A Framework for Identifying Clickbait

Adithya A J*, Krishna J#, Rona Siby&, Roshan Babu@, Mary Priyanka K S+
Department of CSE, College of Engineering Kidangoor, Kottayam, Kerala, India
* aadianeesh@gmail.com, #kj163705@gmail.com, &ronasiby03@gmail.com,
@roshanbabu086@gmail.com, +marypriyanka@ce-kgr.org

Abstract—**The pervasive use of clickbait in YouTube videos negatively impacts user experience by leveraging attention-grabbing headlines, descriptions, and thumbnails to attract views. This deceptive practice has prompted significant research into automated detection frameworks aimed at mitigating its effects. Existing methods for clickbait detection primarily utilize machine learning and deep learning techniques, with approaches focusing on text-based metadata analysis, visual pattern recognition in thumbnails, and sentiment analysis of video titles and descriptions. Transformer models have shown promise in analyzing complex patterns in both textual and visual data, while sentiment analysis provides contextual validation for identifying misleading content. While considerable progress has been made, further research is needed to enhance the robustness and scalability of these frameworks in addressing the dynamic nature of clickbait strategies.**

Index Terms—**Deep learning, Machine learning, clickbait detection**

I. INTRODUCTION

The widespread use of clickbait, particularly in online platforms such as YouTube, has raised significant concerns regarding the quality of user experience and the integrity of content. Clickbait refers to misleading or exaggerated headlines, descriptions, and thumbnails used to attract more views, often at the expense of content accuracy. This phenomenon not only impacts users by leading them to irrelevant or false information but also creates challenges in content moderation and data accuracy. As clickbait tactics evolve, so too must the methods used to identify and combat them.

Several research efforts have focused on addressing this issue by leveraging advanced machine learning algorithms and deep learning models. For instance, Churi and Patil (2020) propose machine learning algorithms to detect clickbait using text-based analysis, while Shaikh et al. (2021) explore the use of deep learning techniques to detect clickbait through a comparative approach. Similarly, Rony et al. (2021) examine how clickbait is used across different topics and its effects on audiences. Other notable contributions include Chakraborty et al. (2021), who explore ensemble learning for clickbait detection in news media, and Varshney and Vishwakarma (2021), who present a unified approach using cognitive evidence for detecting clickbait in YouTube videos.

These studies highlight a variety of techniques, such as sentiment analysis, image recognition, and metadata analysis, to identify misleading content. Machine learning models like Support Vector Machines (SVM), Random Forest, and Extreme Learning Machines have been employed to compare the effectiveness of different detection methods. Moreover, ensemble learning has shown promise in improving the accuracy of clickbait detection by combining multiple classifiers. However, despite significant advancements, challenges remain, including the need for more robust models that can adapt to the constantly evolving tactics of clickbait creators.

This paper reviews the existing literature on clickbait detection, compares various approaches, and discusses the effectiveness of each method. By analyzing the strengths and limitations of these techniques, we aim to provide a comprehensive understanding of the current state of clickbait detection research and outline potential directions for future work.

II. LITERATURE SURVEY

A. Stop Clickbait: Detecting and Preventing Clickbaits in Online News Media

Chakraborty et al. [1] investigates the growing issue of clickbait in online news platforms and proposes methods for its detection and mitigation. Clickbait is defined as headlines or content previews crafted specifically to grab readers' attention and encourage clicks, often at the expense of accuracy or truthfulness. The authors emphasize that the surge in clickbait is driven by the intense competition for online traffic and revenue, leading to a decline in user satisfaction and trust in media outlets. The research focuses on creating automated tools to identify and address clickbait effectively.

The paper identifies key characteristics of clickbait, including curiosity-evoking language, exaggerated claims, and withheld details. It employs machine learning models trained on datasets of clickbait and non-clickbait headlines, analyzing linguistic and stylistic features like excessive punctuation, emotional word choices, and ambiguous pronouns to detect patterns. The authors highlight the challenge of labeling training data due to the subjective nature of defining clickbait.

The paper presents an analysis of syntactic dependencies in clickbait versus non-clickbait headlines, revealing that clickbait tends to use shorter dependencies, resulting in simpler and more direct structures that quickly capture attention. In contrast, non-clickbait headlines exhibit longer dependencies, reflecting greater complexity and richer context. This structural comparison, visualized in a graph, shows clickbait headlines clustering around shorter dependencies, emphasizing how structural simplicity enhances reader engagement.

The paper further explores the performance of various machine learning classifiers, including support vector ma-

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 01 | Jan - 2025 SJIF Rating: 8.448 ISSN: 2582-3930

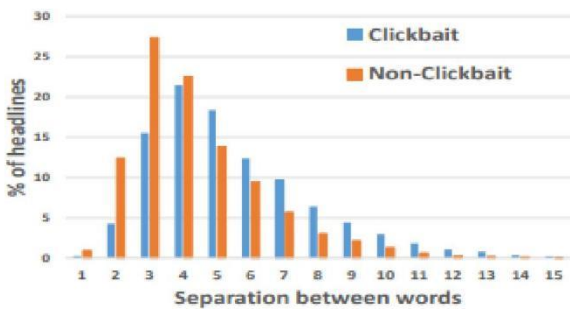

Fig. 1. Distribution of longest syntactic dependencies between all pair of words in clickbait and non-clickbait headlines [1]

chines (SVMs) and neural networks, for clickbait detection. Experimental results demonstrate that combining traditional feature-based methods with deep learning techniques achieves high accuracy in identifying clickbait headlines.

In conclusion, the study offers valuable contributions to the domain of clickbait detection and serves as a foundation for future research. The authors propose that future work could focus on improving these detection models and integrating them into digital platforms to flag or filter clickbait content. Such advancements would enhance the credibility of online media by reducing the spread of sensationalized or misleading information.

B. Clickbait Detection Using Machine Learning Algorithm

Churi et al. [2] conducted an extensive study on clickbait detection using various machine learning techniques. They aimed to classify news headlines as clickbait or non-clickbait by analyzing linguistic differences such as sentence composition, word structure, and lexical nuances. Their dataset comprised 32,000 headlines equally distributed between clickbait and non-clickbait categories, manually labeled to prevent bias. The authors implemented multiple classifiers, including Support Vector Machines (SVM), Logistic Regression, and Decision Trees. Among these, Logistic Regression achieved the highest precision and recall, both at 97% , outperforming the other methods.

The authors reviewed significant contributions from existing literature to provide context for their work. For instance, they referenced Pandey and Kaur's (2018) deep learning approach using BiLSTM and GloVe embeddings, which achieved an accuracy of 98.78% on 80,000 headlines. Additionally, they highlighted the work of Chawda et al. (2019), who employed RCNN combined with GRU and LSTM, achieving 97.76% accuracy. Churi and Patil also discussed Cao et al.'s (2017) use of Random Forest Regression and Chakraborty et al.'s (2016) SVM-based linguistic analysis, demonstrating the effectiveness of traditional machine learning models for feature-based clickbait detection.

Further, the authors emphasized the role of neural networks in multilingual and domain-specific tasks, referencing studies like Klairith and Tanachutiwat's (2018) BiLSTM application for Thai headlines and Fu et al.'s (2017) CNN framework for Chinese and English headlines. These studies demonstrated the adaptability of advanced neural architectures to diverse datasets. Churi and Patil also noted Wongsap et al.'s (2019) Decision Tree-based classifier, which achieved 99.90% accuracy by leveraging features like special characters.

In their own experiments, Churi and Patil's findings underscored the importance of careful feature engineering, as linguistic attributes such as stop words, headline length, and punctuation patterns were found to differ significantly between clickbait and non-clickbait headlines. They concluded that Logistic Regression provided the most balanced performance among the models tested, making it a reliable baseline for further advancements in clickbait detection.

The paper successfully demonstrates how integrating insights from prior research with practical implementations can contribute to developing robust clickbait detection systems. Future work could explore the integration of multimodal data (e.g., images and text), multilingual datasets, and real-time detection capabilities to enhance the efficacy of such systems.

C. An Emotional Analysis of False Information in Social Media and News Articles

Bilal Ghanem et al. [3] delve into the emotional dynamics of false information on social media and online news platforms. The authors explore how false information, including propaganda, hoaxes, clickbait, and satire, manipulates readers by targeting their emotions. Unlike real news, which primarily informs, false information appeals to emotions to influence opinions and behaviors. To address this issue, they propose the Emotionally-Infused Network (EIN), a deep learning model that combines emotional and linguistic features to detect and classify false information effectively.

The authors emphasize the significance of emotional analysis by employing various emotional lexicons, such as EmoLex, LIWC, and Empath, which map textual content to specific emotions like joy, anger, fear, and sadness. These resources allow the detection system to recognize subtle emotional cues embedded in the text. By leveraging these emotional patterns, the authors aim to uncover the manipulation tactics used in false information, distinguishing it from truthful content. This emotional approach enhances the detection process, particularly for types of misinformation that rely on exploiting readers' emotions.

The Emotionally-Infused Network (EIN) is structured into two branches: a content-based branch and an emotional features branch. The content-based branch uses word embeddings processed through LSTM layers and attention mechanisms to understand the context of the text. The emotional branch extracts emotional features from lexicons, providing insights into the text's emotional tone. By combining the outputs of these two branches, the EIN model captures both the linguistic and emotional aspects of false information, enabling it to classify types like clickbait, propaganda, and satire with improved accuracy.

Uses two datasets—news articles and Twitter posts—they evaluate their Emotionally-Infused Network (EIN) model, which integrates emotional features with neural networks. The EIN model outperforms traditional classifiers like SVM and neural networks without emotional features, achieving high accuracy, precision, recall, and F1-scores. Emotional cues, such as ”surprise” and ”fear,” are particularly effective in identifying clickbait, while hoaxes and propaganda remain more challenging due to their resemblance to legitimate content.

D. A Comparative Approach For Clickbait Detection Using Deep Learning

Shaikh et al. [4] examine the growing issue of clickbait in digital news media and present a comparative analysis of deep learning methods for detecting it. Clickbait refers to attention-grabbing headlines designed to maximize user clicks, often by exaggerating or distorting the content they advertise. The increase in clickbait is largely driven by the competitive nature of online traffic and the need for advertising revenue, which raises concerns about user experience and the credibility of digital platforms. The authors propose developing automated systems capable of distinguishing clickbait from legitimate content.

The study begins by providing an in-depth analysis of clickbait characteristics, highlighting techniques such as curiosity-gap phrases, sensationalized language, and the omission of critical details. To capture these patterns, the authors create a labeled dataset containing examples of both clickbait and non-clickbait headlines. They identify linguistic and stylistic features common in clickbait, including frequent punctuation (e.g., multiple exclamation marks), emotionally charged words, and ambiguous or open-ended phrases. Since the definition of clickbait can be subjective, the authors address this challenge by employing a rigorous data collection and labeling methodology.

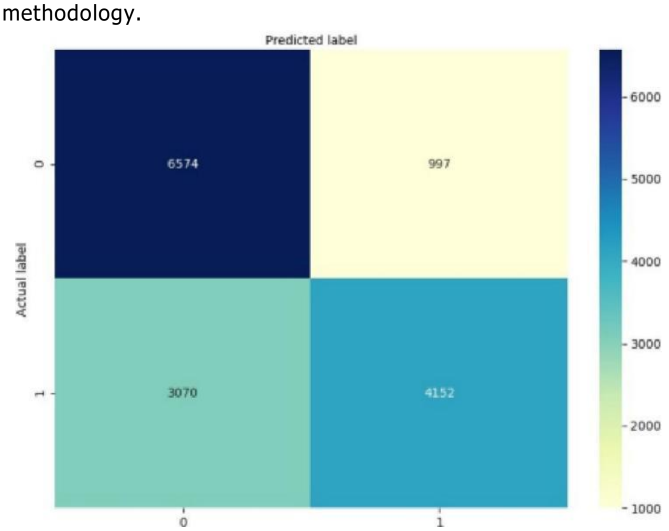

Fig. 2. Confusion Matrix for Random Forest [4]

Fig 2 presents the confusion matrix as a vital tool for assessing the performance of the Random Forest classifier in clickbait detection. This matrix provides a breakdown of true positives (correctly identified clickbait), true negatives (correctly identified non-clickbait), false positives (non-clickbait misclassified as clickbait), and false negatives (clickbait misclassified as non-clickbait). These metrics offer insights into the model’s classification performance. For example, a high count of true positives and true negatives indicates effective classification, while significant false positives or negatives may point to the need for adjustments, such as feature refinement or hyperparameter tuning. Using the confusion matrix, researchers can calculate key performance metrics like accuracy, precision, recall, and F1-score, which provide a comprehensive evaluation of the model’s reliability. This analysis helps uncover biases and ensures that the model achieves balanced and accurate clickbait detection.

In conclusion, the study underscores the potential of deep learning techniques to significantly improve the reliability of clickbait detection systems. These models help reduce the prevalence of sensationalized content and promote the dissemination of trustworthy information on digital platforms. The authors recommend expanding labeled datasets, refining model performance, and improving adaptability to various content types and languages. Future research could focus on real-time implementation and enhancing the interpretability of these models to empower users with better tools for navigating online content.

E. Diving Deep into Clickbaits: Who Use Them to What Extents in Which Topics with What Effects?

Rony et al. [5] address the growing phenomenon of clickbait in online media and analyze its prevalence, user engagement, and impact. The study evaluates the extent of clickbait use by mainstream and unreliable media using a dataset of 1.67 million Facebook posts. To achieve this, the authors develop a clickbait detection model based on distributed subword embeddings, achieving 98.3 % accuracy. This paper investigates which types of media use clickbait the most, what topics are covered, and how clickbait affects user engagement.

The authors compile two datasets: a Headlines Dataset with 32,000 manually labeled headlines (15,999 clickbait and 16,001 non-clickbait), and a Media Corpus with 1.67 million Facebook posts from 153 U.S.-based media organizations. Mainstream media, categorized into print and broadcast, and unreliable media, including clickbait, satire, conspiracy, and junk science outlets, form the basis of the analysis. The clickbait detection model, developed using subword embeddings and Skip-Gramsw, maps text into semantic vector spaces, significantly improving classification performance compared to traditional bag-of-words models.

The experimental results reveal that broadcast outlets (47.56%) and non-news channels frequently use clickbait, while print media shows a lower rate (24.12%). Among unreliable media, satire and clickbait-focused platforms dominate with over 50% clickbait usage, whereas conspiracy outlets have a lower rate (28.7%). These findings indicate that click-bait is widespread, even in mainstream media, particularly in entertainment and lifestyle content.Clickbait headlines in mainstream media focus on sensational topics like celebrity gossip and controversial figures, while non-clickbait headlines emphasize societal issues such as public policy. In unreliable media, clickbait spans diverse topics like politics, health, and sensational narratives. Additionally, non-clickbait headlines align better with their content, whereas clickbait headlines often exaggerate or mislead.

The impact of clickbait on user engagement is significant, with clickbait posts receiving higher reactions, shares, and comments compared to non-clickbait posts in most categories. However, in broadcast media, non-clickbait content occasionally outperforms clickbait in terms of engagement. The authors also observe that unreliable media frequently repost clickbait links to increase their reach and visibility. Even for non-clickbait articles, clickbait techniques are often employed in the accompanying Facebook status messages, further enticing users to engage.

In conclusion, the paper highlights that clickbait is a pervasive practice across both mainstream and unreliable media, driven by its ability to boost user engagement. However, this comes at the cost of media credibility and user trust. The authors propose incorporating headline-body similarity into future clickbait detection models and call for further research into the ethical implications of clickbait practices. This study provides a foundation for addressing clickbait's effects on the media ecosystem and improving content reliability on social platforms.

F. Ensemble Learning Approach for Clickbait Detection Using Article Headline Features

Sisodia et al. [6] proposes a classification model using ensemble learning techniques to distinguish clickbait headlines from genuine ones. Highlighting the negative impacts of clickbaits, such as misinformation and sensationalism, it emphasizes the importance of effective detection mechanisms. By leveraging natural language processing (NLP) to extract headline features, the study evaluates multiple classifiers, achieving high accuracy in clickbait detection.

The research utilizes a dataset comprising 100,000 article headlines, with 60,000 clickbait headlines collected from Buzzfeed articles published in 2014 and 40,000 authentic headlines sourced from Reuters, Associated Press, and The New York Times. These headlines were retrieved using the New York Times Article Search API for the period from January 2013 to March 2016. The data was further divided into balanced and unbalanced training and testing subsets to evaluate classifier performance under different conditions.

Fig. 3 highlights accuracy as a key metric in evaluating machine learning models for clickbait detection, measuring the proportion of correct predictions (true positives and true negatives). The study compares traditional classifiers like SVM with advanced ensemble models such as Random Forest and Gradient Boosting. Ensemble methods outperform individual

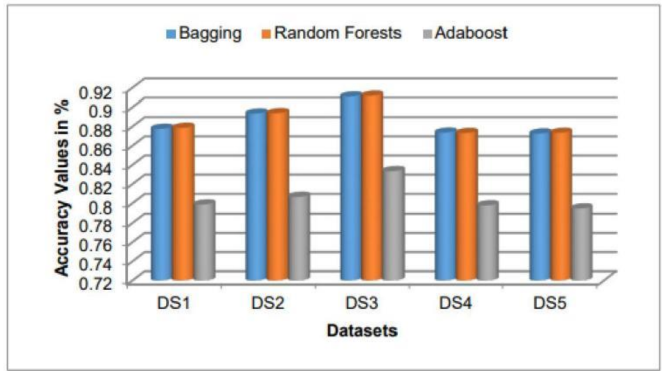

Fig. 3. Accuracy of classifiers [6]

classifiers by capturing the nuances of clickbait headlines and reducing overfitting. However, the authors emphasize that accuracy alone is insufficient; metrics like precision, recall, and F1-score are essential for understanding model performance across both clickbait and non-clickbait instances, particularly on imbalanced datasets.

The study identifies 19 linguistic features from the headlines, such as the presence of exclamation marks, baity words, numbers, and emotional language. These features were manually selected and extracted using NLP tools. Three ensemble learning models—bagging, boosting, and random forests—were trained and tested on the prepared datasets. The classifiers were evaluated based on accuracy, precision, recall, and F-measure to determine their effectiveness in detecting clickbait.

The random forest classifier outperformed other ensemble models, achieving 91.16% accuracy on balanced datasets. However, its performance declined on unbalanced datasets due to the uneven representation of clickbait and genuine headlines. This study emphasizes that balanced datasets yield more reliable performance metrics. *G. Clickbait in YouTube: Prevention, Detection, and Analysis of the Bait Using Ensemble Learning*

Mowar et al. [7] tackles the issue of clickbait in YouTube video titles and presents methods for its detection, prevention, and analysis using ensemble learning techniques. Clickbait refers to user experience and erodes trust in the platform. Their objective is to build an automated system that can effectively detect and mitigate clickbait content on YouTube.

The paper highlights the importance of intrusion detection systems (IDS) in modern cybersecurity and the role of machine learning in automating detection. It reviews the strengths and weaknesses of key classifiers: SVM, known for handling high-dimensional data and excelling in binary classification; RF, an ensemble method that combines decision trees for robust performance, even with noisy data; and ELM, a neural network-based approach valued for its fast training time and simplicity, though less explored in IDS compared to SVM and RF.

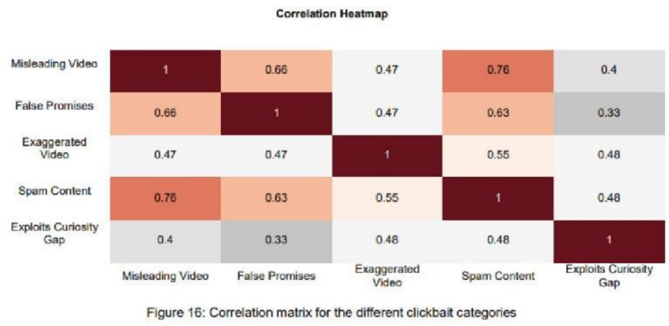

Fig. 4. Correlation matrix for the different clickbait categories [7]

Fig.4 presents a correlation matrix that highlights the relationships between features in clickbait headlines. It reveals how attributes like word frequency, sentiment, and emotional language interact to shape clickbait's structure. The authors identify a strong correlation between hyperbolic language (e.g., exaggerations or sensational phrases) and emotionally charged words, indicating that clickbait often uses these strategies to provoke reactions and attract clicks. This matrix aids in selecting features most indicative of clickbait, enhancing machine learning models' ability to classify headlines accurately.

The authors evaluate classifiers like decision trees, SVM, and neural networks, alongside ensemble methods such as Random Forest and Gradient Boosting. Ensemble learning improves detection accuracy by leveraging multiple models and addressing individual weaknesses. This approach effectively handles the complexities of YouTube video titles, enabling reliable classification of clickbait and non-clickbait content.

The paper suggests integrating the ensemble learning approach into YouTube's content moderation system to automatically flag misleading titles. Future work includes refining feature extraction, using advanced NLP techniques, and incorporating user behavior data to enhance prediction accuracy further, ultimately improving content quality and user satisfaction.

H. Performance Comparison of Support Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection

Ahmad et al. [8] address the challenge of intrusion detection in computer networks, comparing three machine learning algorithms—SVM, Random Forest (RF), and Extreme Learning Machine (ELM)—to determine the most effective intrusion detection system (IDS). IDS plays a critical role in network security by detecting unauthorized access and malicious activities. The study evaluates the algorithms based on accuracy, efficiency, and their ability to detect both known and unknown intrusions.

The paper begins by discussing the significance of IDS in modern cybersecurity and the role of machine learning in automating the detection process. It provides an overview of the strengths and weaknesses of SVM, RF, and ELM in classification tasks. SVM is known for its ability to handle high-dimensional data and is effective in binary classification. RF, an ensemble method, combines multiple decision trees, offering robust performance even with noisy data. ELM, a neural network-based approach, is known for its fast training time and simplicity, though it has been less explored in IDS compared to SVM and RF.

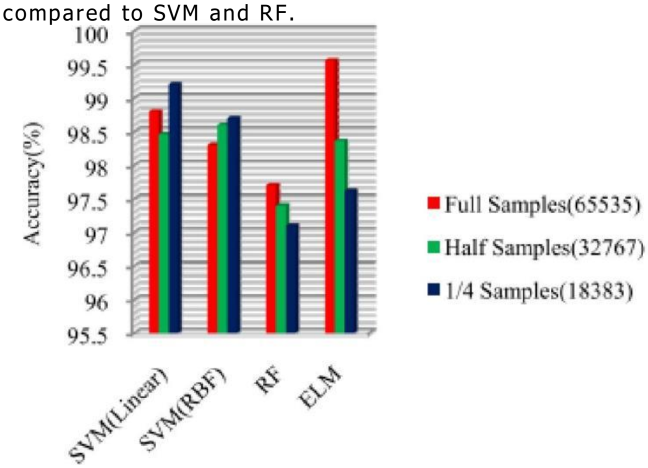

Fig. 5. Accuracy of SVM, RF, and ELM [8]

The fig 5 compares the performance of the Support Vector Machine (SVM), Random Forest (RF), and Extreme Learning Machine (ELM) algorithms in terms of accuracy, using a training set and a separate testing set. The graph shows that Random Forest (RF) achieves the highest accuracy, outperforming both SVM and ELM. This suggests that RF is the most effective model for intrusion detection in the given dataset, performing well across various types of intrusions. Support Vector Machine (SVM) also delivers competitive accuracy but requires more parameter tuning to achieve optimal performance compared to RF. Extreme Learning Machine (ELM), while quick to train, shows the lowest accuracy, indicating that its simplicity does not offer the same level of effectiveness in intrusion detection as the other two models.

The higher accuracy of RF can be attributed to its ensemble nature, which allows it to handle complex data better and resist overfitting. SVM performs reasonably well but may require more careful feature selection and kernel tuning. ELM, while faster to train, shows less effectiveness in terms of accuracy, though it could be suitable for real-time applications where training time is prioritized over detection accuracy.

The authors used a publicly available dataset with labeled instances of network traffic, including both normal and malicious activities. The features for classification include packet characteristics such as size, type, and connection status. The dataset was split into training and testing sets, and the algorithms were evaluated based on metrics like accuracy, precision, recall, and F1-score to measure their effectiveness in detecting intrusions.

The experimental results show that Random Forest (RF) outperforms SVM and ELM in accuracy and precision, ex-

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 01 | Jan - 2025 SJIF Rating: 8.448 ISSN: 2582-3930

celling in handling complex datasets and resisting overfitting. While SVM achieves good accuracy, its reliance on extensive parameter tuning makes it less scalable for large datasets. ELM, though computationally efficient, offers lower accuracy and is better suited for real-time applications prioritizing speed over precision.

In conclusion, RF provides the best balance of accuracy and efficiency, with SVM and ELM remaining useful for specific scenarios. Future research could explore hybrid models, advanced feature extraction, and diverse datasets to enhance detection performance further.

III. CONCLUSION

The reviewed body of research highlights significant advancements in clickbait detection across various platforms, leveraging innovative machine learning, deep learning, and ensemble methods. Techniques such as Convolutional Neural Networks (CNNs), Bidirectional LSTMs, and Random Forests have been pivotal in achieving high accuracy in identifying clickbait across text and video content.

Multi-modal frameworks, integrating textual, visual, and cognitive evidence, have further expanded detection capabilities for platforms like YouTube, addressing challenges posed by misleading thumbnails, titles, and audio content.

Ensemble learning approaches, which combine the strengths of multiple classifiers, have demonstrated superior performance by overcoming individual model weaknesses. For instance, Random Forests consistently outperformed standalone models in detecting nuanced clickbait patterns, highlighting the importance of combining diverse learning strategies.

Emotional analysis has also emerged as a valuable tool, revealing how clickbait exploits human curiosity and emotions to entice user engagement.

A table summarizing the techniques utilized in the reviewed papers has been included, providing a comprehensive overview of the methodologies, datasets, and performance metrics. This comparative analysis emphasizes the diversity and effectiveness of the approaches, aiding in the identification of promising strategies for future development.

Additionally, the introduction of tailored datasets, such as BollyBAIT and the Misleading Video Dataset (MVD), has enabled the creation of robust models capable of both detecting and preventing clickbait before public dissemination. This proactive approach improves user experiences and enhances trust in digital platforms.

Future research should focus on refining multi-modal frameworks by incorporating advanced transformer models, exploring emotional and cognitive features, and broadening the scope of applications in content moderation and digital trust. These efforts aim to create a more transparent and reliable digital media ecosystem.

Paper Title	SVM	Random Forest	CNN	LSTM	Ensemble Methods
Stop Clickbait: Detecting and Preventing Clickbaits in Online News Media	✓	✓	✓	✓	✓
A Comparative Approach for Clickbait Detection Using Deep Learning	✓	✓	✓	✓	
Diving Deep into Clickbaits: Who Use Them to What Extents in Which Topics with What Effects?			✓	✓	
Ensemble Learning Approach for Clickbait Detection Using Article Headline Features		✓			✓
An Emotional Analysis of False Information in Social Media and News Articles				✓	
Clickbait in YouTube: Prevention, Detection and Analysis of the Bait Using Ensemble Learning	✓	✓	✓		✓
Performance Comparison of Support Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection	✓	✓			

TABLE 1
ALGORITHMS USED IN CLICKBAIT DETECTION PAPERS

REFERENCES

[1] A. Chakraborty, B. Paranjape, S. Kakarla, and N. Ganguly, "Stop clickbait: Detecting and preventing clickbaits in online news media," in *2016 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)*, pp. 9–16, IEEE, 2016.

[2] M. Churi and D. Patil, "Clickbait detection using machine learning algorithm," 2022.

[3] B. Ghanem, P. Rosso, and F. Rangel, "An emotional analysis of false information in social media and news articles," *ACM Transactions on Internet Technology (TOIT)*, vol. 20, no. 2, pp. 1–18, 2020.

[4] M. A. Shaikh and S. Annappanavar, "A comparative approach for clickbait detection using deep learning," in *2020 IEEE Bombay Section Signature Conference (IBSSC)*, pp. 21–24, IEEE, 2020.

[5] M. M. U. Rony, N. Hassan, and M. Yousuf, "Diving deep into clickbaits: Who use them to what extents in which topics with what effects?," in *Proceedings of the 2017 IEEE/ACM international conference on advances in social networks analysis and mining 2017*, pp. 232–239, 2017.

[6] D. S. Sisodia, "Ensemble learning approach for clickbait detection using article headline features.," *Informing Sci. Int. J. an Emerg. Transdiscipl.*, vol. 22, pp. 31–44, 2019.

[7] P. Mowar, M. Jain, R. Goel, and D. K. Vishwakarma, "Clickbait in youtube prevention, detection and analysis of the bait using ensemble learning," *arXiv preprint arXiv:2112.08611*, 2021.

[8] I. Ahmad, M. Basheri, M. J. Iqbal, and A. Rahim, "Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection," *IEEE access*, vol. 6, pp. 33789–33795, 2018.