ZKG INTERNATIONAL

ISSN: 2366-1313

SENTIMENT ANALYSIS ON FOOTBALL SPECIFIC TWEETS USING CNN-LSTM METHOD

[1]B. NARESH, [2]S. VENKATESWARA RAO

[1]Lecturer, Dept. of CSE, Megha Institute of Engineering and Technology for Women, naresh.b546@gmail.com

[2]HOD, Dept. of CSE, Megha Institute of Engineering and Technology for Women, sabbineniv@gmail.com

Abstract: Sports fans create many tweets that mimic their opinions and feelings about events that take place during various sports activities. Because of the popularity of football activities, we focus on analyzing the emotions expressed by football fans via Twitter in this table. These tweets reflect a change in fans' sentiment as they watch the game and react to events within the game, such as scoring goals, penalties, etc. This paper uses a global vector word embedding (GloVe) approach that generates a word vector and takes advantage of the facts. In addition to GloVe, we also collect a Sentiment Lexicon as additional statistics. GloVe-assisted word vectors and emotional dictionaries are inputs for the proposed CNN-LSTM hybrid deep mastering model. The proposed version of CNN-LSTM combines the blessings of CNN and LSTM. The CNN derives features from the word embedding that reflect short-term emotional dependence, even when LSTM establishes long-term emotional relationships between words. This article extensively uses random forest with algorithm knowledge, support vector machine, multi-anonymous newbies, K-Nearest Neighbors (KNN), and XG Boost for diagnosis and emotion categories. We reviewed the performance of the proposed CNN-LSTM hybrid with Glow Word Embedding Approach with the 2018 FIFA World Cup Tweet Dataset, our test effects showing 85.46% and 92.56% validation and test accuracy, respectively.

Keywords: Sentiment analysis, Convolutional neural network-Long short term memory, word embedding.

I. INTRODUCTION

Emotion assessment is a growing area of analysis that is currently gaining popularity among researchers. Emotion analysis is finding out a person's values or feelings about an entity [1]. The dramatic increase in emotion diagnosis is compounded by the diversity of importance of social networking packages, such as Twitter, that allow people to share their thoughts and opinions on various topics actively. Activities include, but are not limited to, leaving comments or opinions about products, events, movies, politics, or other presentations. As a result, there is a tremendous amount of emotional data on social media. Removing these valuable logs will benefit many applications [2]. For example, companies want to refine buyers' reviews of their products, which is a good way to gather data on consumer satisfaction levels and choose which product components need improvement. Furthermore can use these facts to compare customer feedback about competitive service or product providers. Also, social media data software can be enhanced as a proxy device for real-time experiments with streaming events. During these events, people create a lot of posts expressing their opinions about the situations that arise during video games. An analysis of the emotions expressed in user posts in football video games can be useful in identifying whether a game has aroused completely negative emotions among the spectators and whether its use has led the government to match. It can be done to warn of possible future riots [3]. As a result, in these paintings, we become familiar with analyzing emotions in conversations about football on Twitter.

Football events with FIFA World Cup and UEFA Champions League attract millions of people on the field. Football is the most popular sport in 8000 different genres, with 3.5 billion spectators worldwide and approximately 265 million players. The success of football in gaining public and media attention is astonishing. During the 2018 FIFA World Cup, fans posted 900 million tweets about Event 3. In addition, Twitter's professional blog pointed out that 115 billion tweets were recorded during the 2018 FIFA World Cup. People reacted to what happened during the competition, such as the desire to score goals, players Predicting the accidents, or the outcome of the next matches. It shows that Twitter has become a place for football fans to talk about their strengths, weaknesses, and events during matches and express opinions and feelings with the praise of their team and opponents. Assessing and summarizing the perception of such a large-scale emotional response will allow us to gain insight into how a person reacts to emotionally severe events, including victory or defeat, and how we communicate when we talk. Events unfold over time. The determination to review this vast amount of statistics stems from the fact that, unlike traditional documents that can be proven and well written, social information is written in informal languages, including words and abbreviations. There are also spelling and grammar errors, and this is short-lived. In addition, football fans chat about sports on social networks. This soccer conversation may seem like a new language to non-fans, especially when combined with everyday phrases.

Furthermore, emotions expressed with the help of football fans are often misused, which makes the analysis more difficult. From the point of view of emotional analysis, the use of fashionable emotional classification in football conversations should lead to confusion in learning. For example, "That tall bomb was sick!" It represents an effective emotion in the realm of football, although the phrases "bomb" and "sick" are associated with negative emotions in the standard context. Similarly, this tweet: "Barcelona did it !!! Holy shit !. With a sense of quality in which the fans are happy for their team, the question remains whether we have an effective way of expressing the emotions expressed by football fans. How to review and adjust settings during events?.

With the rapid development of e-commerce platforms and the Internet, consumers increasingly use e-commerce to purchase merchandise or access services. Online shopping structures allow consumers to choose products from various styles. Consumers buy the product they want, but it causes inconvenience to the buyer due to the product and original product description, unacceptable first-class, merchandise failure after the transaction, and product delivery.

Customers use online social networks (OSN), including Twitter, to share, explain and share their online shopping experience with others. Tweets on Twitter provide users feedback on political events, social events, sports events, and products. It can be very important to examine consumers' sentiments, thoughts, or opinions that have purchased products or received offers and have completed rigorous testing of the product or product or service acquired online or on e-commerce platforms. Currency or emotion detection methods are widely used in the enterprise inventory market to express buyer sentiment on the product, stock, supplier, and market trends.

Emotion analysis is the most common tool for discovering the basic gift of emotion in a message, text, or comment. Emotions can be expressed by the common man or an expert in the field. It can write Emotions in a natural language in an unstructured form or format, and they may be literal and grammatically incorrect. Still, it contains the user's emotions, point of view, and hidden diagnosis. Currency buyers can have voices, reviews, survey responses, online offers, and Twitter tweets.

II. REVIEW OF LITERATURE

Few studies have been conducted towards football-specific sentiment analysis.

Barnaghi et al. [4] utilized a logistic regression algorithm to learn polarity (positive and negative) classifier based on Uni-gram and Bi-gram features. They achieved an accuracy of 72% using the Uni-gram feature. In a similar manner, authors combined N-gram features with external lexicon-based features to improve the performance of the Bayesian logistic regression classifier. The best performance of their proposed method was obtained through combining Unigram and Bi-gram features which resulted in an accuracy of 74%. Both and first used manually labeled tweets (4,162 tweets) to build their sentiment model. Then they collected tweets during the FIFA World Cup 2014 where they utilized the learned sentiment classifier to find correlation between sentiment and major events occurred during the competition.

Alves et al. [5], proposed sentiment analysis method for football related tweets written in Portuguese. In order to train their sentiment models, they collected tweets during the 2013 FIFA Confederations Cup. The tweets are labeled based on two methods: automatically based on the positive and negative emoticons included in the tweets, and a random sample of 1,500 tweets which were manually labeled. The best performance was achieved by SVM (accuracy of 87%) when trained and tested on the automatically labeled data. Yet, this accuracy dropped to 66% when trained and tested on the manually labeled dataset. On other hand, **Gratch et al. [6]** used lexicon-based features to train Naïve Bayes algorithm on SemEval 2014 dataset. They considered the problem of sentiment analysis as classifying a tweet into positive, negative or neutral classes. They proposed training the sentiment classifier on a manuallylabeled general dataset, then using the model to identify sentiment in football related tweets. To evaluate the sentiment model performance on football data, they manually labeled 154 tweets related to the FIFAWorld Cup 2014. Their results showed a strong correlation between the ground truth and the algorithm results. However, the validation set consists of such a small number of tweets, it may not reflect the overall performance of the sentiment model on football related tweets.

Aloufi et al. [7] trained the SVM classifier on the FIFA World Cup 2014 dataset utilizing N-gram features and different lexicon-based features. The dataset used for training is automatically labeled, which means it is vulnerable to incorrect labeling. Their proposed method achieved an accuracy of 85% in classifying tweets into positive, negative, or neutral classes.

ISSN: 2366-1313

To overcome the limitation in previous works, we only suggest increasing football's emotional rating, which can successfully understand the emotions in tweets related to football. To do this, we support a new football dataset that is manually categorized to aid in analyzing the emotions of football tweets. In addition, we develop a new emotional dictionary using the full corpus method. The concept was first incorporated into our previous paintings, which included the construction of an emotional dictionary and a classification using an automatically ranked football dataset. In this article, we take a step further by examining the performance of various learning algorithms using the creation of a manually classified dataset and the selection of functions specifically for the football field.

The meticulous survey of the current state of the art in everyday literature provides more insight into assessing emotions and classification. A complete survey is conducted on the set technique to capture the emotions in this segment, especially in football. Emotional assessment is used in various domains, including recommendation systems, decisions, and opinion support structures. Accurate assessment of emotions plays an important role in e-commerce business success, government policy formulation, stock market trends, etc. Emotion analysis is a point of interest for researchers. Embedded phrases represent text, and word embedding represents textual content in terms of semantics rather than emotions. Text content with emotional statistics is useful for the accurate assessment of emotions. The accuracy of emotional analysis is enhanced by using emotional vocabulary and word embedding, word space, and terms that include ambiguous words or emotions.

III. PROPOSED METHODOLOGY

A. Materials and Methods

To conduct experiments, tweets related to the 2018 FIFA world-cup are gathered using Twitter API. Our dataset include N tweets, it is represented as Dataset = T1 , T2 , T3 , ..TN , where n is 10,007 number of tweets considered for test and training dataset. Each tweet is manually annotated and attributed a sentiment to each tweet. For every tweet following sentiment class is assigned manually:

Positive sentiments: Goal is priceless, amazing player, what a shot, spectacular goalkeeping, Best team in the world right now. What a free-kick, the man for the big occasion, scores the winning penalty to send into the quarter-finals where they will face Russia. Negative sentiments: cowardly coach, Boredom, Hate player who made several mistakes, Sadness of game, the worry of strong opponent team, Only 1 penalty has been saved by the goalkeeper. Neutral sentiment: Register with Coral via the link below to register for Morocco. The dataset is fission into a training, test and validation test. The training dataset is engaged to train the proposed model. A test set is used to execute experiments. The classification report obtained from this dataset. To validate performance parameters such as accuracy, precision, sensitivity, specificity the validation test is conducted. Figure 1 shows the number of tweets considered for our experiments, many positive tweets are more compared to the number of negative tweets. Figure 2 depicts the distribution of tweets in the dataset, it is observed that the average length of a tweet is about 15 to 16 and few tweets have a length of 40 to 50.

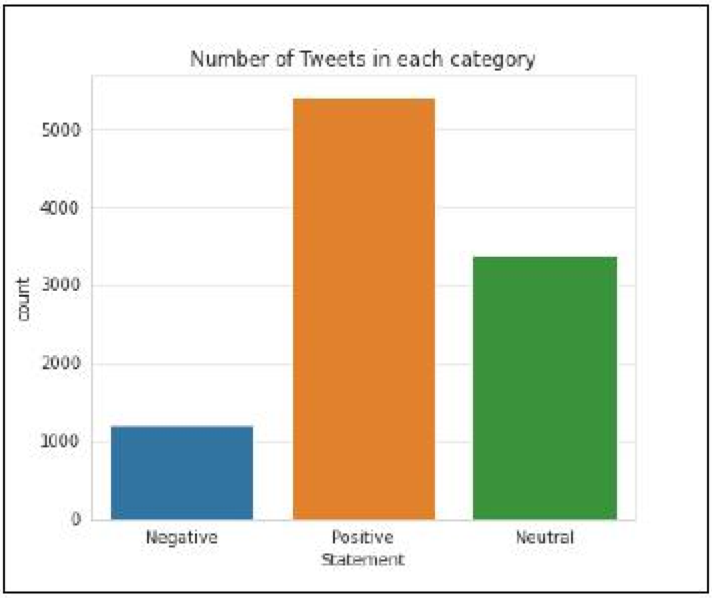

Fig.1 Number of Tweets considered during conduction of experiments

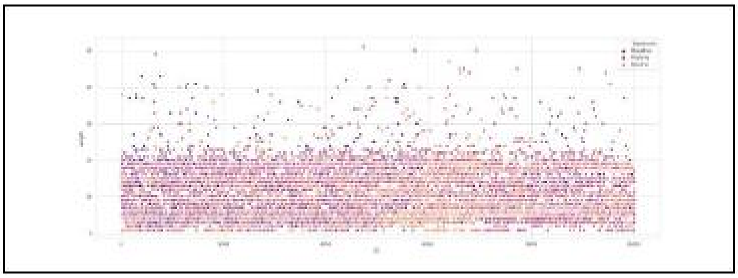

Fig.2 Tweets Distribution Order

B. Data Pre-processing

The tweets posted by user or football fans on Twitter contains hash-tags, abbreviation, slang words, URL, unstructured and incomplete statements. These are noise in tweets posted on

Twitter. The incomplete, unstructured tweets, compressed and noise in tweets are preprocessed using steps shown in figure 3. All uppercase characters were converted to lowercase, discarded link in the tweets, convert @username to username, eliminated multiple spaces, punctuation, number, unrecognized characters, corrected spelling mistakes. These pre-processing steps are performed on our dataset and it helps to increase the performance of our proposed model.

C. Word Embedding

To avoid obfuscate representation of words and apprehend semantic, syntactic, co-occurrence information for words global log bi-linear model is considered in this paper. Global log bi-linear blends the benefits of the word2vec skip-gram model that prodigiously perform the task of word analogy, with a latent semantic analysis that leverage statistical information. The similarity between words is represented as indices in lexicon or vocabulary. Mikolov et al. proposed Skimgram which perform better on analogy task but fail to make use of the statistics of corpus since the model was not trained on global co-occurrence count. Pennington et. al. proposed leastsquare Global vector model, the model is trained on the count of global word-word co-occurrence statistical information. The model produces a word vector with a significant substructure.

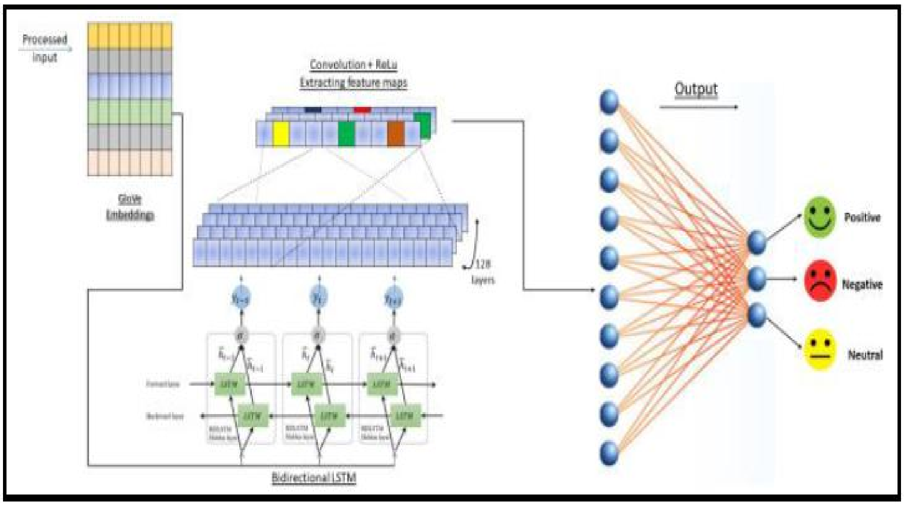

Fig.3 Proposed hybrid CNN+LSTM model with Glove vector

D. Convolutional Neural Network Model

The CNN has the ability to capture close semantic relationship in local regions of text, but LSTM has the capability to capture long-term semantic dependence between the words and

ISSN: 2366-1313

sequences of words. Generally, the convolution layer of CNN performs feature extraction with help of filters; the non-linear features are extracted with the assistance of the ReLU activation function. The feature map dimension is reduced by the pooling layer, which helps to compute cost in subsequent layers and represents the feature of tweet text effectively. A fully connected layer represents the structure between the features of tweets and the anticipated categories. Figure 3 representation functionality flow between the distinct tier or phases of a CNN model and classify tweets into three categories. The new set of tweets serves as an input to CNN. on completion of training, the final layer of CNN categorizes input text into distinct classes and assigns a probability for each class.

E. LSTM

Integration of CNN and LSTM results in an efficient sentiment classifier. CNN has the capability to extract a close semantic relationship between words. While LSTM performs well and solves the problem of long semantic dependence between words. In this research work, the tweets of the dataset are represented using a Global vector. Glove vector produces word vector with significant substructure and it is based on global word co-occurrence statistical information. The word vector produced by Global vector is inputs of CNN and LSTM model. And the extracted features. Figure 4 illustrates the architecture of the proposed hybrid CNNLSTM model. The glove vector word embedding is input for CNN and LSTM model. By integrating the local dependence from CNN and long-term dependence from the LSTM model, the hybrid model can classify tweets according to sentiment. LSTM is a popular and widely accepted RNN model. First, truncate the input sequence and perform padding to have all inputs with the same length. The embedded layer supports a vector length of 100 to represent each word. The next layer of LSTM has 100 memory units and the output layer support 13 output values. The softmax is used as an activation function.

The outputs of the feed-forward are considered as input to the neurons. The output at a neuron depends on present input at time series t and it establishes a relationship between two input values. In LSTM, input data is related to each other. The word embedding (xt) and word sentiment embedding(st) are combined as input to LSTM. Figure 3 shows the standard LSTM with wt, Ot as current the input and output respectively at different time interval t. Figure 5 illustrate proposed LSTM model. The model has three essential switch : input switch it, output switch Ot , and forget switch ft. The input is entered through input switch and output is obtained through output switch at time t. The comparison of present input with prior stature information (h(t1)) and present input(xt) is decided at forget switch. The method of updating memory cell (ci) and current latency values ht is decided by three gates

IV. EXPERMENTAL RESULTS

The results of the proposed hybrid CNN model with the GloVe approach is compared with well-known machine learning algorithms such as Random Forest, Support Vector Machine, Multinomial Nave Bayes, K-Nearest Neighbours(KNN), and XG Boost. These entire machine learning algorithms are known for sentiment analysis. For our experiment, we considered 10007 English tweets specific to the 2018 FIFA world cup. All tweets of the dataset are annotated manually. We carried out experimental on Intel i5, 32GB RAM, GTX 1080 Ti graphic card and Python 3.5 is used as a development programming language. The Scikit-learn, Tensorflow Python libraries are used while developing the proposed model. We evaluated the performance of the proposed model with various performance evaluation metrics such as precision, recall, F1 score, validation accuracy, testing accuracy, sensitivity, specificity, and validation loss are used to assess the performance of the proposed model. The proposed hybrid CNN-LSTM model is executed 20 times and obtained classification results. Figure 6 depicts the results of training, testing classification accuracy for the proposed hybrid CNN-LSTM model. The proposed model achieves an accuracy of about 85% in the classification of tweets into positive, negative, neutralbased on usage CNN- LSTM with GloVe word embedding.

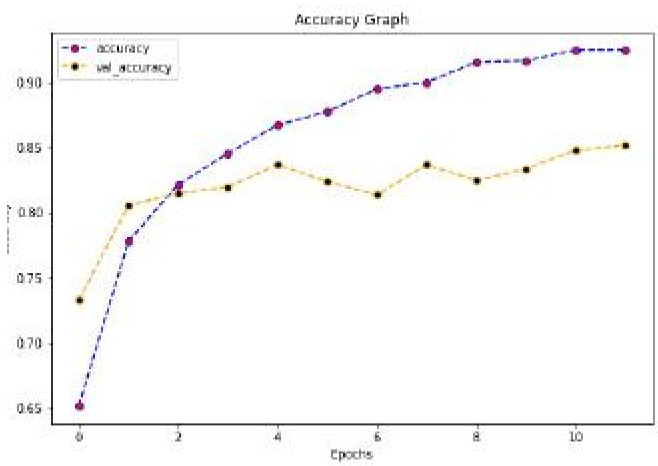

Fig. 4. Classification Accuracy of proposed hybrid CNN-LSTM model

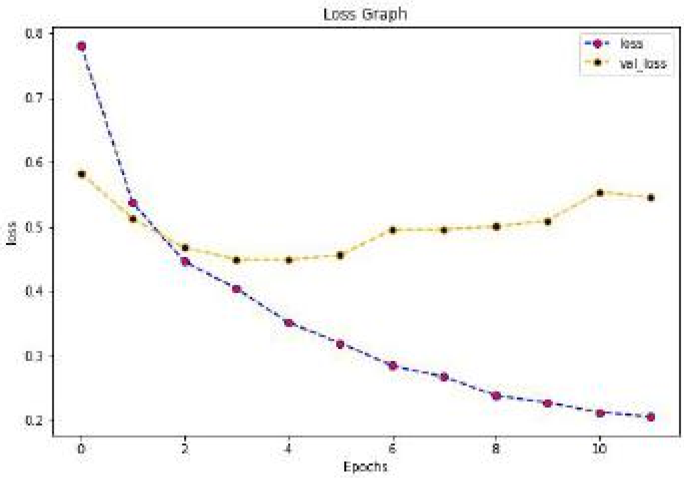

Fig.5 Classification Loss of proposed hybrid CNN-LSTM model

The proposed model extracts sentiment information, short and long-term dependence between the words of tweets. The GloVe construct has high-quality word embedding for a small and large set of unlabelled, incomplete tweets with help of co-occurrence and can confiscation semantic information of words in tweets.

Figure 5 illustrates the loss is the classification of tweets. It is observed in figure 5 that loss has increased slightly because of incomplete sentences in the dataset. For incomplete sentences, GloVe constructs word embedding with help of co-occurrence and semantic information of words in tweets.

Evaluation of Performance Parameters

Category	Precision	Recall	F1-score	Support
Positive	0.90	0.89	0.89	560
Negative	0.78	0.63	0.70	115
Neutral	0.78	0.83	0.81	326
			Accuracy:	0.85

V. CONCLUSION

The traditional word embedding technique leverage semantic information but fails to capture the sentiment of words. With GloVe word embedding approach, sentiment embedding of a word is created. The word vector with semantic and its sentiment embedding are fed as input to the proposed hybrid CNN-LSTM model that performs sentiment analysis. For our experiment we have used 10007 tweets specific to the 2018 FIFA world cup and tweets are manually annotated. From experiment results, it is observed that blended CNN and LSTM

ISSN: 2366-1313

model with GloVe word embedding approach provides a robust and comprehensive and robust model for sentiment classification.

REFERENCES

[1] B. Liu, "Sentiment analysis and opinion mining," Synthesis lectures on human language technologies, vol. 5, no. 1, pp. 1–167, 2012.

[2] T. Al-Moslmi, N. Omar, S. Abdullah, and M. Albared, "Approaches to cross-domain sentiment analysis: A systematic literature review," IEEE Access, vol. 5, pp. 16 173–16 192, 2017.

[3] D. Stojanovski, G. Strezoski, G. Madjarov, and I. Dimitrovski, "Emotion identification in fifa world cup tweets using convolutional neural network," in 2015 11th International Conference on Innovations in Information Technology (IIT), Nov 2015, pp. 52–57.

[4] A. Giachanou and F. Crestani, "Like it or not: A survey of twitter sentiment analysis methods," ACMComputing Surveys (CSUR), vol. 49, no. 2, p. 28, 2016.

[5] X. Fu, J. Yang, J. Li, M. Fang and H. Wang, "Lexicon- Enhanced LSTM With Attention for General Sentiment Analysis", IEEE Access, vol. 6, pp. 71884-71891, 2018, doi: 10.1109/ACCESS.2018.2878425.

[6] H. T. Phan, V. C. Tran, N. T. Nguyen and D. Hwang, "Improving the Performance of Sentiment Analysis of Tweets Containing Fuzzy Sentiment Using the Feature Ensemble Model", IEEE Access, vol. 8, pp. 14630-14641, 2020, doi: 10.1109/ACCESS.2019.2963702.

[7] P. Barnaghi, P. Ghaffari, and J. G. Breslin, "Text Analysis and Sentiment Polarity on FIFA World Cup 2014 Tweet", In Proceeding of ACM Conference on SIGKDD, vol. 15, pp. 10- 13, 2015.

[8] Prasadu Peddi (2020), MINING POSTS AND COMMENTS FROM ONLINE SOCIAL NETWORKS, Turkish Journal of Computer and Mathematics Education, Vol 11, No 3, pp: 1018-1030.

[9] P. Barnaghi, P. Ghaffari and J. G. Breslin, "Opinion Mining and Sentiment Polarity on Twitter and Correlation between Events and Sentiment", In Proceedings of IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService), Oxford, pp. 52-57, 2016.

[10] Prasadu Peddi (2019), Data Pull out and facts unearthing in biological Databases, International Journal of Techno-Engineering, Vol. 11, issue 1, pp: 25-32.