Evaluation of Vader and MultiLingual sentiment analyzers for opinion analytics using graphical illustrations

Taiwo Olaleye[1], Peter Ugege[2], Olalekan Okewale[3], Olatunji Akinade[4], Segun Akintunde[5], Chidi Akparanta[6]

[1, 5] Computer Science Department, Federal University of Agriculture Abeokuta, Nigeria
[2] Information & Communication Technology Centre, Forestry Research Institute of Nigeria
[3] Directorate of Pan African University for Life & Earth Sciences Institute, University of Ibadan, Nigeria
[4] Department of Cybersecurity, University of the West of England, Bristol
[6]Computer Centre & Services, Federal College of Education Abeokuta, Nigeria

Abstract

Sentiment analysis have been prominently employed for opinion analytics across different use cases which seeks to compute polarity values on opinions expressed in textual documents. The Vader-based lexicon approach is widely deployed especially for tweet corpus with emoticon elements. However, there is need to ascertain the peculiarities of the MultiLingual sentiment analyzer alongside the Vader approach towards improving future studies in their use case domain. The Twitter application programming interface was employed in this study to extract public opinions in two corpus, on the Africa Cup of Nations matches involving Nigeria. Experimental results shows no one-to-one mapping between the sentiment scores returned by the two analyzers while the MultiLingual analyzer proves to be reputable for analyzing tweets with shorter phrases. Tokens returned as the most weighted in the two corpus, as analyzed by the two methodologies, likewise shows obvious contrasts in their weights. Two Nigerian players were returned as prominent topics from the two corpus by the topic modelling phase of the study while in MultiLingual analysis, local dialect tokens outweighs other English unigrams, unlike in the Vader-based lexicon approach.

Keywords

Vader analysis, MultiLingual analysis , Football tweets, Opinion mining

1. Introduction

The popularity of footbal across the globe and internet access has provided a platform for cross-fertilization of ideas and opinion moulding through social media platforms, as footbal remains the most watched sport and played in over two hundred countries all over the world [1]. Moreover, the recent lift of ban on Twitter by the Nigerian government on 13th January, 2022 came at a time the bi-annual African Cup of Nations was starting. Young tweets consequently trooped to Twitter to kickstart their support for the nation's national team with different hashtags for different matches. The big opinion corpus generated before, during, and after each group stage matches trends on Twitter for several hours with divergent topical issues encapsulated in individual tweets. As a popular social microbloggin tool employed by nationals and state actors for information dissemination and national integration, Twitter serves as a village square for global, national, and sub-national digital town hall [2]. Opinions freely expressed by users could be subjected to an opinion mining analytics by state actors to gauge the national mood and probably influence state policies. Football being a capital intensive sport, and a tool

COLINS-2022: 6th International Conference on Computational Linguistics and Intelligent Systems, May 12–13, 2022, Gliwice, Poland.
EMAIL: agsobaolaleyetaiwo@gmail.com (T. Olaleye); pugege@yahoo.com (P. Ugege); okewale.akinbosoye@gmail.com (O. Okewale); olatunji2.akinade@live.uwe.ac.uk (O. Akinade); seggzy2001@gmail.com (S. Akintunde); uzochidi2018@yahoo.com (C. Akparanta)
ORCID: 0000-0001-6222-7575 (T. Olaleye); 0000-0002-2722-0076 (P. Ugege); 0000-0003-2272-8660 (O. Okewale); 0000-0002-6126-6801 (S. Akintunde); 0000-0002-9493-1134 (C. Akparanta)

for socioeconomic and national cohession, scrapping of opinions expressed in tweets for analytics would serve stakeholders in different capacities [3]. Consequent upon the rude shock that trails the suden exit of Nigeria from the competition, twitter has been agog with an upsurge of opinions on the remote courses of the goal loss to the Tunisians despite a previous win against Guine Bisau. This study therefore employs the instrumentality of Natural Language Processing to analyse shades of opinions expressed on twitter on the two matches played by Nigeria. The Twitter Application Programming Interface (API) is employed for the acquisition of tweets on the Nigeria match against Guine Bisau, with a favourable outcome, and the match against Tunisia, with a loss. The dual corpus are subjected to sentiment analysis with VADER and the Multilingual sentiment analyzers. Graphical representations of the results are used for the overall analysis of the opinions expressed by Nigerians. The rest of the paper is presented in the following format. Section II discusses related srudies while section III presents materials and methods. Section IV discusses the results while the conclussion and recommendations are presented in section V.

2. Related studies

Studies in the area of sentiment analysis, topic modelling, semantic extraction etc. are numerous in the era of social media and better internet penetration. Employing same Natural Language Processing use cases for football analysis is an emerging genre with global interest owing to the high socioeconomic premium placed on football competitions across the globe. This section aims to discuss similar studies in the area of football opinion mining studies, and sentiment analysis. In [4], an opinion mining analytics aimed at discovering Covid-19 fear trigger was carried out on textual data using NLP. The study's topic modelling approach return *hand washing* as the major bigram that triggers fear of Omicron variant of the pandemic, through the instrumentality of Vader-based lexison approach. The work of [1] was particular about footballing opinion analytics together with prediction of event. A perception model to understudy football players' performances was the thrust of [5] through their online presence. Feedbacks from blogs, editorials, articles, nespapers, and social medias were analyzed which asigns ratings to each player. A football game analysis was conducted by [6] by proposing a framework that predicts results of English Premier League through understudying of past games. A sentiment analyser is incorporated from where a player could gauge the pulse of its fans over a match and likewise has a tool that aids fans on betting. The analysis of women and men sentiments about a match was implemented in the work of [7] and result shows certain similarities in the gender-corpus unexpectedly. Men and women are seen to express similar emotional outbursts like anger, fear, etc. over Premier league match outcomes. A sentiment analysis of tweets discussing football matches was conducted in [8] to identify specific emotions expressed over football-related tweets. The domain-specific approach of the study will further unravel the semantic meaning of expressed emotions and thereby design a sentiment classifier capable of classifying expressed sentiments in football tweet conversations. The prediction of wins and its spread I the Premier league was the main objective of [3] using sentiment analysis. The aim is to identify if sentiments expressed in tweets over matches would be sufficient to poredict future match outcomes and likewise discover if the magnitude of outcome would be correctly predicted based on the degree of sentiments expressed. The study discovered that surge in positive sentiments could net a payout of $3011.20 and also discover that the magnitude of positive sentiment likewise correlates with a point spread. A lexicon-based sentiment analysis study was carried out by [9]. The study tests the feasibility of lexicon-based analyzer for analysing football-based tweets. Over 10000 tweet corpus were acquired related to top-ten football matches for the purpose of the analysis. Experiemntal result shows that sentiment of realistic sets with a proportion of 60% of same polarity can be classified correctly with over 95% accuracy. Sentiment of world cup emotions to study theories of emotions was the thrust of the study carried out in [10]. Contrary to consensus in sport economics, the study revelaed that excitement relates to expression of negative emotions. In [11], a novel approach for multilingual sentiment analysis was implemented on MultiLingual analyzer. Local language of Urdu, English, and Roman-Urdu were analyzed with a novel dictionary on multilingual sentiments. Experimental result shows extreme lexisons with high weights that are used to label the data returned good performance. Sentiment analysis of Indonesian and French lexicons were analyzed by the

SentiSAIL tool in the work of [12]. The multilingual tool efficiently processed the languages with reliable experimental results.

3. Materials & Methods

This section discusses the materials and methods employed in this study which is aimed at unravelling emotions triggered by the outcome of football matches through the analysis of polarities retumned by two sentiment analysers. The framework implemneted for the purpose of this study is as captured in Figure 1.

3.1 Data acquisition

Twitter API was deployed for the acquisition of tweets through the instrumentality of hashtags, which is a promising reference tool with prominent usage on twitter in [2]. Two tweet corpus was captured for the purpose of this study including tweets discussing the outcome of the Nigerian match against Tunisia and the tweets discussing the Nigerian match against Gambia. As represented on Table 1, a total of 141371 unigram of tweet corpus, representing 1000 tweet authors, was captured for the Nigeria/Tunisia corpus while a total of 127072 unigram of tweets, representing 1000 authors, was captured for the Nigeria/GuineBisau match.

3.2 Corpus preprocessing

This study employs the following corpus preprocessing techniques as recommended in [2], to return a clean and tokenized unigram for the subsequent opinion mining analytics:

1. Conversion of the entire N_T and N_G corpus to lower case unigrams
2. Elimination of noises and white spaces from both corpus
3. A *Regexp*-based tokenization approach transforming entire corpuses into unigrams of useful tokens.
4. Filtering of entire corpus for the elimination of stop words like articles, prepositions, and conjunctions that do not add to the emotion-based pattern recognition approach of the analysis

Table 1:
Tweet distribution across the two corpus

Corpus	Corpus_ID	Hashtag	Corpus size	Unigram size
Nigeria/Tunisia	N_T	#NIGTUN	1000	141371
Nigeria/GuineBisau	N_G	#GNBNGA	1000	127072

3.3 Topic Modelling of N_T and N_G corpuses

Abstraction of subjects from textual corpus facilitated by word clusters and their frequency of occurrence is the main thrust of topic modelling [13] as contained in each 1000 documents that makes up each N_T and N_G respectively. It is a use case approach to NLP which unravels topical germane subjects from corpus for information retrieval purposes and opinion mining studies. The topic modelling is employed in this study to further unravel the most significant subject of discuss in each corpus for an interpretative purpose. The Latent Semantic Indexing (LSI) of the Orange data mining toolkit [4] is employed on both N_T and N_G. LSI returns topics with negative and positive keywords with their corresponding weights. Positive weights signify words that are highly illustrative of the topic while for negative weights, the topic rarely occur.

3.4 Sentiment Analysis using VADER-based Lexicon

Opinions expressed in textual corpus are categorized into positive, negative and neutral emotions during sentiment analysis employing emotion polarities. The VADER-approach grades each unigram contained each 1000 documents, as contained in each N_T and N_G, to compute a compound score for each pair of 1000 documents. The methodology calculates the sentiment grades supposing emotion is associated with the existence of definite phrases *(for bi-gram)*, or words (for unigram) as in the case of this study. Opinions are therefore assigned some sentiment values popularly referred to as lexicons. Similar to the work of [14], the occurrence rate of each unigram in the aforementioned dictionary, as used in each tweet document, would determine the calculations of its positive, negative or neutral polarity. The compound score, which is computed from the polarity scores of positive, negative, and neutral sentiment, is used in this study to determine the ultimate sentiment expressed in each 1000 documents of both N_T and N_G. It depicts collected account of the three polarity values and is computed as

$$\mathrm{y} = \frac{\mathrm{y}}{\sqrt{\mathrm{y}^2 + \alpha}} \tag{1}$$

where **y** is the addition of valence values of component unigrams, and $\boldsymbol{\alpha}$ is 15 (default value), which is the regularization constant. The compound score is normalized between -1 and +1 that signifies the positivity and negativity extent of each tweet opinion.

3.5 Sentiment Analysis using MultiLingual analyzer

The other sentiment analyzer employed in this study on the two corpus is the multilingual sentiment analyzer. It is an RNN-based framework that is reputable for limited corpus [15]. The deep learning-based method could employ any of the document embedders like word2vec, LoVe, doc2vec etc. for input layers [16], and then contains the hidden layers, to the output layers, in its methodological framework. The output could then be deployed for sentiment analysis, classification, entity extraction, translations, topic modelling, etc. This approach has proven to return an impressive amount of invariance to the traditional sentiment analysis systems like the Vader variant [16]. This is therefore employed in this study on the acquired N_T and N_G corpus for further analysis.

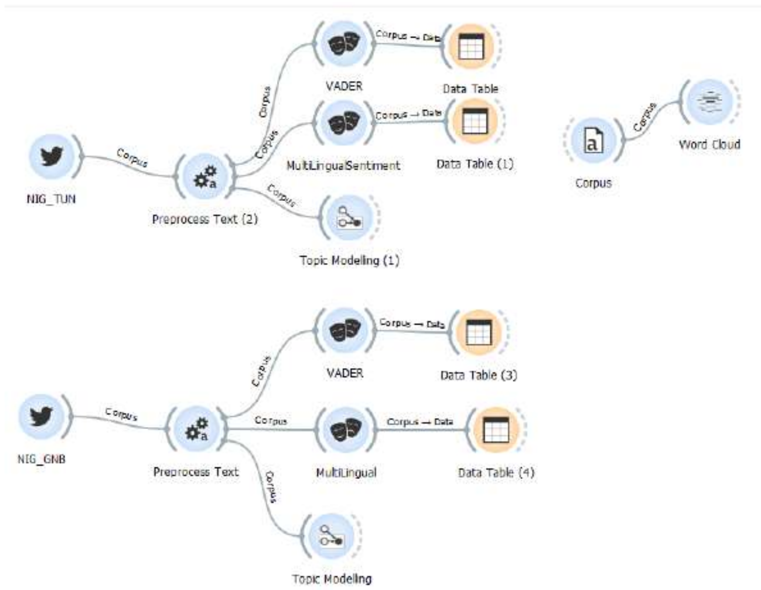

Figure 1: Orange toolkit widgets for study framework

4. Result & Discussion

The outcome of the dual sentiment analysis using Vader-based lexicon and MultiLingual deep sentiment analyzers on the Nigeria/Tunisia (N_T) and the Nigeria/GuineaBisau (N_G) corpus is discussed in this section alongside the output of the LDA Topic modelling on both N_T and N_G. The experiemntal result for Vader approach is sorted with the compound score from the smallest polarity to the highest in order to isolate and cluster the output into their polarity sub-categories. The ten documents with the most positive sentiments expression by Vader approach in the N_T is presented in the clustered bar chart in Figure 2. The corresponding clustered bar chart for the ten most positive sentiment expression as computed by the deep MultiLingual sentiment analyzer on the same N_T is as presented in Figure 3. The Vader-based analysis output for the ten most negative sentiments expressed in N_T is presented in Figure 4 while the Multilingual variant is presented in Figure 5. Analysis executed by Vader approach on N_G for the ten most negatively expressed sentiments is presented in Figure 6 while that of the MultiLingual counterpart is as presented in Figure 7. The positive sentimental versions for both Vader and MultiLingual on N_G is presented in Figure 8 and Figure 9 respectively. Further analysis is done on the ten most positive sentiment opinions as analysed by Vader and MultiLingual on N_T and the result is presented in the word cloud of Figure 10 while that of the ten most negative sentimantal opinion is captured Figure 11. An overall topic modelling by LSI on N_T and N_G respectively is presented in Figure 12 to unravel topical tokens with the element of consensus in each 1000-sized corpus.

On the N_T, notwithstanding the 1:0 loss of Nigeria to Tunisia, some opinions on the match were expressed with positive compound sentiments which may not be unexpected. However, there was no one-to-one mapping between the ten tweet documents that makes each cluster of Vader and Multilingual classes despite belonging to the same category, which is indicative of the sharp contrast between their polarity computations. The same observation goes for the negative polarity category of both analyzers on N_T. However, it can be observed that the tweet documents that makes each positive and negative clusters for MultiLingual on N_T are majorly tweets with limited words unlike the Vader clusters with longer phrases. Similar observations are common with the positive and negative clusters of both Vader and MultiLingual on N_G where no one-to-one mapping exists between the clusters and the documents that makes the MultiLingual clusters are of shorter phrases compared to the Vader counterpart. In positive sentiment cluster MultiLingual on N_T, tokens of local dialets and or pidgin English like 'wahala', 'won' etc., meaning *trouble* and *want,* accrues higher weights with no such traces in the Vader word cloud. The ‘good’, ‘best’, and ‘small’ tokens posses highest weights in the Vader cluster. In the negative category, MultiLingual similarly returned pidgin English dialet tokens like ‘sha’, ‘e’, ‘no’ etc., as against the only local dialet token ‘pass’ noticed in the Vader negative cluster of N_T. The context in which they are used uniquely identifies a proper English from the pidgin local version. The negative clusters of N_T over Vader and MultiLingual however returned highly negative tokens as those with highest weights which is indicative of the mood of tweet posters who expresses their sentiments at the loss of the match. Tokens like ‘bad’, ‘winch’ (meaning *withcraft* in pidgin english), ‘cry’, ‘pikin’ (meaning *child* in pidgin English), ‘die’, and ‘kill’, ‘pain’, ‘sorry’, etc. all encapsulates the mood of the tweet documents in the negative N_T cluster. Prominent tokens like ‘Okoye’ and ‘Moses’ are the most significant topics returned by LSI on N_T and N_G respectively besides other commonplace topics like ‘nigeria’, ‘afcon’, ‘ngatun’, and ‘gnbnga’. Experimental result of the topic modelling shows that the name of the Nigeria goal keeper, Okoye, is prominent in the 141371 unigram that makes up the N_T while the name of the Nigeria striker, Moses, was the most prominent in the 127072 unigram as contained in the N_G corpus.

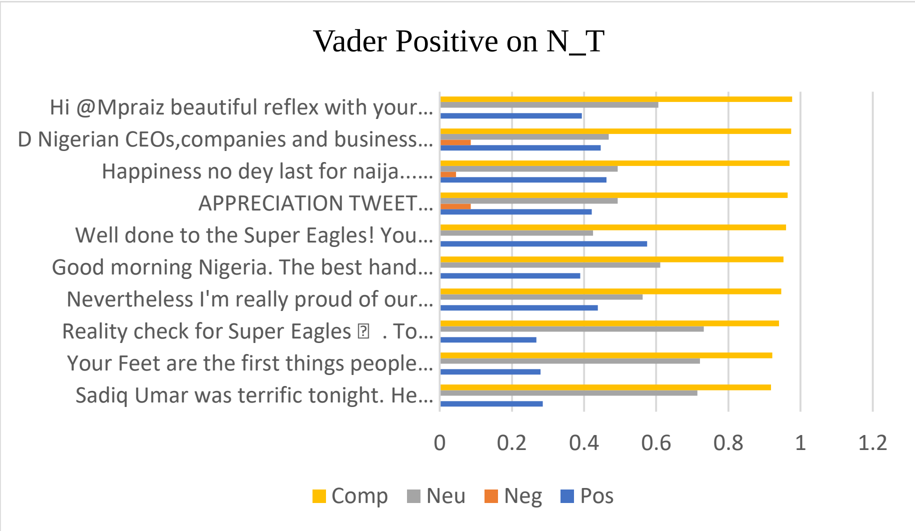

Figure 2: Clustered bar chart for ten most positive polarities by Vader on N_T

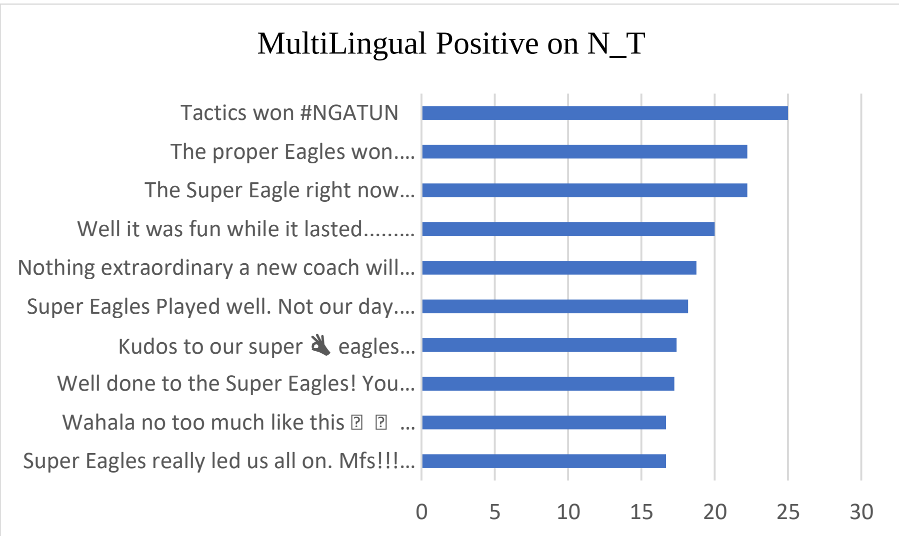

Figure 3: Clustered bar chart for ten most positive polarities by MultiLingual on N_T

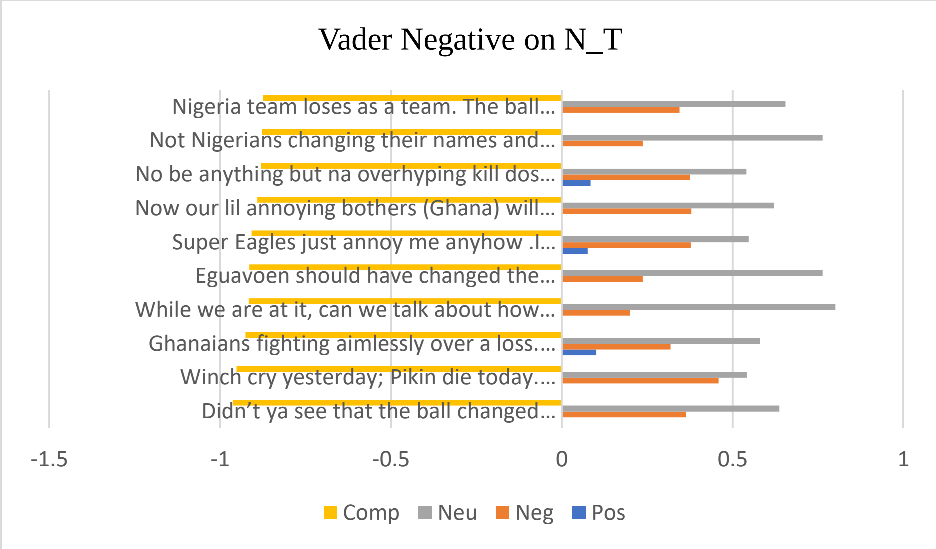

Figure 4: Clustered bar chart for ten most negative polarities by Vader on N_T

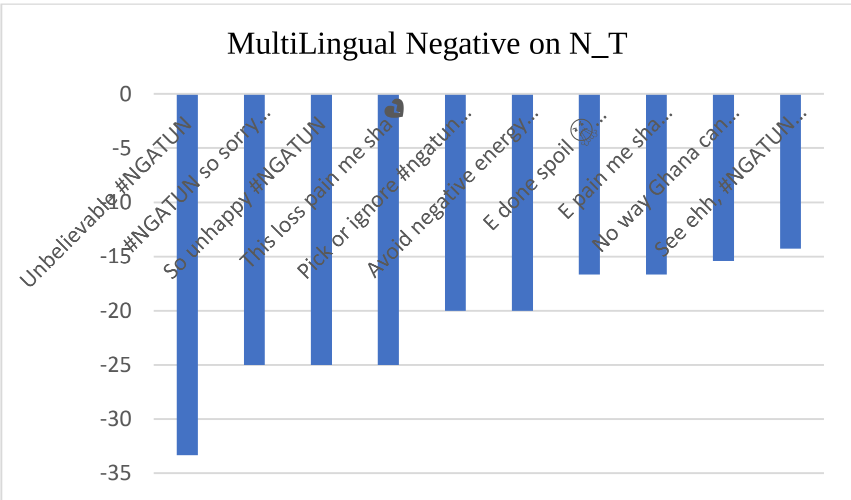

Figure 5: Clustered bar chart for ten most negative polarities by MultiLingual on N_T

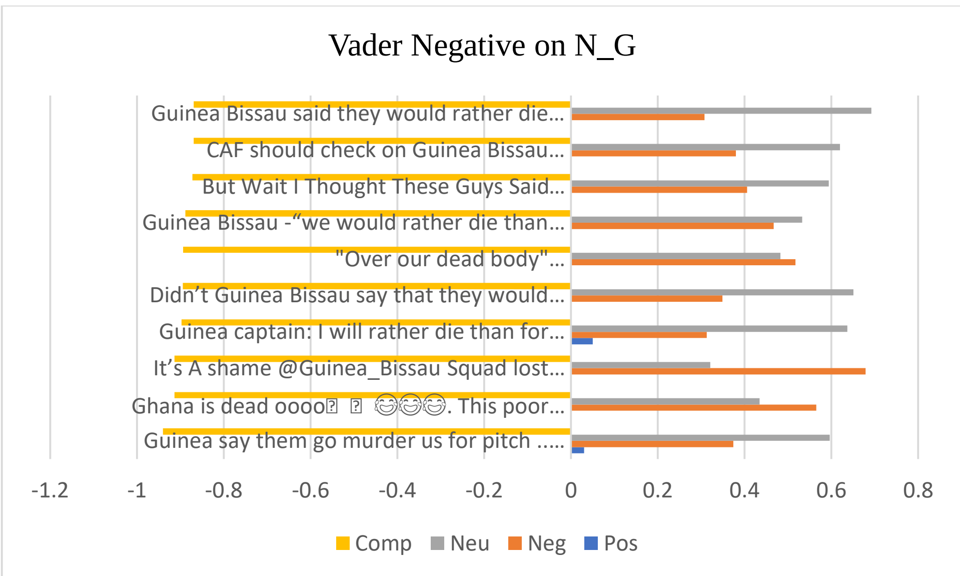

Figure 6: Clustered bar chart for ten most negative polarities by Vader on N_G

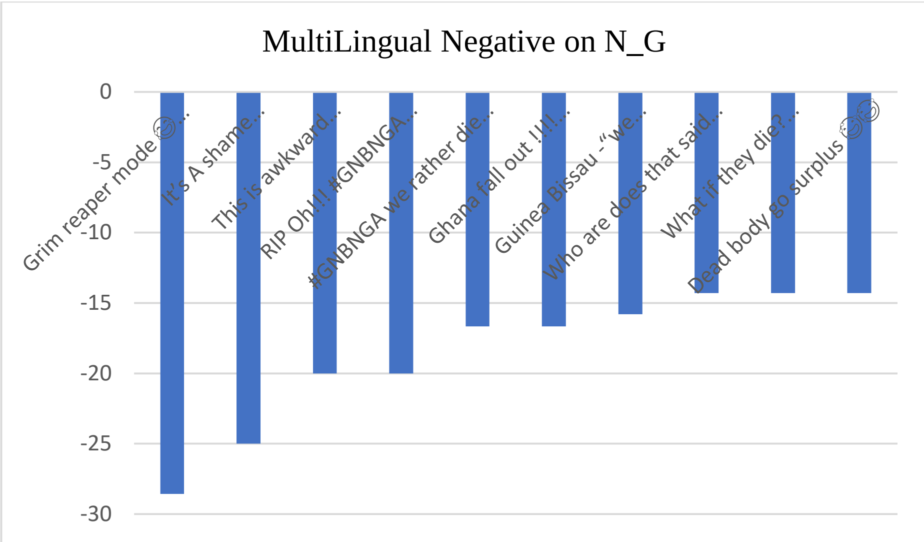

Figure 7: Clustered bar chart for ten most negative polarities by MultiLingual on N_G

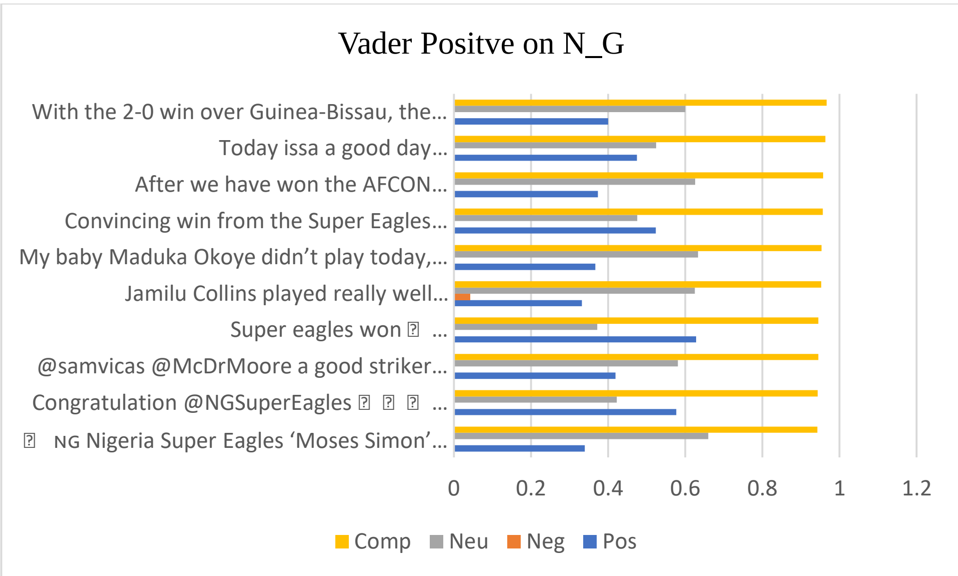

Figure 8: Clustered bar chart for ten most positive polarities by Vader on N_G

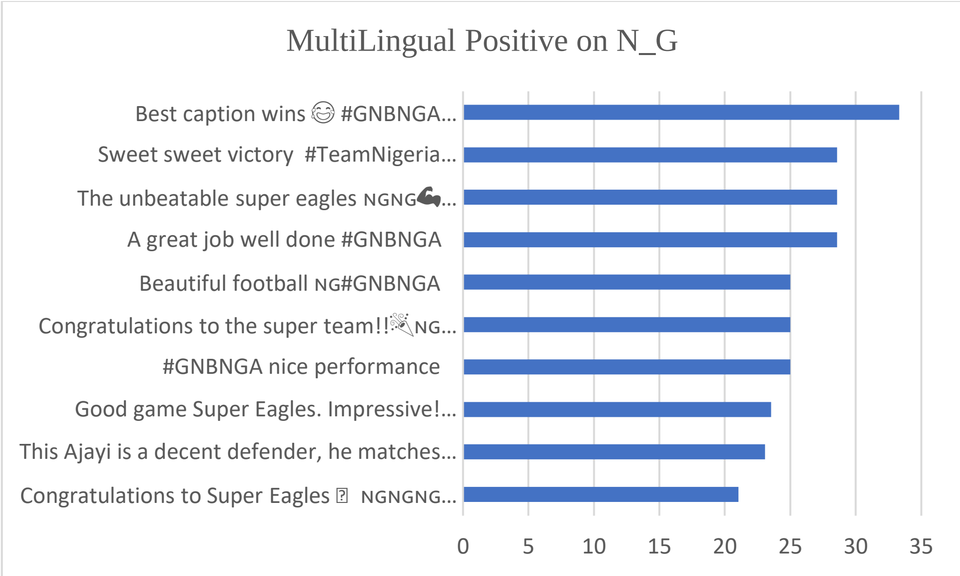

Figure 9: Clustered bar chart for ten most positive polarities by MultiLingual on N_G

(a) *(b)*

Figure 10: Word cloud of ten most positive N_T tokens for (a) Vader and (b) MultiLingual

(a) *(b)*

Figure 11: Word cloud of ten most negative N_T tokens for (a) Vader and (b) MultiLingual

Topic	Topic keywords
1	#_#, buhari_RB, ngatun_RB, okoye_NN, buhari_VBP, buhari_JJ, http_JJ, afcon2021_RB, ngatun_JJ, afcon2021_JJ
2	nigeria_VBP, nigeriavstunisia_NN, afcon2021_RB, nigeriavstunisia_RB, nigeria_NNS, nigeria_RB, nigeria_JJ, okoye_RB, ngatun_NN, ngat
3	#_#, ngatun_NN, ngatun_JJ, co_NN, http_NN, afcon2021_JJ, nigeria_RB, okoye_RB, nigeriavstunisia_JJ, okoye_JJ
4	afcon2021_NN, afcon_JJ, nigeria_NNS, http_VBP, nigeria_VBP, http_JJ, afcon_VBP, buhari_VBP, buhari_NN, buhari_JJ
5	afcon_VBP, http_JJ, nigeria_JJ, okoye_MD, okoye_VB, ngatun_VB, afcon_NN, nigeria_NNS, buhari_JJ, buhari_VB

(a)

Topic	Topic keywords
1	#_#, co_NN, http_NN, gnbnga_JJ, gnbnga_NN, nigeria_RB, afcon2021_RB, teamnigeria_NNS, afcon2021_JJ, guinea_NN
2	co_NN, http_NN, gnbnga_JJ, nigeria_RB, #_#, gnbnga_NN, guinea_NN, afcon2021_JJ, bissau_NNS, moses_NNS
3	gnbnga_NN, gnbnga_JJ, moses_NNS, guinea_NN, nigeria_RB, bissau_NNS, afcon2021_JJ, afcon2021_NN, co_NN, http_NN
4	nigeria_RB, guinea_NN, bissau_NNS, gnbnga_JJ, http_NN, co_NN, afcon2021_JJ, teamnigeria_NNS, afcon2021_NN, gnbnga
5	nigeria_RB, guinea_NN, afcon2021_JJ, bissau_NNS, teamnigeria_NNS, http_VBP, co_NN, moses_NNS, gnbnga_NN, gnbnga_

(b)

Figure 12: Result of five prominent topics for (a) N_T and (b) N_G respectively

5. Conclusion & Recommendation

This study evaluatyes the performances of the Vader-based lexicon sentiment analyzer and the MultiLingual sentiment analyzers through an opinion mining analytics of football tweets. Two tweet corpus of 1000 tweet each, comprising of 141371 and 127072 from the hashtag #NGRTUN and #GNBNGA respectively. Experiemntal results of sentiment analysis and topic modelling reveals that while Vader-based approach targets tweets of longer phrases, MultiLingual analyzer works with tweets of shorter phrases. The MultiLingual analyzer likewise returned unigram of local dialets than the Vader approachin the word cloud analysis. Two players were returned as the most discussed topics in the two corpus by LSI topic modelling, including Okoye and Moses, who might have played significant roles in the two matches where Nigeria lost the N_T match and won the N_G match. Further studies could expand the scope of the textual corpus above 1000 to ascertain if MultiLingual would retain its penchant for short phrases.

6. Acknowledgement

The authors appreciates the painstaking efforts of the reviwers which made this work turn out better.

7. References

[1] N. J. D. Souza, H. N. Samrudh, S. Gautham, B. U. S. Bhat and N. Nagarathna, "Football Game Analysis and Prediction," in *Soft Computing and Signal Processing*, 2021.

[2] T. Olaleye, T. Arogundade, A. Abayomi-Alli and K. Adesemowo, "An Ensemble Predictive Analytics of COVID-19 Infodemic Tweets using bag of words," in *Data Science for COVID-19*, Singapore, Elsevier, 2021.

[3] R. P. Schumaker, A. T. Jarmoszko and C. S. L. Jr., "Predicting wins and spread in the Premier

League using a sentiment analysis of twitter," *Decision Support Systems*, pp. 76-84, 2016.
[4] T. O. Olaleye, S. M. Akintunde, C. Akparanta, T. A. Avovome, O. F. Oluyen and A. O. Akparanta, "Opinion Mining Analytics for Spotting Omicron Fear-Stimuli Using REPTree Classifier and Natural Language Processing," *International Journal for Research in Applied Science & Engineering Technology*, pp. 995-1005, 2022.
[5] Z. Kabir, "Perception Model to Analyze Football Players' Performances," *PACIS*, 2019.
[6] n. Souza, H. Samrudh, S. Gautham, B. S. Bhat and N. Nagarathna, "Football Game Analysis and Prediction," in *Soft Computing and Signal Processing* , 2021.
[7] M. B. Babac and V. Podobnik, "A sentiment analysis of who participates, how and why, at social media sport websites: how differently men and women write about football," *Online Information Review*, 2016.
[8] S. Aloufi and A. E. Saddik, "Sentiment identification in football-specific tweets," *IEEE Access* , 2018.
[9] F. Wunderlich and D. Memmert, "Innovative approaches in sport science-lexicon-based sentiment analysis as a tool to analyze sports-related Twitter communication," *Applied Sciences*, 2020.
[10] G. M. Lucas, J. Gratch, N. Malandrakis, E. Szablowski and E. Eessler, "GOALLL:: Using sentiment in the world cup to explore theories of emotion," *Image and Vision Computing*, pp. 58-65, 2017.
[11] M. Asif, A. Ishtiaq, H. Ahmad, H. Aljuaid and J. Shah, "Sentiment analysis of extremism in social media from textual information," *Telematics and Informatics*, 2020.
[12] G. Shalunts, G. Backfried and H. S. Alam, "Sentiment analysis in Indonesian and French by SentiSAIL," in *017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, 2017.
[13] T. Sherstinova, O. Mitrofanova, T. Skrebtsova, E. Zamiraylova and M. Kirina, "Topic Modelling with NMF vs. Expert Topic Annotation: The Case Study of Russian Fiction," in *Mexican International Confeence on Artificial Intelligence*, 2020.
[14] H. Hota, D. Sharma and N. Verma, "Lexicon-based sentiment analysis using Twitter data: a case of COVID-19 outbreak in India and abroad," in *Data Science for COVID-19*, London, Elsevier, 2021, pp. 275-293.
[15] E. F. Can, A. Ezen-Can and F. Can, "Multilingual Sentiment Analysis: An RNN-Based Framework for Limited Data," *arXiv:1806.04511*, 2018.
[16] P. Ghaffari, "Leveraging Deep Learning for Multilingual Sentiment Analysis," LinkedIn, 2016.