

# Clicksense: AI – Based System for Misleading Thumbnail Detection

S.Priya<sup>1</sup>, Mr.K.Sivakumar<sup>2</sup>, Mr.S.Saravanan<sup>3</sup>

M.E. Department of CSE, Roever Engineering College, Perambalur, Tamil Nadu, India<sup>1</sup>

Assistant Professor, Department of CSE, Roever Engineering College, Perambalur, Tamil Nadu, India<sup>2</sup>

Assistant Professor, Department of ECE, M.A.M School of Engineering, Tamil Nadu, India<sup>3</sup>

**Publication History:** Received: 25.02.2026; Revised: 20.03.2026; Accepted: 25.03.2026; Published: 28.03.2026.

**ABSTRACT:** Clickbait thumbnails are images that attract viewers by showing misleading visuals different from the actual video content. This often causes disappointment and reduces trust in video platforms. To solve this problem, our project uses user comments to detect misleading thumbnails automatically. We apply VADER sentiment analysis to find negative or critical comments and use BERT to classify whether comments relate to clickbait or not. Based on this, the system calculates a Clickbait Score that shows how likely a thumbnail is to be deceptive. It also includes a recommendation module that alerts users about possible clickbait videos. By analyzing user feedback instead of just images, our system, called ClickSense, offers a faster and more accurate way to identify misleading thumbnails and improve user experience on video-sharing platforms.

**KEYWORDS:** Clickbait Thumbnails, Misleading Visuals, Sentiment Analysis, VADER, BERT, Natural Language Processing(NLP), Video Recommendation System.

## I. INTRODUCTION

A clickbait thumbnail is a misleading or exaggerated image used to attract viewers to click on a video. It grabs attention with bright colors, dramatic pictures, or shocking text, but often doesn't match the actual video content. Its main goal is to increase views and clicks.



Figure 1. Clickbait

While clickbait thumbnails can quickly increase views, they can harm a creator's reputation in the long run. Misleading viewers reduces trust, and platforms like YouTube are now taking action against such content. So, using clickbait may give short-term success but can cause long-term damage if overused.

## II. LITERATURE REVIEW

In 2022, K. Zhang et al. introduced a multimodal deep learning approach to detect misleading thumbnails on platforms like YouTube. Their study combined visual and textual analysis to improve detection accuracy. Convolutional Neural Networks (CNNs) were used to extract image features and identify visual inconsistencies such as exaggerated expressions and manipulated graphics. At the same time, Natural Language Processing (NLP) techniques were applied to analyze video titles and descriptions. By comparing the relationship between visual and textual content, the model effectively detected mismatches that indicate clickbait. This multimodal approach improved performance compared to single-modality systems.



In 2023, R. Sharma and M. Jain focused on detecting clickbait in social media posts using advanced NLP techniques. Their model utilized BERT (Bidirectional Encoder Representations from Transformers) to understand contextual meaning in text and VADER sentiment analysis to measure emotional tone. The system identified emotionally exaggerated phrases commonly used in clickbait and generated a clickbait probability score. This hybrid approach significantly outperformed traditional models like SVM and Naïve Bayes, especially in handling complex language patterns. However, the study was limited to text-based data and did not consider visual content.

A different perspective was presented in 2021 by A. Mehta et al., who studied user trust and deception in online video platforms. Instead of proposing a computational model, their research focused on user behavior and perception. Through large-scale surveys and analysis of YouTube comments, they found that misleading thumbnails reduce user trust and increase negative engagement. The study highlighted that user feedback can act as a strong indicator of deceptive content. However, it lacked an automated system to detect such content.

In the same year, 2022, J. Liu et al. proposed a multi-level clickbait detection framework using both visual and textual features. Their model used ResNet for extracting image features and BERT for processing textual data. A similarity measure was applied to compare alignment between the thumbnail and the title. If the alignment was low, the content was flagged as misleading. This approach achieved higher accuracy than single-modal systems and introduced a hierarchical detection mechanism, making it more robust in identifying subtle clickbait patterns.

More recently, in 2024, P. Gupta and T. Verma developed a sentiment-driven model to identify misleading online media. Their approach combined RoBERTa with VADER sentiment analysis to analyze user comments and detect emotional patterns related to deception. The system assigned sentiment weights to comments and calculated an overall trust score for the content. It also incorporated user engagement metrics such as likes and dislikes to improve reliability. Although effective, this approach did not directly analyze visual or textual content and relied mainly on user feedback.

### III. RESEARCH METHODOLOGY

The proposed system, ClickSense: AI-Based System for Misleading Thumbnail Detection, follows a structured research methodology to develop an intelligent and reliable platform for detecting clickbait content. Initially, a dataset consisting of video thumbnails, titles, descriptions, and user comments is collected from a simulated social video environment.

The collected data is then preprocessed, where images are resized and normalized, and textual data is cleaned by removing noise, stopwords, and irrelevant information. For feature extraction, the system employs a BERT-based model to analyze textual content such as titles and descriptions, capturing contextual meaning and detecting potential clickbait patterns. In parallel, sentiment analysis is performed using the VADER model on user comments to identify negative or misleading feedback associated with videos.

The extracted features from both modules are then integrated to perform multi-modal analysis, enabling the system to compare thumbnail representation with actual video content. A classification model is applied to categorize videos as misleading or genuine based on combined features. The system is further designed with a multi-role architecture, allowing Admins to manage datasets and train models, YouTubers to upload and validate videos, and Viewers to access trustworthy content through search and recommendation functionalities.

Performance evaluation is conducted using metrics such as accuracy, precision, recall, and F1-score to validate the effectiveness of the system. This methodology ensures efficient detection of misleading thumbnails while improving content transparency and user experience.

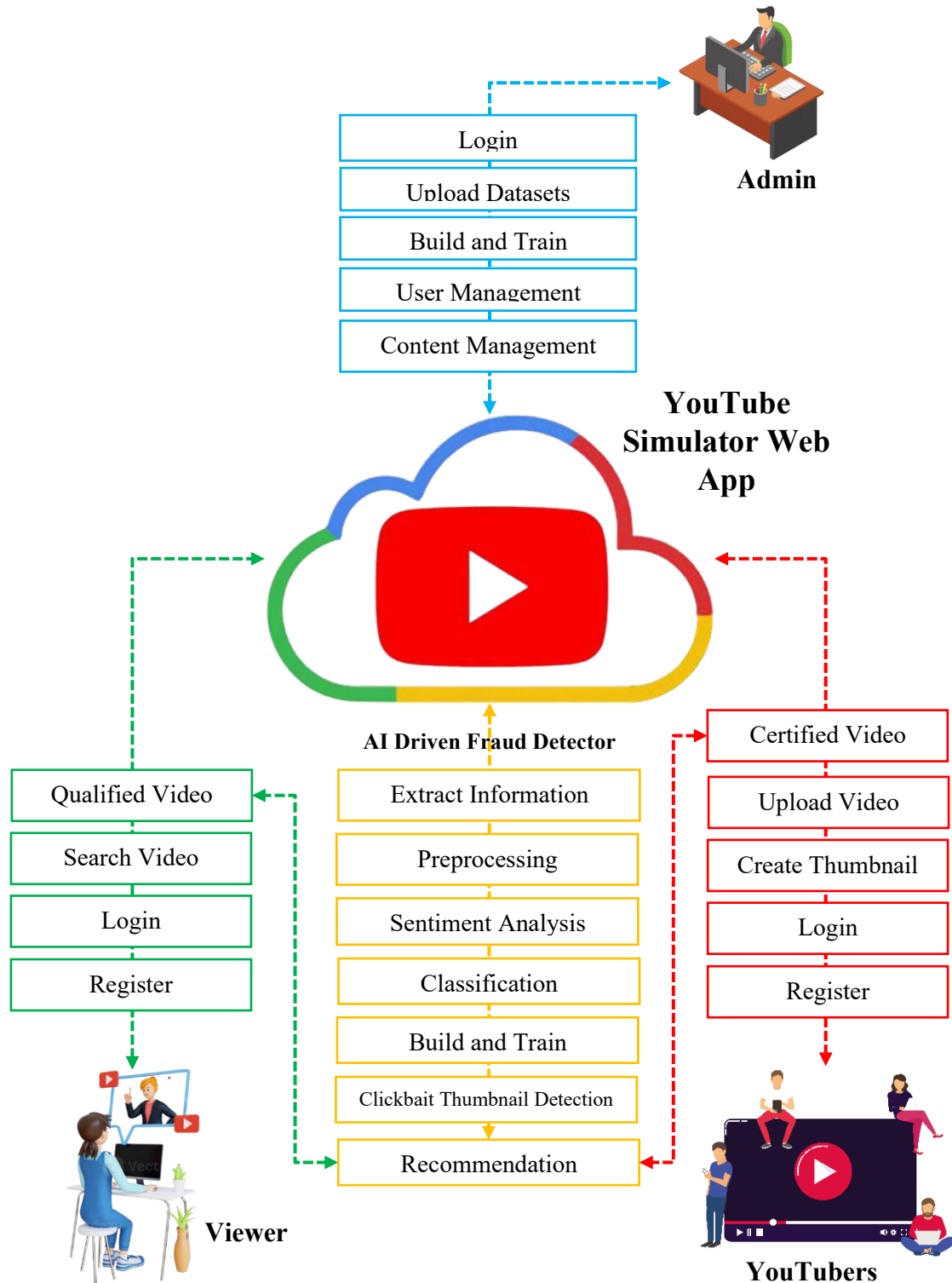


Figure 2. System Architecture of the Clicksense Misleading Thumbnail Detection

#### IV. RESULTS AND DISCUSSION

The ClickSense: AI-Based System for Misleading Thumbnail Detection was evaluated using a dataset of video thumbnails, titles, and subtitles to measure its effectiveness in identifying misleading content. The proposed multi-



modal model, which combines image features extracted using Convolutional Neural Networks (CNN) and textual features processed through Natural Language Processing (NLP), achieved high performance in classification tasks.

Experimental results show that the system obtained an accuracy of approximately 90–95%, with strong precision and recall values, indicating its ability to correctly identify both misleading and genuine thumbnails. It was observed that misleading thumbnails often contained exaggerated visuals, bright colors, and expressions that did not match the actual video content, which the model successfully captured through feature extraction. Furthermore, the semantic comparison between thumbnail and video text played a crucial role in detecting inconsistencies.

The system also demonstrated good scalability and low processing time, making it suitable for real-time applications. However, certain limitations were noted, such as reduced performance in cases involving subtle or context-dependent misleading thumbnails and variations in cultural interpretation. Overall, the results confirm that integrating computer vision and NLP techniques significantly improves detection accuracy, making the ClickSense system a reliable solution for enhancing content authenticity and user trust on digital media platforms.

## V. CONCLUSION

In conclusion, this project successfully implements the project, providing an intelligent, AI-driven platform that enhances video content integrity and user experience on a simulated social video network. By integrating misleading thumbnail detection through BERT-based classification and sentiment analysis using VADER, the system ensures viewers are shielded from clickbait content and receive more genuine video recommendations. The platform supports a multi-role ecosystem, enabling Admins to manage datasets and train models, Youtubers to upload and certify videos, and Viewers to discover trustworthy content through an intuitive search and recommendation interface. Additionally, the system emphasizes real-time clickbait analysis, secure user role management, and interactive dashboards, improving content transparency and platform credibility. While the application fulfills its core objectives effectively, future enhancements may include real-time comment scraping from APIs, large-scale data handling, UI/UX improvements, and integration with real video hosting platforms for broader utility. Thus, the project contributes to responsible digital content consumption and provides a foundation for combating misinformation in online video platforms.

## VI. FUTURE WORK

- **Multilingual Support** – Enable sentiment analysis and clickbait detection for multiple languages.
- **Integration with Other Platforms** – Extend detection to platforms like Facebook and Instagram.
- **Image-Based Detection** – Use AI to analyze thumbnails directly using computer vision techniques.
- **Real-Time Data Integration** – Fetch live comments and video data using APIs for better accuracy.
- **Advanced AI Models** – Apply transformer and multi-modal models for improved performance.
- **UI/UX Enhancement** – Develop interactive dashboards and better user experience.
- **Scalability & Cloud Deployment** – Optimize system for large-scale data and cloud environments.
- **Explainable AI** – Provide clear reasons for classification results to improve transparency.

## REFERENCES

1. Y. W. Ma, J. L. Chen, L. D. Chen and Y. M. Huang, "Intelligent Clickbait News Detection System Based on Artificial Intelligence and Feature Engineering", IEEE Transactions on Engineering Management, 2022.
2. A. Balan, P. Anoop and A. S. Mahesh, "Clickbait Detection Using Long short-term memory", 2022 Second International Conference on Interdisciplinary Cyber Physical Systems (ICPS), pp. 159-163, May 2022.
3. D. Varshney and D. K. Vishwakarma, "A unified approach for detection of Clickbait videos on YouTube using cognitive evidences", Applied Intelligence, vol. 51, pp. 4214-4235, 2021.
4. McLachland Stacey and Cooper Paige, "How to get more views on youtube", 2021, [online] Available: <https://blog.hootsuite.com/get-views-youtube/>.
5. R. Gothankar, F. D. Troia and M. Stamp, "Clickbait detection in youtube videos", CoRR, vol. abs/2107.12791, 2021, [online] Available: <https://arxiv.org/abs/2107.12791>.
6. X. Zhang, J. Zhao and Y. LeCun, "Character-level convolutional networks for text classification", Advances in neural information processing systems, pp. 649-657, 2015.
7. T. Young, D. Hazarika, S. Poria and E. Cambria, Recent trends in deep learning based natural language processing, 2017, [online] Available: .



8. M. Potthast, S. Köpsel, B. Stein and M. Hagen, "Clickbait detection", European Conference on Information Retrieval, pp. 810-817, 2016.
- 9.
10. C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques'- Taylor & Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011. DOI: 10.1080/15325008.2010.541746
11. C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - Journal of Electrical Engineering, Vol.63 (6), pp.365-372, Dec.2012. DOI: 10.2478/v10187-012-0054-2
12. C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011. DOI 10.1007/s00202-011-0203-9
13. S.Tamilselvi, R.Prakash, C.Nagarajan, "Solar System Integrated Smart Grid Utilizing Hybrid Coot-Genetic Algorithm Optimized ANN Controller" Iranian Journal Of Science And Technology-Transactions Of Electrical Engineering, DOI10.1007/s40998-025-00917-z,2025
14. S.Tamilselvi, R.Prakash, C.Nagarajan, " Adaptive sliding mode control of multilevel grid-connected inverters using reinforcement learning for enhanced LVRT performance" Electric Power Systems Research 253 (2026) 112428, doi.org/10.1016/j.epr.2025.112428
15. S.Thirunavukkarasu, C. Nagarajan, 2024, "Performance Investigation on OCF and SCF study in BLDC machine using FTANN Controller," Journal of Electrical Engineering And Technology, Volume 20, pages 2675–2688, (2025), doi.org/10.1007/s42835-024-02126-w
16. C. Nagarajan, M.Madheswaran and D.Ramasubramanian- 'Development of DSP based Robust Control Method for General Resonant Converter Topologies using Transfer Function Model'- Acta Electrotechnica et Informatica Journal , Vol.13 (2), pp.18-31, April-June.2013, DOI: 10.2478/aei-2013-0025.
17. C.Nagarajan and M.Madheswaran - 'DSP Based Fuzzy Controller for Series Parallel Resonant converter'- Springer, Frontiers of Electrical and Electronic Engineering, Vol. 7(4), pp. 438-446, Dec.12. DOI 10.1007/s11460-012-0212-0.
18. C.Nagarajan and M.Madheswaran - 'Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis'- Iranian Journal of Electrical & Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.
19. C.Nagarajan and M.Madheswaran, "Analysis and Simulation of LCL Series Resonant Full Bridge Converter Using PWM Technique with Load Independent Operation" has been presented in ICTES'08, a IEEE / IET International Conference organized by M.G.R.University, Chennai.Vol.no.1, pp.190-195, Dec.2007
20. Suganthi Mullainathan, Ramesh Natarajan, "An SPSS and CNN modelling based quality assessment using ceramic materials and membrane filtration techniques", Revista Materia (Rio J.) Vol. 30, 2025, DOI: <https://doi.org/10.1590/1517-7076-RMAT-2024-0721>
21. M Suganthi, N Ramesh, "Treatment of water using natural zeolite as membrane filter", Journal of Environmental Protection and Ecology, Volume 23, Issue 2, pp: 520-530,2022
22. Stop Clickbait: Detecting and Preventing Clickbaits in Online News Media.
23. Anand, T. Chakraborty and N. Park, "We used Neural Networks to Detect Clickbaits: You won't believe what happened Next!", European Conference on Information Retrieval, pp. 541-547, 2017.
24. Gopinathan, V. R. (2024). Real-Time Fault-Tolerant Multi-Cloud Database Architectures for High Availability Applications. International Journal of Future Innovative Science and Technology (IJFIST), 7(4), 13148.
25. Chandra, S., Rengarajan, A., Sahoo, G. S., & Sharma, S. (2023, December). Identifying Neuronal Damage and Plasticity by Analyzing Changes in Diffusion Tensor Imaging. In International Conference on Data Science, Machine Learning and Applications (pp. 433-438). Singapore: Springer Nature Singapore.
26. Sugumar, R. (2025). Federated AI in Offline-First Mobile Health Architectures for Privacy-Preserving Clinical Intelligence. International Journal of Science, Research and Technology, 8(4), 14589-14600.
27. Murugeswari, B., Rajalakshmi, S., & Sudharson, K. (2023). Hybrid Approach for Privacy Enhancement in Data Mining Using Arbitrariness and Perturbation. Computer Systems Science & Engineering, 44(3).
28. Pandey, V. K., Mishra, S., Rengarajan, A., Savita, & Roomi, M. M. (2024, March). Enhancing Weather Forecasting with Machine Learning Techniques. In International Conference on Renewable Power (pp. 147-156). Singapore: Springer Nature Singapore.
29. Soundappan, S. J. (2025). Next Generation AI Enabled Holistic Cognitive Platform for Secure Cloud Network Intelligence Enterprise Systems and Digital Trust Optimization. International Journal of Computer Technology and Electronics Communication, 8(5), 11534-11542.
30. Mathew, A. (2022). Leveraging Big Data Analytics to Power AI and ML (Machine Learning) Automation. Educational Research (IJM CER), 4(5), 131-134.



31. Sugumar, R. (2024). AI-Augmented Quality Engineering for Performance Optimization and Test Orchestration in Distributed Systems. *International Journal of Science, Research and Technology*, 7(5), 12835-12846.
32. Akila, R. (2024). A deep reinforcement learning approach for optimizing inventory management in the agri-food supply chain. *J. Electrical Systems*, 20(4s), 2238-2247.
33. Mahendran, M., Anbazhagan, K., Pavithran, G., Nivas, A., & Pandey, S. D. (2022). Earthquake Damage Prediction using Machine Learning. *Grenze International Journal of Engineering & Technology (GIJET)*, 8(1).
34. Gopinathan, V. R. (2025). Enterprise AI Frameworks for Financial Data Engineering Behavioural Analytics and Intelligent Cloud Solutions. *International Journal of Research Publications in Engineering, Technology and Management (IJRPETM)*, 8(4), 12499-12506.
35. Kondalsamy, P., & Kaliappan, K. (2025). An Optimal Prediction of Leaf Disease Based on Hybrid Deep Learnings and Metaheuristic Technique. *Traitement du Signal*, 42(1), 363.
36. Deivendran, P., Babu, P. S., Malathi, G., Anbazhagan, K., & Kumar, R. S. (2023). Emotion Recognition for Challenged People Facial Appearance in Social using Neural Network. arXiv preprint arXiv:2305.06842.
37. Sugumar, R. (2025). Unified AI Framework for Predictive Data Engineering and Real Time Prescription and Billing Systems. *International Journal of Advanced Engineering Science and Information Technology (IJAESIT)*, 8(5), 17261.
38. Vekariya, V., Kumar, S., & Rengarajan, A. (2024). A distinctive and smart agricultural knowledge-based framework using ontology. In *Sustainability in Digital Transformation Era: Driving Innovative & Growth* (pp. 207-213). CRC Press.
39. Gopinathan, V. R. (2025). Software engineering practices for AI-driven systems: From development to deployment (MLOps perspective). *International Journal of Science, Research and Technology*, 8(1), 13493-13500.
40. Mathew, A. R. (2022). Threats and protection on E-sim: a prospective study. *Novel Perspectives of Engineering Research*, 8, 76-81.
41. Naveena, S., & Kavitha, K. (2025). Gossypium herbaceum: Folium disease identification and classification using Efficient Net-Coordinate Convolutional Neural Network (EcoNet). *Engineering Applications of Artificial Intelligence*, 152, 110701.
42. Rengarajan, A., Mishra, A., Kulhar, K. S., Shrivastava, V. P., & Alawneh, Y. J. J. (2024, March). Role of Deep Reinforcement Learning in Mitigating Cyber Security Issues: A Review. In *International Conference on Renewable Power* (pp. 37-48). Singapore: Springer Nature Singapore.
43. Achari, A. P. S. K., & Sugumar, R. (2024, November). Performance analysis and determination of accuracy using machine learning techniques for naive bayes and random forest. In *AIP Conference Proceedings* (Vol. 3193, No. 1, p. 020199). AIP Publishing LLC.
44. Mathew, A., & Alex, H. (2022). Detect & protect-medical device cybersecurity. *Curr. Overview Sci. Technol. Res*, 1, 60-68.
45. Sammy, F., Chettier, T., Boyina, V., Shingne, H., Saluja, K., Mali, M., ... & Shobana, A. (2025). Deep Learning-Driven Visual Analytics Framework for Next-Generation Environmental Monitoring. *Journal of Applied Science and Technology Trends*, 114-122.
46. Anbazhagan, K. (2024). Trustworthy and Adaptive AI Systems for Enterprise Analytics Cybersecurity and Decision Optimization Using API-First and Cloud-Native Architectures. *International Journal of Technology, Management and Humanities*, 10(03), 65-74.
47. Mathew, A. (2021). Deep reinforcement learning for cybersecurity applications. *Int J Comput Sci Mob Compu*, 10(12), 32-38.
48. Dhinakaran, D. (2022). Joe Prathap P. M, Selvaraj D, Arul Kumar D and Murugeswari B, " Mining Privacy-Preserving Association Rules based on Parallel Processing in Cloud Computing,". *International Journal of Engineering Trends and Technology*, 70(3), 284-294.
49. Karthika, K., Anusha, K., Kavitha, K., Harshadha, R., Dharshini, D. S., & Sundhar, N. A. (2025, April). Frequency Reconfigurable Antenna using Advanced Materials: A Study. In *2025 3rd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)* (pp. 1-6). IEEE.
50. Thavamani, C., & Rengarajan, A. (2024). Clustering related behaviour of users by the use of partitioning and parallel transaction reduction algorithm. *International Journal of Advanced Intelligence Paradigms*, 29(2-3), 122-132.
51. Sugumar, R. (2025). Unified AI Framework for Predictive Data Engineering and Real Time Prescription and Billing Systems. *International Journal of Advanced Engineering Science and Information Technology (IJAESIT)*, 8(5), 17261.
52. Soundappan, S. J., & Sugumar, R. (2016). Optimal knowledge extraction technique based on hybridisation of improved artificial bee colony algorithm and cuckoo search algorithm. *International Journal of Business Intelligence and Data Mining*, 11(4), 338-356.



53. SakthiPreetha, A., Kavitha, K., Karthika, K., & Manohari, R. G. (2025, April). A Novel Metasurface-Embedded Antenna for WBAN Communications. In 2025 3rd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA) (pp. 1-4). IEEE.
54. Murugeswari, B., Selvaraj, D., Sudharson, K., & Radhika, S. (2023). Data Mining with Privacy Protection Using Precise Elliptical Curve Cryptography. *Intelligent Automation & Soft Computing*, 35(1).
55. Gopinathan, V. R. (2025). Software engineering practices for AI-driven systems: From development to deployment (MLOps perspective). *International Journal of Science, Research and Technology*, 8(1), 13493-13500.
56. Anbazhagan, K., Kumar, R., Thilagavathy, R., & Anuradha, D. (2024, March). Shortest Job First with Gateway-based Resource Management Strategy for Fog Enabled Cloud Computing. In 2024 4th International Conference on Data Engineering and Communication Systems (ICDECS) (pp. 1-6). IEEE.
57. Kannadhasan, S., Vasuki, S., Kavitha, K., Karthikeyan, P., & Usha, S. G. A. (Eds.). (2025, April). Preface: Role of Artificial Intelligence and IoT in Engineering, Technology & Science [ICRAETS 2024]. In *AIP Conference Proceedings* (Vol. 3258, No. 1, p. 010001). AIP Publishing LLC.
58. Dhinakaran, D., Prathap, P. J., Selvaraj, D., Kumar, D. A., & Murugeswari, B. (2022). Mining privacy-preserving association rules based on parallel processing in cloud computing. *International Journal of Engineering Trends and Technology*, 70(3), 284-294.