Social Media Analysis Humberside Fire and Rescue

Abstract

Sentiment analysis has become a driving force into improving the performances of organizations and businesses in recent years. This research was conducted on behalf of the Humberside fire and rescue aiming to identify the community's perspective of the organizations performance, in order to address the shortcomings and improve the general services provided by the organization. The second phase of the study concentrated on the creation of an artificial intelligence model designed to detect risks mentioned in social media but potentially not reported through official channels. The goal was to empower the Humberside Fire and rescue team to promptly address and mitigate these potential risks.

A variety of approaches were employed to gain the best results and insights, this included experimenting with different deep learning techniques that has been proven to provide the best outcomes such as Bidirectional Long Short-Term Memory (BiLSTM), as well as transformer models like Robustly optimized Bidirectional Encoder Representations from Transformers (RoBERTa) and eXtreme Learning NETworks (XLNET).

Through the experiments, it was concluded that the RoBERTa Cardiff pretrained model yielded the most accurate results in classifying different sentiments. The analysis revealed that the majority of tweets directed at Humberside Fire and Rescue were of a positive sentiment. Additionally, the RoBERTa Base model demonstrated the highest accuracy in predicting possible risks in given text.

1.Introduction

Major organizations are currently reliant on collecting data from their customers or the communities they are serving in order to ensure in delivering the best services or even maximizing their gains. In recent years people have been turning to social media to express their opinions whether it be related to views on businesses, government, or even discussing their concern with different aspects for local community [1]. Therefore, this study was aimed at capturing the East Yorkshire community sentiments towards the Humberside Fire and Rescue services. The other focus of the study was developing an artificial intelligence model that would be able to classify a tweet as a risky behaviour to be investigated and handled by the Humberside Fire and Rescue. While sentiment analysis has been extensively explored by various researchers [2] , the goal of this study was to experiment with different deep learning models and identify the one that would more suitable and deliver the most accurate results for this task. On the other hand, research involving the detection of risky behaviour over social media platforms has been primarily focused on identifying national threats [3], and is often conducted by law enforcement agencies. Therefore, this work opens doors for organizations to leverage advancements in artificial intelligence to anticipate potential risks and effectively manage them, thereby enhancing community safety and optimizing resource utilization.

The concept of opinion analysis traces back to the 20th century, the start of computer-based sentiment analysis came in the early 2000 coinciding spread of subjective text on the internet [4]. A fundamental component of the current sentiment analysis is Natural Language Processing (NLP). Over the years, this concept developed exponentially in the year 1990 with the rise of machine learning algorithm led to huge advancements in the NLP in areas such as machine translation and speech recognition [5].

Sentiment analysis is divided into three major parts: sentence level, document level and aspect level [6], the study conducted will be focused on the sentence level sentiment analysis, where it will aim to identify the sentiment or emotional tone of each sentence. This involves determining whether a sentence expresses a positive, negative, or neutral sentiment using different deep learning algorithms. Transformer models have revolutionized the field of deep learning, particularly in NLP [7]. They emerged as a breakthrough in 2017, overcoming the limitations of Recurrent Neural Networks (RNNs) by introducing the revolutionary concept of self-attention[8]. Unlike RNNs, which process sequential data in a linear fashion, transformers rely on self-attention mechanisms, enabling them to capture long-range dependencies and contextual relationships within the input data[8]. This innovation has enabled transformers to achieve state-of-the-art performance on a wide range of NLP tasks, including machine translation, text summarization, sentiment analysis, and question answering.

One of these noteworthy developments in transformer models is the RoBERTa Base model, the model is a variant of the BERT model [9]. While both models share a similar architecture, RoBERTa exhibits several key distinctions that contribute to it superior performance. For instance, the RoBERTa model is trained on a larger dataset of text measuring at 160GB compared to BERT’s 16GB [9]. RoBERTa model also implements dynamic masking in its pre-training, offering varied masked word positions for each instance, in comparison with BERT's static masking. With a 10-fold larger vocabulary than BERT, RoBERTa excels in handling a broader range of linguistic complexities [9]. As a result, the RoBERTa model can capture intricate contextual relationships within the language, leading to heightened performance in various NLP tasks.

Given the unavailability of predefined sentiment labels on the data collected from Twitter, the Cardiff’s pre-trained RoBERTa model was utilized for sentiment analysis. This model was crafted by a team from Cardiff university, they used a RoBERTa Base model and trained it on a 124 million tweets over the years, the project commenced in January 2018 [10]. This model classifies text into three different categories which are: Negative, Neutral, and Positive. The developers state that the accuracy of the model outperforms that of the BERTweet and RoBERTa Base in a tweet evaluation test achieving 73.7% in sentiment classification [10].

To ensure a broader exploration of various methodologies, we assessed the performance of a BiLSTM, a recurrent neural network (RNN) variant extensively employed in NLP projects. The BilSTM has the capability of processing input data in both forward and backward directions simultaneously [11]. This bidirectional approach empowers BiLSTM to effectively capture contextual dependencies in sequential data, rendering it particularly valuable for NLP tasks such as sentiment analysis, named entity recognition, and language translation [11]. By considering both preceding and succeeding words, BiLSTM enhances the model's capacity to comprehend relationships within a sequence, making it a potent tool for tasks where context plays a crucial role.

Considering the various aforementioned factors, it was deemed that these cutting-edge models will be utilized for the study, thereby guaranteeing a robust and accurate analysis.

2.Methodology

2.1 Data collection

For this research, data collection was facilitated through the Twitter API, focusing on tweets related to @HumbersideFire and the hashtag #HumbersideFire. The search queries covered the timeline from 2019 to 2023. This approach yielded a total of 5,527 tweets, providing a dataset relevant to the Humberside Fire and Rescue for analysis.

Recognizing the challenge posed by unlabelled data obtained from Twitter, an additional dataset was sourced from Kaggle dedicated for Twitter sentiment analysis [12]. This supplementary dataset comprised 16,130 tweets with sentiment labels, enhancing the scope for experimentation with various sentiment analysis models.

2.2 Data Exploration and preparation

The initial phase of the research involved identifying key contributors within the dataset. Notably, the Humberside Fire and Rescue official account emerged as the primary contributor, with a contribution of 1,415 tweets. Recognizing the potential influence of such an official account on sentiment analysis results, a decision was made to exclude tweets originating from this account to ensure a more insightful analysis of community perception. Further analysis of interactions revealed that the Humberside Fire and Rescue official account garnered significant engagement, accumulating over 25,000 likes and 6,000 retweets.

To ensure unbiased sentiment analysis results, tweets from key figures and organizations affiliated with Humberside Fire and Rescue were excluded. This included tweets from Chris Blacksell the former Chief, Phil Shillito the current Chief Officer [13], and Vicky Shakesby Head of Emergency Response [14]. Additionally, tweets from official entities like National Highways: Yorkshire and Lincolnshire Fire and Rescue were omitted. These exclusions were implemented due to the significant influence these accounts hold, aiming to provide a sentiment analysis that better represents the broader community's opinions regarding the Humberside Fire and Rescue services.

To enhance the relevance and quality of the tweet data for subsequent sentiment analysis, a thorough preprocessing approach was employed. This involved a series of steps to refine the text content. First, mentions and hashtags were systematically removed to focus on the core content, excluding user references and minimizing potential noise. Additionally, web links or URLs were eliminated to prevent external references from influencing sentiment analysis, ensuring data integrity. Non-alphabetic characters, symbols, and numeric characters were also removed, maintaining clean and standardized text for sentiment evaluation. The tokenization process, facilitated by TweetTokenizer, broke down tweets into individual words or tokens, a crucial step in preparing the data for analysis. Common English stop words, with little semantic value, were excluded to filter out less meaningful words and enhance text quality. Lemmatization, reducing words to their base or root form, further standardized language in the tweets. This preprocessing procedure was consistently applied to various transformer models in the study, including RoBERTa Cardiff pretrained sentiment and emotions models, RoBERTa Base, and the XLNET model. This preprocessing procedure was carried out identically to all the transformer models used in this study which are the RoBERTa Cardiff pretrained sentiment model, RoBERTa Cardiff pretrained emotions model, RoBERTa Base, and the XLNET model.

2.3 Model Designs

In this stage of our methodology, we will delve into the various designs adopted for the models. The first model is the pretrained RoBERTa model for sentiment utilized to perform the classification on the pre-processed data. This model has been fine-tuned specifically for sentiment classification on Twitter data [10]. The process begins with the loading of the model using the Hugging Face Transformers library. The AutoTokenizer is employed to convert the cleaned tweet text into a format suitable for the model, while the AutoModelForSequenceClassification class configures the model for sequence-based sentiment classification. Once loaded, each cleaned tweet undergoes tokenization using the tokenizer. This involves converting the text into a sequence of numerical tokens, facilitating model interpretation. The RoBERTa model processes these tokenized sequences, providing probability scores for different sentiment classes. Post-inference, the sentiment label is assigned based on the class with the highest probability. This integrated approach ensures that sentiment analysis is conducted with a high level of accuracy, benefiting from the model's extensive training on Twitter data and its ability to understand contextual relationships within language. The model's fine-tuning for sentiment-specific tasks enhances its capacity to discern nuanced sentiments expressed in our dataset related to the Humberside Fire and Rescue services.

The risk detection model is built using the RoBERTa Base, with tokenization and Data Loader preparation orchestrated to facilitate training. The process involves tokenizing the text, encoding it for input, and generating Data Loader instances for both training and validation sets. This iterative training process, spanning multiple epochs, involves feeding batches of tokenized and labeled tweet data to the RoBERTa model. The model adapts its parameters to optimize risk detection, guided by the manually labeled tweets. The same model architecture was also implemented in use for the XLNET model.

The BiLSTM model was utilized for both the sentiment analysis and risk detection task. Initially the model underwent initial training on labeled Twitter sentiment data sourced from Kaggle [12]. This training phase equipped the model with the ability to distinguish nuances in sentiment. Subsequently, the trained BiLSTM model was employed for sentiment classification within the Twitter dataset for the Humberside Fire and Rescue. The BiLSTM model architecture for tweet sentiment analysis comprises several key components and hyperparameters. The tokenized tweet data undergoes preprocessing, including padding for uniform sequence length and one-hot encoding for label representation. The model consists of an embedding layer with a dimensionality of 100, followed by two bidirectional LSTM layers, each with 100 units, and dropout layers with a rate of 0.5 after each bidirectional LSTM layer to prevent overfitting. The final bidirectional LSTM layer is followed by another dropout layer and a dense output layer with a SoftMax activation function, producing probabilities for three sentiment classes. The model is compiled using the Adam optimizer with a learning rate of 0.0001, categorical cross-entropy loss function, and accuracy as the evaluation metric. During training, an Early Stopping callback is employed, monitoring the validation loss and halting training if no improvement is observed for five consecutive epochs. This mechanism restores the model weights to those yielding the lowest validation loss. The training process involves 50 epochs with a batch size of 16. The model underwent experimentation by the addition and removal of layers, ultimately it was identified that the three layers configuration best outcomes. Despite exploring regularizers, such as L1 and L2, the model's performance did not show improvement, hence, it was excluded from the final model.

To address the task of detecting risks in tweets, a specific subset was chosen from the overall dataset. This subset comprised tweets classified as having negative sentiments. This selection was deemed appropriate after reviewing the outcomes of sentiment classification, as these negative sentiment tweets were observed to represent community concerns related to developing situations with potential risks. In contrast, positive and neutral tweets were predominantly expressions of appreciation or support for the services provided by Humberside Fire and Rescue. At times, they also consisted of shared news and announcements without significant risk-related factors suitable for training the model. Manual labelling was conducted by inspecting the text in each tweet to identify whether it represented a risk-related tweet or was otherwise classified as non-risk. Subsequently, the RoBERTa Base, XLNET, and BiLSTM models were trained and validated on specific segments of this dataset. They were then tested on other segments of the same dataset that were not part of the training or validation sets. The results of these tests were compared with the actual labels to evaluate the models' performance.

2.4 Model Evaluation

Several metrics were employed to measure the performance of the models, and among these, the classification report and confusion matrix. The utilization of a classification report, allowed to gauge the model's performance by employing four metrics accuracy, precision, recall, and F1-score [15]. Additionally, a confusion matrix was employed as another evaluative approach, to better understand the predictions made by the model in comparison to the actual assigned labels of the text. This method involved four key indicators: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Furthermore, an additional test aimed to identify the most suitable model for risk detection. This involved subjecting a sample of custom sentences to each model and comparing their predictions on these sentences.

3.Results

In the initial findings of sentiment analysis, utilizing the RoBERTa pre-trained Cardiff model, a breakdown of sentiments within tweets related to Humberside Fire and Rescue was observed. After excluding official accounts that could potentially obscure sentiment analysis results, the distribution of sentiments revealed 1893 tweets classified as positive, 1255 as neutral, and 211 as negative.

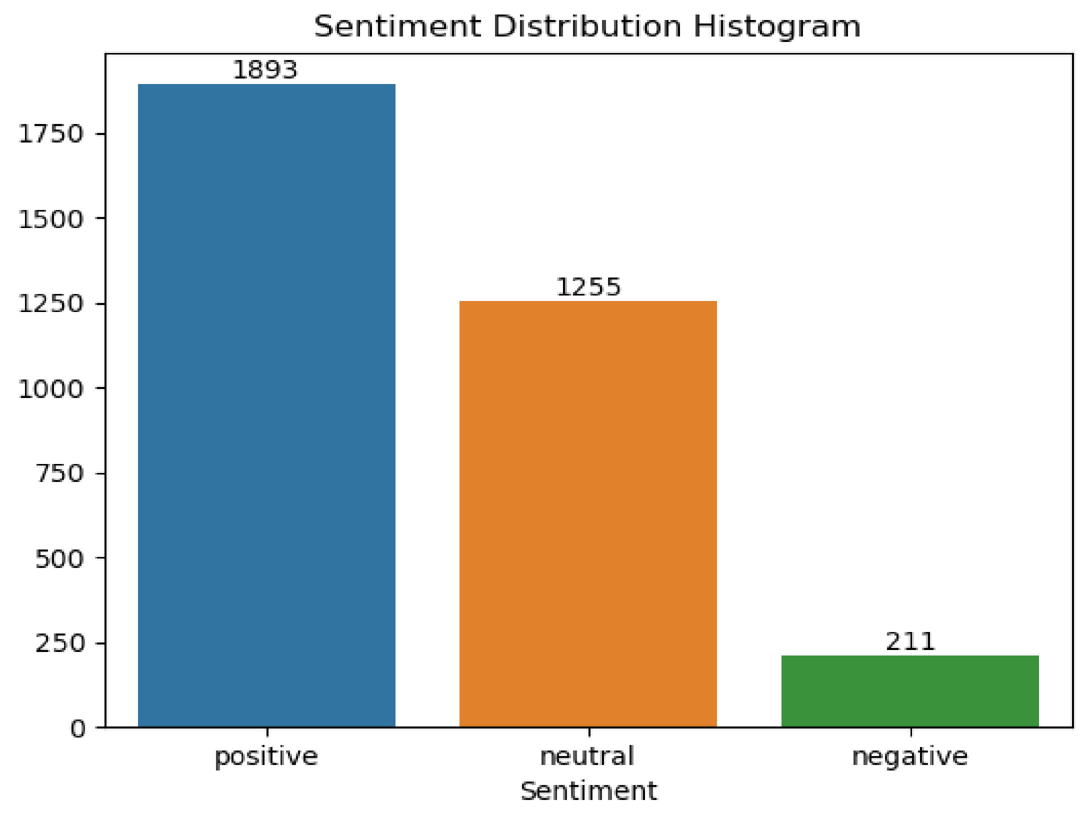
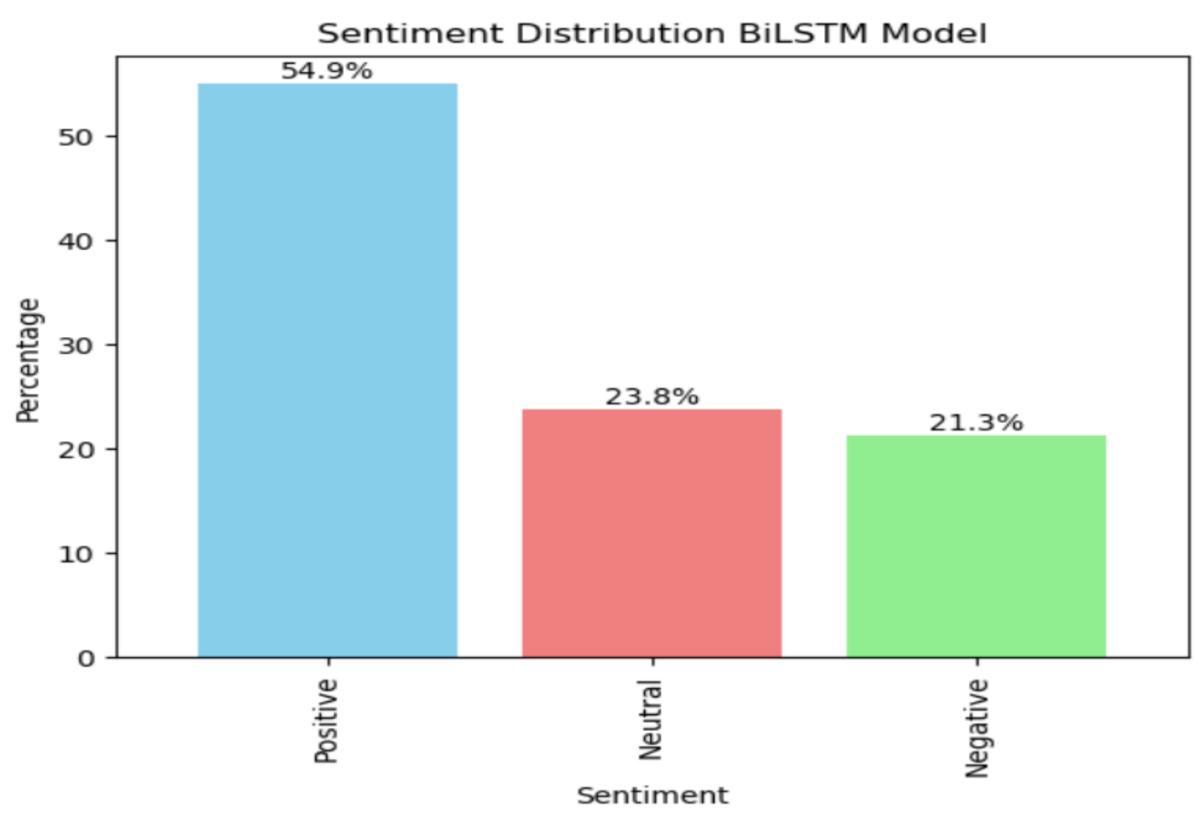
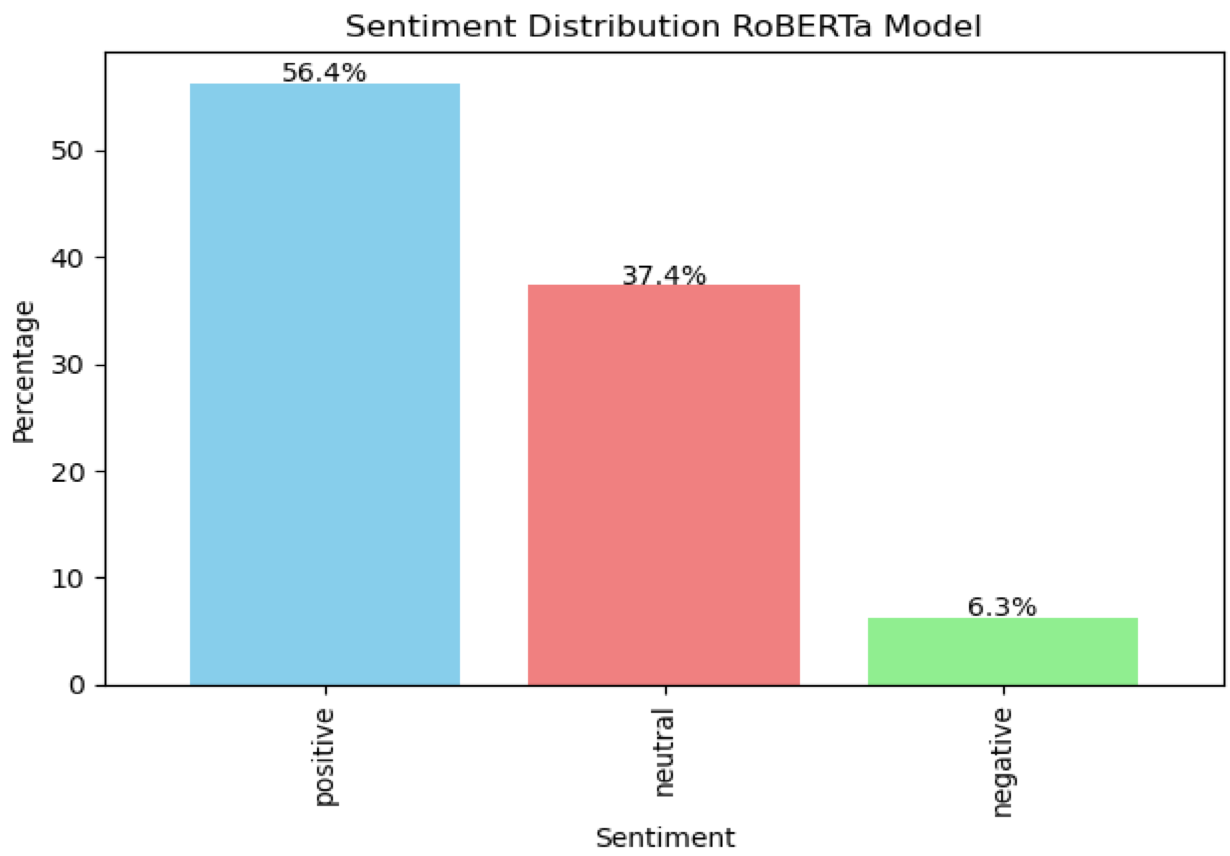


Figure 1 The distribution of sentiments the result of the pretrained RoBERTa sentiment model. As it can be observed the highest ranked sentiment is the positive sentiment.

As previously mentioned during the study conducted experiments with different NLP deep learning models, the distribution percentages of different sentiments are visually depicted in Figure 2.





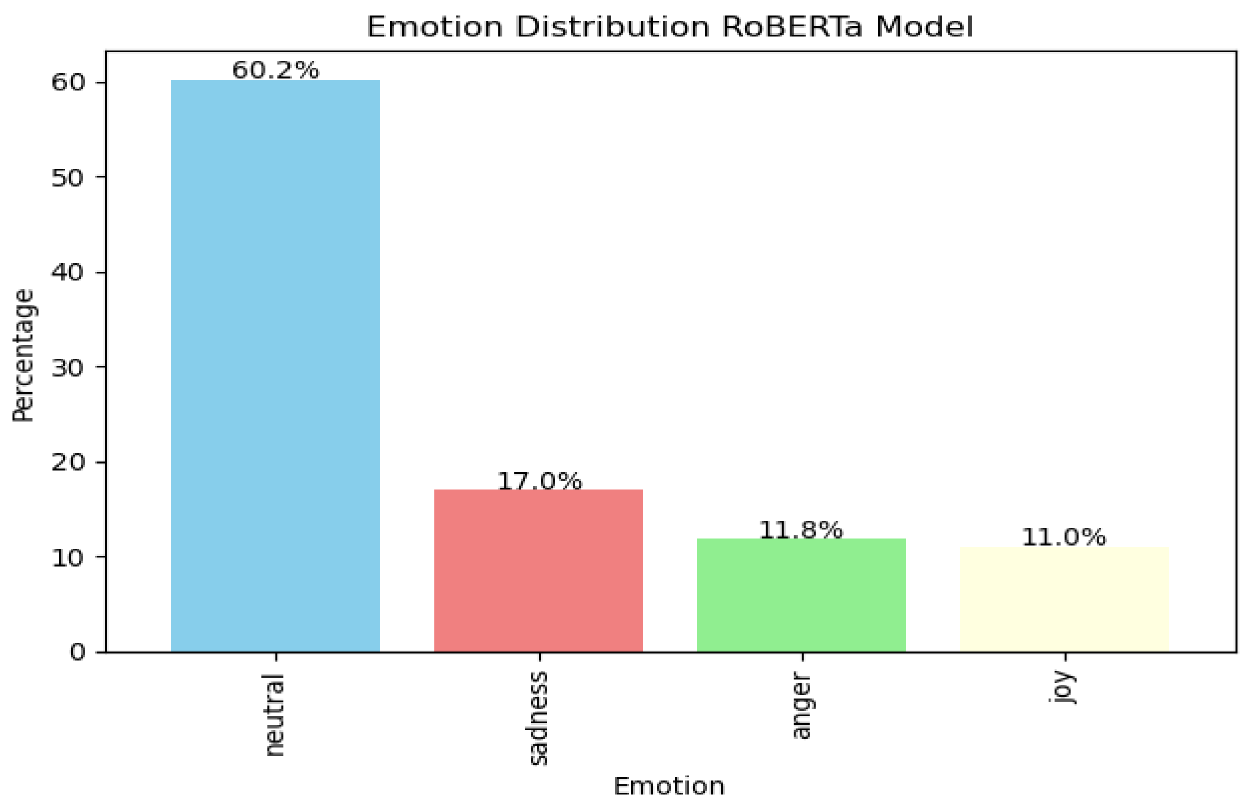


Figure 2 These charts represent the different models results for the sentiment classification.

The RoBERTa pretrained sentiment model classified 56.4% of the tweets as positive, 37.4% as neutral, and 6.3% as negative sentiments. Conversely, the RoBERTa pretrained emotions model categorized tweets as 60.2% neutral, 17.0% sadness, and nearly equally distributed anger and joy at 11.8% and 11.0%, respectively. Moving on to XLNET and BiLSTM, both models exhibited approximately similar predictions for positive sentiments, with 55.4% and 54.9%, respectively. However, the BiLSTM reported 21.3% negative sentiment and 23.8% neutral sentiments, whereas the XLNET model predicted 17.7% negative and 27% neutral sentiments. The results of the various sentiment models were meticulously to identify the best-performing model. A subset of these results is presented in Table 1. Notably, the pretrained RoBERTa emotions model exhibited a higher rate of misclassifying different tweets. Most of the confusion in the misclassification arose from the model predicting the text as containing Anger.

Tweet	RoBERTa pretrained sentiment model prediction	RoBERTa pretrained emotions model prediction
Another community action day raising awareness of fire risks. Strong partnership working by your Hull North team @HumbersideFire @FortemCares @Hullccnews @NNetworkhull & Cllrs Randall Around Stroud Crescent today - say hello! @humberbeat #buildingcommunityresilience	Positive	Anger
Great collaborative meeting today with @HumbersideFire . @NaturalEngland & @SYFR .joint exercises planned for later in the year and local moorland issues discussed that effect both areas . And new equipment tested out	Positive	Anger
@southendroad @Humberbeat_NL @BBCLookNorth @NorthLincsCNews @NajModakBBC @peter_levy @ABPHumber @KarlTurnerMP @BBCCountryfile @piersmorgan @OPCCHumberCEO Our local police force @Humberbeat don't think this is acceptable but our local authority @NorthLincsCNews say nothing to see here. What's your view @HumbersideFire ? You fancy navigating this with your vehicles, or should residents just learn to put their own fires out ?	Negative	Joy
BritishBarley @AuntBessies @BirdsEye @HumbersideFire @TEAL_Trust @joel_for @AtaxiaUK Good work Tony	Positive	Neutral
@HumberMD @HumbersideFire @CFOBlacksell @hullmensayno this was wrapped with funding from the community safety partnership in 2017. It has been used to raise awareness in schools as well as being an operational appliance. There is also a @Humberbeat car with similar messages.	Neutral	Anger

Table 1 This table represents sample of the misclassification made by the RoBERTa pretrained emotions model in comparison with the sentiment model.

In order to understand the sentiment trends throughout the time period of the collected data and identify potential correlations between sentiment spikes and events that may have caused these sentiments, a stacked plot was employed. This is illustrated in Figure 3. In 2020, sentiment activity appeared generally consistent, except for a slight uptick in recorded negative sentiments in August. Moving to 2021, January marked the peak in expressed negative sentiments, whereas July exhibited the highest count of positively classified sentiments without any recorded negative ones. In 2022, October stood out with the highest count of positively classified sentiments, and August experienced a notable increase in negative sentiments. In 2023, overall sentiment activity seemed ordinary, with May recording the highest registered negative sentiments, while June witnessed the pinnacle of positive sentiment.

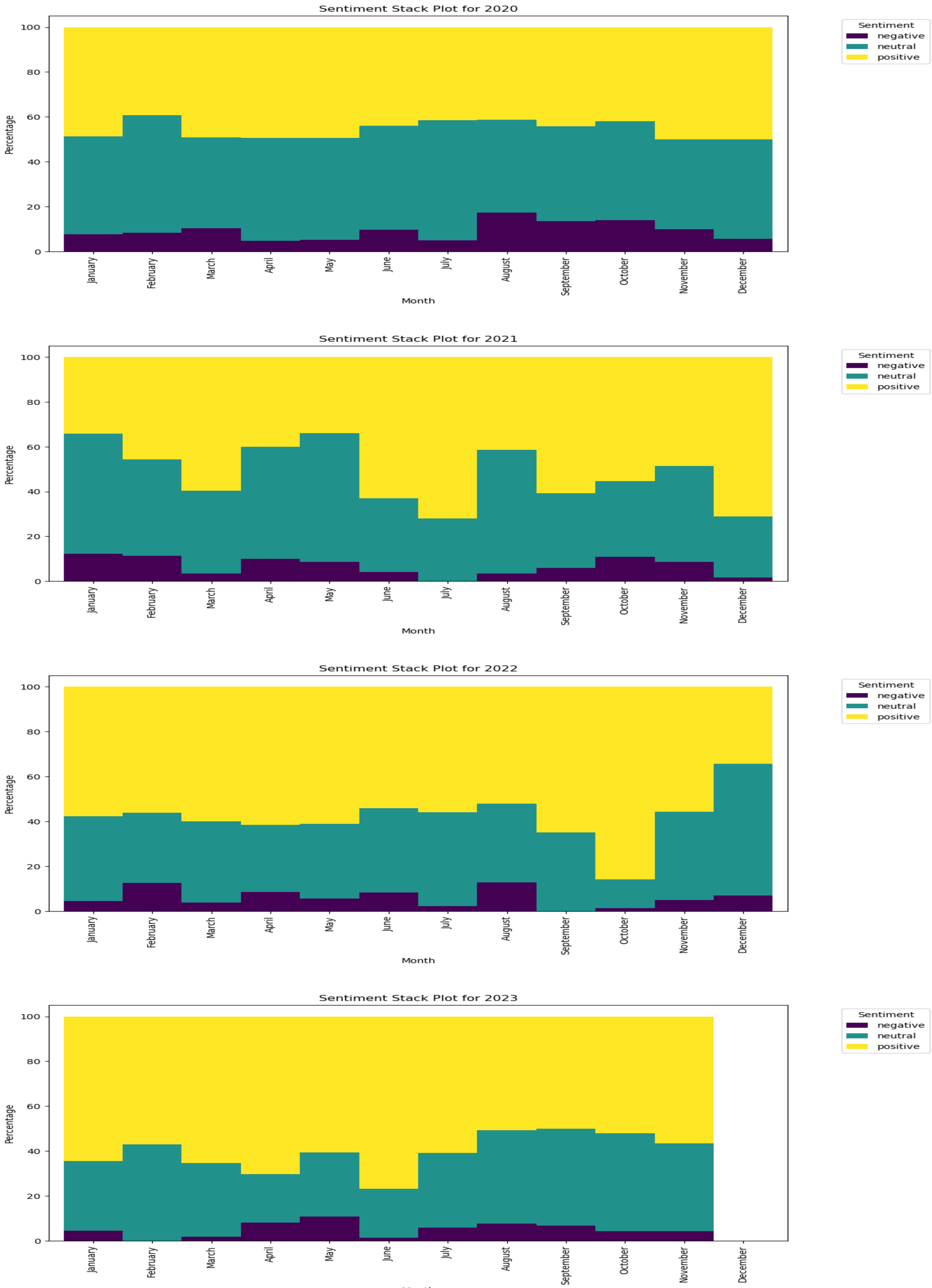


Figure 3 This stacked plot represents the different sentiment distribution between the year 2020 to 2023.

In the process of developing the risk detection model, three models were experimented with, all trained on similar subsets of manually labelled data and tested on unseen data for validation. Table 2 presents the results of these experiments.

Model	Validation Accuracy	Prediction Accuracy	Number of Wrong Predictions
RoBERTa Base	94%	93%	4
XLNET	94%	86%	8
BiLSTM	81%	88%	5

Table 2 These are the results of different tested model for risk detection.

In the validation set, the RoBERTa Base model demonstrated a 93% accuracy, predicting 14 tweets as potential risks. However, it made 2 incorrect predictions and failed to identify 2 actual risk-related tweets. On the other hand, the XLNET risk detection model achieved an 86% accuracy. It correctly predicted 5 instances of risk in unlabelled tweets but misclassified 1 tweet and overlooked a total of 7 tweets, resulting in a total of 8 incorrect predictions. The final model assessed for risk detection was a BiLSTM, with an 81% validation accuracy and an 88% prediction accuracy. This model accurately identified 8 tweets as risks but misclassified 5 tweets. These findings collectively underscore the efficacy of these models in detecting and assessing risks within the dataset.

Custom Text	RoBERTa Base Model Prediction	XLNET Model Prediction	BiLSTM Model Prediction
Had a wonderful time in hull today.	No risk detected	No risk detected	No risk detected
There is a fire in the south street we need your assistance @HumbersideFire.	Risk detected	No risk detected	Risk detected
The kids are lighting fireworks in Pearson Park it is really dangerous.	Risk detected	No risk detected	Risk detected
I see smoke coming from the paragon station.	Risk detected	No risk detected	Risk detected

Some teenagers are jumping of the bridge into the water.	Risk detected	No risk detected	No risk detected
There is no incident in the Beverly Road.	Risk detected	No risk detected	No risk detected

Table 3 Comparison of the the predictions made by different models on

4.Discussion

After conducting multiple experiments, the RoBERTa pretrained sentiment model displayed higher accuracy in predicting tweets compared to the RoBERTa pretrained emotion model. The latter misclassified numerous sentiments expressed in the tweets, as indicated in the sample provided in Table 1. Although the XLNET and BiLSTM models showed similar results to the RoBERTa pretrained sentiment model in classifying positive sentiments, their outcomes differed when classifying neutral and negative sentiments.

The results of the sentiments classification produced by the RoBERTa pretrained was closely examined, this revealed that positive tweets were largely expressions of gratitude or support for the organization's services or activities. Neutral tweets were predominantly consisting of public announcements, including those made by the Humberside Fire and Rescue and is retweeted by members of the community, as well as news shared by various media outlets. While a closer examination of negative tweets showed that a significant portion originated from community members reporting incidents or issuing warnings. These tweets may have mentioned the Humberside Fire and Rescue account, but they were not necessarily directed as a criticism of its services. Acknowledging the robustness of sentiment analysis models, it's crucial to recognize inherent limitations, such as the challenge of discerning sarcasm and potential bias in the dataset. Understanding these constraints ensures a better interpretation of results and fosters a comprehensive understanding of community sentiments. Integrating sentiment analysis results with other data sources, such as incident reports and community surveys, enriches the understanding of community dynamics.

While the temporal representation of sentiment distributions in Figure 3 aimed to identify trends or expressions of sentiment, it did not reveal a direct correlation between sentiment spikes and specific events. However, one notable exception was the increase in neutral sentiment during March 2020, coinciding with the flooding of the River Aire in Snaith and East Cowick [16]. Closer examination of these tweets revealed community discussions about the flooding, announcements from the Humberside Fire and Rescue, and initiatives to support those affected by the disaster. By actively monitoring social media conversations, in the occurrence of similar events could help pinpoint the most affected regions and prioritize their rescue and relief efforts. As well as, it could aid in crafting proactive community engagement strategies during similar events.

The research also sought to create a model capable of identifying risks shared by community members on social media. Three models underwent training and testing, as detailed in the methodology section. The most effective model proved to be the RoBERTa Base, which made 4 errors when predicting on new data. Another test was conducted to simulate future scenarios and assess the models' performance by inputting custom sentences into each model and observing the predictions, as shown in Table 3. The XLNET model performed poorly in this test, consistently predicting no risk. In contrast, the RoBERTa Base and BiLSTM demonstrated more promising performances in detecting fire risks in the provided text. Although the RoBERTa model made an incorrect prediction when the text indicated no risk, it showed the ability to identify a wider range of risks. Conversely, the BiLSTM model primarily focused on fire-related risks and may not be as versatile in risk detection, therefore, based on the obtained results, it was concluded that deploying a RoBERTa model would be more beneficial for the intended purpose.

5.Conclusion

In conclusion, the examination of community attitudes toward the Humberside Fire and Rescue services underscores the significance of leveraging advanced natural language processing models. The RoBERTa Base model demonstrated superior accuracy in sentiment analysis and risk detection. Overall, the community highly values the services of the Humberside Fire and Rescue, as evidenced by most tweets classified as positive sentiment, with the 6% classified as negative sentiment mostly comprising concerns or worries rather than actual negative feelings towards the organization.

The success of the risk detection model, particularly the RoBERTa Base's superior performance, highlights its potential for proactive emergency response. Effectively identifying potential risks mentioned on social media enables the Humberside Fire and Rescue team to swiftly address concerns, ultimately contributing to improved community safety.

While acknowledging the accomplishments of this research, ongoing refinement of the model and exploration are crucial to adapt to changing dynamics in social media and community viewpoints. Future efforts could focus on enhancing the risk detection model, involving the collection of a more extensive dataset with counterexamples to refine the model's ability to differentiate between various types of risk, thus improving overall accuracy. Additionally, a larger dataset might enable the classification of risks based on its urgency allowing a better utilization of resources.

References

[1] Cano-Marin E, Mora-Cantallops M, Sánchez-Alonso S. Twitter as a predictive system: a systematic literature review. Journal of Business Research. 2023 Mar 1;157:113561.

[2] Cui J, Wang Z, Ho SB, Cambria E. Survey on sentiment analysis: evolution of research methods and topics. Artificial Intelligence Review. 2023 Jan 6:1-42.

[3] Florea M, Potlog C, Pollner P, Abel D, Garcia O, Bar S, Naqvi S, Asif W. Complex project to develop real tools for identifying and countering terrorism: real-time early detection and alert system for online terrorist content based on natural language processing, social network analysis, artificial intelligence and complex event processing. InChallenges in Cybersecurity and Privacy-the European Research Landscape 2022 Sep 1 (pp. 181-206). River Publishers.

[4] Mäntylä MV, Graziotin D, Kuutila M. The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. Computer Science Review. 2018 Feb 1;27:16-32.

[5] Richie G, Thompson H. Natural language processing. AI tools, techniques & Applications, Harper Row, New York. 1984:358-80.

[6] Nandwani P, Verma R. A review on sentiment analysis and emotion detection from text. Social Network Analysis and Mining. 2021 Dec;11(1):81.

[7] Amatriain X. Transformer models: an introduction and catalog. arXiv preprint arXiv:2302.07730. 2023 Feb 12.

[8] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. Advances in neural information processing systems. 2017;30.

[9] Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, Levy O, Lewis M, Zettlemoyer L, Stoyanov V. Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692. 2019 Jul 26.

[10] Loureiro D, Barbieri F, Neves L, Anke LE, Camacho-Collados J. Timelms: Diachronic language models from twitter. arXiv preprint arXiv:2202.03829. 2022 Feb 8.

[11] Hameed Z, Garcia-Zapirain B, Ruiz IO. A computationally efficient BiLSTM based approach for the binary sentiment classification. In2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT) 2019 Dec 10 (pp. 1-4). IEEE.

[12] Kaggle. Arash Keissami, "Twitter Sentiment Analysis: Hatred Speech". Kaggle, 2021. Available from: https://www.kaggle.com/datasets/arkhoshghalb/twitter-sentiment-analysis-hatred-speech/data [Accessed 10 Dec 2023].

[13] Humberside Fire and Rescue Service. Strategic Leadership Team. Humberside Fire and Rescue Service, 2023. Available from: https://humbersidefire.gov.uk/about-us/strategic-leadership-team [Accessed 01 Dec 2023].

[14] Humberside Fire and Rescue Service. Senior management structure. Humberside Fire and Rescue Service, 2023. Available from: https://humbersidefire.gov.uk/about-us/senior-management-structure [Accessed 01 Dec 2023].

[15] Grandini M, Bagli E, Visani G. Metrics for multi-class classification: an overview. arXiv preprint arXiv:2008.05756. 2020 Aug 13.

[16] Hull Daily Mail. Devastating Snaith and East Cowick floods: The story so far. Hull Daily Mail; 2021 Jan 28. Available from:

https://www.hulldailymail.co.uk/news/hull-east-yorkshire-news/devastating-snaith-east-cowick-floods-4930942 [Accessed 10 Dec 2023].