ENGINEERING IN ADVANCED RESEARCHSCIENCE AND TECHNOLOGY

UGC Care Group-I
Vol.03, Issue.01
June-2022
Pages: -675-680

OPTIMIZED CONVOLUTIONAL NEURAL NETWORK FOR FAKE NEWS DETECTION

[1]MOGILI SIVA, [2] P SUBBARAO

[1]M. Tech, Computers Science in Artificial Intelligence, Kakinada Institute of Engineering and Technology - II, Korangi, A.P

[2]Associate Professor, Dept. of CSE, Kakinada Institute of Engineering and Technology - II, Korangi, A.P

ABSTRACT: Recently, there has been a significant and widespread increase in fake news, which is defined as provably false material disseminated with the intent of defrauding people. Because it enhances political polarisation and people's suspicion of their leaders, the spread of this type of misinformation poses a serious threat to social cohesion and well-being. As a result, false news is an issue that is affecting our social lives, notably in politics. To solve this problem, this study provides unique methodologies based on Machine Learning (ML) and Deep Learning (DL) for a false news identification system. The major goal of this research is to discover the best model for achieving high accuracy. As a result, we offer a Convolutional Neural Network model that is optimized for detecting bogus news (OPCNN-FAKE).Using fake news benchmark datasets, we compare the performance of the OPCNN-FAKE with the Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and the six regular ML techniques: Decision Tree (DT), logistic Regression (LR), K Nearest Neighbor (KNN), Random Forest (RF), Support Vector Machine (SVM), and Naive Bayes (NB). The parameters of ML and DL were optimised using grid search and hyperopt optimization approaches, respectively. In addition, standard ML models employed N-gram and Term Frequency Inverse Document Frequency (TF-IDF) to extract features from the benchmark datasets, while DL models used Glove word embedding to encode features as a feature matrix. To assess the OPCNN-performance, FAKE's accuracy, precision, recall, and the F1-measure were used to validate the results. In comparison to other models, the OPCNN-FAKE model has attained the best performance for each dataset. Furthermore, the OPCNN-False outperforms the other models in terms of cross-validation and testing outcomes, indicating that the OPCNN-FAKE is substantially better at detecting fake news than the other models.

Index terms: Convolutional nueral network, passive aggressive classifier, svc, multinomialnb

INTRODUCTION: In recent years, the ability of a user to write anything on online news platforms such as social media and news websites newspapers has led to the propagation of misleading information. Online social media platforms (Twitter, Facebook, Instagram, YouTube, etc.) have become the primary source of news for people around the world, particularly in developing nations. Therefore, anyone from anywhere in the world can use popular social media and social networking as platforms to publish any statement and spread fake news through various networking sites to achieve various goals, which may be illegitimate. We are currently experiencing significant ramifications for society, business, and culture as a result of the increasing use of social media, which have the potential to be both detrimental and beneficial. Fake news is widely regarded as one of the most severe dangers to global commerce, journalism, and democracy, with significant collateral harm. The stock market suffered a $130 billion loss as a result of a false news story saying that the US President Barack Obama had been injured in an explosion. According to statistics published by Stanford University academics, 72.3 percent of fake news originates from official news outlets and online social media platforms. Because of the negative impact of fake news on society, and as fake news is widely regarded as one of the most serious challenges to global commerce, media, and democracy, posing signi_cant societal harm to them, it is critical to build effective fake news detection systems. With the rapid advances in Arti_cial Intelligence (AI),a signi_cant number of experiments are being undertaken to tackle issues that were never addressed in the framework of computer science, such as fake news detection . Automatic detection approaches based on Machine Learning (ML) have been studied to combat the emergence and dissemination of false news. The majority of fake news detection systems utilize ML approaches to help consumers in _ltering the content they are seeing and determining if a given news piece is misleading or not. Deep Learning (DL) techniques recent accomplishments in difficult natural language processing tasks make them viable for detecting fake news effectively and efficiently. Creating automatic, trustworthy, and accurate systems for identifying fake news on social media is a trending topic of research. The process of determining if a certain news item on any _eld, from any social media domain, is purposefully or inadvertently misleading might be characterized as fake news detection. Convolutional Neural Network (CNN) has been prominent in many _elds with the best performance, including computer vision, smart building structures, and natural language processing. CNN uses convolution layers, pooling layers, and fully connected layers to extract more features with high-level and low-level features. Therefore, we have proposed an Optimal CNN model for Fake news detection (OPCNN-Fake) that can extract high-level and low-level features from the dataset to detect fake news, and it has registered the best performance compared with others models.

EXISTING SYSTEM: Classification of any news item /post / blog into fake or real one has generated great interest from researchers around the globe. Several research studies have been carried out to find effect of falsified and fabricated news on masses and reactions of people upon coming through such news items. Falsified news or fabricated post news is any textual or non-textual content that is fake and is generated so the readers will start believing in something which is not true. For Example recently a news item was floated on social media networking platform Facebook by a accredited Journalist of Srinagar J&K Titled " Beasts in White Aprons " regarding mismanagement and carelessness of doctors in a local Pediatric hospital of Srinagar with a Image

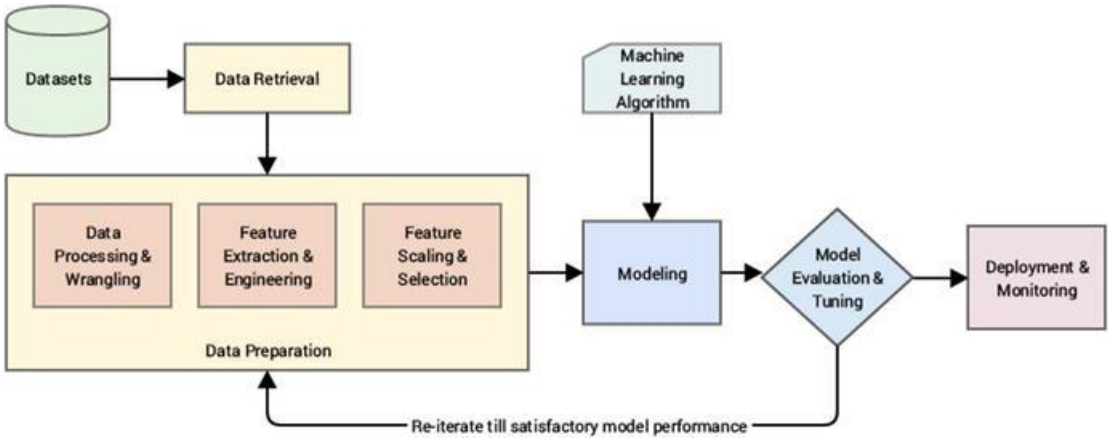
DISADVANTAGES

- The easy access and exponential growth of the information available on social media networks has made it intricate to distinguish between false and true information
- . The easy dissemination of information by way of sharing has added to exponential growth of its falsification.
- The credibility of social media networks is also at stake where the spreading of fake information is prevalent.

- Thus, it has become a research challenge to automatically check the information viz a viz its source, content and publisher for categorizing it as false or true. Machine learning has played a vital role in classification of the information although with some limitations

PROPOSED SYSTEM The easy dissemination of information by way of sharing has added to exponential growth of its falsification. The credibility of social media networks is also at stake where the spreading of fake information is prevalent. Thus, it has become a research challenge to automatically check the information viz a viz its source, content and publisher for categorizing it as false or true. Machine learning has played a vital role in classification of the information although with some limitations. This paper reviews various Machine learning approaches in detection of fake and fabricated news. The limitation of such and approaches and improvisation by way of implementing deep learning is also reviewed.

SYSTEM ARCHITECHTURE



Modules:

Load Dataset:

- Load data set using pandas read_csv() method.

.Split Data Set:

- Split the data set to two types. One is train data test and another one is test data set.

.Train data set:

- Train data set will train our data set using fit method.

.Test data set:

- Test data set will test the data set using algorithm.

.Predict data set:

- Predict() method will predict the results.

RESULT:

COMPARISION TABLE

Ref	The number of dataset	Dataset name	Models' performance
[33] (2020)	One dataset	FakeNewsNet (PolitiFact topic)	CNN Accuracy= 62.9% and F1-score =58.3%
		FakeNewsNet (GossipCop topic)	CNN Accuracy= 72.3% and F1-score =72.5%
[27] (2020)	One dataset	The dataset from Kaggle	hybrid CNN-LSTM Accuracy =97.5%
[29] (2021)	Two datasets	ISOT dataset	hybrid CNN-RNN Accuracy =100% of the training set
		FA-KES dataset	hybrid CNN-RNN Accuracy = 60% of training set
[28] (2020)	One dataset	The dataset from Kaggle	The CNN only (Accuracy of 91.50%, Precision = 90.74%, Recall = 92.07%, F1-Score = 91.40%) The FNDNet (deeper CNN) model (Accuracy of 98.36%, Precision =99.40%, Recall =96.88%, F1-Score =98.12%)
[30] (2019)	Two dataset	Weibo	Accuracy = 88.82%
		NewsFN	Accuracy=90.10%
Our work	Four datasets	Dataset from Kaggle (dataset1)	The OPCNN-FAKE model Cross-validation results (Accuracy = 99.99%, Precision = 100%, Recall = 99.97%, and F1-score = 99.97%). Testing set (Accuracy = 97.84%, Precision = 97.86%, Recall = 97.84%, and F1-score = 97.84%).
		FakeNewsNet	The OPCNN-FAKE model Cross-validation results (Accuracy = 98.65%, Precision = 99.34%, Recall = 98.87%, and F1-score = 99.1%). Testing results (Accuracy = 95.26%, Precision = 95.28%, Recall = 95.26%, and F1-score = 95.27%).
		FA-KES5	The OPCNN-FAKE model Cross-validation results (Accuracy = 97.23%, Precision = 97.37%, Recall = 97.58%, and F1-score = 97.26%) Testing results (Accuracy = 53.99%, Precision = 53.86%, Recall = 53.91%, and F1-score = 53.99%).
		ISOT	OPCNN-FAKE Cross-validation results (Accuracy = 100%, Precision = 100%, Recall = 100%, and F1-score = 100%). Testing results (Accuracy = 99.99%, Precision = 99.99%, Recall = 99.99%, and F1-score = 99.99%).

CONCLUSION

This paper has introduced a fake news detection system using two approaches, namely, regular ML and DL. In DL, we have proposed the OPCNN-FAKE model that has achieved the best performance. The proposed OPCNN-FAKE model consists of six layers: an embedding layer, a dropout layer, a convo-lutional layer, a pooling layer, a atten layer, and an output layer. Also, it has been optimized using hyperopt optimiza-tion technique; the different values of parameters for each layer have been adapted, and the best values that achieved the best performance have been selected. Also, n-gram with TF-ID and word embedding feature extraction methods have been used for ML and DL, respectively. We compared the OPCNN-FAKE with RNN, LSTM, and the six regular ML techniques: DT, LR, KNN, RF, SVM, NB using four fake news benchmark datasets. Each dataset has been split into 80% training dataset and 20% testing dataset. The training datasets have been used to optimize and train the models, while the testing datasets were used to evaluate the mod-els.

Also, cross-validation and the testing results have been registered showing that the OPCNN-FAKE model has achieved the best performance for each dataset compared with the other models. For dataset1, the OPCNN-FAKE model achieved the best performance for the testing results (Accuracy D 97.84%, precision D 97.86%, recall D 97.84%, and F1-score D 97.84%). For dataset2 (Fake Newsnet), in the testing results, OPCNN-FAKE model has achieved the high-est performance for the testing results (Accuracy D 95.26%, precision D 95.28%, recall D 95.26%, and F1-score D 95.27%). For dataset3, OPCNN-FAKE model has achieved the highest performance for the testing results (Accuracy D 53.99%, precision D 53.86%, recall D 53.91%, and F1-score D 53.99%). For dataset4, the OPCNN-FAKE model has achieved the highest performance for the testing results (Accuracy D 99.99%, precision D 99.99%, recall D 99.99%, and F1-score D 99.99%). In future, we will use our proposed model to detect COVID-19 fake news. Also, we plan to apply multimodel-based methods with recently pre-trained word embeddings (i.e., Elmo, XLNet, etc.) to handle visual information like video and images. In addition, we may use knowledge-based and fact-based approaches to detect fake news. We will also expand our planned dataset to include data from additional languages.

References:

[1] M. H. Goldani, R. Safabakhsh, and S. Momtazi, ``Convolutional neural network with margin loss for fake news detection,'' *Inf. Process. Manage.*, vol. 58, no. 1, Jan. 2021, Art. no. 102418.

[2] A. Bondielli and F. Marcelloni, ``A survey on fake news and rumour detection techniques,'' *Inf. Sci.*, vol. 497, pp. 38_55, Sep. 2019.

[3] K. Rapoza, ``Can 'fake news' impact the stock market,'' Tech. Rep., 2017. [Online]. Available: https://www.forbes.com/sites/kenrapoza/2017/02/26/ can-fake-news-impact-the-stock-market/?sh=187125e22fac

[4] H. Allcott and M. Gentzkow, ``Social media and fake news in the 2016 election,'' *J. Econ. Perspect.*, vol. 31, no. 2, pp. 211_236, 2017 .

[5] L. Bozarth and C. Budak, ``Toward a better performance evaluation framework for fake news classi_cation,'' in *Proc. Int. AAAI Conf. Web Social Media*, vol. 14, 2020, pp. 60_71.

[6] Y. Liu and Y.-F. B. Wu, ``Early detection of fake news on social media through propagation path classi_cation with recurrent and convolutional networks,'' in *Proc. 32nd AAAI Conf. Artif. Intell.*, Apr. 2018, pp. 354_361.

[7] F. Monti, F. Frasca, D. Eynard, D. Mannion, and M. M. Bronstein, ``Fake news detection on social media using geometric deep learning,'' 2019, *arXiv:1902.06673*. [Online]. Available: http://arxiv.org/abs/1902.06673

[8] S. Girgis, E. Amer, and M. Gadallah, ``Deep learning algorithms for detecting fake news in online text,'' in *Proc. 13th Int. Conf. Comput. Eng. Syst. (ICCES)*, Dec. 2018, pp. 93_97.

[9] S. Ghosh and C. Shah, ``Towards automatic fake news classi_cation,'' *Proc. Assoc. Inf. Sci. Technol.*, vol. 55, no. 1, pp. 805_807, Jan. 2018.

[10] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, ``Convolutional neural networks: An overview and application in radiology," *Insights Imag.*, vol. 9, pp. 611_629, Aug. 2018.