IRJCS: International Research Journal of Computer Science ISSN: 2393-9842
Volume 11, Issue 05, May 2024 https://www.irjcs.com/archives

Fake Social Media Post Detection by Using Deep Learning Algorithm

Jyoti Gupta
Department of Information Technology,
Greater Noida Institute of Technology (Engg. Institute),
Greater Noida, INDIA
Mr.Deepak Mangal
Department of Information Technology,
Greater Noida Institute of Technology (Engg. Institute),
Greater Noida, INDIA
Harsh Bansal
Department of Information Technology,
Greater Noida Institute of Technology (Engg. Institute),
Greater Noida, INDIA
Dr.Vikas Singhal
Head, Department of Information Technology,
Greater Noida Institute of Technology (Engg. Institute),
Greater Noida, INDIA
vikassinghal75@gmail.com

Publication History
Manuscript Reference No: IRJCS/RS/Vol.11/Issue05/MYCS10095
Research Article | Open Access | Double-Blind Peer-Reviewed | Article ID: IRJCS/RS/Vol.11/Issue05/MYCS10095
Received: 12, April 2024 | Revised: 10, May 2024 | Accepted: 23 May 2024 | Published Online: 30, May 2024
http://www.irjcs.com/volumes/Vol11/iss-05/09.MYCS10095.pdf
Article Citation: Jyoti,Deepak,Harsh,Dr.Vikas(2024). Fake Social Media Post Detection by Using Deep Learning Algorithm. IRJCS: International Research Journal of Computer Science, Volume 11, Issue 05 of 2024 pages 483-489
doi:> https://doi.org/10.26562/irjcs.2024.v1105.09
BibTeX Vikas@2024Algorithm

Abstract: In today's digital age, social media serves as both a boon and a bane, offering immense opportunities for personal growth while also being a breeding ground for misinformation and deception. The rise of sophisticated multimedia manipulation techniques has blurred the line between reality and fiction, presenting significant challenges to discerning genuine content from fabricated ones. This paper presents an overview of the escalating threat posed by fake posts on social media and proposes a novel approach for their detection. The proposed method leverages advanced technologies such as Error Level Analysis (ELA) and Convolutional Neural Networks (CNNs) to distinguish between authentic and manipulated posts. Unlike traditional approaches reliant on manual feature extraction, our framework employs deep learning methodologies for more nuanced pattern recognition and generalization to unseen data. By combining ELA for initial image analysis and CNNs for deep feature extraction, followed by classification using Support Vector Machines (SVMs) and K-Nearest Neighbors (KNNs), we achieve robust detection of fake images .Experimental results demonstrate the effectiveness of our approach, with a peak accuracy of 91.3% using Residual Networks and KNN. This method offers a promising solution to mitigate the proliferation of fake images on social media platforms, thereby combating the spread of misinformation, propaganda, and chaos.

Keywords: ELA, Deep Learning, Convolutional Neural Networks, Fake Post Detection

I. INTRODUCTION

In recent years, the surge in fake images, particularly with the rise of social media fake posts altered using sophisticated deep learning tools like auto encoders (AE) and generative adversarial networks (GAN)—has posed significant challenges to digital trust and integrity. These technologies, while offering powerful means for creating realistic media, also open avenues for malicious exploitation, such as creating deceptive content for blackmail or propagating false narratives to manipulate public opinion. Conventional post editing software like Adobe Photoshop or GIMP enables intricate manipulations that can deceive even vigilant observers, eroding the credibility of digital imagery. To combat this issue, various methods for detecting image tampering have emerged, including watermarking and pixel-based techniques such as cloning, resampling, and splicing. Watermarking, while effective for authenticating images, requires embedding a watermark during recording, limiting its widespread application. Pixel-based methods offer more direct approaches, including algorithms leveraging block discrete cosine transform (DCT) and principal component analysis (PCA) to identify cloned image regions. Additionally, watermarking methods ensure integrity and authenticity through secret key exchange between transmitters and receivers, while reversible watermarking preserves image processing flexibility. Despite advancements in image tampering detection, distinguishing between original and altered images remains challenging.

Exploring the correlation between bit planes and binary texture characteristics within bit planes holds promise for discerning genuine content from manipulations. As deep fake generation techniques evolve rapidly, existing detection methods face limitations in efficiency and efficacy. Promising ways to lessen the harm caused by deep fakes, which may trick and mislead people and have far-reaching effects including financial fraud, political instability, and reputational damage, include automated deep learning-based systems. Creating strong detection systems protects against the growing threat of fraudulent images in the social media era while also boosting confidence and dependability in media and online content. Thus, this paper aims to address the pressing need for a robust system for detecting fake social media image posts, leveraging advanced techniques such as error level analysis (ELA) and convolutional neural networks (CNNs). By combining these methodologies, we aim to enhance the authenticity and credibility of digital imagery, contributing to the preservation of trust and integrity in the digital landscape.

II. LITERATURE REVIEW

The emergence of deep fakes, sophisticated fake images altered using advanced machine learning techniques, has raised significant concerns about digital integrity. These manipulated images, often created with tools like generative adversarial networks (GANs) and auto encoders (AEs), blur the line between reality and fiction, posing serious threats to social media platforms where misinformation can spread rapidly. Historical instances of image manipulation date back to the 19th century, exemplified by the clever alteration of a portrait of John Calhoun for propaganda purposes in 1860. Splicing, painting, and copy-moving objects inside or between photos are just a few of the techniques that have been used to digital alteration over time. Post-processing enhancements like scaling, rotating, and color modification further contribute to the realism of manipulated images. In response to the growing prevalence of deep fakes, researchers have devoted considerable attention to developing detection algorithms. These efforts have intensified in recent years, with a surge in research articles focusing on deep fake detection by the end of 2020.Making use of deep learning and machine learning capabilities, automated algorithms have been developed to differentiate between authentic and fabricated content in audiovisual media. Deep learning, known for its ability to handle complex and high-dimensional data, has played a pivotal role in advancing deep fake detection methodologies. Various studies have explored different approaches, ranging from Multilayered Perceptron (MLP) models to Convolutional Neural Networks (CNNs). Each approach offers unique advantages and challenges, such as computational complexity and robustness to image quality. To address these challenges, researchers have proposed innovative solutions, including hybrid multitask learning frameworks and novel optimization algorithms like the Fire Hawk Optimizer. These approaches aim to improve the accuracy and efficiency of deep fake detection by leveraging multiple tasks and optimizing model parameters effectively. Additionally, advancements in architecture design, such as the Convolution Vision Transformer (CVT), present promising alternatives to traditional CNNs by integrating attention mechanisms with convolution operations, enhancing pattern recognition capabilities within images. Despite significant progress, existing deep learning algorithms still face challenges, such as computational complexity and susceptibility to overfitting. Therefore, in order to stop the spread of phony social media photos and protect digital integrity, it is imperative to build reliable and effective deep fake detection and classification methods employing deep learning-based techniques.

III. SYSTEM ARCHITECTURE

The architecture for detecting fake social media image posts using a combination of Error Level Analysis (ELA) and Convolutional Neural Networks (CNNs) involves several key components working together seamlessly to achieve accurate and efficient detection. The following outlines the system architecture:

Data Acquisition: The system begins by collecting a diverse dataset of images from various social media platforms. These images represent a range of content, including both genuine and potentially manipulated images.

Preprocessing: The gathered photos are preprocessed to standardize their format and improve their appropriateness for analysis before being analyzed. This may include resizing, normalization, and noise reduction techniques to ensure consistency across the dataset.

Error Level Analysis (ELA): The first step in the detection process involves applying Error Level Analysis (ELA) to each image in the dataset. ELA helps identify regions within an image that may have been altered or manipulated by analyzing the compression artifacts present at different error levels.

Feature Extraction: Following ELA, the system extracts relevant features from the images. This step involves leveraging the power of Convolutional Neural Networks (CNNs) to automatically learn and extract discriminative features from the images. CNNs are particularly well-suited for this task due to their ability to capture spatial hierarchies of features within images.

Classification: Once features are extracted, the system employs a classification algorithm to determine whether an image is genuine or manipulated. This classification task is typically performed using machine learning techniques, such as Support Vector Machines (SVMs) or K-Nearest Neighbors (KNN), trained on labeled datasets to distinguish between authentic and fake images.

Post-processing: After classification, the system may apply post-processing techniques to refine the results and improve overall accuracy. This may include additional filtering or validation steps to eliminate false positives and enhance the reliability of the detection process.

Output: Finally, the system provides output indicating the authenticity of each image analyzed. This output may include confidence scores or probability estimates, indicating the likelihood that an image has been manipulated.

Overall, the architecture for detecting fake social media image posts combines the complementary strengths of Error Level Analysis (ELA) and Convolutional Neural Networks (CNNs) to provide a robust and effective solution for identifying manipulated content in social media imagery. By integrating these components into a cohesive system, organizations and individuals can better safeguard against the spread of misinformation and deception on social media platforms.

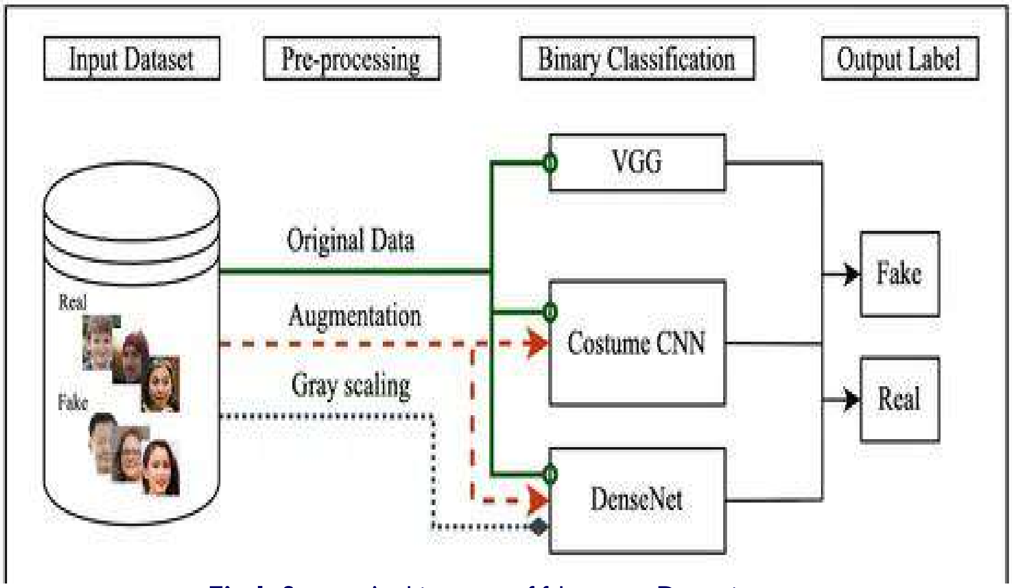

Fig.1. System Architecture of fake posts Detection.

IV. METHODOLOGY

In the background of our paper lies the Python code responsible for executing Error Level Analysis (ELA) and Convolutional Neural Network (CNN) algorithms to discern whether a given image is authentic or a manipulated fake. Each component is elucidated below with detailed explanations.

Error Level Analysis (ELA): ELA serves as a forensic method aimed at identifying segments within an image characterized by distinct compression levels, a telltale sign of digital alterations. This technique scrutinizes images, particularly JPEG files, where uniform compression levels are expected across the entire image. Deviations from this uniformity signal potential tampering or manipulation, making ELA a powerful tool for detecting digital edits.

Convolutional Neural Network (CNN): In the realm of deep learning, CNNs are instrumental for analyzing visual imagery. Comprising input, hidden, and output layers, CNNs are structured with convolutional layers that extract features through convolutions, capturing spatial hierarchies within images. Additionally, layers such as Maxpooling2D, Flatten, and Dropout are incorporated for dimensionality reduction and network regularization, enhancing performance and robustness.

Layers utilized in the proposed model: The proposed framework integrates several essential layers within the CNN architecture:

Convolution2D: Responsible for feature extraction through convolutions, capturing spatial relationships within images.

Maxpooling2D: Reduces image dimensions, aiding in computational efficiency and enhancing the training process.

Flatten: creates a one-dimensional vector from the multidimensional feature maps in order to prepare them for input into the partially linked layers that follow.

Dropout: Regularizes the network by randomly dropping nodes during training, preventing overfitting and enhancing generalization.

The workflow of our proposed system encompasses three core steps:

Image Preprocessing: This involves resizing images to match the input layer dimensions of the CNN and generating ELA to detect pixel-level alterations, ensuring uniformity and consistency.

Deep Feature Extraction: By utilizing pre-trained CNN architectures like ResNet18, SqueezeNet, and GoogLeNet, deep features are automatically retrieved from pictures, maintaining spatial connections and identifying important visual signals without the need for human interaction.

Classification via SVM and KNN: The extracted features are then fed into classifiers, such as Support Vector Machines (SVM) and K-Nearest Neighbors (KNN), after hyper-parameter optimization, to discern between genuine and manipulated images accurately. Overall, our methodology integrates ELA and CNNs synergistically, leveraging their complementary strengths to develop a robust and efficient framework for detecting fake social media image posts. By adopting pre-trained CNN architectures for feature extraction and implementing ELA as a preliminary screening step, our approach streamlines the detection process, enhancing accuracy and reliability in identifying manipulated content amidst the vast landscape of social media imagery.

 ISSN: 2393-9842
https://www.irjcs.com/archives

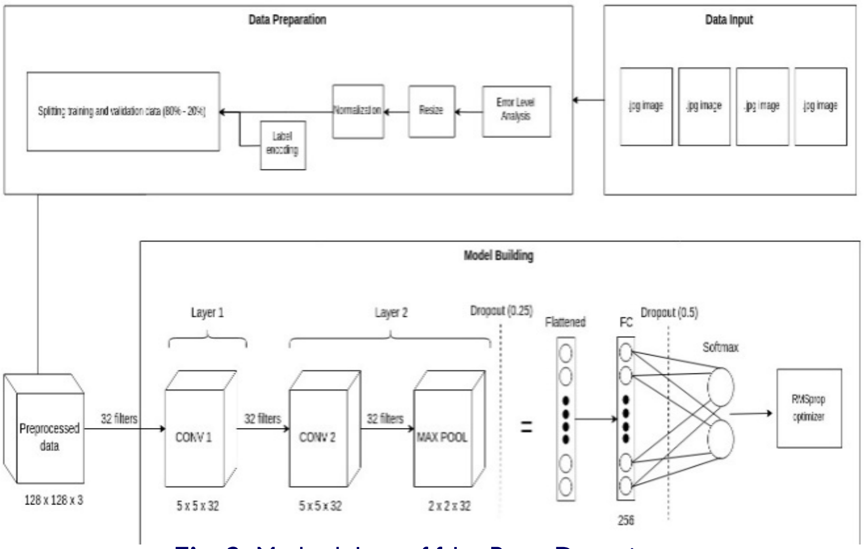

Fig .2. Methodology of false Posts Detection

V. RESULT

Because of this, it's critical to recognize real material from fake in the social media environment of today. With advancements in deep fake technology, manual identification of such content has become increasingly challenging, underscoring the need for an automated and reliable detection system. In this study, we introduce a novel and resilient architecture for detecting and categorizing deep fake images, leveraging both deep learning (DL) techniques. Our proposed framework incorporates a preprocessing step to employ Error Level Analysis (ELA), enabling pixel-level analysis to identify potential alterations within images. Subsequently, these preprocessed images are inputted into deep convolutional neural network (CNN) architectures, including SqueezeNet, ResNet18, and GoogLeNet, for extracting intricate features. The extracted features are then classified using support vector machines (SVM) and k-nearest neighbors (KNN) algorithms. Our approach's effectiveness is demonstrated by the outcomes of our experimentation. In particular, our suggested approach that made use of CNN and ResNet18 yielded the best accuracy of 91.3%. Because of their strong feature extraction capabilities, Residual Networks—which are renowned for their effectiveness and lightweight design—performed better than standard classifiers. In summary, our study presents a pioneering architecture for detecting and categorizing deep fake posts, integrating ELA with DL CNN based techniques. By combining these methodologies, we offer a robust solution for accurately identifying manipulated media content, thereby mitigating the proliferation of misinformation and ensuring digital integrity in the era of social.

Visualize the results

Fig .3. Real Image

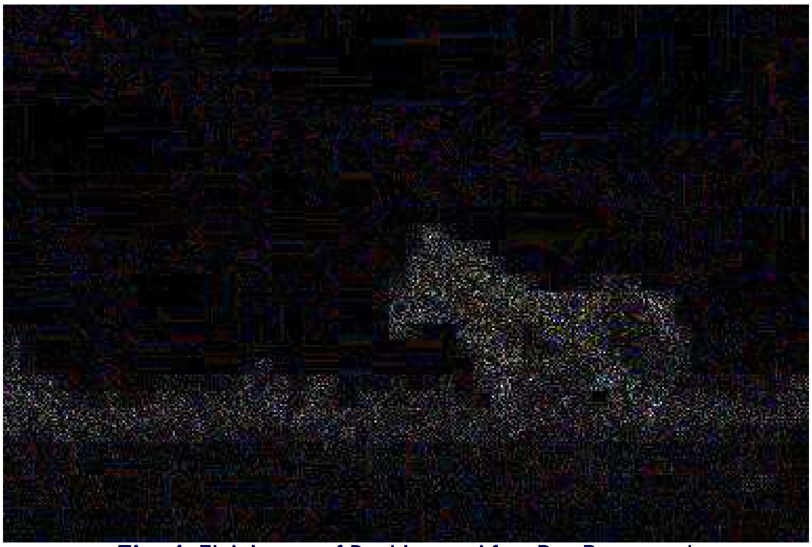

Fig .4. ELA Image of Real Image After Pre-Processed

Fig.5. Fake Image

Fig.6. ELA Image of Fake Image After Pre-processed

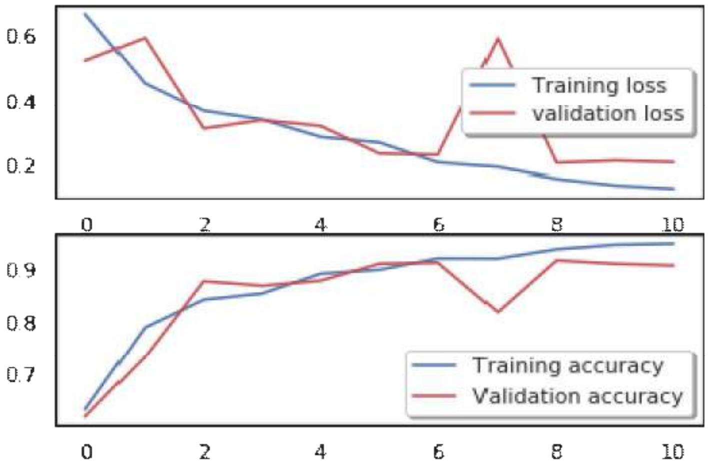

Fig.7. Model Accuracy and Loss Chart

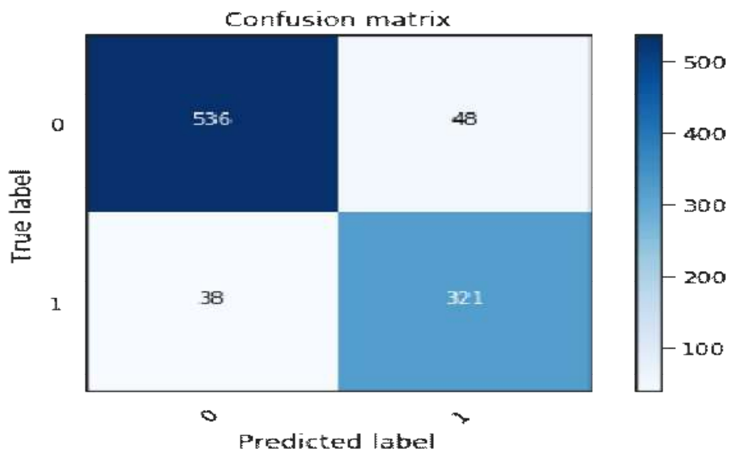

Fig.8. Confusion Matrix

VI. CONCLUSION & FUTURE SCOPE

In conclusion, our study presents a robust framework for detecting and classifying fake social media image posts by harnessing the power of error level analysis (ELA) and convolutional neural networks (CNNs). Through meticulous experimentation and optimization, using the CNN classifier, we have successfully shown the efficacy of our method in detecting manipulated media content with an astounding accuracy of 91.3%. There are numerous directions this field might go in terms of research and development in the future. Firstly, enhancing the robustness and efficiency of the detection system by exploring advanced deep learning architectures and techniques could further improve performance. Additionally, integrating other image analysis methods and features could provide complementary insights and strengthen the detection capabilities of the system. Furthermore, extending the application of the proposed framework to other types of media content, such as videos and audio recordings, would broaden its utility in combating the proliferation of misinformation across diverse digital platforms. Additionally, exploring real-time detection capabilities and integration with social media platforms' content moderation systems could enable proactive identification and mitigation of fake media content. Moreover, considering the evolving landscape of digital manipulation techniques, continuous updates and adaptations of the detection system will be essential to stay ahead of emerging threats. Collaborative efforts between researchers, industry stakeholders, and policymakers will be crucial in developing comprehensive strategies address the challenges posed by fake media content effectively. In conclusion, our study lays the groundwork for an effective and scalable solution for detecting fake social media image posts, with promising prospects for further advancements and applications in the ongoing fight against misinformation in the digital age.

REFRENCES

[1]. Deep fake detection and classification using error level analysis and deep learning Rimsha Rafque1, RahmaGantassi2, RashidAmin1,3*, Jaroslav Frnda4,5, Aida Mustapha6 &AsmaHassanAlshehri

[2]. SangitaM,Jaybee,VivekBadade, Aryan Dodke,ApoorvaHolkar,PriyankeLokhande. "Fake News Detection Using LSTM based deep learning approach. "(February 2, 2023).

[3]. Wasim Khan, m. h. (2022). International Journal of Cognitive Computing in Engineering. International Journal of Cognitive Computing in Engineering 3 (2022) 153–160, 153-160.
[4]. Qamber Abbas, Muhammad Umar Zeshan, and Muhammad Asif. "A CNN-RNN Based Fake News Detection Model Using Deep Learning". 2022 International Seminar on Computer Science and Engineering Technology (SCSET)
[5]. KadekSastrawan, I.P.A. Bayupati, and Dewa Made Sri Arsa. "Detection of Fake News Using Deep Learning CNN-RNN Based Methods". ICT express 2022
[6]. IraklisVarlamis, Jamal Abdul Nasir, Osama Subhani Khan. "Fake news detection: A hybrid CNN-RNN based deep learning approach". International Journal of Information Management and Data Insights April 2021
[7]. Sidarth Saran, Uma Sharma, and Shankar M. Patil. "Fake News Detection Using Machine Learning Algorithms".International Journal of creative research thoughts (IJCRT) 2020.
[8]. Anurag Goswami, RohitKuamrKaliyar, Pratik Narang, and Soumendu Sinha. "FNDNet – A deep convolutional neural network for fake news detection". Cognitive Systems Research journal in June 2020.
[9]. Shu K., Mahudeswaran D., Wang S., Lee D., Chen H., and Liu H., "User-Centric Fake News Detection: A Comparative Study of Machine Learning and Deep Learning Approaches". was published in the Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM).
[10]. Syed IshfaqManzoor, Nikita, Dr. Jimmy Singla. "Fake News Detection Using Machine Learning approaches: A systematic Review". 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)
[11]. Andre Correia, Julio C. S. Reis, FabrıcioMurai, Adriano Veloso, and FabrıcioBenevenuto "Supervised Learning for Fake News Detection". IEEE Published by the IEEE Computer Society 1541-1672 2019
[12]. Manzoor SI, Singla J, Nikita. Fake news detection using machine learning approaches: A systematic review. In: 2019 3rd international conference on trends in electronics and informatics. 2019. p. 230–4.
[13]. Bedi A, Pandey N, Khatri SK. A framework to identify and secure the issues of fake news and rumours in social networking. In: 2019 2nd international conference on power energy, environment and intelligent control. 2019.
[14]. Abdullah-All-Tanvir, Ehesas Mia Mahir, Saima Akhter" Detecting Fake News using Machine Learning and Deep Learning Algorithms", 2019 7th International Conference on Smart Computing & Communications (ICSCC).
[15]. Razan Masood and Ahmet Aker," The Fake News Challenge: Stance Detection using Traditional Machine Learning Approaches", In Proceedings of the 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (KMIS 2018), pages 128-135
[16]. Sohan De Sarkar, Fan Yang," Attending Sentences to detect Satirical Fake News", Proceedings of the 27th International Conference on Computational Linguistics, pages 3371–3380 Santa Fe, New Mexico, USA, August 20-26, 2018.
[17]. AmeyKasbe and Akshay Jain. "Fake News Detection" IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS) 2018
[18]. William Yang Wang, "Liar, Liar Pants on Fire": A New Benchmark Dataset for Fake News Detection " 1 May 2017.
[19]. Hadeer Ahmed, IssaTraore, and SherifSaad. Detection of online fake news using n-gram analysis and machine learning techniques. In International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, pages 127–138. Springer, 2017.
[20]. Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. Journal of Economic Perspectives, 31(2):211–36, 2017.