International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307

Available Online at : www.ijsrcseit.com

doi : https://doi.org/10.32628/CSEIT23903146

A Novel Fake-News Dataset and Detection System to Mitigate Cyber War with Emphasis on Nigerian News Events

Samera Uga Otor*[1], Beatrice Obianiberi Akumba[2], Joseph Sunday Idikwu[3]

*Department of Mathematics and Computer Science, Benue State University, Makurdi, Benue State, Nigeria

ARTICLEINFO

Article History:

Accepted: 20 June 2023

Published: 15 July 2023

Publication Issue

Volume 9, Issue 4

July-August-2023

Page Number

24-32

ABSTRACT

Fake-news refers to a cyber-weapon launched through the social media, as, its consequence can result to the breakdown of law and order in the society both physically and on the cyber-social-space. In Nigeria, there is currently no established law that guides the use of social media. Therefore, the rate at which fake-news propagates is alarming. This paper presents a new dataset, with focus on Nigeria's trending news such as EndSARS and Herdsmen attacks, which was further used to simulate Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) machine learning models to detect fake-news. The data were extracted from twitter using twitter Application Package Interface (API) and from facebook using a scraping tool. The dataset was encoded using Unicode escape function in python to make all characters accessible by the algorithm and tokenised using Global Vectors for Word Representation. The dataset was used to train CNN and RNN models built in python on google colab platform to detect fake-news using accuracy, sensitivity, recall and F1 score as evaluation metrics. Results showed that RNN performed better in terms of accuracy and precision, at 82.34% and 93.19% compared to 81.96% and 79.65% for CNN, F1 scores are approximately the same for both models and CNN performed better than RNN in terms of recall at 98.03% to 50.61% for RNN.

Keywords: Fake-news, EndSARS, Herdsmen, Cyber-Social-Space, Social media, Machine Learning

I. INTRODUCTION

The term fake-news consists of two words "fake" which implies something that is not genuine, but portrayed to be (Apuke & Omar, 2020) and "NEWS" an abbreviation which stands for "Notable Events, Weather and Sports" (Ikenwa, 2019). There is no agreed definition for fake-news. However, it is attributed to information or news that is false (Zhou & Zafarani, 2020; Shu, et al., 2018; Allcott & Gentzkow, 2017). Therefore, fake-news is a false, hoax, vague, misconstrued news claimed to be real news, deliberately created and circulated through the social media for the purpose of propaganda, misinformation, political gain among others, to mislead and cause panic among the public of interest. This news has been in existence for as long as the printing press originated in 1439 (Shu, et al., 2017) and has been in circulation before the era of digital technology (de-Beer & Matthee, 2021). However, it is on a wider spread currently due to the alarming growth rate of electronic media, electronic gadgets, availability of internet, introduction of several social media applications/platforms and their ease of use. In a highly populated country like Nigeria, estimate showed that there are around 33 million users active on the social media daily (Kemp, 2021). There is even high competition amongst people as to who can produce more news to generate traffic on the internet. This has led to deliberate propaganda and hate speech resulting to physical clashes, cyber-misconducts and cyber-war. Hence, fake-news can be said to be a cyber-weapon which must not be undermined as it disrupts the individual's right to truth, abuses human right and the right to democracy, stimulates cynicism in an organisation, encourages political enormity and rift, undermines constituted authorities and its effect can foster breakdown of law and order. Clarke & Knake, (2010), gave examples of cyber-weapons as releasing of information bombs, dumping of information garbage, disseminating propaganda, applying information deception, releasing cloned information among others, which constitutes what fake-news is all about. These certify that fake-news is a cyber-weapon. Therefore, there is need to rid the society of these news items.

A piece of news item termed to be fake, may contain some elements of truth but may be exaggerated, may not conform to the date reported or may attribute a scenario or personality that is real to a fake scenario or personality. To distinguish intuitively between what is false and what is true in order to get rid of it, the reader must read past the headlines to learn the source, author, and date. Furthermore, it is necessary to determine whether the news is a joke, overblown, or ascribed to the wrong scene/person, as well as to examine users' interaction, comments, and social behaviour in relation to the news (Shu et al., 2018; Limeng et al., 2019), among other things. Another way to combat fake-news, is to verify manually from fact checking sites such as factcheck.org, snopes.com, politifact (Baly, et al., 2018), open secrets, the sunlight foundation, pointer institute, flack check, truth or fiction, hoax slayer and verifiable news sites such as Channels Television, African Independent Television, Nigerian television authority among others. However, a lot of time is spent checking these sites manually as a result, debunking comes too late (Baly, et al., 2018). Hence, this paper proposed a model to detect fake-news using text mining technique such as information retrieval methods to spool out news content relevant to users' query on Twitter, Facebook, Channels television, African independent television and British broadcasting corporation with focus on the Nigeria's Endsars and Fulani herdsmen to generate a dataset. This dataset was further processed using Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) machine learning algorithms to detect fake-news.

The rest of the paper is organised as follows. Section two reviews existing literature in the subject area; section three describes the methodology used to retrieve the dataset, develop and simulate the proposed model while section four presents the results obtained from the simulated model and section five concludes by pointing out limitations and future research areas in this work.

II. Literature Review

The concern raised by the spread of fake-news recently has prompted the need for research in this area. More so, that research in this area is still limited. Generally, in a highly populated country like Nigeria with almost everyone having a smart phone pre-installed with at least one of the social media applications such as

WhatsApp, Facebook, Twitter among others, one cannot imagine how fast news will traverse.

Fake-news is always presented as factual and accurate however in reality it is not. Therefore, people tend to believe what they see on the websites or social media and do not try to validate if the information provided is true or false. Since people are often unable to spend enough time to crosscheck and be sure of the credibility of news, intelligent detection of fake-news is essential. Therefore, it has continued to receive high attention from the government and research community globally.

In the year 2020, the Nigeria's Minister of Information and Culture, Lai Mohammed suggested that the National Assembly should formulate a National policy on the use of social media to control fake-news, misinformation and hate-speech (Premium Times, 2020). He opined that social media is a platform of choice for those who propagate fake-news, and that there is an urgent need for a National policy to be put in place to curb excesses and the misuse of technology and resources that dominates the social media space to mitigate the spread of fake-news and hate-speech. Despite all efforts put in place by the existing society, people, technology and processes, there is evidence of fake-news in some form every day.

Nigeria is a place where fake-news is becoming more dominant, creating a gesture among people and making sure that the original news is reduced completely. For example, during the outbreak of Ebola virus, some people died and others were hospitalised due to excessive intake of salt water as a result of a fake message which emanated on the social media that bathing and drinking of salt water could prevent being infected (Apuke & Omar, 2020). Another major instance of fake-news that dominated the Nigeria's social media space was the rumoured health status of Bola Ahmed Tinubu one of the 2023 Nigerian presidential candidates, now the president, who went on a medical trip. Several pictures were dispersed on social media of how critically ill he was even though those pictures could not be verified as real. This was later debunked by circulating another video of him ridding on a bicycle hale and hearty.

Furthermore, there was news about Bishop Mathew Hassan Kukah the Catholic Bishop of Sokoto Diocese, alleging that his house was burnt down by Muslim mobs during the crisis that erupted in Sokoto state due to the killing of a Christian student Deborah Samuel in May 2022. The Bishop debunked the news on social media. Initiators of these fake-news items may have devised them to discredit or destabilize the Nigerian Government, cause religious crisis, inter-tribal sentiments and incite people to violence.

Available fake-news detection techniques are either manual or automated. Baly, et al., (2018), describes manual fake-news detection as a technique that habitually employ all the methods and procedures a person can use to verify news by visiting fact checking sites such as crowd sourcing of real news to compare with unverified news. However, this can be cumbersome due to the overwhelming amount of data generated online daily. More so, that, cognitive psychologists Chen et al., (2016) opined that it is difficult for humans to accurately point out between what is fake and what is real since humans by chance can do so at only 4 percent. Therefore, manual fact checking has quickly become ineffective as it struggles to scale with the volume of data generated online.

Automated detection systems provide value in terms of automation and scalability (Ayuya, 2020). Most researches on automated detection are based on machine learning techniques.

One of the earliest works on generating a dataset for fact checking was by Vlachos and Riedel (2014). A task was defined using statements fact-checked by journalists available online which were compared to statements from 2 fact checking sites to construct a dataset which was analysed using k-Nearest Neighbour classifier to handle fact-checking as a classification task. This is a semi-automated approach with some limitations of manual approaches. In the same vein, Ferreira & Vlachos, (2016), proposed the "Emergent" dataset. The dataset consists of 300 rumoured claims and 2,595 associated news articles, collected and labelled by journalists with an assessment of their truthfulness as true, false or unverified. Each associated article was summarized into a headline and labelled to indicate whether its stance is for, against, or observing the claim. Where, observing indicates that the article merely repeats the claim. The dataset was further analysed using a logistic regression classifier to develop features that examine the headline and its agreement with the claim. The dataset is limited in size and concentrated on only the headlines. Furthermore, Wang (2017) also collected the LIAR dataset, which consists of 12.8K manually, annotated short PolitiFact assertions, and used multiple classifiers to identify them. Despite the fact that the dataset is larger than prior ones, classification results using CNN revealed accuracy as low as 0.270. Another, as seen in Salem et al. (2019), used semi-supervised machine learning approaches on datasets from a data centre such as the Syrian Violation Documentation Centre (news events concerning the Syrian war) to detect fake news. This dataset focuses solely on the Syrian conflict.

Some of the strategies described above focused on short statements and headlines rather than paying attention to more detailed news. As a result, they are limited in the sense that some news may be true yet contain twisted, unclear, or exaggerated content. Allowing the reader to receive the content from multiple sources (both reliable/verifiable and unreliable/non-verifiable) and reading them thoroughly will enable the reader better verify and defend the news item. Hence, fake-news datasets focusing on aspects such as news content, social context, spatiotemporal information, and user comments, among others, where proposed to solve this restriction (Shu et al., 2018; Anushaya et al., 2020). These are comprehensive data repository, however, results still showed low accuracy for fake-news detection using both classical feature-based supervised machine learning models such as K-Nearest Neighbour, Support Vector Machine, Naïve Bayes, Passive Aggressive Classifiers and deep learning model such as CNN.

Kaliyar et al., (2021), opined that deep learning models have immense advantage over existing classical feature-based approaches in that they do not require any handwritten features; instead, they identify the best feature set on their own. Therefore, recent fake-news detection techniques employed RNN and CNN deep learning techniques to analyse existing datasets to detect fake-news (Popat et al., 2018; Reddy, 2019; Kaliyar et al., 2021; Jamal et al, 2021). However, the Nigerian circumstance still needs further exploration, as there are no available datasets from Nigeria social media space. Hence, this research also provides a system for detecting fake-news from news content and headlines using both CNN and RNN Models. First, a new dataset consisting of Nigerian news events was generated and the dataset was further analysed using the models to determine which performs better in detecting fake-news.

III. METHODS AND MATERIAL

A. Data Extraction

The focus of this paper is to generate a dataset that is based on news events in Nigeria from social media and analyze the dataset to detect fake-news. Hence, the Endsars and Fulani herdsmen attack news events were selected with focus on Facebook and Twitter. Tweets on the EndSars protest that occurred around October 2020 were extracted from the twitter archive API (Application Program Interface). A total of 750710 tweets were extracted using the keywords "#EndSARS,EndSARS, #ENDSARS, ENDSARS, #endsars, endsars, #Sorosoke, Sorosoke, Police Brutality, police brutality, Lekki Toll Gate, lekki toll gate, massacre, Masacre" from the general twitter archive platforms of channelsTV, AIT and BBC twitter handles into a coma separated values(CSV) file. In the same vein, a total of 1169 Facebook posts on EnSars and Herdsmen were extracted using a scraping tool known as Web Scraper into a CSV file. The total dataset consists of 751876 rows and 74 columns. The keywords in the dataset are as presented in Figure 1.

B. Data Preparation and Pre-processing

The dataset was pre-processed in the following ways to adhere to the machine learning methods used: First, it was encoded in Python using the unicode escape function. Unicode is an international character encoding system that assigns a unique number to each prediction class field used for the learning method were determined. The author-verified feature of the dataset, whose values are either TRUE or FALSE, was chosen to be the predicting class, while the text field was used as the learning feature, after thorough examination. The dataset is as shown in Figure 2. False values in the author-verified field were replaced with zeros, while true values were replaced with ones. Similarly, a section of the Facebook dataset was manipulated by arbitrarily introducing new write-ups to some fraction of the content to generate the fake dataset, and the predicting class was renamed the label field. Additionally, special characters and null values were removed from the dataset using Python's clean text function, and the dataset was tokenized using Global Vectors for Word Representation, an unsupervised learning algorithm used to obtain vector representations for words (glove.6B.100d.txt) in preparation for neural network models.

Figure1. Keywords in the scrapped dataset

C. Model Building and Validation

Convolutional neural network (CNN) and recurrent neural network (RNN) models were created in Python using the Google Colab platform and the parameters listed in Tables 1 and Table 2. Furthermore, the models were trained with the processed dataset, focusing on only the text and label fields at a 30% to 70% division for training and validation. The model was evaluated to predict whether or not news items are fake, utilizing accuracy, precision, recall, and F1-Score as evaluation measures. The Algorithm is as presented in Figure 3.

	K	M	N	AU	BC	BD
1	text	source	public_metrics.lil	author.location	author.url	author.verified
2	Fallen Unarmed Peaceful Nigerian Proteste	Twitter Web App	285	Abuja, Nigeria	https://t.co/3qRo	TRUE
3	#EndSARS: Gov. Zulum commends Borno	Twitter Web App	61	Abuja, Nigeria	https://t.co/3qRo	TRUE
4	EndSARS: Enugu police arrest suspects ov	Twitter Web App	28	Abuja, Nigeria	https://t.co/3qRo	TRUE
5	EndSARS Protest: U.S. Congressman War	Twitter Web App	67	Abuja, Nigeria	https://t.co/3qRo	TRUE
6	EndSARS Protest: Police parade 31 suspe	Twitter Web App	23	Abuja, Nigeria	https://t.co/3qRo	TRUE
7	EndSARS Protests: Nigerian youths comm	Twitter Web App	3439	Abuja, Nigeria	https://t.co/3qRo	TRUE
8	EndSARS Protests: Edo police arrest 116 s	Twitter Web App	9	Abuja, Nigeria	https://t.co/3qRo	TRUE
9	EndSARS Protests: Bauchi Governor inauc	Twitter Web App	19	Abuja, Nigeria	https://t.co/3qRo	TRUE
10	EndSARS Protests: Oba of Benin appeals t	Twitter Web App	45	Abuja, Nigeria	https://t.co/3qRo	TRUE
11	EndSARS Protests: Bayelsa Govt. Inaugur	Twitter Web App	22	Abuja, Nigeria	https://t.co/3qRo	TRUE
12	EndSARS: Abia Police arrest 18 suspects i	Twitter Web App	20	Abuja, Nigeria	https://t.co/3qRo	TRUE
13	Police Brutality: Lagos judicial panel of inqu	Twitter Web App	42	Abuja, Nigeria	https://t.co/3qRo	TRUE

Figure 2. Sample Dataset

TABLE I. CNN PARAMETERS

Model: “sequential”		
Layer (type)	**Output Shape**	**Parameter**
embedding (Embedding)	(None, None, 100)	4000000
dropout (Dropout)	(None, None, 100)	0
lstm (LSTM)	(None, 250)	351000
dense (Dense)	(None, 250)	62750
dropout_1 (Dropout)	(None, 250)	0
dense_1 (Dense)	(None, 1)	251
Total parameters	4,414,001	
Trainable parameters	414,001	
Non-trainable parameters	4,000,000	

TABLE 2. RNN PARAMETERS

Model: “sequential”		
Layer (type)	**Output Shape**	**Parameters**
embedding (Embedding)	(None, None, 100)	4000000
dropout (Dropout)	(None, None, 100)	0
conv1d (Conv1D)	(None, None, 250)	75250
max_pooling1d(MaxPooling1D)	(None, None, 250)	0
conv1d_1 (Conv1D)	(None, None, 250)	312750
global_max_pooling1d(GlobalMaxPooling1D)	(None, 250)	0
dense (Dense)	(None, 250)	62750
dropout_1 (Dropout)	(None, 250)	0
dense_1 (Dense)	(None, 1)	251
Total parameters:	4,451,001	
Trainable parameters:	451,001	
Non-trainable parameters:	4,000,000	

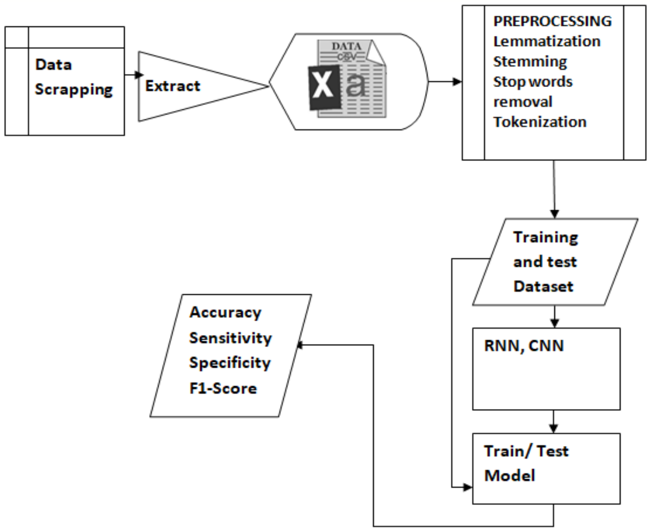

Figure 3. Model Flowchart

IV. RESULTS AND DISCUSSION

The simulation was run in the Google Colaboratory Python Environment on an Intel Core i5 HP laptop with 2GB of RAM, yielding the following results: Table 3 displays the accuracy, F1 score, precision, and recall of RNN and CNN. Figures 4, 5, 6, and 7 are the outcomes of RNN and CNN training loss versus epochs and training accuracy versus epochs, respectively. The results reveal that RNN outperforms CNN in terms of accuracy and precision; the F1 score is roughly the same for both models; and CNN outperforms RNN in terms of recall.

TABLE 3
RESULTS OF EVALUATION METRICS

Metrics	RNN(%)	CNN(%)
Accuracy	0.8234	0.8196
F1	0.7685	0.7627
Precision	0.9319	0.7965
Recall	0.5061	0.98027

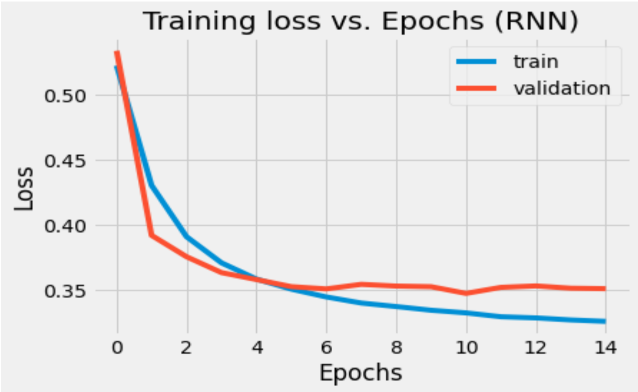

Figure 4. RNN Training Loss vs Epoch

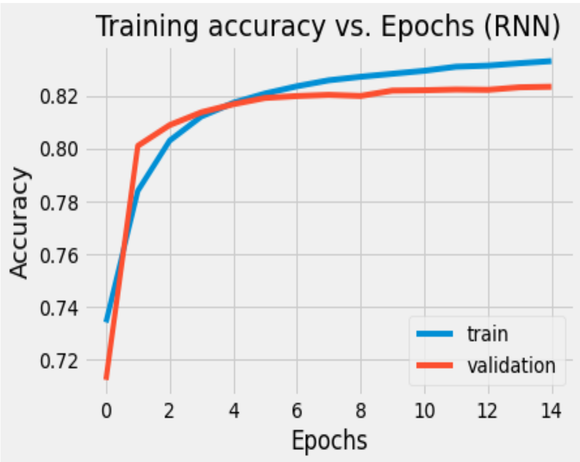

Figure 5. RNN Training Accuracy vs Epoch

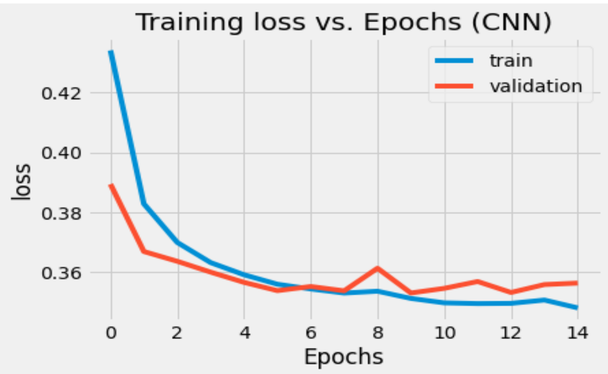

Figure 6. CNN Training Loss vs Epoch

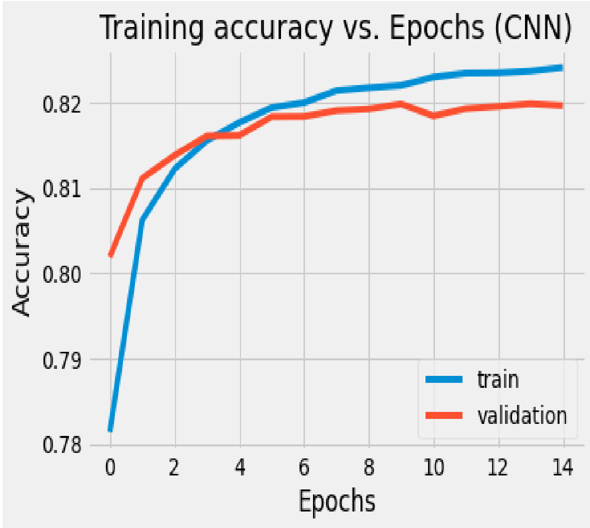

Figure 7. CNN Training Accuracy vs Epoch

V. CONCLUSION

This study created a new dataset based on Nigerian news articles, with a focus on the EndSARS demonstration and headsmen attacks. To detect fake-news, the dataset was further analyzed using CNN and RNN machine learning models. The dataset may be used to detect fake news with a minimum accuracy of 82.34%, which can be increased with additional parameter adjustment. The Full dataset will be made available upon request. However, part of the dataset generated can be accessed via this link sammy-stack/Research-Repo: for my research (github.com). It can be used as a benchmark dataset for sentiment analysis, fake-news mitigation, and fake-news detection programs. In the future, we plan to expand the dataset by including automatic updating mechanisms to keep it current for real-time detection. We plan to use it for sentiment analysis as well.

Although a conclusion may review the main points of the paper, do not replicate the abstract as the conclusion. A conclusion might elaborate on the importance of the work or suggest applications and extensions. Authors are strongly encouraged not to call out multiple figures or tables in the conclusion these should be referenced in the body of the paper.

VI. REFERENCES

[1]. Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 .Journal of Economic Perspectives- Springer, 31(2), 211-236.

[2]. Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 .Journal of Economic Perspectives- Springer, 31(2), 211-236.

[3]. Anushaya, P., Aisuwariya, T., Vamsee, K. K., & Vasudevan, S. K. (2020). An Innovative and Implementable Approach for Online Fake News Detection Through Machine Learning. Journal of Computational and Theoritical Nanoscience, 17(1) , 130-135. https://doi.org/10.1166/jctn.2020.8639.

[4]. Apuke, O. D., & Omar, B. (2020). Fake news proliferation in Nigeria:Consequences, motivations, and prevention through awareness strategies, Humanities & Social Sciences Reviews, 8(2), 318-327. DOI:10.18510/hssr.2020.8236

[5]. Ayuya, C. (2020, September 18). Introduction to Automated Fake News Detection. https://www.section.io/engineering-education/automated-fake-news-detection/

[6]. Baly, R., Karadzhov, G., Alexandrov, D., Glass, J., & Nakov, P. (2018). Predicting Factuality of Reporting and Bias of News Media Sources,Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, pages 3528–3539, oct31-nov 4. Brussels, Belgium: Association for Computational Linguistics.

[7]. de-Beer, D., &Matthee, M. (2021). Approaches to Identify Fake-News: A Systematic Literature Review. In: Antipova, T. (eds).Integrated Science in Digital Age 2020. ICIS 2020. Lecture Notes in Networks and Systems, 136,13-22. Springer, Cham. https://doi.org/10.1007/978-3-030-49264-9_2

[8]. Chen, Y., Conroy, N. J., & Rubin, V. L. (2016). News in an online world: The need for an "automatic crap detector." Proceedings of the Association for Information Science and Technology, 52(1), 1–4. https://doi.org/10.1002/pra2.2015.145052010081

[9]. Clarke, R. A., & Knake, R. (2010). Cyber War : the Next Threat to National Security and What to Do About It. Harpercollins E-Books.

[10]. Ferreira, W., & Vlachos, A. (2016). Emergent: a novel data-set for stance classification, Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: Human language technologies (NAACL-HLT 2016) San Diego, California, June 12-17, 2016(pp. 1163–1168)

[11]. Ikenwa, C. (2019, December 17). What Is The Full Meaning Of NEWS Acronym? Nigerian Infopedia. https://nigerianinfopedia.com.ng/what-is-the-full-meaning-of-news-acronym/

[12]. Jamal A. N., Subhani O. K., & Varlamis, I. (2021). Fake news detection: A hybrid CNN-RNN based deep learning approach. International Journal of InformationManagement Data Insights 1(1), 100007. DOI:10.1016/J.JJIMEI.2020.100007

[13]. Kemp, S. (2021, February 11). Digital in Nigeria: All the Statistics You Need in 2021. DataReportal – Global Digital Insights. https://datareportal.com/reports/digital-2021-nigeria

[14]. Kaliyar, R.K., Goswami, A. & Narang, P. (2021) FakeBERT: Fake news detection in social media with a BERT-based deep learning approach. Multimedia Tools and Applications , 80, 11765–11788 . https://doi.org/10.1007/s11042-020-10183-2

[15]. Limeng, C., Kai, S., Suhang, W., Dongwon, L., & Huan, L. (2019). dEFEND: A System for

Explainable Fake News Detection in The 28th ACM International Conference on Information and Knowledge Management (CIKM '19), November 3–7, 2019, Beijing, China. ACM, New York, NY, USA, 4 pages. Beijing, China: ACM, New York, NY, USA https://doi.org/10.1145/3357384.3357862.

[16]. Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. Neurocomputing, 234, 11–26.

[17]. Okoro, N., & Emmanuel, N. O. (2018). Beyond Misinformation: Survival Alternatives for Nigerian Media in the "Post-Truth" Era. African Journalism Studies, 39(4), 67-90. https://doi.org/10.1080/23743670.2018.1551810.

[18]. Popat, K., Mukherjee, S., Yates, A., & Weikum, G. (2018). Declare: Debunking fake news and false claims using evidence-aware deep learning. arXiv:1809.06416.

[19]. Premium Times (2020).Lai Mohammed: We need to 'dominate', control what Nigerians post on Social media. Retrieved December 20, 2022, fromhttps://www.premiumtimesng.com/news/headlines/423463-lia-mohammed-we-need-to-domiante-what-nigerians-postin-social-media

[20]. Reddy, B., Y. (2019). Detection of Fake News from Social Media by Deep Learning Structure DOI:10.13140/RG.2.2.16161.84324

[21]. Salem, F. K. A., Feel, R. A., Elbassuoni, S., Jaber, M., & Farah, M. (2019). FA-KES: A Fake News Dataset around the Syrian War. Proceedings of the International AAAI Conference on Web and Social Media, 13, 573–582. https://doi.org/10.1609/icwsm.v13i01.3254

[22]. Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. ACM SIGKDD Explorations Newsletter, 19(1), 22–36. https://doi.org/10.1145/3137597.3137600

[23]. Shu, K., Mahudeswaran, D., Wang, S., Lee, D., & Liu, H. (2018). FakeNewsNet: A Data Repository with News Content, Social Context and Dynamic Information for Studying Fake News on Social Media. arXiv preprint arXiv:1809.01286 , 22-36.

[24]. Ugwuanyi, S. U. (2017). Influence of Fake News on Public Perception of Nigeria's Online Newspapers. Global Journal of Human-Social Science: An Arts & Humanities -Psychology, 17(5), 4–12.

[25]. Vlachos, A., & Riedel, S. (2014). Fact checking: Task definition and dataset construction. Proceedings of the ACL workshop on language technologies and computational social science, Baltimore, MD, USA (pp. 18–22).

[26]. Wang, W. Y. (2017). Liar, liar pants on fire: A new benchmark dataset for fake news detection. arXiv:1705.00648.

[27]. Zhou, X., & Zafarani, R. (2020). A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. ACM Computing Surveys, 53(5). https://doi.org/10.1145/3395046

Cite this article as :

Samera Uga Otor, Beatrice Obianiberi Akumba, Joseph Sunday Idikwu, "A Novel Fake-News Dataset and Detection System to Mitigate Cyber War with Emphasis on Nigerian News Events", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 4, pp.24-32, July-August-2023. Available at doi : https://doi.org/10.32628/CSEIT23903146
Journal URL : https://ijsrcseit.com/CSEIT23903146