JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

BAITAWARE : A CLICKBAIT DETECTION SYSTEM USING DEEP LEARNING

Fardeen Kachawa[1], Sujal Bhatt[1], Kaif Siddique[1], Bhavesh Choudhary[1] , Neelam Phadnis[2]

[1]Computer Engineering Department, Student, Shree L.R. Tiwari College of Engineering, Mira Road, Thane, Maharashtra, India
[2]Computer Engineering Department, Assistant Professor, Shree L.R. Tiwari College of Engineering, Mira Road, Thane, Maharashtra, India

ABSTRACT

In the age of information proliferation on the internet, the ubiquitous presence of clickbait content poses a serious challenge to content consumers, affecting digital literacy and cybersecurity. This project introduces the development of a Clickbait Detection System, leveraging advanced deep learning techniques to empower users with the means to discern and categorize clickbait effectively. This system is driven by the urgent need to counter the dissemination of deceptive and sensationalized content, providing users with a tool to make informed decisions about the content they engage with. The core objectives of the project are twofold: First, to create a user-friendly web-based application using the Python programming language and the Flask framework, allowing users to easily upload images and receive real-time clickbait detection results. Second, to implement advanced deep learning algorithms that process and analyze images for clickbait elements, utilizing Optical Character Recognition (OCR) for text extraction and Convolutional Neural Networks (CNNs) for image recognition. These deep learning models, enhanced through extensive training, are poised to provide accurate clickbait detection results. In conclusion, this initiative offers a thorough and user-centered solution to the urgent problem of clickbait material. Modern deep learning algorithms are incorporated into a user-friendly web application to give people the tools they need to safely traverse the digital world while promoting digital literacy and boosting online security.

Keywords: Clickbait, Deep Learning, Machine Learning, CNN, Optical Character Recognition(OCR)

I. INTRODUCTION

As clickbait has grown in popularity, it has become a major obstacle to online content consumption in the age of information overload and click value. Clickbait attempts to trick people into clicking on information that frequently falls short of expectations by using dramatic titles and deceptive images. It follows that there is now a greater demand than ever for reliable clickbait detection tools.

Building a successful detection system requires a fundamental understanding of the dynamics of clickbait. As a result, we enter into the examination of large-scale datasets from earlier research that include a variety of online platforms and content categories. Carefully selected samples of both clickbait and non-clickbait situations are included in these databases.

Analyzing the trends and characteristics found in previous clickbait data serves as the basis for our methodology. We take into account a number of factors, including headline structure, subtle language, picture qualities, and interaction metrics, in order to obtain an understanding of the changing tactics that clickbait producers utilize.

Furthermore, we examine the difficulties encountered by conventional techniques for identifying clickbait, recognizing their constraints in adjusting to the constantly evolving domain of web content.

Large datasets are used to train the algorithm to discriminate between samples that are and are not clickbait. The system is continuously improved and gets better at identifying misleading information thanks to continuous improvement techniques and real-time monitoring, which frequently involve user input. This method makes use of deep learning's advantages to develop reliable and flexible clickbait detection systems that can keep up with the constantly changing online content market.

Clickbait identification is only one of the many difficult problems that deep learning, a subset of machine learning, has shown promise in resolving. Artificial neural networks are used in deep learning based clickbait detection to automatically discern between sensationalized and authentic content by analyzing various textual and visual features associated with online content. By helping consumers avoid false information, this technology also helps content platforms maintain their integrity.

The area of deep learning-based clickbait detection is examined in this introduction. We will look at the motivations behind these systems' development, the techniques and approaches employed, the challenges faced, and the potential impacts these systems could have on the online ecosystem.

II. LITERATURE REVIEW

Clickbait detection systems now in use have developed through a variety of approaches. To improve detection accuracy, the LSACD model incorporates lure and similarity variables together with an adaptive weighted sum technique. Language obstacles are overcome via a CNN-based strategy, which reliably produces results without depending on language-specific factors. In terms of accuracy, another CNN approach outperforms other machine learning algorithms. The RNN model, on the other hand, benefits greatly from contextual awareness and uses both left and right context to boost categorization. Bait Radar's multi-model deep learning method exhibits great accuracy and quick real-time inference, while deep learning models like as the Bidirectional LSTM obtain impressive accuracy rates.

Some models use deep learning frameworks to outperform state-of-the-art methods, while others expand into identifying other forms of clickbait by utilizing data from social media sites. In a seminal work, the misleading nature of clickbait is emphasized, and a deep learning model is shown, highlighting the critical need to solve this urgent online issue. Together, these algorithms improve accuracy, real-time capabilities, and the classification of various clickbait kinds, contributing to the constantly changing environment of clickbait detection.

The author Amol Agarwal [11] presents a model for detection of clickbait by utilizing convolutional neural networks and presents a compiled clickbait corpus. We create a corpus using multiple social media platforms and utilize deep learning for learning features rather than undergoing the long and complex process of feature engineering. Our model achieves high performance in identification of clickbaits. Index Terms—Clickbait, convolutional neural networks, deep learning. Clickbaits, in social media, are exaggerated head-lines whose main motive is to mislead the reader to “click” on them. They create a nuisance in the online experience by creating a lure towards poor content. Online content creators are utilizing more of them to get increased page views and thereby more ad revenue without providing the backing content.

The authors of this paper, Bilal Naeem, Aymen Khan, Mirza Omer Beg, Hasan Mujtaba[12], presents a study on deep learning framework for clickbait detection, utilizing a modified version of LSTM model and linguistic analysis .The framework focuses on understanding the semantics of clickbait headlines and classifying them as either clickbait or legitimate news .The dataset for training and evaluation is collected from Reddit, a social media platform, ensuring relative consistency and authenticity. The proposed framework outperforms state-of-the-art techniques with a classification accuracy of 97% .The paper contributes to the field by proposing a modified LSTM deep learning framework, generating a labeled dataset, and providing a thorough evaluation of the technique. The approach considers the intent and semantics of text headlines, detecting key characteristics of clickbait such as curiosity and sensationalism The paper highlights the importance of scalable solutions for analyzing sentence structures and grammar within high volumes of social media data.

The authors, Bhanuka Gamage, Adnan Labib, Aisha Joomun, Chern Hong Lim, and KokSheik Wong[13], discuss the issue of clickbait on popular online platforms like youtube, which provokes users to click on videos using attractive titles and thumbnails. As a result, users ended up watching a video that does not have the content as publicized in the title. This issue is addressed in this study by proposing an algorithm called BaitRadar, which uses a deep learning technique where six inference models are jointly consulted to make the final classification decision. These models focus on different attributes of the video, including title, comments, thumbnail, tags, video statistics and audio transcript. The final classification is attained by computing the average of multiple models to provide a robust and accurate output even in situations where there is missing data. The proposed method is tested on 1,400 YouTube videos. On average, a test accuracy of 98% is achieved with an inference time of ≤ 2s.

The authors, Saumya Pandey and Gagandeep Kaur[14], put forward some interesting studies on the clickbait content that is omnipresent in our day to day life.Massive outreach of the online media along with changing information consumption patterns have revealed the dark side of digital media. The intentional use of tempting, eye-catching, exaggerated and misleading content to capitalize on the voracious appetite of the readers by creating an information gap has flooded the news websites. Thus, we were motivated to develop deep learning models that utilize the lexical as well as semantic features of the headline and the corresponding text to effectively detect clickbait. The Bidirectional Long Short-Term Memory model using GloVe embedding achieves an accuracy of 98.78% that outperforms the previous work. Furthermore, our study for using the Genetic algorithm for hyperparameter optimization also gave promising results with an accuracy of 95.61%.

The authors of this paper, Sarjak Chawda, Aditi Patil, Abhishek Singh, Prof. Ashwini Save[15], provides in-depth information about clickbait which can be found in sensational headlines that often exaggerate facts, usually to entice readers to click on them. Many researchers have proposed different techniques involving various Machine Learning algorithms such as Support Vector Machine (SVM), Decision Tree, Random Forest, and Deep Learning techniques such as Recurrent Neural Network (RNN), Long Short Term Memory (LSTM) and Convolutional Neural Network (CNN). Although there have been previous attempts by many researchers on detection of Clickbait titles, very few have taken into consideration the context of the title. Context plays a vital role in capturing the semantics of the text. Misclassification of Clickbait titles can be avoided using context. The Recurrent Convolutional Neural Network (RCNN) considers the context for text classification. In this system, clickbait classification is done using RCNN model, and later enhanced with LSTM and Gated Recurrent Unit (GRU) to capture long term dependencies and provide better accuracy than the previous state-of-the-art techniques.

The authors, Mohammed Adil Shaikh and Sneha Annappanavar [6], propose a method for using deep learning algorithms namely Convolution Neural Network (CNN) for detecting the clickbaits on the social media platforms. The used method focuses on the textual features which consider the word sequence information and also learns the word meanings from the entire dataset. Our Results obtained a high accuracy of 0.82% comparatively better than different Machine Learning algorithms. We also did comparative analysis with the classification algorithm called Random Forest (RF). Click baits are essentially attention-grabbing headlines or titles that embellish the truth to get readers to "click" on them. These clickbaits can take many different forms, including pictures, videos, and adverts. These links will take you to anonymous websites that are a nuisance on the internet and provide very little information.

III. RELATED WORK

Potthast, Kopsel, Stein, and Hagen [1] gathered over 3000 tweets from the top 20 publishers on Twitter among computer scientists. The teaser message or title, the linked web page, and the meta information were the three fields from which they assembled handmade features to develop a model. There were some simple text and dictionary features in the teaser message. The meta information contained aspects pertaining to the tweets themselves, while the associated web page had text and readability characteristics. A supervised classification mechanism that received these characteristics and produced results of 0.79 ROC-AUC at 0.76 accuracy and 0.76 recall was used. They discovered that category one features alone performed better than features from any other category, with word 1-gram and character n-gram features making the largest contributions because they are known to accurately represent writing styles. Eight categories of clickbait were established by Biyani, Tsioutsiouliklis, and Blackmer [2]. They collected 1349 clickbait and 2724 non-clickbait web pages from the Yahoo home page. Three main categories comprise its handmade characteristics. Features that are content-based are those that come from the titles' wording and the pages' content. The use of quotation marks, exclamations, questions, etc. was considered a feature in addition to classic elements like bigrams and unigrams. Features that rely on similarity determine how similar the title and first five lines are from the webpage's body separately. While Forward Reference features made use of characteristics developed after the four categories of forward references provided by Blom and Hansen [3], Informality features assessed the formality and caliber of the pages.

A problem with recurrent neural networks is that their gradients disappear. With an accuracy score of 0.774, Gradient Boosted Trees was utilized for Clickbait identification, which included feature extraction from the body, title, content, and degree of informality [2]. A. Chakraborty et al. [3] offered a dataset of 32,000 titles, both clickbait and non-clickbait, and suggested an SVM-based method for clickbait detection that produced a 93% accuracy rate, outperforming other machine learning models like Random Forests, Decision Trees, and Radial Basis Function Kernels. CNN and RNN networks were used in Deep Learning-based techniques for classification.

The majority of the work done thus far has been on a single social media site, which hinders the development of a general model for spotting clickbait across several platforms. Furthermore, the majority of the models created simply consider lexical subtleties, ignoring the semantics, which gives each model language its uniqueness. As a result, after carefully analyzing the numerous characteristics that were examined in each research, these studies assisted in constructing and deriving the crucial features for generating models. An extensive overview of the method used to identify clickbait may be found in the section that follows.

IV. DATA SET

Three well-known social media platforms—Reddit, Facebook, and Twitter—were the sources of the data that was used [3].Every social networking site has its own restrictions. For example, Twitter only permits 140 characters per tweet. As a result, we trained our deep learning model for clickbait identification using data from a variety of sources. We also receive some data from http://www.clickbaitchallenge.org/, which provides us with sufficient information for our corpus. The clickbait comes from well-known news aggregation websites like Upworthy, BuzzFeed, and others. The non-clickbait comes from articles on Wikinews. Additionally, we annotated a Chinese corpus of clickbait from UC Headlines, a well-known news aggregator [4].

TABLE 1
STATISTICS OF GIVEN DATA

	Clickbait	Legit
Text-based Data	17264	21253
Image-based Data	15402	3023

TABLE 2
EVALUATION METRICS

Evaluation Metrics	Formula
Precision	TP/(TP+FP)
Recall	TP/(TP+FN)
Accuracy	(TP+TN)/N
F1 Score	$\frac{(2\times Recall\times Precision)}{(Recall+Precision)}$

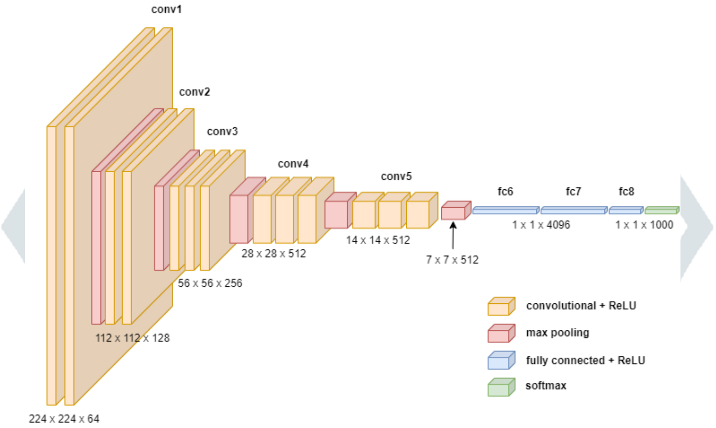
Where,

TP : True Positive
FP : False Positive
TN: True Negative
FN: False Negative
N : No of data

V. ALGORITHMS

A. Convolutional Neural Network

CNNs algorithm has good accuracy on data with greater position. For our classification, we are trying to get the best position in textual data given their short length and their tendency to focus on cyber-crimes. We used network that receive input
text in the form of sequences of numerical representation of stemmed unigrams [5]. CNNs are a class of deep learning models designed to process and analyze grid-like data, such as images and sequences, by learning hierarchical representations of patterns within the data. The core architectural feature of a CNN is the convolutional layer, which applies filters (kernels) to local regions of the input data, enabling the network to recognize spatial patterns and features.
Agrawal [11] shows the performance of using CNN to detect Clickbait. However, there is no optimized model of CNN to detect Clickbait until now.



B. Random Forest Algorithm

Random Forest algorithm used to make the comparative analysis with our proposed which is based on Deep learning model CNN. Random Forest algorithm is used for both classification as well as regression. To be precise, it can also be called the collection of decision trees classifiers. Random Forest is a category of Supervised Machine Learning algorithm which depends on Ensemble Learning. This is a type of Learning in which different algorithms can be connected, or the same algorithm can be used numerous times to form a more powerful prediction model. Random forest has an advantage over decision trees as it helps in correcting the habit of overfitting. The Random Forest algorithm gives an efficient estimation of the generalized error [6].

D. Support Vector Machine

SVM is a practical supervised machine learning technique that may be applied to the resolution of classification and regression issues [7]. The support vectors are any training samples that lie on the marginal hyperplanes [8]. Figure 9 shows the marginal hyperplanes as H1 and H2. A hyperplane is a line that divides content into clickbait and non-clickbait categories. The distance between the marginal hyperplanes is known as the margin.
The SVM divides all of the highlighted data items into two classes after determining the best hyperplane from the training dataset: Clickbait and Non-Clickbait. The features $x1$ ...xn are the TFIDF of title, TFIDF of body and cosine similarity and the class label, yi is either clickbait or non-clickbait. The output class yi is classified into two classes: clickbait $(yi = +1)$ and non-clickbait $(yi = -1)$.

The Hyperplane H and Marginal hyperplane $H1$ and $H2$ equations are:

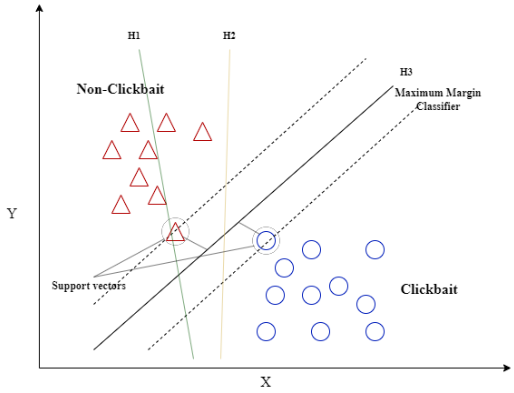
$$H : w^T xi + b = 0$$

$$H1 : w^T xi + b = -1$$

$$H2 : w^T xi + b = 1$$

where, represents transpose of weight vector and b represents bias [24]. The data points that were correctly classified should satisfy the inequality:

$$yi\,(w^T xi + b) \geq 1 \quad for \qquad xi\,,\ i = 1,\ 2,\ .\ .\ .\ .\ [9]$$



E. Logistic Regression

Logistic regression is a statistical model used for binary classification problems, where the outcome variable is categorical with two possible classes. Despite its name, logistic regression is a classification algorithm, not a regression algorithm. It is widely used in various fields, including medicine, finance, and machine learning. It is a straightforward and interpretable algorithm, but it is mainly suitable for linearly separable problems. For more complex relationships in the data, more advanced algorithms like support vector machines or neural networks may be more appropriate.

F. XGBOOST

XGBoost, short for eXtreme Gradient Boosting, is a powerful and efficient machine learning algorithm that belongs to the class of ensemble learning methods. It has gained popularity for its high performance and versatility in various types of predictive modeling tasks. XGBoost is particularly known for its effectiveness in structured/tabular data and has been a winning algorithm in numerous machine learning competitions. It is a combination of ensemble learning, regularization, and optimization techniques that makes it a robust algorithm that is widely used and appreciated in the machine learning community. Its ability to handle complex relationships in data, avoid overfitting, and provide interpretable feature importance makes it a popular choice for many practical applications.

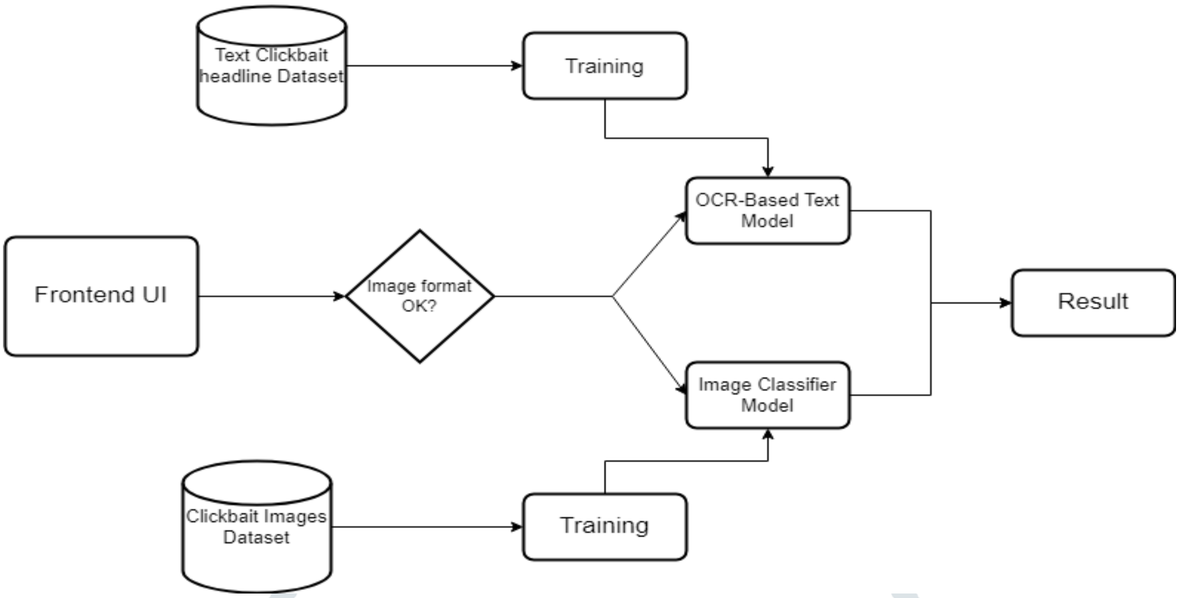
G. Recurrent Convolutional Neural Network

Recurrent Neural Networks experience the issue of vanishing gradients. Gradient Boosted Trees has been used for Clickbait detection consisting of feature extraction from the body, title, content and degree of informality, and achieved the precision score of 0.774 [2] . A. Chakraborty et al. [9] provided a clickbait dataset comprising of 32000 clickbait and non-clickbait titles and proposed a SVM based approach for clickbait detection which resulted in an accuracy of 93% and it provided a better accuracy in comparison to the other machine learning models such as Radial Basis Function Kernel, Decision Trees and Random Forests. Deep Learning based approaches utilized CNN and RNN networks for classification. A RNN and LSTM based model for clickbait detection in Filipino and English language managed to acquire an accuracy of 91.5% [10]. Another model based on CNN for clickbait detection had a F1 Score 0.86 [12].

VI. PROPOSED SYSTEM

BaitAware, our advanced Clickbait Detection System, employs a hybrid model for precise and efficient clickbait identification. The system integrates both text and image analyses to offer a comprehensive solution. The text analysis utilizes a Multinomial Naive Bayes model, focusing on language patterns within headlines to discern clickbait content. Simultaneously, image analysis leverages Convolutional Neural Networks (CNNs) for effective feature extraction from images, capturing visual cues indicative of clickbait. The two analyses converge through an ensemble model, synergistically enhancing the overall accuracy of clickbait detection. This fusion of text and image models results in a robust and versatile system.

The text model is trained to recognize linguistic nuances associated with clickbait, while the image model employs CNN architecture, including convolutional layers, activation functions, and max-pooling layers for effective image feature extraction. The ensemble model combines the strengths of both analyses, creating a more nuanced and accurate detection mechanism. This comprehensive solution is encapsulated in a Flask web application, ensuring user-friendly accessibility. Users can effortlessly upload images or input links, initiating a seamless process of clickbait analysis. The technical sophistication of the hybrid model and its integration into an accessible web interface position BaitAware as an effective tool in the ongoing battle against clickbait proliferation.



V. RESULT

The results of our clickbait detection system using deep learning are in, and they're promising. Our text-based model boasts an impressive accuracy of 87.34% using Multinomial Naïve Bayes algorithm, demonstrating its ability to discern clickbait from genuine content with notable reliability. Meanwhile, our image-based model shines even brighter, achieving a remarkable accuracy of 91.65% using Convolutional Neural Network. These results underscore the efficacy of leveraging deep learning techniques in identifying clickbait, offering a potent tool for enhancing online content quality and user experience. With further refinement and integration, such systems hold significant potential in combating misinformation and fostering a more trustworthy digital landscape.

IV. CONCLUSION

In conclusion, the development of a clickbait detection system utilizing deep learning marks a significant stride in combating deceptive content proliferation across text and image mediums. By harnessing the power of neural networks, this system effectively discerns between authentic, informative content and clickbait, thereby empowering users to make informed decisions while navigating online platforms. Its ability to analyze both textual and visual cues ensures comprehensive coverage, enhancing its utility in diverse digital environments. As we continue to refine and optimize this technology, we move closer to fostering a healthier online ecosystem characterized by transparency and trustworthiness.

V. REFERENCES

[1] M. Potthast, S. Köpsel, B. Stein and M. Hagen, "Clickbait Detection," in *European Conference on Information Retrieval*, Weimar, 2016.

[2] P. Biyani, K. Tsioutsiouliklis and J. Blackmer, "8 Amazing Secrets for Getting More Clicks": Detecting Clickbaits in News Streams Using Article Informality," in *Proceedings of the AAAI Conference on Artificial Intelligence*, California, 2016.

[3] J. N. Blom and K. R. Hansen , "Click bait: Forward-reference as lure in online news headlines," *Journal of Pragmatics*, vol. 76, pp. 87-100, 2015.

[4] J. Fu, L. Liang, X. Zhou and J. Zheng, "A Convolutional Neural Network for Clickbait Detection," in *International Conference on Information Science and Control Engineering*, Nanchang, 2017.

[5] H.-T. Zheng, J.-Y. Chen, X. Yao, A. K. Sangaiah and Y. Jiang, "Clickbait Convolutional Neural Network," *Symmetry*, vol. 10, no. 5, pp. 1-12, 2018.

[6] M. A. Shaikh and S. Annappanavar, "A Comparative Approach For Clickbait Detection Using Deep Learning," in *IEEE Bombay Section Signature Conference (IBSSC)*, Mumbai, 2021.

[7] S. Tamrakar, B. K. Bal and R. B. Thapa, “ASPECT BASED SENTIMENT ANALYSIS OF NEPALI TEXT USING SUPPORT VECTOR MACHINE AND NAIVE BAYES,” *Technical Journal,* vol. 2, no. 1, pp. 22-29, 2020.

[8] “Support Vector Machines,” [Online]. Available: https://scikit-learn.org/stable/modules/svm.html. [Accessed 2024].

[9] A. Chakraborty, B. Paranjape, S. Kakarla and N. Ganguly, “Stop Clickbait: Detecting and preventing clickbaits in online news media,” in *ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM),* San Francisco, 2016.

[10] P. K. Dimpas, R. V. Po and M. J. Sabellano, “Filipino and english clickbait detection using a long short term memory recurrent neural network,” in *International Conference on Asian Language Processing (IALP),* Singapore, 2017.

[11] A. Agrawal, “Clickbait detection using deep learning,” in *International Conference on Next Generation Computing Technologies (NGCT)*, Dehradun, 2016.

[12] B. Naeem, A. Khan, M. O. Beg and H. Mujtaba, “A deep learning framework for clickbait detection on social area network using natural language cues,” *Journal of Computational Social Science,* vol. 3, no. 1, pp. 231-243, 2020.

[13] B. Gamage, A. Labib, A. Joomun, C. H. Lim and K. Wong, “Baitradar: A Multi-Model Clickbait Detection Algorithm Using Deep Learning,” in *International Conference on Acoustics, Speech and Signal Processing (ICASSP),* Toronto, 2021.

[14] S. Pandey and G. Kaur, “Curious to Click It?-Identifying Clickbait using Deep Learning and Evolutionary Algorithm,” in *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Bangalore, 2018.

[15] S. Chawda, A. Patil, A. Singh and A. Save, “A Novel Approach for Clickbait Detection,” in *International Conference on Trends in Electronics and Informatics (ICOEI),* Tirunelveli, 2019.

[16] S. ALBAWI, T. A. MOHAMMED and . A.-Z. Saad , “Understanding of a convolutional neural network,” in *International Conference on Engineering and Technology (ICET)*, Antalya, 2017.

[17] R. Chauhan, K. K. Ghanshala and R. Joshi, “Convolutional Neural Network (CNN) for Image Detection and Recognition,” in *International Conference on Secure Cyber Computing and Communication (ICSCCC),* Jalandhar, 2018.

[18] A. Bajaj, H. Nimesh, R. Sareen and D. K. Vishwakarma, “A Comparative Analysis Of Classifiers Used For Detection of Clickbait In News Headlines,” in *International Conference on Intelligent Computing and Control Systems (ICICCS),* Madurai, 2021.

[19] K. Shu, S. Wang, T. Le, D. Lee and H. Liu, “Deep Headline Generation for Clickbait Detection,” in *International Conference on Data Mining (ICDM),* Singapore, 2018.

[20] G. Fumera, I. Pillai and F. Roli, “Spam Filtering Based On The Analysis Of Text Information,” *Journal of Machine Learning Research,* vol. 7, no. 12, pp. 2699-2720, 2006.