2023 26th International Conference on Computer and Information Technology (ICCIT)
13-15 December, Cox's Bazar, Bangladesh

2023 26th International Conference on Computer and Information Technology (ICCIT) | 979-8-3503-5901-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCIT60459.2023.10441156

Analyzing Public Sentiment on Social Media during FIFA World Cup 2022 using Deep Learning and Explainable AI

Shafakat Sowroar Arnob, M. A. Ahad Shikder, Tashfiq Alam Ovey,
Ehsanur Rahman Rhythm and Annajiat Alim Rasel
Department of Computer Science and Engineering (CSE)
School of Data and Sciences (SDS)
BRAC University
66 Mohakhali, Dhaka - 1212, Bangladesh
{shafakat.sowroar.arnob, m.a.ahad.shikder, tashfiq.alam.ovey, ehsanur.rahman.rhythm}@g.bracu.ac.bd, annajiat@gmail.com

***Abstract*—Analysis of public sentiment is extremely useful for comprehending the responses of the general public during important events, and the FIFA World Cup 2022 was no exception. Within the scope of this study, we used deep learning models such as roBERTa, distilBERT, and XLNet to conduct an analysis of the views that were stated on Twitter during the first day of the tournament. These models were fine-tuned using a comprehensive dataset consisting of 30,000 tweets, which had been preprocessed. The performance of these models was assessed using measures such as accuracy, F1-score, precision, recall, etc. In addition, we used an Explainable AI known as Local Interpretable Model-Agnostic Explanations (LIME) so that we could better understand how model decisions were made in sentiment classification. Our research has shown that roBERTa is an excellent model for classifying sentiment, and it has also shown the significance of interpretability achieved using LIME. Our research enhances the understanding of sentiment analysis during major sports events and suggests future directions for research in this domain.**

***Index Terms*—Sentiment Analysis, Deep Learning, roBERTa, distilBERT, XLNet, LIME**

I. Introduction

Sporting events have long since evolved beyond the concept of straightforward competition to become worldwide gathering places for individuals of many backgrounds and ways of life. Among these noteworthy sports events, the FIFA World Cup stands out as a sight that can't be matched. All countries join together to enjoy the wonderful game of football around this time. The 2022 FIFA World Cup, held in Qatar, marked a significant turning point in the long and illustrious history of this prestigious sport. Aside from the importance it has for sports, the World Cup also serves as a unique lens through which we are able to evaluate the general sentiment of mind of audiences all around the world in real time.

In this digital age, social media platforms have become the most significant for public engagement and discussion. Particularly on Twitter, people from diverse backgrounds share their thoughts and reactions to major events. On the first day of the 2022 World Cup, an extraordinary increase in the amount of online conversation was seen. Moreover, thousands of individuals are communicating with one another about their views, feelings, and even criticisms on Twitter. Twitter posts reflect the current mood and provide an enormous amount of data for sentiment analysis. They represent the public's emotional response to the tournament which shows sentiment overall.

The discipline of sentiment analysis, which is a subfield of Natural Language Processing, gives us the ability to methodically explore and understand the emotions conveyed in written material. We are able to go deeper into the intricate network of tweets pertaining to the World Cup by using the capabilities of deep learning models like roBERTa [1], [2], distilBERT [3], and XLNet [4]. By this, we are able to identify the sentiments that are driving the worldwide discourse around the tournament. These models have shown a remarkable capacity, after being pre-trained on massive text corpora, to accurately express nuanced feelings in a range of different environments. Our dataset of 30,000 tweets from the opening day of the 2022 World Cup was used to do fine-tuning on three different models: roBERTa, distilBERT, and XLNet. Their performance was meticulously evaluated using a broad variety of assessment indicators, which enabled us to compare how well each one captured the sensory nuances.

To provide a global perspective on how people feel, we studied how individuals reacted to the World Cup's opener. We used Explainable AI like LIME [5], to address the intuitive basis for our deep learning models' choices. Interpretability illuminates our models' categorization decisions and the features underneath them.

This research expands our knowledge of the convergence of big sports events, social media, and natural language processing. Moreover, it highlights the relevance of sentiment analysis as a tool for assessing the collective sentiment of global audiences at unusual periods. Our study will help policymakers and stakeholders gain insights and extract necessary information from future big events like FIFA World Cup and other sports events.

979-8-3503-5901-5/23/$31.00 ©2023 IEEE

II. Related Works

Many sentiment analysis studies have examined the importance of user-generated content, notably on Twitter, in connection to major events. The opening day of the FIFA World Cup 2022's emotional impact has been a hot topic for months. These studies illustrate that social media data may reveal popular opinions and attitudes during crises. Analysing massive amounts of user-generated material may reveal trends and reactions.

The author in [6] discusses a Twitter-based study on global sentiment towards Qatar's 2022 World Cup hosting. The research included three stages: before Qatar was chosen as host, after the selection, and during the event. This research found 84% positive Twitter sentiment and 16% negative. The model's 87% accuracy in predicting sentiment in unannotated data is noteworthy.

[7] examined a new dataset using Logistic Regression, Random Forest, Naive Bayes, and SVM classifiers. Logistic regression had the best accuracy (93%), followed by random forest classifier (92%). Using Naive Bayes, the accuracy averaged 88%, whereas SVM-Classifier averaged 93% This shows that machine learning can properly evaluate Qatar World Cup 2022 Arabic-speaking Twitter sentiment.

[8] used Twitter data from the 2014 World Cup in Brazil to assess worldwide opinion. An algorithm extracted emotional words using WordNet's part of speech and context. Sentiment polarities were analysed using naive Bayes, SVM, KNN, and random forest. The greatest AUC was 0.97 for random forest, while naive Bayes had the maximum accuracy of 88.17%. This research illustrates that mining 2014 World Cup Twitter data using NLP and machine learning may disclose people' feelings.

The study analysed football fan sentiment on Twitter [9]. Sentiment analysis collected semantic and syntactic data from GloVe and other word vectors. An emotional lexicon offered context. RBF, SVM, MNB, K-Nearest Neighbours, and XG Boost were used for sentiment analysis. Football supporters' sentiments were best classified by Random Forest in their 2018 FIFA World Cup Twitter experiment.

The authors in [10] studied aspect-level sentiment analysis using transformer-based pre-trained models like BERT and RoBERTa. The aspect-based technique was applied to BERT and RoBERTa models after they had tested them without it. The aspect-based method outperformed standard models by approximately 1%. Among the models tested, the aspect-based BERT model had the highest accuracy and performance.

In addition to sports, elections have been subject to sentiment research. [11] analysed 2022 election-related Twitter exchanges. They used sentiment analysis and topic modeling to determine voter preferences and discussion themes. They used Naive Bayes and Support Vector Machines for sentiment analysis, with Naive Bayes ranking higher at 73% than 69%.

The studies in [12] have attempted to demonstrate connections between shifts in sentiment and big events such as lockdown announcements, demonstrating the concrete impact that significant developments have on public sentiment, which frequently takes the form of mood deteriorations in response to certain occurrences.

Existing word embedding approaches failed to capture sentiment polarity, but sentiment-specific word embeddings (SSWE) have solved this issue. [13] introduced sentiment-specific word embeddings to overcome this limitation. The study shows that SSWE is competitive in sentiment analysis, suggesting that these systems might enhance sentiment categorization.

The author introduces TWITA in [14], the first Italian tweet corpus produced automatically and transferable to any language. On general and topic-specific TWITA datasets, they test sentiment analysis. Their initial Italian polarity database was automatically matched from three resources, using simply a polarity lexicon.

These studies demonstrate the relevance of social media data, sentiment analysis methods, and developing assessment tasks in assessing public feelings, responses, and mood fluctuations, notably during big events and on Twitter.

III. Dataset and Preprocessing

Our dataset for sentiment analysis contains all relevant discourse regarding the FIFA World Cup 2022 on Twitter [15]. It is a large collection that includes 30,000 tweets covering the opening day of the FIFA World Cup 2022. The dataset only includes English tweets that consistently utilise the #WorldCup2022 hashtag. Tweet, Sentiment, Number of Likes, Source of Tweet and Date Created are in this dataset.

TABLE I: Dataset Examples Before Preprocessing

Tweet	**Sentiment**
Ecuador players after 1-0 against Qatar...	neutral
Morgan Freeman too old damn can't believe...	negative
So, I guess we have a #WorldCup2022 opener...	neutral

TABLE II: Dataset Examples After Preprocessing

preprocessed_tweet	**labels**
ecuador players losing qatar tonight fifaworld...	0
morgan freeman old damn cant believe...	1
guess world cup opener ahead...	0

In order to guarantee the accuracy and consistency of the data, we utilized systematic steps in our preprocessing pipeline. There was substantial preprocessing of the tweets. The data was preprocessed by removing user mentions, URLs, stop words, emojis etc. Each word was lowercased to provide uniformity and simplify text matching. Also, text cleaning removed unnecessary letters, symbols, and formatting to improve data quality. Tokenization, which segments text into discrete pieces, was done to prepare it for our deep-learning models. We added a new column to the database called "preprocessed_tweet" so that we could save the text of tweets that had been preprocessed and standardized. Each data point was carefully annotated with sentiment labels, and expressions were sorted into one of three basic groups: neutral (0), negative (1), and positive (2). In table I, we can see the Tweet and Sentiment before preprocessing, and in table II, we can see the preprocessed and labeled tweet data. In Fig. 1 we can visualize that the negatively labeled sentiment data portion is a bit lesser than the positive and neutral data in the dataset.

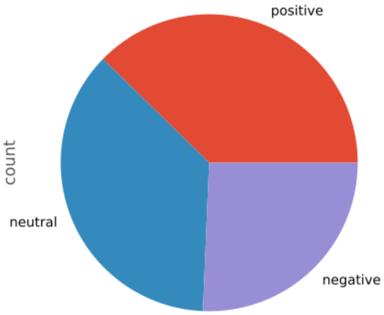


Fig. 1: Pie chart of Total Sentiment Classes in the Dataset.

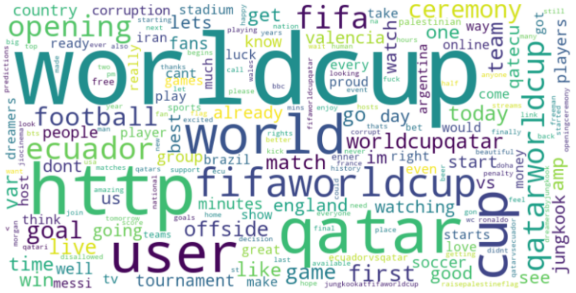


Fig. 2: Word Cloud of the dataset.

To simplify the dataset, we created a word cloud and highlighted the most frequently occurring terms in Fig. 2. This thoroughly created and preprocessed dataset is the foundation for our sentiment research, allowing us to dive deep into user sentiment during the FIFA World Cup 2022.

IV. Methodology

This section discusses sentiment analysis using deep learning models like roBERTa, distilBERT, and XLNet. This part covers justifying our model selection, model architectures, training process, evaluation metrics, and model interpretation utilising Explainable AI like LIME (Local Interpretable Model-agnostic Explanations).

A. *Deep Learning Models Selection*

The exceptional performance of deep learning models such as roBERTa [1], [2], distilBERT [3], and XLNet [4] in a variety of natural language processing (NLP) tasks, such as sentiment analysis, was a primary factor in their selection. Because they were pre-trained on large text corpora, these models can pick up subtle contextual information and nuances. Their accurate representations and transfer learning make them ideal for social media sentiment analysis.

B. *Model Architectures and Fine-Tuning*

To analyse FIFA World Cup 2022 tweet sentiment, each model was fine-tuned. In fine-tuning, pre-trained models were updated by adding a sentiment classification layer. A brief overview of each model's architecture follows:

- **roBERTa:** roBERTa is a modification of the BERT architecture that includes improved training methods. It does this by using a bidirectional transformer architecture and by capturing information about the context from both directions. We began the process of fine-tuning the model by first training it on our preprocessed dataset, then we added a classification layer on top of the roBERTa model that had been pre-trained.
- **distilBERT:** distilBERT is a distilled version of BERT that was developed to maximize productivity without sacrificing performance. It employs a structure that is similar to BERT's but has fewer parameters overall. DistillBERT was enhanced by the addition of a classification layer, which was then fine-tuned using our data.
- **XLNet:** XLNet presents a permutation-based training strategy to capture relationships between all input locations. XLNet is an extension of the convolutional neural network. It makes use of an architecture known as a transformer. We fine-tuned XLNet with a classification layer in order to do sentiment analysis, just as we did with the previous models.

C. *Training Process and Evaluation Metrics*

The preprocessed text was tokenized. The sentences were then encoded into tensor slices for deep learning models. We fine-tuned each model using preprocessed data. We changed the models' hyper-parameters to boost sentiment analysis. To maximise model performance, we carefully selected hyperparameters including learning rates, batch sizes, training epochs etc.

To analyse our models' efficacy, we employed accuracy, F1 score, precision, recall, loss, etc. The percentage of correctly categorised attitudes represents classification accuracy. The F1 score integrates accuracy and recall, making it useful for unbalanced datasets. A model's ability to precisely classify answers as positive or negative is called "Precision". Model training loss reflects how well the model fits data. We can exhaustively analyse the models' sentiment classification performance.

D. *LIME for Interpretability*

We used LIME (Local Interpretable Model-agnostic Explanations), an Explainable AI tool, to explain our deep learning model predictions in addition to performance measurements [5]. LIME illuminates the factors that affect predictions and these models' classification decisions. LIME highlights the relevance of words or phrases during classification to help us understand how the model draws its results.

Fine-tuning cutting-edge deep learning models, rigorous training and testing, and LIME for model interpretation help us understand FIFA World Cup 2022 tweets' sentiments.

V. Experiments and Results

Our sentiment analysis experiments using roBERTa, distilBERT, and XLNet will be discussed in this section. To compare performance, we analyse accuracy, F1-score, precision, and recall. We also evaluate the significance of these metrics in sentiment analysis and how our conclusions may affect FIFA World Cup 2022 sentiment analysis.

Our rigorous experimental technique began with preprocessing the dataset, as described in the "Dataset and Preprocessing" section. For the model input, texts were tokenized and encoded tensor slices were made. We then fine-tuned three distinct models—roBERTa, distilBERT, and XLNet—utilizing specific pre-trained models from Huggingface tailored to each, namely 'cardiffnlp/twitter-roberta-base-sentiment-latest' for roBERTa [1], [2], 'distilbert-base-uncased' for distilBERT [3], and 'sshleifer/tiny-xlnet-base-cased' for XLNet [4]. To improve model performance, hyperparameter tuning was done. This includes warmup steps, weight decay, epochs etc. adjustments. A rigorous examination using accuracy, F1-score, precision, and recall assessed the models' ability to categorise people's sentiments.

TABLE III: Comparison of Evaluation Metrics

Model	Accuracy	F1 Score	Precision	Recall
roBERTa	84.8%	0.83	0.83	0.82
distilBERT	83.5%	0.82	0.81	0.82
XLNet	79.6%	0.72	0.78	0.67

Table III presents the evaluation metrics for roBERTa, distilBERT, and XLNet on the sentiment analysis task. These results showcase the performance of the models in terms of accuracy, F1-score, and precision.

A. Discussion of Metrics

- **Accuracy:** Accuracy can be defined as the proportion of a total number of predictions in which sentiments were accurately identified. During all of our evaluations, roBERTa demonstrated the highest level of precision, with an accuracy level of 84.8%. distilBERT finished in second place with 83.5% of the vote, while XLNet took third place with 79.6%.
- **F1-Score:** In order to determine the F1-score, one must first determine the harmonic mean of accuracy and recall scores. This creates a score that strikes a healthy balance between the two components of the examination. The fact that roBERTa was able to get the maximum possible score on the F1 test (0.83) indicates that it carried out its tasks quite capably in general. The value of distilBERT's F1 score was 0.82, which indicates that it performed very well. XLNet's F1 score was 0.72, which was lower than the competition's score despite the fact that it had a lower accuracy.
- **Precision and Recall:** Precision is the capacity of the model to properly categorize positive or negative feelings, while recall is the ability of the model to detect actual positive situations. Both the accuracy (0.83) and recall (0.82) that roBERTa displayed were quite high, which is an indication of a well-balanced categorization. distilBERT also displayed balanced accuracy and recall, in contrast to XLNet, which had somewhat lower values for both parameters. This may be because of the use of the tiny version of XLNet pre-trained model.

B. Implications and Insights

The selection of evaluation metrics is crucial in sentiment analysis since it determines how effectively a model captures positive, negative, and neutral thoughts about a subject. Our study has led to some important FIFA World Cup 2022 findings. RoBERTa outperforms other models in accuracy and F1-score, proving its supremacy in sentiment analysis for this event. High accuracy provides precise sentiment categorization, and high recall identifies real positives, making sentiment analysis reliable. Precision and recall are crucial, especially when working with unbalanced datasets since precision enables correct sentiment categorization and recall helps detect true positives. Our findings show that optimising models by adjusting hyperparameters and employing additional tactics like warmup steps and weight decay improves model performance. LIME for model interpretation will also help us understand sentiment categorization logic and model decision-making. Despite a great performance by roBERTa. These findings show that model selection and optimisation tactics may significantly impact sentiment categorization, enriching our understanding of global sentiment during major events.

VI. LIME Interpretability For Models

We used Local Interpretable Model-Agnostic Explanations (LIME) [5], an Explainable AI method, to better understand the behaviour of our deep learning models. The main objective was to illuminate the factors that influence our models' predictions to better understand their decision-making processes.

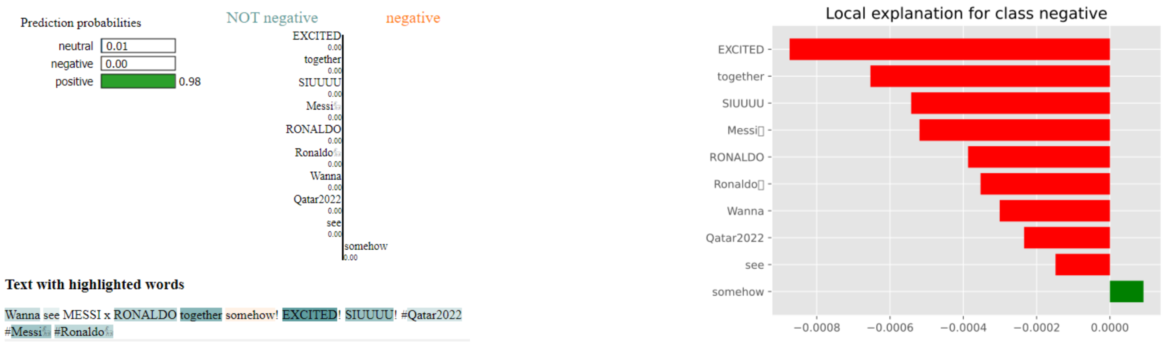
To illustrate the application of LIME, we selected a sample instance for classification by our three models — roBERTa, distilBERT, and XLNet. The instance tweet was:"Wanna see MESSI x RONALDO together somehow! EXCITED! SIUUUU! #Qatar2022 #Messi #Ronaldo". From figures in 3a, 4a and 5a, the classifications by the models for this instance were as follows:

- **roBERTa:** Positive(0.98), Neutral(0.01), Negative(0.0)
- **distilBERT:**Positive(0.96), Neutral(0.02), Negative(0.02)
- **XLNet:** Positive(0.34), Neutral(0.34), Negative(0.32)

Importantly, in 3b, 4b and 5b, values are in the negative (Red) side mean "Not Negative" and values are in the positive (Green) side mean "Negative" as the plot is made for Class Negative.

From fig. 3a and 3b, we can see that it highlighted "somehow" as the only Negative word in that instance and highlighted all the other words as Not Negative, and "EXCITED" carries the highest probability value for being Not Negative and in second position comes to the word "together".

From fig. 4a and 4b, it also highlighted "somehow" as the only Negative word in that instance and highlighted all the



(a) LIME Explanations for roBERTa Model

(b) roBERTa: Probability Plot of words in the Instance

Fig. 3: LIME interpretation and plot of words for roBERTa

other words as Not Negative, and “EXCITED” carries the highest probability value for being Not Negative and second position comes the word “RONALDO”.

From fig. 5a and 5b, we can see that it highlighted 3 words - “SIUUUU”, “EXCITED”, and “x” as the Negative words in that instance and highlighted all the other words as Not Negative, and “#Ronaldo” carries the highest probability value for being Not Negative and second position comes to the word “Qatar2022”.

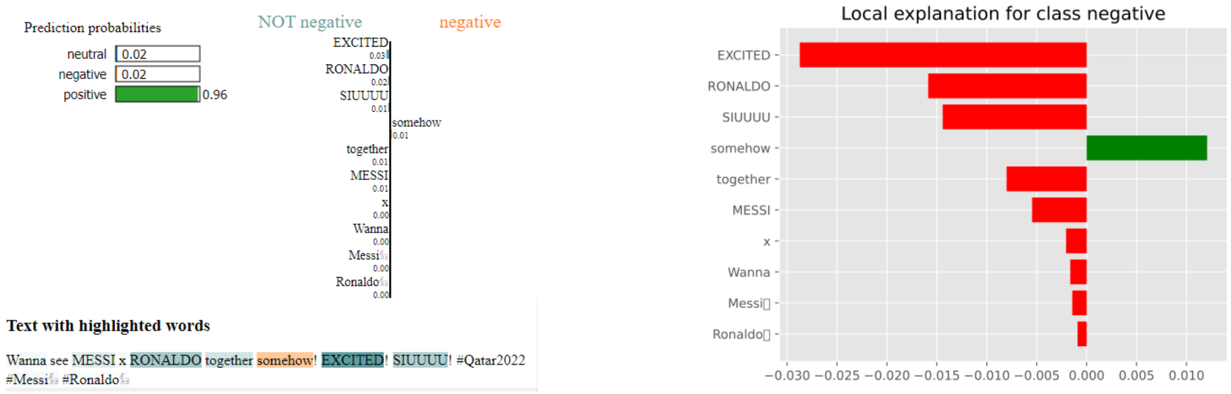
LIME included explanations for the classifications it produced for each model, drawing attention to the significant role that individual words have in determining the outcome of the prediction. Notably, the word “somehow” was consistently recognized as the biggest driver of negative emotion across roBERTa and distilBERT models. This is an interesting finding. On the other hand, words such as “EXCITED” and “together” held the greatest probability of not being connected with a negative mood in either roBERTa or distilBERT. One more finding is that XLNet model identifies the “EXCITED” as Negative word while this word carries the highest probability values for being Not Negative in both roBERTa and distilBERT.

These LIME explanations illuminate our models’ delicate decisions and how words affect sentiment predictions. Additionally, frequent identification of “somehow” as a negative feature demonstrates the context-sensitive nature of the models. LIME’s interpretability has helped us understand model predictions’ logic and behaviour.

VII. Conclusion, Limitation and Future Work

In conclusion, the findings of our study have offered significant insights into the analysis of sentiment during important sporting events, with a particular emphasis on the FIFA World Cup 2022. In order to conduct an analysis of the feelings that were shared on Twitter in the course of the event, we made use of the capabilities offered by deep learning models, in particular roBERTa, distilBERT, and XLNet. In addition, we used LIME, which stands for Local Interpretable Model-Agnostic Explanations, in order to improve the interpretability of the judgments that these models produced.

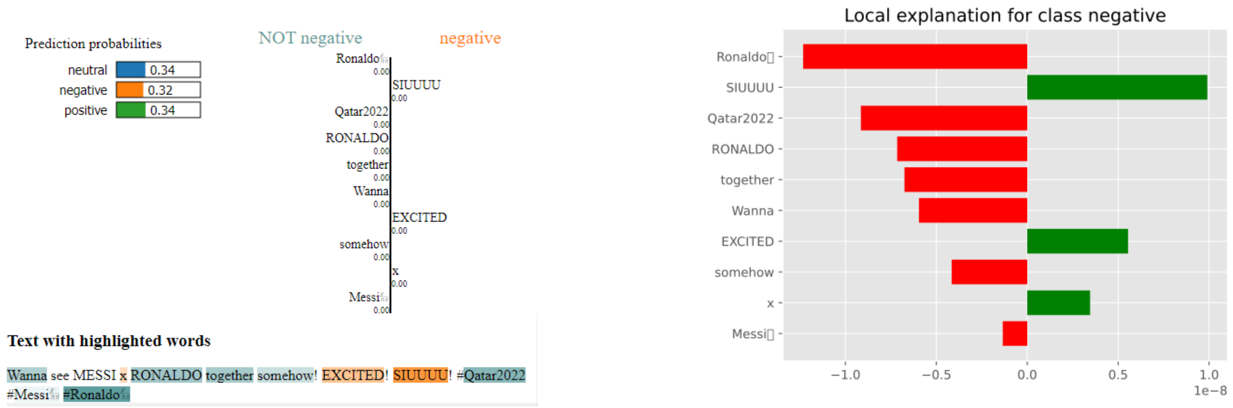
Our most important discoveries illustrate the efficacy of deep learning models, in particular roBERTa, in effectively collecting attitudes in an environment characterized by constant change and activity on social media. These models achieved excellent levels of accuracy and F1 scores, demonstrating their capacity to comprehend and categorize the feel-



(a) LIME Explanations for distilBERT Model.

(b) distilBERT: Probability Plot of words in the Instance.

Fig. 4: LIME interpretation and plot of words for distilBERT



(a) LIME Explanations for XLNet Model.

(b) XLNet: Probability Plot of words in the Instance.

Fig. 5: LIME interpretation and plot of words for XLNet

ings stated by users in the course of a significant international athletic event.

Additionally, LIME has illuminated these models' decision-making processes, providing vital insights into the main features and terms that categorise emotion. Both academics and practitioners need interpretability to understand model behaviour and make appropriate assessments.

However, our study faced several limitations. The dataset is limited to English tweets from the FIFA World Cup 2022 opening day. The global response to the incident may be better understood if the dataset included tweets in several languages and was gathered over a longer period of time. LIME provides valuable insights, but further research is needed to determine how interpretability approaches might be used in sentiment analysis.

For future research, researchers may generalise deep learning models across sports and languages. During events, real-time sentiment analysis may reveal audience emotions. More advanced interpretability tools and methodologies may help us understand model decisions.

Our work illuminates sentiment analysis's utilisation during major sports events and the need for cross-contextualizing models. In the digital age, we may enhance our understanding of global opinion by overcoming boundaries and exploring new avenues.

References

[1] J. Camacho-Collados, K. Rezaee, T. Riahi, A. Ushio, D. Loureiro, D. Antypas, J. Boisson, L. Espinosa-Anke, F. Liu, E. Martínez-Cámara *et al.*, "Tweetnlp: Cutting-edge natural language processing for social media," in *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. Abu Dhabi, UAE: Association for Computational Linguistics, Dec. 2022, pp. 38–49. [Online]. Available: https://aclanthology.org/2022.emnlp-demos.5

[2] D. Loureiro, F. Barbieri, L. Neves, L. Espinosa Anke, and J. Camacho-collados, "TimeLMs: Diachronic language models from Twitter," in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 251–260. [Online]. Available: https://aclanthology.org/2022.acl-demo.25

[3] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, "Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter," *ArXiv*, vol. abs/1910.01108, 2019.

[4] Z. Yang, Z. Dai, Y. Yang, J. G. Carbonell, R. Salakhutdinov, and Q. V. Le, "Xlnet: Generalized autoregressive pretraining for language understanding," *CoRR*, vol. abs/1906.08237, 2019. [Online]. Available: http://arxiv.org/abs/1906.08237

[5] M. T. Ribeiro, S. Singh, and C. Guestrin, ""why should I trust you?": Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*, 2016, pp. 1135–1144.

[6] S. Dewi and D. B. Arianto, "Twitter sentiment analysis towards qatar as host of the 2022 world cup using textblob," *Journal of Social Research*, vol. 2, no. 2, pp. 443–455, 2023.

[7] M. Faisal, Z. Abouelhassan, F. Alotaibi, R. Alsaeedi, F. Alazmi, and S. Alkanadari, "Sentiment analysis using machine learning model for qatar world cup 2022 among different arabic countries using twitter api," in *2023 IEEE World AI IoT Congress (AIIoT)*. IEEE, 2023, pp. 0222–0228.

[8] R. Patel and K. Passi, "Sentiment analysis on twitter data of world cup soccer tournament using machine learning," *IoT*, vol. 1, no. 2, p. 14, 2020.

[9] S. U. Hegde, S. B. Basapur *et al.*, "Distilbert-cnn-lstm model with glove for sentiment analysis on football specific tweets." *IAENG International Journal of Computer Science*, vol. 49, no. 2, 2022.

[10] G. R. Narayanaswamy, "Exploiting bert and roberta to improve performance for aspect based sentiment analysis," 2021.

[11] R. E. Demillo, G. Solano, and N. Oco, "Philippine national elections 2022: Voter preferences and topics of discussion on twitter," in *2023 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*. IEEE, 2023, pp. 724–729.

[12] A. Kruspe, M. Häberle, I. Kuhn, and X. X. Zhu, "Cross-language sentiment analysis of european twitter messages duringthe covid-19 pandemic," *arXiv preprint arXiv:2008.12172*, 2020.

[13] D. Tang, F. Wei, N. Yang, M. Zhou, T. Liu, and B. Qin, "Learning sentiment-specific word embedding for twitter sentiment classification," in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2014, pp. 1555–1565.

[14] V. Basile and M. Nissim, "Sentiment analysis on italian tweets," in *Proceedings of the 4th workshop on computational approaches to subjectivity, sentiment and social media analysis*, 2013, pp. 100–107.

[15] T. AI, "Fifa world cup 2022 tweets," Dec 2022. [Online]. Available: https://www.kaggle.com/datasets/tirendazacademy/fifa-world-cup-2022-tweets