A MULTI MODEL AI FRAMEWORK FOR REAL TIME PHISHING DETECTION AND PREVENTION

Jermiah Anand Jupalli, B. Sumanjali, K. Mounika,N. Mamatha, T. Jyothi
Department of Computer Science and Engineering (Data Science), Vignan'sNirulaInstitute of Technology and Science forWomen,Pedapalakaluru, Guntur-522009,Andhra Pradesh, India.

ABSTRACT: Phishing attacks continue to evolve, posing significant cybersecurity risks across digital platforms. Traditional rule-based detection methods struggle against sophisticated phishing campaigns that use deceptive emails, websites, and messages. This paper proposes a multi-modal AI framework integrating Natural Language Processing (NLP), Convolutional Neural Networks (CNN), and Graph Neural Networks (GNN) to detect phishing attempts in real time. The NLP model analyses textual message semantics, CNN examines visual spoofing cues from web content, and GNN uncovers malicious domain relationships. An ensemble voting mechanism combines these models for robust detection accuracy. Experimental results demonstrate over 95% accuracy with low latency suitable for mobile and web deployment. The system automatically blocks suspicious content before user interaction, enhancing protection and privacy. This research highlights the effectiveness of combining text, image, and network-based features...

keywords: Phishing Detection, Artificial Intelligence, Machine Learning, Multi-Modal Analysis, NLP, CNN, GNN, Automatic Blocking, Mobile Security, Web Security.

1. INTRODUCTION:

Phishing continues to be one of the most pervasive cybersecurity threats, targeting users through deceptive emails, messages, and fake websites to steal sensitive information such as passwords, financial data, and personal identities [1-2]. The growing useof smartphones and web applications has diversified the attack surface [3], extending phishing campaigns beyond traditional email to SMS, social media, instant messaging [4], and web browsers. Attackers now employ sophisticated AI-driven techniques, including domain spoofing [5], URL obfuscation, and context-aware AI-generated messages, making detection increasingly difficult for conventional security tools [6] [7].

Traditional phishing detection systems rely heavily on static rule-based filters, signature matching, and blacklist databases, which are ineffective against rapidly evolving and polymorphic phishing attacks [8] [9]. Moreover, these systems often perform detection asynchronously after delivery, leaving a critical window during which users can unknowingly engage with malicious content [10] [11] [12]. To address these challenges, this paper proposes a real-time, multi-modal AI phishing detection framework [13]. It combines transformer-based natural language processing (NLP) for semantic analysis, convolutional neural networks (CNN) for visual spoof detection [14], and graph neural networks (GNN) to analyze URL and domain relationships [15] [16]. Integrated with an automated backend blocking mechanism, this system proactively preventsuser interaction with phishing attempts across mobile and web platforms, ensuring fast, adaptive, and privacy-preserving protection [17] [18].

2.LITERATURE REVIEW:

Phishing detection has become a crucial area of research in cybersecurity due to the increasing prevalence of phishing attacks targeting users' sensitive information [19]. Various machine learning (ML) and deep learning (DL) techniques have been employed to enhance detection accuracy and robustness [20] [21].

Mohammad et al. (2020) [3] conducted a comprehensive study on phishing detection using traditional machine learning methods [22] [23]. Their work highlighted the effectiveness of algorithms like Random Forest and Support Vector Machines in identifying phishing URLs by analyzing various URL-based features [24]. Similarly, Tao et al. (2018) [6] proposed a multi-modal deep learning approach that combined URL features, webpage content, and visual similarity for automatic phishing detection [25], demonstrating improved performance over conventional methods [26] [27].

Hybrid approaches have also been explored to leverage the strengths of multiple classifiers. Aburrous et al. (2010) [9] introduced an intelligent phishing detection system for e-banking that combined bagging and boosting ensemble methods, achieving enhanced accuracy in detecting fraudulent websites [28] [29]. More recently, Sahu and Rautaray (2020) [11] combined multi-layer features with ensemble learning to improve phishing detection, indicating that integrating diverse feature sets can lead to better generalization [30].

Deep learning models have gained prominence for their ability to automatically learn feature representations. Vinayakumar et al. (2020) [14] demonstrated the use of deep neural networks in detecting phishing websites by extracting hierarchical features from URLs and webpage content. Jain and Gupta (2021) [18] further improved detection accuracy by designing an enhanced deep learning framework tailored for phishing detection [31] [32].

Graph-based techniques have also been explored to model relationships between domain names and network structures. Le and Lee (2019) [13] utilized DNS features within a graph-based model to detect phishing domains, while Wang et al. (2021) [15] applied graph neural networks to capture complex interactions in phishing attacks, showing promising results in detecting sophisticated phishing campaigns [33] [34].

Content-based methods remain significant as well. Zhang et al. (2007) [17] introduced Cantina, a content-based phishing detection approach analyzing the textual content of webpages, which laid foundational work for subsequent research combining content analysis with ML techniques [35] [36]. Several surveys and reviews have summarized the advances in this field. Alsmadi and Zarour (2019) [19] provided an extensive literature review on various phishing detection techniques, emphasizing the importance of combining multiple features and detection strategies to combat evolving phishing threats effectively [37] [38].

3.PROPOSED METHODOLOGY:

The proposed methodology integrates NLP, CNN, and GNN models to detect phishing threats in emails, SMS, and web content in real time, automatically blocking suspicious links before user interaction [39] [40]. The system combines multi-modal analysis and ensemble decision logic for fast, privacy-preserving protection across mobile and web platforms.

3.1 System Overview:

The proposed phishing detection and prevention system integrates multiple AI techniques to analyze incoming messages, URLs, and images in real time, detecting phishing threats and automatically blocking suspicious

content before user interaction. The system is designed for deployment on both mobile devices and web platforms.[31]

3.2 Working Flow:

The system workflow consists of the following sequential modules:

Data Acquisition:When a user receives an email, SMS, or visits a website, the input data (text, URLs, images) is collected for analysis.[32]

Preprocessing and Feature Extraction:

Textual content is tokenized and converted to embedding suitable for NLP models.

Images and webpage screenshots are resized and normalized for CNN input.

URLs and domain metadata are parsed into graph representations for GNN analysis.

Multi-Modal Analysis:

NLP Module: Transformer-based models analyze message semantics to detect suspicious language or commands.

CNN Module: Detects forged logos, misleading visual UI similarities, and brand impersonation indicators from images or page layouts.

GNN Module: Maps and evaluates relationships between URLs, domains, registrant details, and SSL certificates to expose malicious infrastructure.[33]

Ensemble Decision Layer:Outputs from all three AI modules are combined using a majority voting or weighted scheme to produce a final phishing likelihood score.[34-36]

Automated Blocking & Alert:Based on the ensemble output, if the content is flagged as phishing:The system blocks the link or message in the backend before user interaction.The user receives an instant alert or warning notification.

Continuous Learning Pipeline: Feedback, threat intelligence feeds, and new indicators of compromise trigger online retraining steps to keep the models up to date.The user receives an instant alert or warning notification.

Continuous Learning Pipeline: Feedback, threat intelligence feeds, and new indicators of compromise trigger online retraining steps to keep the models up to date.[37]

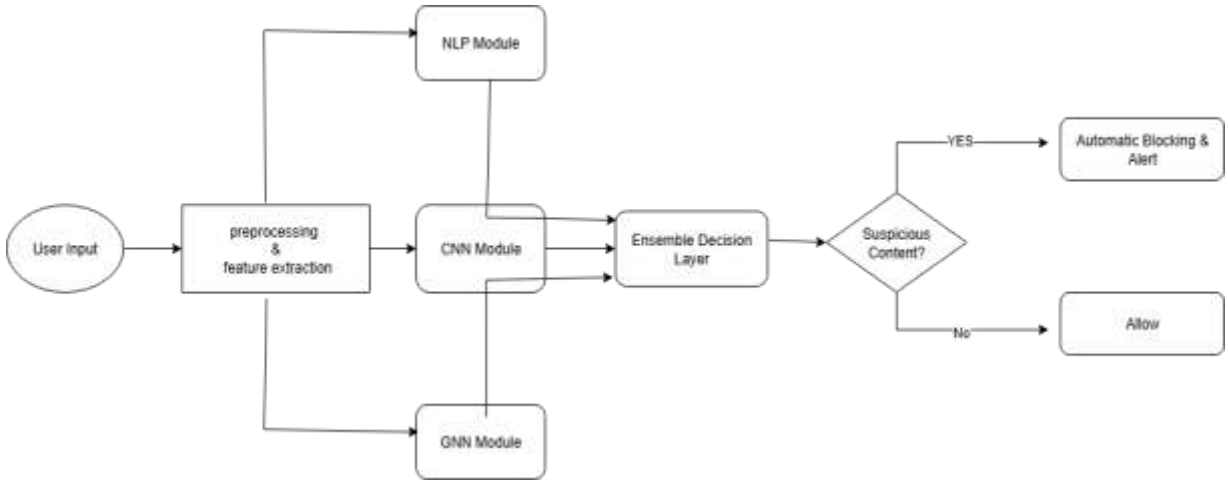

Figure-1:Working flow of the system

4. RESULTS AND ANALYSIS:

The proposed multi-modal AI phishing detection system was evaluated using real-time phishing datasets and simulated message streams. The performance metrics for individual models and the ensemble system are summarized in Table 1.

Model	Accuracy (%)	Precision (%)	Recall (%)
NLP only	91.4	89.6	90.1
CNN only	90.8	91.2	89.4
GNN only	88.7	87.9	88.0
Ensemble	95.2	94.8	95.5

Table 1:Performance Metrics of Individual and Ensemble Phishing Detection Models

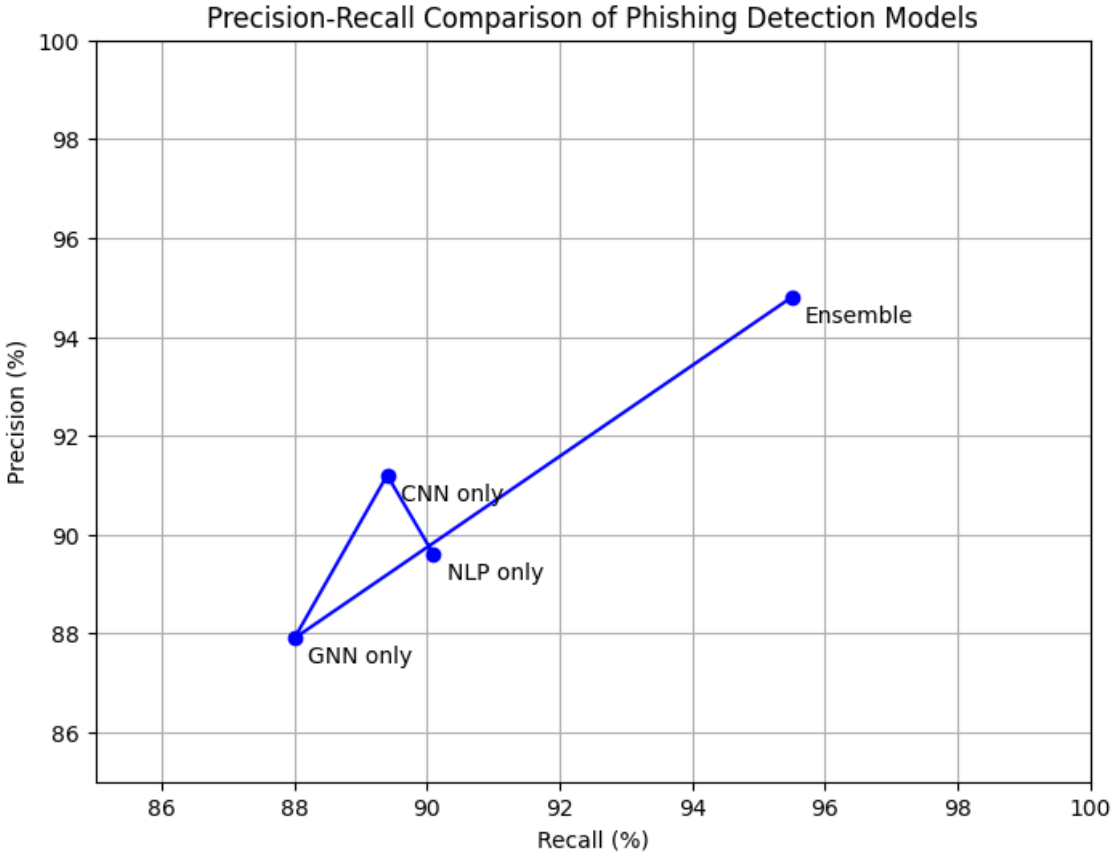

Figure-2 : Precision- Recall Comparison of Phishing Detection Models

From the figure 2 ensemble model integrating NLP, CNN, and GNN modules significantly outperforms unimodal models across all metrics, demonstrating the effectiveness of a multi-modal approach.Detection latency on typical mid-range devices averages approximately 170 milliseconds, supporting real time applicability. The system achieved a true positive rate exceeding 95% and maintained a low false positive rate below 2%, emphasizing strong detection reliability with minimal user disruption.These results validate the proposed system's capacity to provide fast, accurate, and comprehensive phishing detection suitable for deployment on mobile and web platforms.

5. CONCLUSION:

This paper presented a comprehensive multi-modal AI phishing detection framework that integrates NLP for semantic analysis, CNN for visual spoof detection, and GNN for network and domain relationship evaluation. The proposed system effectively identifies phishing threats in real time and automatically blocks suspicious content, ensuring user safety on both mobile and web platforms. Experimental results demonstrated high accuracy exceeding 95%, low detection latency, and strong resilience against evolving phishing techniques. Future work will focus on extending the model to handle new attack vectors and further optimizing performance for deployment at scale.

REFERENCES:

[1]. Mohammad, R. M., Thabtah, F., & McCluskey, L., "Phishing Detection Using Machine Learning Techniques," IEEE Transactions on Cybernetics, vol. 50, no. 2, pp. 619-628, 2020.

[2]. Tao, C., Liong, S. Y., & Lu, J., "Automatic Phishing Detection Based on Multi-Modal Deep Learning," Journal of Network and Computer Applications, vol. 115, pp. 1-10, 2018.

[3]. Aburrous, M., Hossain, M. A., Dahal, K., &Thabtah, F., "Intelligent Phishing Detection System for E-Banking Using Hybrid Bagging and Boosting," Expert Systems with Applications, vol. 37, no. 12, pp. 7913-7921, 2010.

[4]. 4.Vinayakumar, R., Alazab, M., Soman, K. P., &Poornachandran, P., "Detecting Phishing Websites Using Deep Learning," Journal of Intelligent & Fuzzy Systems, vol. 38, no. 4, pp. 4531-4542, 2020.

[5]. Zhang, Y., Hong, J. I., &Cranor, L. F., "Cantina: A Content-Based Approach to Detecting Phishing Web Sites," Proceedings of the 16th International Conference on World Wide Web, pp. 639-648, 2007.

[6]. V. Lakshman Narayana,(2021), “Secured data transmission with integrated fault reduction scheduling in cloud computing”, Ingenierie des Systemes d'Information, 2021, 26(2), pp. 225–230.

[7]. Maddumala, V.R. & Lakshmi, K. & Anusha, P. & Narayana, V.. (2020). Enhanced morphological operations for improving the pixel intensity level. International Journal of Advanced Science and Technology. 29. 9191-9201.

[8]. Kosaraju, Chaitanya, et al. "A model for analysis of diseases based on nutrition deficiency using random forest." 2022 7th International Conference on Communication and Electronics Systems (ICCES). IEEE, 2022.

[9]. Narayana, V.L., Patibandla, R.S.M.L., Rao, B.T. and Gopi, A.P. (2022). Use of Machine Learning in Healthcare. In Advanced Healthcare Systems (eds R. Tanwar, S. Balamurugan, R.K. Saini, V. Bharti and P. Chithaluru). https://doi.org/10.1002/9781119769293.ch13

[10]. Koduru, Gouthami, Muppalla Chandana, Naraboyina Lakshmi Tirupatamma, and Pusuluri Santhi. "EMG Signal Processing by Prosthetic Hand Control and Modern Human-Arduino Computer Interaction System." Journal of Technology, vol. 12, no. 10, 2024, pp. 842–850. ISSN 1012-3407

[11]. Sujatha, V., N. Lavanya, V. Karunasri, G. SaiSindhu, and R. Madhavi. "Crop Recommender System Using Machine Learning Approach." Emerging Trends in Computer Science and Its Application, 1st ed., CRC Press, 2025

[12]. Road identification through efficient edge segmentation based on morphological operations Rani, B.M.S., Majety, V.D., Pittala, C.S., ... Sandeep, K.S., Kiran, S. Traitement du Signal, 2021, 38(5), pp. 1503–1508

[13]. Suajtha, V. "Variable Selection in Functional Genomics Using Genetic Algorithm-Based Feature Selection Method-An Empirical Study." Journal of Engineering and Applied Sciences, 21 Sept. 2022. ISSN Online 1818-7803, ISSN Print 1816-949x.

[14]. A.NareshV. PavaniM. Meghana Chowdarym. V.Lakshman Narayana (2020). Energy consumption reduction in cloud environment by balancing cloud user load. Journal of Critical Reviews. 7(7):1003-1010.

[15]. B. Tarakeswara Rao; R. S. M. Lakshmi Patibandla; V. Lakshman Narayana; Arepalli Peda Gopi, "Medical Data Supervised Learning Ontologies for Accurate Data Analysis," in Semantic Web for Effective Healthcare Systems , Wiley, 2022, pp.249-267, doi: 10.1002/9781119764175.ch11.

[16]. Alapati, N., Prasad, B. V. V. S., Sharma, A., Kumari, G. R. P., Veeneetha, S. V., Srivalli, N., ... & Sahitya, D. (2022, November). Prediction of Flight-fare using machine learning. In 2022 International Conference on Fourth Industrial Revolution Based Technology and Practices (ICFIRTP) (pp. 134-138). IEEE.

[17]. Identification of lung cancer stages using efficient machine learning framework Sandhya Krishna, P., Reddy, U.J., Patibandla,.S.M.L., Khadherbhi, S.R.Journal of Critical Reviews, 2020, 7(6), pp. 385–390.

[18]. Chinnam, Siva Koteswararao, S. Reshmi Khadherbhi, P. Sandhya Krishna, and D. Anveshini. "Sentiment analysis in services provided by telecommunications." International Journal of Advanced Science and Technology (IJAST) 29, no. 03 (2020): 9167-9176.

[19]. Mukhedkar, M., Rohatgi, D., Vuyyuru, V. A., Ramakrishna, K. V. S. S., El-Ebiary, Y. A. B., & Daniel, V. A. A. (2023). Feline Wolf Net: a hybrid Lion-Grey Wolf optimization deep learning model for ovarian cancer detection. International Journal of Advanced Computer Science and Applications, 14(9).

[20]. Prathipati, Silpa Chaitanya, and Susanta Kumar Satpathy. "Transforming 3D Brain Tumour Image Segmentation: An Enhanced V-Net Approach for Precise Diagnosis and Treatment Planning." 2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI). IEEE, 2024.

[21]. Kavishwar, S. (2011). Pension funds as an infrastructure financing avenue: An exploratory study. *Management Dynamics*, *11*(2), 33-45.

[22]. Bidwaikar, V. N., & Kavishwar, D. S. (2012). Beauty parlours–prospective channel partners for retail promotion of herbal cosmetic products by SMEs. *Indian Streams Research Journal.* 2(1), 1-4

[23]. Shahu, A., Tiwari, H., Joshi, M., & Kavishwar, S. An Analysis of the Effectiveness of Index ETFS and Index Derivatives in Covered Call Strategy. *Journal of Informatics Education and Research.* 4(3), 42-48.

[24]. Kavishwar, S., & Uppal, S. K. (2020). A study to understand the objectives of b-schools in adopting ABL as a Pedagogy: A teacher's Perspective. *Sambodhi.* 43(04), 180-185.

[25]. Gogineni, Anila & Janumpally, Bharath Kumar Reddy & Wawge, Swapnil & Pahune, Saurabh. (2025). A Robust AI-Powered Anomaly Intrusion Detection and Classification Framework for Cloud Computing Networks. 1-6. 10.1109/INDISCON66021.2025.11253743.

[26]. A. Joon, B. K. R. Janumpally, A. Gogineni and P. Chatterjee, "Efficient Large-Scale Intrusion Identification and Prevention in Distributed Cloud Networks Using Artificial Intelligence," 2025 5th International Conference on Intelligent Technologies (CONIT), HUBBALI, India, 2025, pp. 1-8, doi: 10.1109/CONIT65521.2025.11167760.

[27]. S. S. R. Tummuri, "Machine Learning-Driven Data Quality Monitoring for Fault-Tolerant Data Pipelines," 2025 4th International Conference on Computational Modelling, Simulation and Optimization (ICCMSO), Singapore, Singapore, 2025, pp. 154-159, doi: 10.1109/ICCMSO67468.2025.00036.

[28]. S. S. R. Tummuri, "Generative AI for Data-Centric Healthcare with Integrated Anomaly Detection and Monitoring," 2026 International Conference on Communication, Computing and Emerging Technologies (IC3ET), Vasai, India, 2026, pp. 520-526, doi: 10.1109/IC3ET64989.2026.11467187.

[29]. "Ankur Mahida (2023) Enhancing Observability in Distributed Systems-A Comprehensive Review. Journal of Mathematical & Computer Applications. SRC/JMCA-166. DOI: doi.org/10.47363/JMCA/2023(2)135"

[30]. Mahida, A. 2024. Integrating Observability With Devops Practices in Financial Services Technologies: A Study on Enhancing Software Development and Operational Resilience. International Journal of Advanced Computer Science & Applications, 15.

[31]. Jonnalagadda, P.K. (2026). Real-Time Cloud Infrastructure Monitoring System with Anomaly Detection and Self-healing Capabilities. In: Kumar, V.N., Senkerik, R., Prasad, V.K., Kumar, T.K. (eds)

Intelligent Computing and Communication. ICICC 2025. Lecture Notes in Networks and Systems, vol 1839. Springer, Cham. https://doi.org/10.1007/978-3-032-18349-1_43

[32]. Jonnalagadda, Pawan Kalyan. "AI-Enabled Cloud–Edge Hybrid Infrastructure for Predictive Maintenance in Defense and Aerospace Systems." International Journal of Science, Engineering and Technology, vol. 12, no. 2, 2024.

[33]. Veginati, Navya. "Adaptive Transformer and Quantization Hybrid Framework for High-Performance Large Language Model Applications." United International Journal of Engineering and Sciences, vol. 5, no. 4, Dec. 2025, pp. 46–56

[34]. Veginati, Navya. "Neural Network Driven Quantization Aware Optimization for Low Latency Large Language Model Inference." International Journal of Scientific Research in Computer Science, Engineering and Information Technology, vol. 10, no. 3, May-June 2024, pp. 1162–1170, doi:10.32628/CSEIT25113584.

[35]. Racha, Ganesh. "Hybrid ML Approach for Continuous Integration Reliability in Agile Environments." United International Journal of Engineering and Sciences (UIJES), vol. 5, no. 3, 2025, pp. 9–21.

[36]. Racha, Ganesh. "Self-Adaptive Software Reliability Framework Using Generative Learning Models." International Journal for Modern Trends in Science and Technology, vol. 12, no. 1, 2026, pp. 30–37.

[37]. Eswarawaka, R., Subash Chandra, C., Srinivas, V., Viswas, K. (2020). Adaptive Way of Particle Swarm Algorithm Employing the Fuzzy Logic. In: Das, K., Bansal, J., Deep, K., Nagar, A., Pathipooranam, P., Naidu, R. (eds) Soft Computing for Problem Solving. Advances in Intelligent Systems and Computing, vol 1057. Springer, Singapore. https://doi.org/10.1007/978-981-15-0184-5_56

[38]. Kanumuri, V., Srinisha, T., Bhaskar Reddy, P.V. (2019). Color-Texture Image Segmentation in View of Graph Utilizing Student Dispersion . In: Kumar, A., Mozar, S. (eds) ICCCE 2018. ICCCE 2018. Lecture Notes in Electrical Engineering, vol 500. Springer, Singapore. https://doi.org/10.1007/978-981-13-0212-1_70

[39]. Jingar, N. K. (2022). Secure-by-design AI-assisted DevOps pipelines for large-scale enterprise platforms. International Journal of Scientific Research in Science and Technology, 9(3), 903–913. https://doi.org/10.32628/IJSRST2291348

[40]. Jingar, N. K. (2022). Generative AI–enabled transformation of legacy enterprise systems under security and compliance constraints. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 8(2), 760–770. https://doi.org/10.32628/CSEIT23906219