrophe tweet dataset, we fine-tune the pre-trained BERT model and assess its performance using common assessment measures including accuracy, precision, recall, and F1-score. We are going to estimate whether a specific tweet is about a real disaster or not using a competition dataset that we have collected.

Key words: BERT, Transformers, Exploratory Data Analysis (EDA), Natural Language Processing (NLP) and Disaster Tweets.

Introduction:

Social media sites like Twitter are essential for spreading information and giving real-time updates during disasters. However, it can be difficult to find pertinent information for disaster response due to the large volume of tweets posted during disasters. As a result, efficient techniques for automatically categorising tweets on disasters are crucial to aid disaster management efforts. Deep learning-based NLP methods, such BERT with Transformers, have recently demonstrated astounding performance in a range of NLP applications, including tweet classification. To better understand the data's features and distribution, a thorough analysis of the data must be done before using such cutting-edge methods. In order to increase the efficacy and performance of current approaches.

This study employs natural language processing (NLP) methods to analyse tweets about disasters. The purpose of the paper is to use EDA to investigate the features of tweets about disasters, and then to apply BERT with Transformers, a deep learning model, to classify tweets. The goal of the research is to demonstrate how BERT performs better at analysing disaster-related tweets than more established techniques like TD-IDF and other machine learning models. The expected outcomes are all outlined in the introduction, which also serves as an overview of the study. It is highlighted the importance of precise tweet analysis for disaster management initiatives. The introduction also makes reference to the sophisticated features of BERT, such as bidirectional contextual representations, deep learning, and self-attention processes, that help it be effective at understanding the meaning of language used in catastrophe situations and capturing contextual nuances.

Twitter has grown to be a crucial communication tool in emergency situations. Due to the increasing use of cellphones, people can instantly report an emergency if they are there when it occurs. As a result, more businesses are considering adopting language programming strategies to track data from Twitter. Understanding texts and speeches is the main objective of the computer science field of natural language processing (NLP), more especially artificial intelligence.
The growing accessibility of digital information has greatly aided NLP's ascent to success. Today's social media platforms enable a variety of ideas to be expressed online by actual users on a range of subjects, from their comments on a single grocery store item to their political stances.

The increasing availability of digital information has significantly contributed to the emerging success of NLP. Nowadays, social media allows a large number of different opinions to be posted online by real users discussing a wide range of topics, from their thoughts on a single supermarket product to their political positions.We will provide an overview of our approach to the competition in this paper, which included a combination of exploratory data analysis (EDA) and the use of the BERT model with the Transformers library. We will go over the steps we took for data preprocessing, such as cleaning the text and tokenizing the text for input into the model.

Background Study:

According to Gregoire Burel and Harith Alani's , "Crisis Event Extraction Service(CREES) - Automatic Detection and Classification of Crisis-related Content on Social Media," Crisis Event Extraction Service (CREES) is an open-source web API that automatically categorises posts during crisis situations. Annotations for crisis-related documents, event kinds, and information categories are provided by the service. Convolutional neural networks (CNNs) support it, and it has been tested against conventional machine learning models.

Stowe, Paul's in their paper "Developing and Assessing Annotation Processes for Twitter Data during Hazard Events" explains how social media may be used to communicate about natural disasters and offers an annotation framework for finding trends in users' social media activity. An annotation method to consistently identify hazard-related themes in Twitter, an assessment of agreement rates and challenges in finding annotation categories, and a public distribution of both the dataset and guidelines created from this scheme are its three contributions.

According to Prashant, Burel, Gregoire, Khare, Alani, and Harith's article "Classifying Crises-Information Relevancy using Semantics," social media platforms are important channels for disseminating information during emergency circumstances. Humanitarian organisations find it difficult to go through the massive amounts of data. Prior research centred on the use of statistical aspects, however this strategy can occasionally be unsuitable for certain types of crises. In this essay, we investigate how semantics affect the classification of Twitter postings.

According to Matti Wiegmann's "Analysis of Detection Models for Disaster-Related Tweets," social media is seen as a rich resource for disaster management and relief efforts, but the high class imbalance between disaster-related and non-disaster-related messages makes it difficult to make a trustworthy detection. In this context, we provide the Disaster Twitter Corpus 2020, an expanded collection of available materials that contains 123,166 tweets from 46 catastrophes that include 9 different categories of disasters.

According to Nilani Algiriyage's article, "Identifying Disaster-related Tweets: A Large-Scale Detection Model Comparison," social media sites like Twitter and Facebook are becoming important for acquiring situational awareness during catastrophes. In recent years, automated tasks employing machine learning (ML) or deep learning (DL) approaches have been developed to analyse social media posts connected to disasters.

According to Naveen Venkat and his colleagues' article, "Comparing Learning-Based and Matching-Based Approaches for Detecting Disaster Related Tweets," social networking platforms like Twitter are being used by individuals all over the world. They are used in both natural and man-made catastrophes because of how easily they can convey a lot of information in a short period of time. Finding tweets with geotags might be essential for connecting victims with aid.

Sreenivasulu Madichetty and Sridevi M's "Detection of Informative Tweets in Crisis Events" Convolutional neural networks (CNN) and artificial neural networks are the foundation of the proposed method (ANN). The suggested technique outperforms the current methods in terms of accuracy, precision, recall, and F1-score. A real-time Twitter dataset is used to evaluate the suggested strategy.

According to Anna Kruspe and Jens Kersten's article, "Detection of Informative Tweets in Crisis Events," social media messages can be a valuable source of information in emergency circumstances. The accurate recognition of educational messages in a deluge of data is one difficulty. Here, we offer techniques for Twitter tweets that are automatically tied to a crisis.

According to Parilla-"Automatic Ferrer's Classification of Disaster-Related Tweets," Twitter has emerged as one of the quickest sources of news and other information. They create some learning models in their work that can recognise disaster-related tweets automatically. The classifier models were created using a dataset of tweets gathered during the 2012 Habagat floods in Manila.

“Acerbo and Rossi's study, "Filtering informative tweets during emergencies: A machine learning method," makes a machine learning proposal for building a classifier that can differentiate between useful and non-informative posts. Also, they look at word patterns that are common to three major natural disasters—fire, earthquake, and flood—and compare and contrast them.

In their paper "Fine-tuned BERT Model for Multi-Label Tweets Classification," Hamada M. Zahera explained how to divide disaster-related tweets into several sorts of information (i.e, labels). They try to prioritise the tweets that are most pertinent during emergencies. Then, we categorise tweets into pertinent information kinds. Examples of information kinds are Volunteer, MovePeople, and SearchAndRescue. Ten BERT layers make up our enhanced BERT model. Also, we entered the TREC-IS 2019 competition with our technique, and the assessment results showed that our approach surpasses the F1-score of median score in finding useful information.

In their paper, "Applying Modified TF-IDF with Collocation in Classifying Disaster-Related Tweets," Gleen A. Dalaorao and Ariel M. Sison explained that the term "Disaster-related tweets classification" refers to tweets that are posted on Twitter during time-critical events and are grouped together based on pre-defined categories. Although it aids classifiers in extracting valuable features, the Term Frequency - Inverse Document Frequency (TF-IDF) can be misleading and hinder classification effectiveness. By using an upgraded TF-IDF with collocation, this research seeks to analyse the efficacy of tweet categorization. The confusion matrix, accuracy, recall, and F1 score are taken into account while evaluating performance.

Motivation:

During disasters, social media platforms such as Twitter are frequently used for real-time information sharing, communication, and situational awareness. Disaster prevention, resource allocation, and emergency response can all benefit from disaster-related tweet analysis.

The outcomes of this project could have a big impact on the real world. It can assist emergency responders, policymakers, and other stakeholders in making knowledgeable decisions during disasters by obtaining knowledge from the EDA and creating an accurate categorization model using BERT. It can also advance the field of NLP research by showing how well BERT and transformer-based models work in applications for disaster management.

Challenges:

There are benefits and drawbacks to using social media data in catastrophe situations. Positively, social media may offer real-time information that may be essential for emergency action in the initial hours, minimising both human loss and financial harm. Also, informative information may be categorised and filtered using machine learning algorithms, which makes it simpler for crisis managers to comprehend the situation and make wise judgements. Furthermore, the performance and classification accuracy for each catastrophe may be enhanced by the use of several models, such as CNN and BERT.

There are certain drawbacks to take into account, though. For instance, social media reactions to calamities may have detrimental societal and financial effects. Social media usage may also affect how people make judgements depending on the information they see, which can result in inaccurate or misleading information. Also, it might be difficult to appropriately categorise useful tweets due to the language barrier and the usage of slang and jargon in tweets. Last but not least, localising information from Twitter may be tricky, and finding helpful tweets after a crisis might be tough.

Software Requirements:

Python: As this project calls for Python coding, you must have the Python programming language installed on your PC, preferably version 3.x.

NLP Libraries: NLP libraries such as NLTK (Natural Language Toolkit), spaCy, TextBlob, and gensim will be required.

Libraries: To complete the various activities in this project, you will need to install a number of Python libraries, including:
Pandas: used for data analysis and manipulation.
NumPy: for array operations and numerical processing.
Matplotlib and Seaborn are tools for displaying data.
Scikit-learn: for machine learning evaluation metrics and algorithms.
TensorFlow: for BERT and other deep learning models.
Transformers :A well-liked library for working with pre-trained language models, such as BERT.

System Support:

Device name: Sanjay
Processor: Intel(R) Core(TM) i7-10750H CPU @ 2.60GHz 2.59 GHz
Installed RAM: 8.00 GB (7.87 GB usable)
GPU PhysY: NVIDlA GeForce GTX 1650 4GB
CUDA Cores: 896
System type:64-bit operating system, x64-based processor
Edition: Windows 11 Home Single Language
Version: 22H2
OS build: 22621.1413
Experience: Windows Feature Experience Pack 1000.22639.1000.0

System Design:

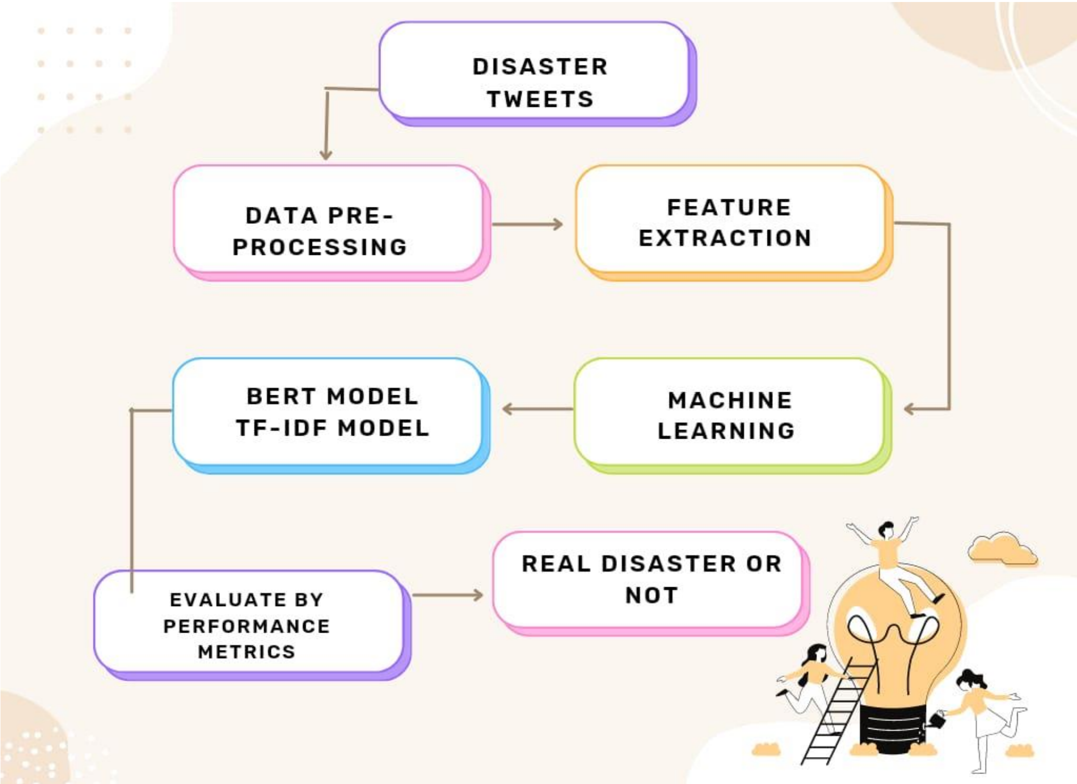

Fig.1

Methodology:

Data Preprocessing and Data Exploration: We began by loading the dataset, exploring its structure, and preprocessing it. We removed unnecessary columns, handled missing values, and cleaned up the text by removing special characters and stop words.

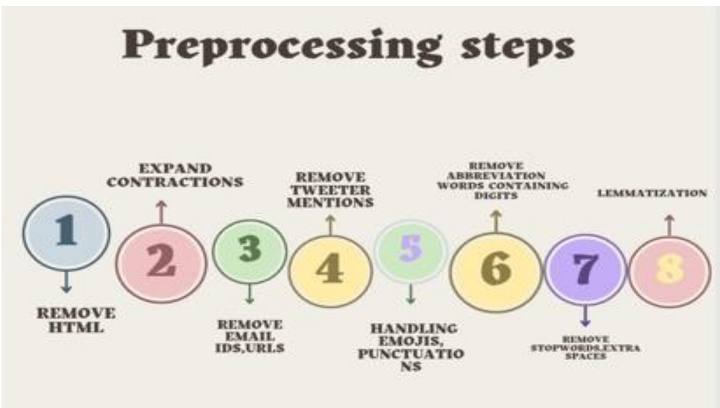

Fig.2

Pre - Processing Steps

- 1) Remove HTML
- Expand Contractions

  There are many contractions of words used in informal communication such as can't: cannot, they've: they have or even modern contractions such as sux: sucks.

- Remove URLs

  URL stands for Uniform Resource Locator, which is used to locate resources on the web. However, they generally do not provide any additional information in the NLP task

- Remove Email IDs

  Email ids have become ubiquitous over the years and appear everywhere. As they do not provide any additional information (unless you are specifically extracting the emails from the text for specific usecase) we need to remove them. Similar to the previous case, email id can be completely removed or replaced with a common word such as "email"

- Remove Tweeter Mentions

  The text contains mentions using @. This generally appears in Tweeter and online forums. We need to remove these mentions before removing the punctutions otherwise they will be hard to find without the @ attached to it.

- Handling Emojis
- Abbreviation/Acronym Disambiguation

  There are large number of abbreviations and acronyms used in the text. These abbreviations can contain meaningful information for the classification task and might get removed or destorted during other preprocessing steps and hence they need to be expanded earlier in the preprocessing.

- Remove Punctuations

  Punctuations are used for defining the structure of the text such as full stops for terminating the sentences. They can be used for sentence tokenization. However, in some NLP tasks, punctuations do not provide any relevant information and need to be removed.

- Remove Digits or Words Containing Digits

  This might not be appropriate in many cases. For example "MH370" mentioned in the tweets corresponds to Malaysia Airlines Flight 370 which went missing. In this case, keeping this number in the text might be useful in the disaster tweet classification

- Remove Stopwords Stopwords are words like 'a, and, the, is, can' which are removed to only keep information rich words in the text.
- Removing Extra Spaces-While performing different preprocessing steps, additional spaces are introduced in the text at the start, end or in-between words which need to be removed.

- Stemming or Lemmatization Stemming uses the stem of the word, while lemmatization uses the context in which the word is being used.

Exploratory Data Analysis (EDA) was then performed to better understand the data distribution and identify any patterns. To gain insights into the dataset, we used various visualisation techniques such as bar plots, word clouds, and histograms.

Feature Extraction: The TF-IDF vectorizer will be used to convert the preprocessed text data into a feature matrix. The importance of each word in each tweet will be represented by this matrix.

Model Building: To build our prediction model, we used the BERT model in conjunction with the Transformers library. We used a pre-trained BERT model and fine-tuned it using the Adam optimizer and binary cross-entropy loss function on our dataset. As input to the BERT model, we will use the feature matrix generated by the TF-IDF vectorizer.

Results & Disussions:

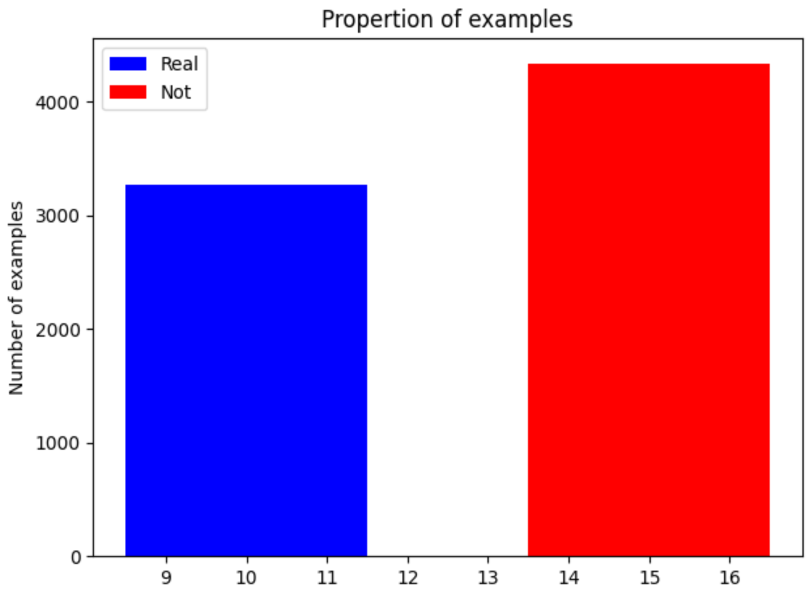

Fig.3

The above bar chart represents the computing the number of instances of each class. The blue color represents whether the disaster is real and the red color represents the disaster is not a real.

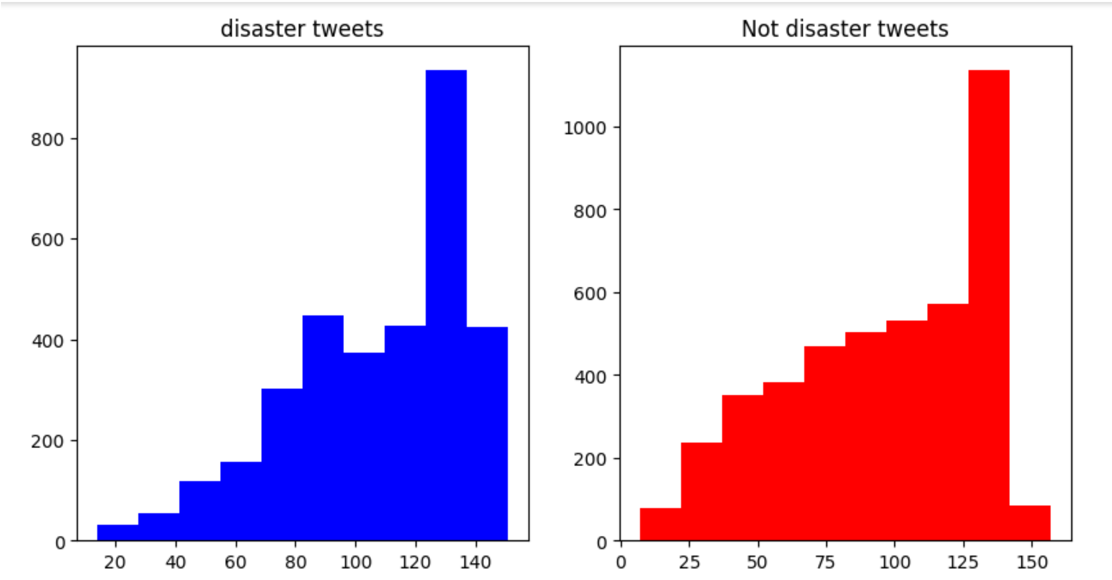

Fig.4

The above Histogram represents the disaster tweets and Not disaster tweets. Disaster tweets represents in blue color and Not disaster tweets represents in red color.

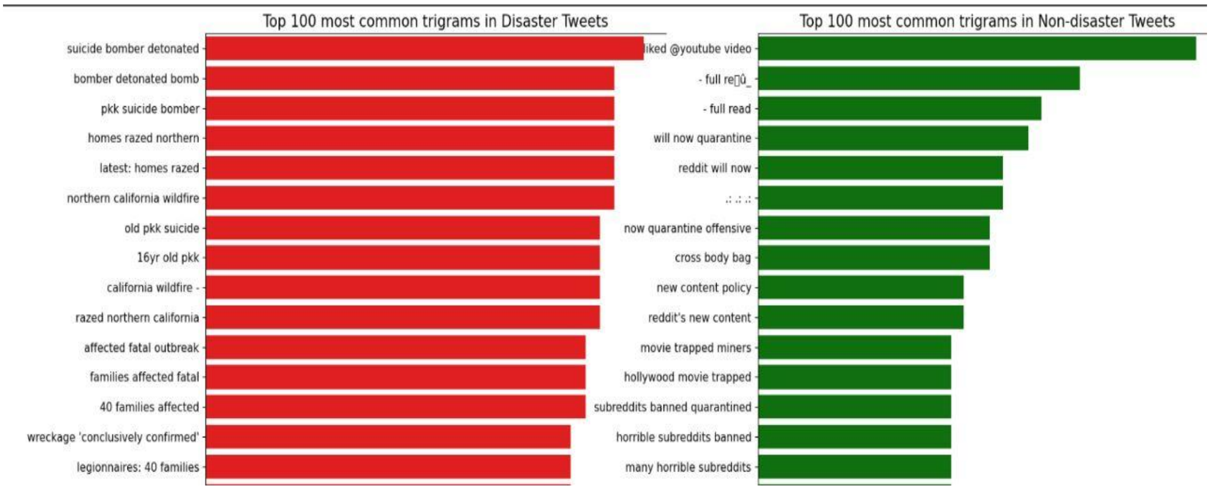

Fig.5

The top 100 most prevalent trigrams in disaster tweets included terms about specific occurrences like suicide bombers, wildfires, and deadly breakouts. These tweets employed language that conveyed urgent requests for assistance and information, as well as particular descriptions about the occurrences. The top 100 most prevalent trigrams in non-disaster tweets, on the other hand, included a broader range of themes, including Reddit's new content policy, cross body bags, and a Hollywood film about stranded miners. These tweets used more generic language and did not focus on urgent or specific occurrences.

Overall, the differences in vocabulary used in disaster and non-disaster tweets show the need of developing NLP models capable of classifying and responding to catastrophe-related tweets in a timely and correct manner.Due to the more obvious context, there are no shared trigrams between the two classes.Bigrams and the most frequent trigrams in tweets about disasters are remarkably similar. They provide a lot of information regarding disasters, yet they might only offer bigrams and no further details.The most frequent trigrams in non-disaster tweets resemble bigrams a lot and have much more punctuation.

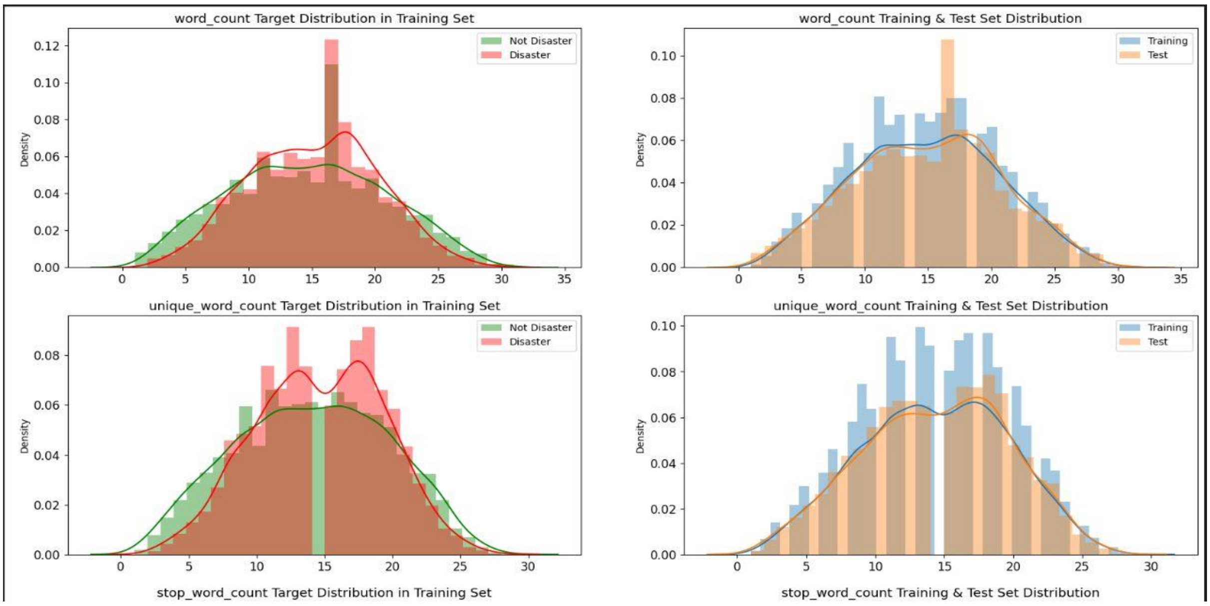

Fig.6

Once again, disaster-related tweets appear to contain a bit more words than non-disaster-related ones. Now, check the average word count per tweet. They both appear to be distributed relatively normally. If word complexity varies from tweet class, we might investigate this further to learn more. It appears that tweets about disasters generally have lengthier words than tweets about non-disasters.

	Algorithm	ROC AUC	Accuracy	Precision	F1 Scores	recall
4	Multinomial NB	80.67	74.79	79.90	65.49	55.75
0	Logistic Regression	80.04	73.80	73.15	67.21	62.65
1	SVC	79.96	74.72	74.60	68.13	62.98
2	Kernel SVM	79.81	73.54	77.80	63.46	53.81
6	Random Forest	76.74	62.97	98.75	24.35	14.12
3	KNN	71.66	69.86	70.86	58.94	50.83
5	Decision Tree Classifier	60.43	64.48	73.20	39.88	27.73

Fig.7

The table displays the results of seven different machine learning algorithms on a specific dataset using the five assessment metrics ROC AUC, accuracy, precision, F1 score, and recall.

The Multinomial NB algorithm has the greatest ROC AUC score of 80.67, which shows that it performs the best overall among the seven algorithms, as seen in the table. Unfortunately, Multinomial NB's accuracy is only 74.79%, which is lower than that of SVC and Logistic Regression.

Random Forest has the lowest false positive rate of all the algorithms with a precision score of 98.75%, making it the algorithm with the greatest precision score. Unfortunately, Random Forest's F1 score and recall are quite poor, suggesting that it has a large false negative rate.

For most evaluation measures, the performance of Logistic Regression, SVC, and Kernel SVM are comparable, with Logistic Regression marginally outperforming the other two in terms of accuracy and F1 score.As opposed to this, Decision Tree Classifier performs the worst out of all the methods, scoring poorly in ROC AUC, precision, F1 score, and recall.
In conclusion, the top-performing algorithms, according to the provided assessment metrics, are Multinomial NB, Logistic Regression, SVC, and Kernel SVM, while Random Forest and Decision Tree Classifier exhibit some performance flaws. But, the particular environment and objectives of the machine learning project will determine which method is most suitable.

Fig.8

This is the graph for training and validation loss with the epoch 3.

```
The BERT model has 393 different named parameters.

==== Embedding Layer ====

bert.embeddings.word_embeddings.weight                   (30522, 1024)
bert.embeddings.position_embeddings.weight                 (512, 1024)
bert.embeddings.token_type_embeddings.weight                 (2, 1024)
bert.embeddings.LayerNorm.weight                               (1024,)
bert.embeddings.LayerNorm.bias                                 (1024,)

==== First Transformer ====

bert.encoder.layer.0.attention.self.query.weight          (1024, 1024)
bert.encoder.layer.0.attention.self.query.bias                 (1024,)
bert.encoder.layer.0.attention.self.key.weight            (1024, 1024)
bert.encoder.layer.0.attention.self.key.bias                   (1024,)
bert.encoder.layer.0.attention.self.value.weight          (1024, 1024)
bert.encoder.layer.0.attention.self.value.bias                 (1024,)
bert.encoder.layer.0.attention.output.dense.weight        (1024, 1024)
bert.encoder.layer.0.attention.output.dense.bias               (1024,)
bert.encoder.layer.0.attention.output.LayerNorm.weight         (1024,)
bert.encoder.layer.0.attention.output.LayerNorm.bias           (1024,)
bert.encoder.layer.0.intermediate.dense.weight            (4096, 1024)
bert.encoder.layer.0.intermediate.dense.bias                   (4096,)
bert.encoder.layer.0.output.dense.weight                  (1024, 4096)
bert.encoder.layer.0.output.dense.bias                         (1024,)
bert.encoder.layer.0.output.LayerNorm.weight                   (1024,)
bert.encoder.layer.0.output.LayerNorm.bias                     (1024,)

==== Output Layer ====

bert.pooler.dense.weight                                  (1024, 1024)
bert.pooler.dense.bias                                         (1024,)
classifier.weight                                            (2, 1024)
classifier.bias                                                   (2,)
```

Fig.9

The BERT model is a deep learning technique for natural language processing that was created by Google. It is an effective text classification and language translation tool. The BERT model is used to categorise tweets as disaster or non-disaster tweets. To adapt to the specific classification task, the BERT model is fine-tuned on training data.

The TF-IDF vectorizer (Term Frequency-Inverse Document Frequency) is a popular feature extraction tool in natural language processing. It is a statistical method for determining the significance of a word in a document based on its frequency and the number of documents in which it appears. The TF-IDF vectorizer is used to convert tweets into numerical feature vectors. The TF-IDF vectorizer assigns a score to each word in the corpus based on its frequency in a specific tweet and its overall frequency in the corpus.

The BERT model and the TF-IDF vectorizer are both effective text classification tools. The BERT model is more complex and requires more computational resources than the TF-IDF vectorizer, but it can be more accurate.

```
  Accuracy: 0.83
  F1: 0.78
  Validation Loss: 0.43
  Validation took: 0:00:13

Training complete!
Total training took 0:08:45 (h:mm:ss)
```

Fig.10

In Bert model we got accuracy as 83% and F1-score as 0.78 and we got average training loss as 0.33 and validation loss as 0.43.

Conclusion:

We conclude that the BERT model outperforms with 82% than TD-IDF and other ML models for processing tweets about disasters. This is probably due to the BERT model's improved ability to capture linguistic context-specific nuance and comprehend the meaning of particular words and phrases. Our finding suggest that disaster-related tweets can be analysed and addressed using NLP models like BERT.

Future work:

We will further improve our model's performance even more by tuning its hyperparameters. We could also try combining different models to improve our overall accuracy. Cross-validation techniques could be used to ensure that our model is robust and generalizable. To address the imbalance in the dataset, we will try different techniques such as oversampling or undersampling.

References :

[1] "Social media and disasters: a functional framework for social media use in disaster planning, response, and research," Disasters, J. B. Houston, J. Hawthorne, M. F. Perreault, E. H. Park, M. Goldstein Hode, M. R. Halliwell, S. E. Turner McGowen, R. Davis, S. Vaid, and others.

[2] F. Alam, F. Ofli, and M. Imran, "Crisismmd: Multimodal twitter datasets from natural disasters," in AAAI Conference on Web and Social Media (ICWSM), AAAI, Stanford, California, USA, June 2018.

[3] C. Caragea, A. Silvescu, and A. H. Tapia, "Identifying informative messages in disaster events using convolutional neural networks," in the 2016 International Conference on Information Systems for Crisis Response and Management

[4] Processing social media messages in mass emergencies: A survey by M. Imran, C. Castillo, F. Diaz, and S. Vieweg, ACM Computing Surveys (CSUR)

[5] "Twitter under crisis: Can we trust what we rt?" in Proceedings of the First Workshop on Social Media Analytics by M. Mendoza, B. Poblete, and C. Castillo. ACM, 2010

[6] "Characterizing communal microblogs during disaster events," Advances in Social Networks Analysis and Mining (ASONAM), 2016 IEEE/ACM International Conference on. IEEE, 2016. K. Rudra, A. Sharma, N. Ganguly, and S. Ghosh.

[7] "Mapping moods: Geo-mapped sentiment analysis during hurricane Sandy," ISCRAM, 2014. C. Caragea, A. C. Squicciarini, S. Stehle, K. Neppalli, and A. H. Tapia.

[8] "Extracting situational information from microblogs during disaster events: a classification-summarization approach," in Proceedings of the 24th ACM International Conference on Information and Knowledge Manage- ment. K. Rudra, S. Ghosh, N. Ganguly, P. Goyal, and S. Ghosh. ACM, 2015

[9] "Classifying information from microblogs during epidemics," in Proceedings of the 2017 International Conference on Digital Health, by K. Rudra, A. Sharma, N. Ganguly, and M. Imran. ACM, 2017

[10] M. Sridevi and M. Sreenivasulu, "Mining informative words from the tweets for detecting the resources during disaster," International Conference on Mining Intelligence and Knowledge Exploration. Springer, 2017

[11] Convolutional neural networks for phrase classification, Y. Kim, 2014 arXiv preprint arXiv:1408.5882.

[12] Adadelta: An Adaptive Learning Rate Method, by M. D. Zeiler, arXiv preprint 1212.5701, 2012.

[13] "Dropout: a straightforward method to prevent neural networks from over-fitting," The Journal of Machine Learning Research, vol. 15, no. 1, pp. 1929–1958, 2014. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhut-dinov.

[14] "A logical calculus of the ideas immanent in nervous activity," by W. S. McCulloch and W. Pitts. the mathematical biophysics bulletin.

[15] "The perceptron: a probabilistic model for information storage and organisation in the brain," F. Rosenblatt. Psychological examination.

[16] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv:1704.04861, 2017.

[17] P. Werbos, Unpublished Doctoral Dissertation, Harvard University, 1974, "Beyond Regression: New Tools for Prediction and Analysis in Behavior Science."

[18] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M.

[19] F. Chollet et al. (2015) "Keras," available at https://keras.io. 2015.

[20] Filtering emergency-related information from tweets using machine learning C. Rossi and F. S. Acerbo Posted on December 11, 2017.

[21] Automatic Disaster-Related Tweet Classification Proceso L. Fernandez Jr., PhD, Jaime T. Ballena IV, PhD, and Beverly Estephany Parilla-Ferrer (ICIET'2014) 28-29 December 2014 Bangkok (Thailand)

[22] detection of educational tweets during crises Jens Kersten, Friederike Klan, and Anna Kruspe each have one.Jena, Germany's German Aerospace Center (DLR), Preprint. 10 July 2020, 10 July 2020, Author(s) 2020. 4.0 CC BY License.

[23] Using Deep Neural Networks to Scan for Informative Tweets During Disaster 2019 11th International Conference on Communication Systems & Networks, January.

[24] contrasting Learning-Based and Matching-Based Techniques for Recognizing Disaster-Related Tweets,Project: A Dive into Machine Learning, November 2018.

[25] Detecting Tweets about Disasters: A Large-Scale Model Comparison 18th International Conference on Information Systems for Crisis Response and Management, July 2021, Conference Proceedings,Located at Virginia TechMultimodal data fusion project.

Analysis of Detection Models for Tweets Related to Disasters DOI:10.5281/zenodo.3713920, May 2020A conference called the 17th International Conference on Information Systems for Crisis Response and Management was held at Blacksburg, Virginia, in the United States.

[27] Relevancy of Crises-Information Classification Using Semantics Item for a conference or workshop How to cite: Alani, Harith; Burel, Gregoire; and Khare, Prashant (2018). Semantic classification of crisis-information relevance. in Lecture Notes in Computer Science: ESWC 2018: Proceedings of the 15th International Conference, Springer.

[28] Proceedings of the 15th ISCRAM Conference - Rochester, NY, USA, Crisis Event Extraction Service (CREES) - Automatic Detection and Classification of Crisis-related Content on Social Media May, 2018.
[29] Creating and Testing Twitter Data Annotation Procedures for Hazardous Events Rebecca Morss3, Julie Demuth3, Heather Lazrus3, Kevin Stowe1, Martha Palmer1, Jennings Anderson1, Leysia Palen1, Kenneth M. Anderson1, Marina Kogan2, University of New Mexico, Albuquerque, NM, 87131; University of Colorado, Boulder, CO, 80309;Boulder, CO, 80307; National Center for Atmospheric Research.

[30] In May 2017, Tonon, Cudr-Mauroux, Blarer, Lenders, and Motik published their study. ArmaTweet: Semantic Tweet Analysis for Event Recognition. Conference on the Semantic Web in Europe (pp. 138-153). Cham Springer.

Youtube link: https://youtu.be/GNK3YTu62EI