TRUTH SEEKER: DETECTING AND COUNTERING FAKE NEWS TO PREVENT MISINFORMATION

Team Members with Roll No.
Reshma Yasmin M. A (211501080)
Shruthi G (211501099)

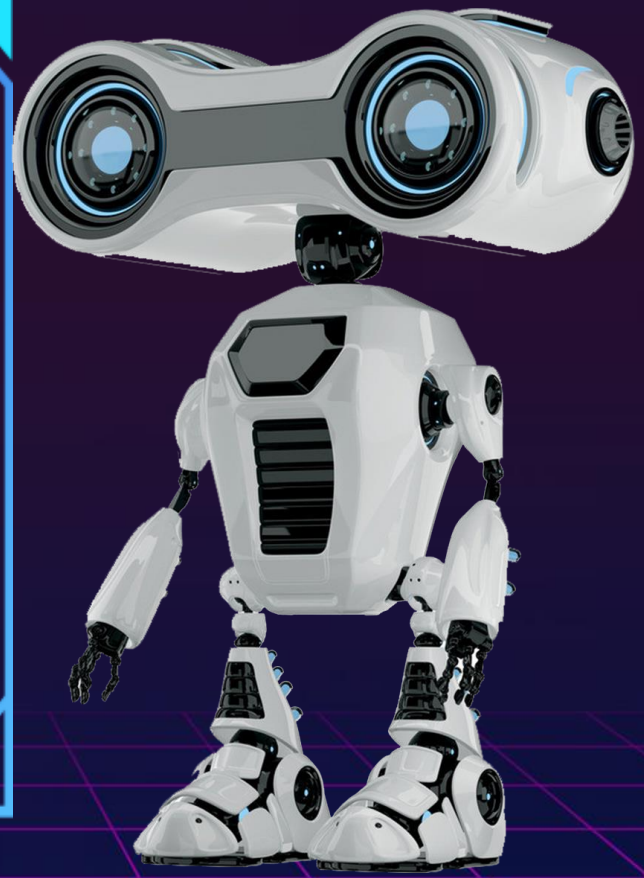
Mentor Name:
Dr. Sekar K

Department of Artificial Intelligence & Machine Learning

AI19811 – Phase II Project

ABSTRACT

- In the digital era, misinformation spreads rapidly through both text and images, threatening public trust and decision-making.
- The system employs Natural Language Processing (NLP) and Vision Transformers (ViT) to detect fake news in both textual and visual formats.
- Text-based detection uses TF-IDF vectorization with a Passive Aggressive Classifier with hybrid model for classifying news articles.
- Image-based detection leverages a fine-tuned ViT-B/16 model for binary classification of news images as real or fake.
- A Flask web interface enables real-time user interaction, allowing uploads of news text or images for instant classification.



MOTIVATION

- **Rapid Spread of Misinformation:** The rise of social media accelerates the spread of fake news, impacting public opinion and safety.

- **Multi-Modal Misinformation:** Fake news isn't limited to text—images are increasingly manipulated to mislead audiences.

- **Lack of Reliable Detection Tools:** Existing tools often focus only on text or lack real-time capabilities.

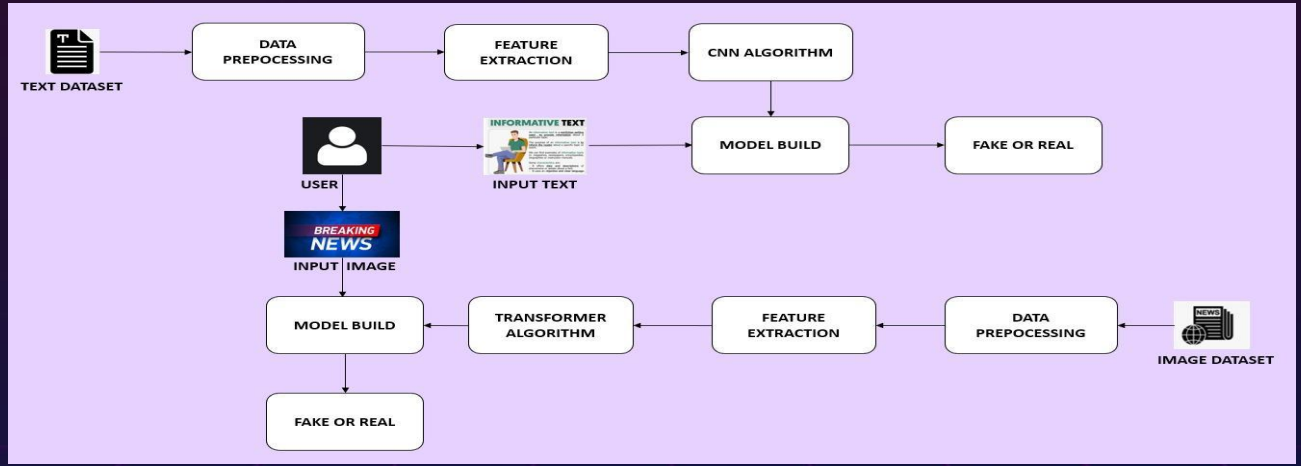
- **Need for Truthful Content Verification:** Ensuring the authenticity of news is crucial, especially in sensitive domains like politics and health.

- **Inspire Critical Thinking:** Empower users to question and verify content before trusting or sharing it.

OBJECTIVES

- **Detect Fake News in Text and Images:**Use NLP techniques (TF-IDF, PassiveAggressiveClassifier) ,hybrid model(Naïve Bayes,SVM,CNN) and ViT-B/16 for accurate classification.
- **Develop a Unified Detection Platform:** Build a Flask-based web interface combining both models for real-time news verification.
- **Enable Domain-Based Categorization:** Classify news into Politics, Weather, Sports, and Others for enhanced contextual relevance.
- **Improve User Interaction & Experience:** Design an intuitive UI/UX for easy uploads and instant, clear results.
- **Promote Digital Media Literacy:** Empower users to recognize misinformation and encourage responsible content consumption.

ARCHITECTURAL DESIGN / FRAMEWORK OF PROPOSED SYSTEM



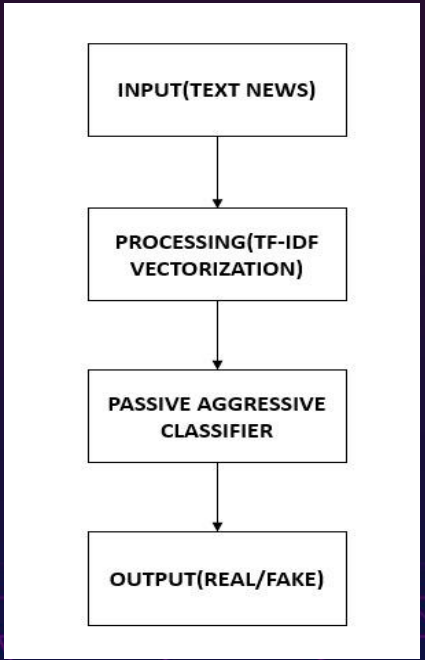
LIST OF MODULE

1. Text-Based Fake News Detection
2. News Categorization (Domain Classification)
3. Image-Based Fake News Detection
4. UI/UX and File Upload
5. Integration and Result Display

MODULE 1: DESCRIPTION

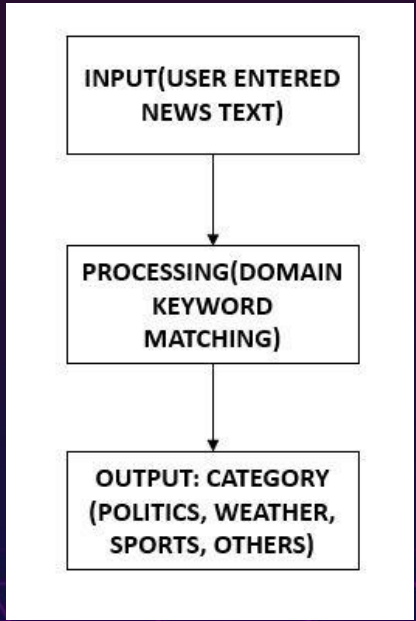
- This module accepts news text as input and uses NLP techniques to convert it into numerical vectors using TF-IDF.
- The Passive Aggressive Classifier with hybrid model(Naïve Bayes,SVM,CNN) is then used to classify the news as Real or Fake.
- It is lightweight, fast, and ideal for large-scale text classification.

DESIGN



MODULE 2: DESCRIPTION DESIGN

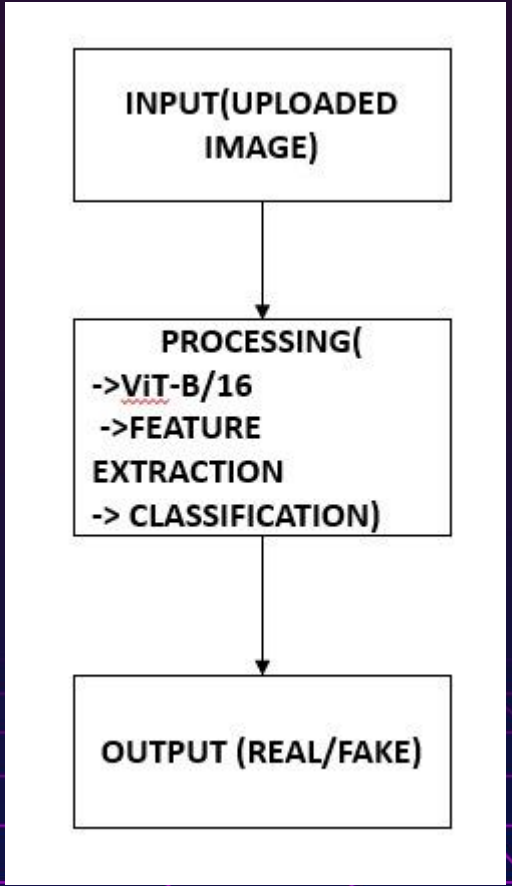
- This module analyzes the input news content to classify it into categories like Politics, Weather, or Sports, based on the presence of relevant keywords.
- It enhances the understanding of news context and supports domain-specific analysis.



MODULE 3: DESCRIPTION

- This module handles news image detection using Vision Transformer (ViT-B/16).
- The image is transformed, passed through the pretrained ViT model, and classified into Real or Fake based on deep feature representations. It is especially useful for verifying memes, posters, and visual misinformation.

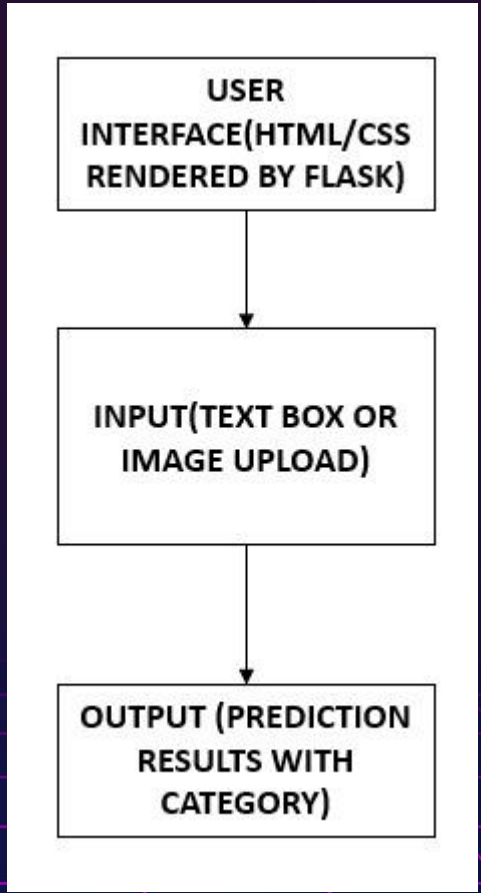
DESIGN



MODULE 4: DESCRIPTION

- This module allows users to interact with the system via a web-based UI. Users can either enter text manually or upload an image.
- The results are shown in a clear, responsive design with predicted output and news category.

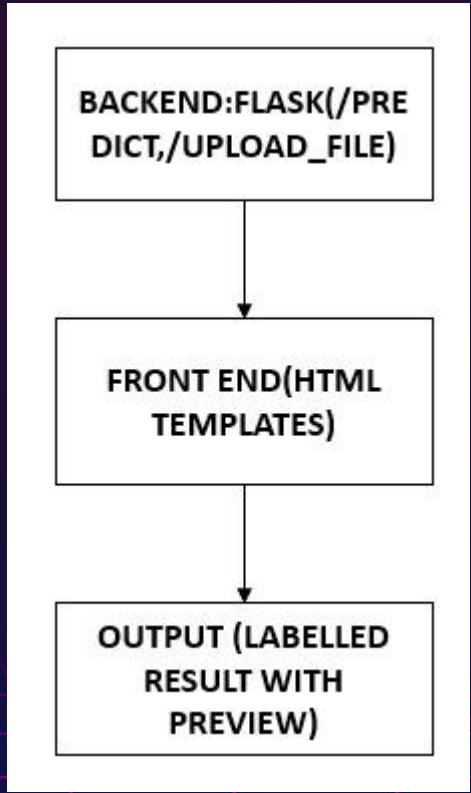
DESIGN



MODULE 5: DESCRIPTION

- This module connects all components together. It handles POST requests, invokes the correct detection logic (text/image), and displays the result to the user.
- This ensures a smooth end-to-end experience from input to prediction.

DESIGN



IMPLEMENTATION RESULTS

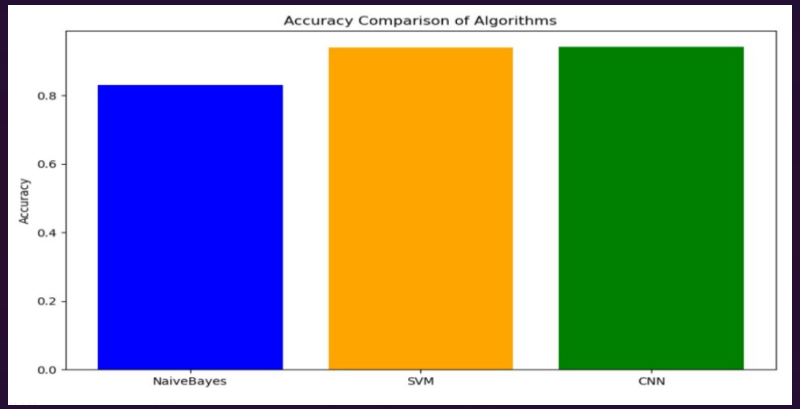


Figure 1 :Accuracy Comparison of Different Algorithms for Fake News Detection.

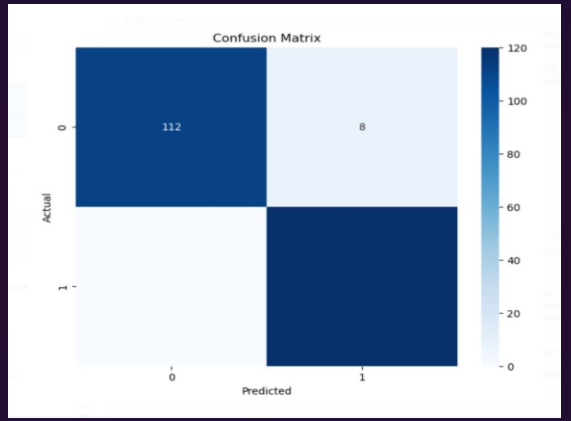


Figure 2 : Confusion Matrix of the Fake News Detection Model

Classification Report:

	precision	recall	f1-score	support
0	1.00	0.93	0.97	120
1	0.94	1.00	0.97	120
accuracy			0.97	240
macro avg	0.97	0.97	0.97	240
weighted avg	0.97	0.97	0.97	240

Figure 2 : Model Classification Performance

Figure 4: Fake news detection system classifying input text as real news.

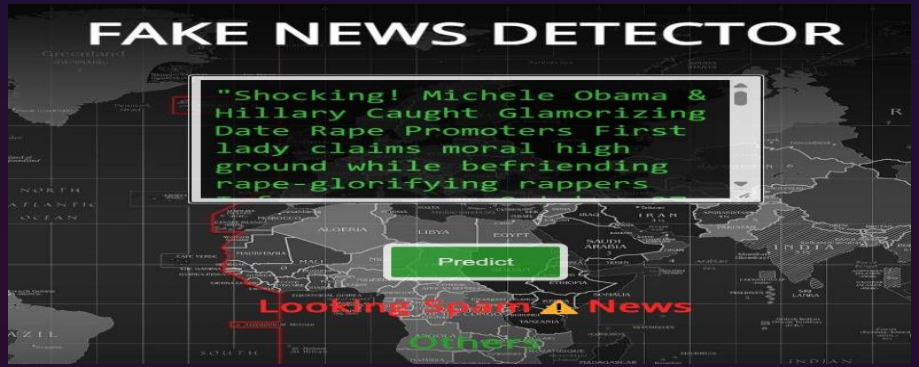


Figure 5: Fake news detection system classifying input text as fake news.

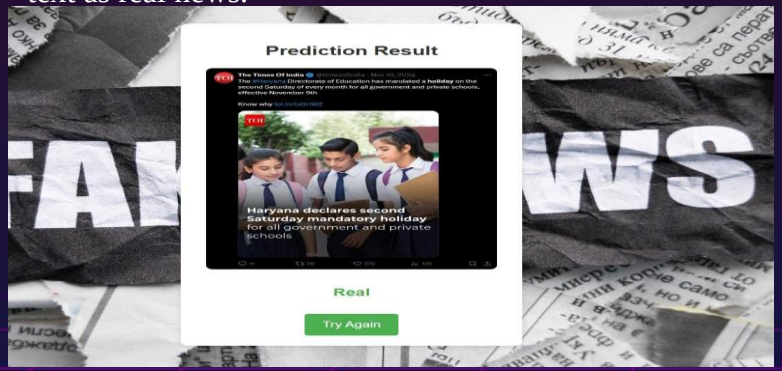


Figure 5: Fake news detection system classifying input image as real news.

Figure 7: Fake news detection system classifying input text as fake news.

PERFORMANCE ANALYSIS

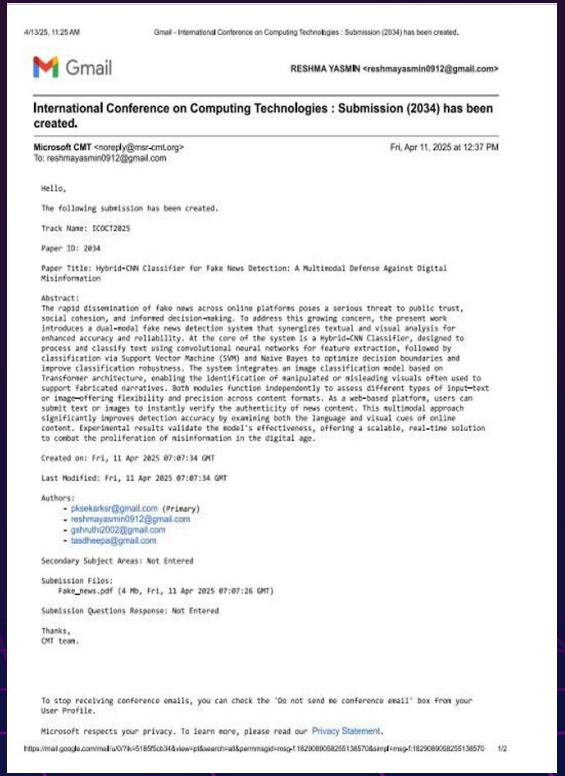
- High Accuracy: Achieved 98.8% training and 97% testing accuracy over 10 epochs, with steadily decreasing loss.
- Low Misclassification: Confusion matrix shows only 8 misclassifications out of 240 samples.
- Strong Metrics: Precision, recall, and F1-score all exceed 0.94, indicating well-balanced classification.
- Model Reliability: CNN-based model shows excellent learning and generalization capability.
- Outperforms Baselines: Surpasses Naive Bayes (83%) and SVM (94%) with 94.5% accuracy in image-based detection.

CONCLUSION

- Developed an effective **fake news detection system** combining **textual** and **image-based** analysis.
- Achieved high accuracy using:
  - **Text-based models** like Naive Bayes, SVM, and CNN with TF-IDF.
  - **Image-based CNN** for visual misinformation classification.
- Built an integrated **Flask web application** for real-time prediction.
- System helps combat the spread of misinformation by verifying both **news text and images**.
- Lays the foundation for building more robust, hybrid misinformation detection systems.

PATENT/PUBLICATION/COMPETITION DESIGN STATUS

- **Team ID:** B014
- **Team Member:** 211501080 Reshma Yasmin MA 211501099 Shruthi G
- **Team Supervisor:** Dr. K. Sekar M.E., Ph.D.,
- **Publication Status:** Communicated
- **Particulars of Conference:** The International Conference on Computing Technologies (ICOCT 2025)
- **Host Institution:** Jyothi Institute of Technology
- **Conference Date:** 13-14 June 2025



4/13/25, 11:25 AM Gmail - International Conference on Computing Technologies : Submission (2034) has been created.

Gmail RESHMA YASMIN <reshmayasmin0912@gmail.com>

International Conference on Computing Technologies : Submission (2034) has been created.

Microsoft CMT <noreply@msr-cmt.org> Fri, Apr 11, 2025 at 12:37 PM
To: reshmayasmin0912@gmail.com

Hello,

The following submission has been created.

Track Name: ICOCT2025

Paper ID: 2034

Paper Title: Hybrid-CNN Classifier for Fake News Detection: A Multimodal Defense Against Digital Misinformation

Abstract:
The rapid dissemination of fake news across online platforms poses a serious threat to public trust, social cohesion, and informed decision-making. To address this growing concern, the present work introduces a dual-modal fake news detection system that synergizes textual and visual analysis for enhanced accuracy and reliability. At the core of the system is a Hybrid-CNN Classifier, designed to process and classify text using convolutional neural networks for feature extraction, followed by classification via Support Vector Machine (SVM) and Naive Bayes to optimize decision boundaries and improve classification robustness. The system integrates an image classification model based on Transformer architecture, enabling the identification of manipulated or misleading visuals often used to support fabricated narratives. Both modules function independently to assess different types of input—text or image—offering flexibility and precision across content formats. As a web-based platform, users can submit text or images to instantly verify the authenticity of news content. This multimodal approach significantly improves detection accuracy by examining both the language and visual cues of online content. Experimental results validate the model's effectiveness, offering a scalable, real-time solution to combat the proliferation of misinformation in the digital age.

Created on: Fri, 11 Apr 2025 07:07:34 GMT

Last Modified: Fri, 11 Apr 2025 07:07:34 GMT

Authors:
- pksekarkar@gmail.com (Primary)
- reshmayasmin0912@gmail.com
- gshruthi2002@gmail.com
- tasdheepa@gmail.com

Secondary Subject Areas: Not Entered

Submission Files:
Fake_news.pdf (4 Mb, Fri, 11 Apr 2025 07:07:26 GMT)

Submission Questions Response: Not Entered

Thanks,
CMT team.

To stop receiving conference emails, you can check the 'Do not send me conference email' box from your User Profile.

Microsoft respects your privacy. To learn more, please read our Privacy Statement.

https://mail.google.com/mail/u/0/?ik=5185fcb34&view=pt&search=all&permmsgid=msg-f:1829089058255138570&simpl=msg-f:1829089058255138570 1/2

REFERENCES

1. A. Ahmed, A. Aljabouh, P. K. Donepudi, and M. S. Choi, “Detecting Fake News Using Machine Learning: A Systematic Literature Review,” arXiv preprint arXiv:2102.04458, 2021.
2. Atrey, P. K., Hossain, M. A., El Saddik, A., & Kankanhalli, M. S. (2010). Multimodal Fusion for Multimedia Analysis: A Survey. Multimedia Systems, 16(6), 345-379.
3. Castillo, C., Mendoza, M., & Poblete, B. (2011). Information Credibility on Twitter. Proceedings of the 20th International Conference on World Wide Web (WWW), 675-684.
4. Chen, H., Kazerooni, D., & Hsu, K. (2021). Deep Learning for Fake News Detection using CNN and RNN. IEEE International Conference on Big Data (Big Data), 5378-5383. Kumar, A., & Singh, A. (2023).
5. Fake News Detection Using Multimodal Deep Learning Models: A Review. Journal of Information Science, 49(2), 172-185.
6. Li, Y., Jia, K., & Wang, Q. (2024). Multimodal Fake News Detection Based on Contrastive Learning and Similarity Fusion. IEEE Access, 12, 155351-155364.
7. Malanowska, A., Mazurczyk, W., Araghi, T. K., Megías, D., & Kuribayashi, M. (2024)Digital Watermarking—A Meta-Survey and Techniques for Fake News Detection. IEEE Access, 12, 36311-36345.
8. Oshikawa, R., Qian, J., & Wang, W. (2018). A Survey on Natural Language Processing for Fake News Detection. Proceedings of ACL, 74-84.
9. Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345-1359.
10. Park, M., & Chai, S. (2023). Constructing a User-Centered Fake News Detection Model by Using Classification Algorithms in Machine Learning Techniques. IEEE Access, 11, 71517-71527.

THANK YOU