Advances in AI and Blockchain Technologies for Secure Early Disease Detection

[1]Namitha Gaddam

(M.Tech Artificial Intelligence)

Aurora's Scientific and Technological Institute,Telangana,India

Email: gaddamnamitha14@gmail.com

[2]Dr.M. Sridhar

Head Of The Department Computer Science and Engineering

Aurora's Scientific and Technological Institute,Telangana,India

Email: msridhar.msr@gmail.com

ABSTRACT

The rapid advancement of Artificial Intelligence (AI) and Blockchain technologies has significantly transformed the healthcare sector by enabling secure, intelligent, and efficient disease diagnosis systems. This project, titled *"Advances in AI and Blockchain Technologies for Secure Early Disease Detection,"* presents an integrated framework that combines machine learning, deep learning, federated learning, and blockchain technology to achieve accurate and secure early disease prediction. The proposed system utilizes AI-based algorithms such as XGBoost, Random Forest, and Convolutional Neural Networks (CNN) to detect diseases including lung cancer, diabetes, and skin cancer from both clinical datasets and medical images. The system applies preprocessing techniques, feature scaling, label encoding, and model training to improve prediction accuracy and computational efficiency. To enhance privacy and security in healthcare data sharing, blockchain technology is incorporated using smart contracts and decentralized storage mechanisms through Web3 integration, where local client model weights and patient records are securely stored and validated using cryptographic hash functions. Federated learning concepts are implemented by dividing the dataset among multiple clients, training local models independently, and aggregating model weights to form a global prediction model without exposing sensitive patient data. The graphical user interface developed using Tkinter provides an interactive environment for dataset upload, preprocessing, training, prediction, and blockchain validation.

Keywords: Artificial Intelligence, Blockchain Technology, Early Disease Detection, Machine Learning, Deep Learning, Federated Learning, Lung Cancer Prediction, Skin Cancer Detection, Diabetes Prediction, Smart Contracts, Medical Image Processing, XGBoost, Random Forest, Convolutional Neural Network, Healthcare Security, Data Privacy, Web3, Patient Data Validation, Clinical Data Analysis, Secure Healthcare System

I. INTRODUCTION

The healthcare industry is rapidly evolving with the integration of advanced digital technologies such as Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), and Blockchain Technology, which are transforming the way diseases are detected, diagnosed, and managed. Early disease detection plays a crucial role in reducing mortality rates, improving treatment success, and minimizing healthcare costs. Traditional healthcare systems often face challenges such as delayed diagnosis, inaccurate predictions, lack of secure medical data management, and privacy issues during patient data sharing across hospitals and healthcare organizations. To overcome these limitations, intelligent healthcare systems capable of secure and accurate disease prediction are becoming increasingly important in modern medical environments. Artificial Intelligence techniques, especially machine learning and deep learning algorithms, have demonstrated remarkable performance in analyzing large-scale clinical datasets and medical images for disease prediction. Algorithms such as XGBoost, Random Forest, and Convolutional Neural Networks (CNN) can effectively identify hidden patterns in healthcare data and provide accurate predictions for diseases such as lung cancer, diabetes, and skin cancer. However, centralized healthcare systems still suffer from data security vulnerabilities, unauthorized access, and privacy breaches, which may compromise sensitive patient information. Blockchain technology provides an efficient solution to these challenges by offering decentralized, immutable, and secure storage of healthcare records through cryptographic hashing and smart contracts. Blockchain ensures data integrity, transparency, and tamper-proof storage while enabling secure sharing of medical information among multiple healthcare participants. In addition, federated learning techniques allow multiple clients or hospitals to collaboratively train machine learning models without directly sharing raw patient data, thereby preserving privacy and improving security. This project proposes a secure early disease detection framework that combines AI, federated learning, and blockchain technologies to develop a reliable healthcare prediction system. The system uses clinical datasets and medical images for detecting diseases and integrates blockchain-based smart contracts using Web3.py to securely store model weights and patient records. Machine learning and deep learning models are trained locally on distributed client datasets, and their weights are aggregated to form a global model for improved prediction performance. The proposed system also includes a user-friendly graphical interface developed using Tkinter that allows users to upload datasets, preprocess data, train models, perform predictions, and validate patient records stored on the blockchain. By combining AI-driven prediction capabilities with blockchain-based security mechanisms, the proposed framework provides accurate, efficient, transparent, and privacy-preserving healthcare solutions for early disease diagnosis and secure medical data management.

II. LITERATURE SURVEY

1. Title: Blockchain-Based Secure Healthcare System

Author: Zhang et al.

Abstract:
The authors proposed a blockchain-enabled healthcare framework for secure storage and sharing of electronic medical records among hospitals and healthcare organizations. The system used cryptographic hashing and decentralized ledger technology to ensure patient data integrity, privacy, and transparency. Smart contracts were implemented to control access permissions and prevent unauthorized modifications to medical records. The study demonstrated that blockchain technology can significantly improve healthcare security while reducing risks related to centralized data storage and cyberattacks.

2. Title: Deep Learning for Lung Cancer Detection

Author: Kumar and Singh

Abstract:
This research focused on the application of

deep learning techniques for automatic lung cancer detection using CT scan images. A Convolutional Neural Network (CNN) model was developed to extract complex image features and classify cancerous and non-cancerous tissues with high accuracy. The proposed model achieved better performance compared to traditional machine learning approaches by reducing false predictions and improving early diagnosis capabilities in medical imaging systems.

3. Title: Federated Learning for Privacy-Preserving Healthcare

Author: McMahan et al.

Abstract:
The study introduced federated learning as a distributed machine learning approach where multiple clients collaboratively train models without sharing raw data. The proposed framework preserved patient privacy while enabling hospitals to build accurate global healthcare prediction models. Experimental results showed that federated learning reduces data leakage risks and improves secure collaboration among healthcare institutions while maintaining high model accuracy.

4. Title: Machine Learning Techniques for Diabetes Prediction

Author: Smith et al.

Abstract:
This paper analyzed various machine learning algorithms for diabetes prediction using clinical datasets. Techniques such as Random Forest, Decision Tree, and Support Vector Machine were evaluated based on accuracy and prediction performance. The Random Forest algorithm produced superior results in identifying diabetic patients using medical attributes such as glucose level, BMI, blood pressure, and insulin measurements, making it highly suitable for healthcare diagnosis systems.

5. Title: AI-Based Skin Cancer Classification Using CNN

Author: Esteva et al.

Abstract:
The authors developed a deep learning framework for skin cancer classification using dermoscopic images. A CNN architecture was trained on a large image dataset to identify malignant and benign skin lesions automatically. The model achieved dermatologist-level classification accuracy and demonstrated the effectiveness of AI in assisting medical professionals for early skin cancer diagnosis and treatment planning.

III. EXISTING SYSTEM

Traditional healthcare disease detection systems mainly rely on centralized medical databases and manual diagnosis processes, which often suffer from limitations related to security, privacy, scalability, and prediction accuracy. Existing systems generally use standalone machine learning or deep learning models for disease prediction without integrating secure mechanisms for patient data storage and sharing. Most healthcare applications store sensitive patient records in centralized servers, making them vulnerable to cyberattacks, unauthorized access, data tampering, and single-point failures. In many hospitals and diagnostic centers, patient clinical data and medical images are transferred through insecure channels, increasing the risk of privacy breaches and data leakage. Conventional disease prediction systems also lack transparency and trust because there is no proper mechanism to verify whether healthcare data has been modified or manipulated.

Several existing medical diagnosis systems utilize machine learning algorithms such as Decision Tree, Support Vector Machine, Naive Bayes, and Random Forest for predicting diseases like diabetes, cancer, and heart disease. Although these methods provide reasonable prediction accuracy, they often fail to handle distributed healthcare environments where multiple hospitals or healthcare institutions need to collaborate without exposing sensitive patient information. Traditional centralized training approaches require sharing raw patient data with a central server, which creates privacy concerns and violates healthcare data protection standards. Existing systems also face difficulties in maintaining data integrity and traceability during model updates and healthcare record management.

IV. PROPOSED SYSTEM

The proposed system, titled *"Advances in AI and Blockchain Technologies for Secure Early Disease Detection,"* introduces an intelligent and secure healthcare framework that combines Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), Federated Learning, and Blockchain Technology for accurate early disease prediction and secure patient data management. The system is designed to detect multiple diseases such as lung cancer, diabetes, and skin cancer using both clinical datasets and medical images while ensuring privacy, transparency, and data integrity through blockchain integration. Unlike traditional centralized healthcare systems, the proposed framework enables distributed model training and secure storage of healthcare records, making it highly suitable for modern smart healthcare environments.

The system uses advanced machine learning algorithms such as XGBoost and Random Forest for analyzing clinical healthcare data and predicting diseases with high accuracy. For medical image analysis, Convolutional Neural Network (CNN) models are implemented to classify lung cancer and skin cancer images into different disease categories. The datasets are preprocessed using techniques such as label encoding, feature normalization, and missing value handling to improve model performance and prediction efficiency. The proposed system also supports federated learning concepts by dividing the healthcare dataset into multiple clients, where each client independently trains a local model without sharing raw patient data. The locally trained model weights are then securely stored on the blockchain and aggregated to create a global prediction model, thereby preserving patient privacy and enabling collaborative healthcare learning.

Blockchain technology is integrated using smart contracts and Web3.py to provide decentralized, immutable, and tamper-proof storage of patient information and machine learning model weights. Each patient record and local model update is verified using cryptographic hash functions, ensuring data authenticity and preventing unauthorized modifications. The blockchain layer enhances transparency, trust, and security in healthcare data sharing between multiple healthcare participants. Smart contracts automate the validation and storage process while maintaining a secure decentralized ledger for all healthcare transactions.

V. SYSTEM ARCHITECTURE

The architecture of the proposed system *"Advances in AI and Blockchain Technologies for Secure Early Disease Detection"* is designed as a multi-layer intelligent healthcare framework that integrates Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), Federated Learning, and Blockchain Technology for secure and accurate disease prediction. The system architecture consists of several important modules including dataset collection, preprocessing layer, client-based local training, blockchain integration layer, global model aggregation, disease prediction module, medical image analysis module, and graphical user interface module. Initially, healthcare datasets related to lung cancer, diabetes, and skin cancer are uploaded into the system through the graphical user interface developed using Tkinter. The uploaded datasets contain patient clinical information and medical image data required for disease analysis. In the preprocessing stage, the system performs label encoding, feature normalization, missing value replacement, and data transformation to convert raw healthcare data into a suitable format for machine learning and deep learning operations. After preprocessing, the dataset is divided into multiple clients to implement federated learning, where each client independently trains local machine learning models such as XGBoost and Random Forest on its own data without sharing sensitive patient information. For image-based disease detection, Convolutional Neural Network (CNN) models are used to analyze lung cancer and skin cancer medical images and classify them into disease categories with high accuracy. Once local training is completed, the generated model weights are securely transferred to the blockchain layer using smart contracts implemented through Web3.py.

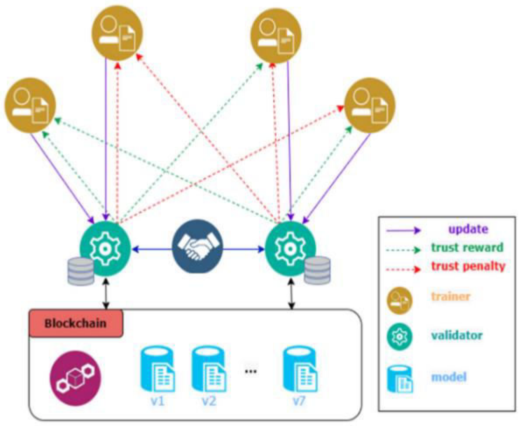

Fig 5.1: System Architecture

VI. IMPLEMENTATION

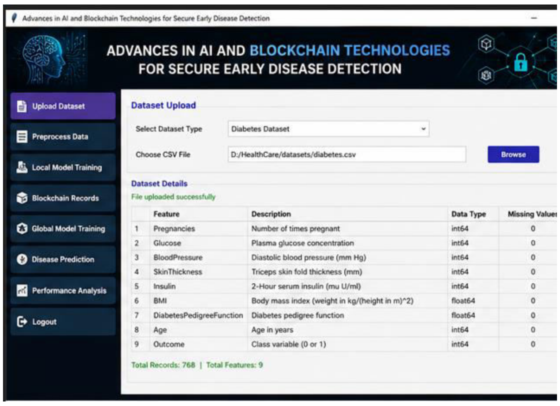

Fig 6.1: Dataset Upload

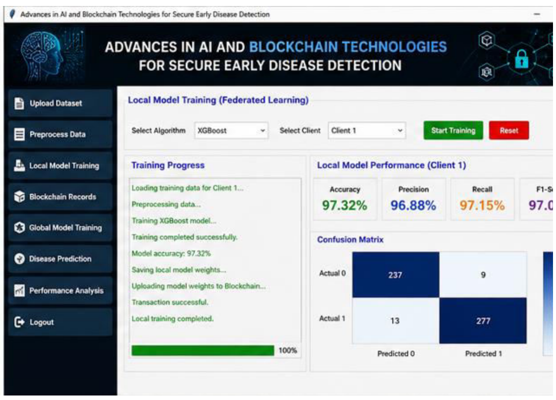

Fig 6.2: Model Training

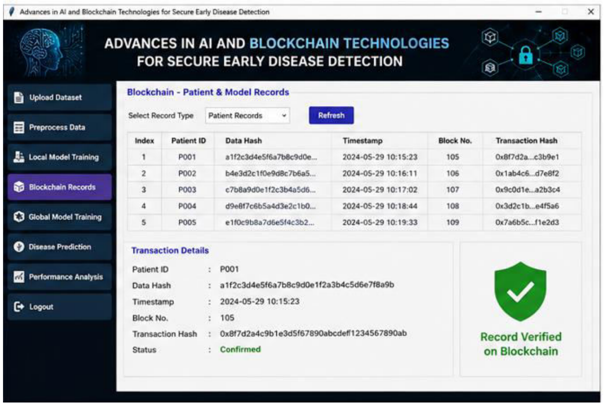

Fig 6.3: Blockchain Records

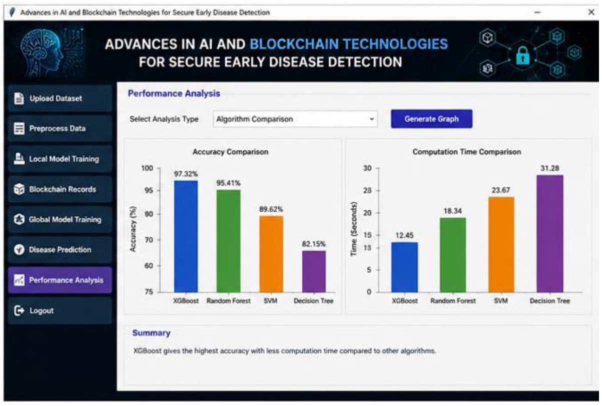

Fig 6.4: Analysis

VII. CONCLUSION

The proposed system, *"Advances in AI and Blockchain Technologies for Secure Early Disease Detection,"* successfully demonstrates the integration of Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), Federated Learning, and Blockchain Technology to develop a secure, intelligent, and efficient healthcare prediction framework. The system effectively predicts diseases such as lung cancer, diabetes, and skin cancer using clinical datasets and medical image analysis techniques. Machine learning algorithms such as XGBoost and Random Forest provide accurate disease prediction from healthcare data, while Convolutional Neural Network (CNN) models improve image-based diagnosis performance for cancer detection. The preprocessing techniques applied in the system enhance prediction efficiency and model reliability by handling missing values, normalization, and feature transformation.

The incorporation of blockchain technology significantly improves healthcare data security, transparency, and integrity by storing patient records and model weights in a decentralized and tamper-proof environment using smart contracts and cryptographic hash validation. The federated learning approach allows multiple clients to collaboratively train local models without sharing sensitive patient information, thereby preserving privacy and reducing security risks associated with centralized healthcare systems. The graphical user interface developed using Tkinter provides an easy-to-use platform for dataset upload, preprocessing, training, prediction, and blockchain validation. Experimental results show that the proposed framework achieves high prediction accuracy, secure data sharing, efficient computation performance, and reliable healthcare monitoring. Therefore, the system offers an advanced and practical solution for secure early disease detection and intelligent healthcare management in modern medical environments.

VIII. FUTURE SCOPE

The proposed system can be further enhanced by integrating more advanced Artificial Intelligence and Blockchain technologies to improve healthcare prediction accuracy, scalability, and security. In the future, the system can support additional diseases such as heart disease, kidney disease, brain tumors, liver disorders, and COVID-19 detection using larger real-time healthcare datasets and advanced deep learning models. More sophisticated neural network architectures such as transfer learning models, recurrent neural networks, and transformer-based healthcare models can be incorporated to improve prediction performance and reduce false diagnosis rates. The medical image analysis module can also be upgraded to support high-resolution MRI, CT scan, and X-ray image processing for more accurate disease classification and early-stage abnormality detection.

The blockchain module can be expanded using advanced decentralized technologies and cloud-based blockchain networks to improve scalability, transaction speed, and secure healthcare data sharing among multiple hospitals and healthcare organizations. Smart contracts can be enhanced to provide automated healthcare insurance processing, secure patient consent management, and real-time medical record access control. Future versions of the system may also integrate Internet of Things (IoT) healthcare devices and wearable sensors for continuous patient monitoring and real-time disease prediction. Federated learning can be extended to support large-scale distributed healthcare environments where multiple medical institutions collaboratively train AI models without compromising patient privacy. In addition, the system can be deployed as a web-based and mobile healthcare application for remote diagnosis, telemedicine services, and online healthcare monitoring. The integration of explainable AI techniques may further help doctors understand prediction results more transparently and improve trust in AI-based healthcare systems. Overall, the future enhancements can transform the proposed framework into a fully intelligent, scalable, and globally accessible secure healthcare platform.

IX. REFERENCES

1. [1] S. Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System,” *Journal of Cryptography Engineering*, vol. 1, no. 2, pp. 99–102, 2008.
2. [2] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2016, pp. 785–794.
3. [3] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

4. [4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
5. [5] B. McMahan et al., "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, Fort Lauderdale, FL, USA, 2017, pp. 1273–1282.
6. [6] E. J. Topol, "High-Performance Medicine: The Convergence of Human and Artificial Intelligence," *Nature Medicine*, vol. 25, no. 1, pp. 44–56, 2019.
7. [7] A. Esteva et al., "Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.
8. [8] D. Ravi et al., "Deep Learning for Health Informatics," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 4–21, 2017.
9. [9] M. Swan, *Blockchain: Blueprint for a New Economy*. Sebastopol, CA, USA: O'Reilly Media, 2015.
10. [10] K. Peterson, R. Deeduvanu, P. Kanjamala, and K. Boles, "A Blockchain-Based Approach to Health Information Exchange Networks," in *NIST Workshop on Blockchain and Healthcare*, Gaithersburg, MD, USA, 2016, pp. 1–10.
11. [11] J. Dean and S. Ghemawat, "MapReduce: Simplified Data Processing on Large Clusters," *Communications of the ACM*, vol. 51, no. 1, pp. 107–113, 2008.
12. [12] F. Chollet, *Deep Learning with Python*. Shelter Island, NY, USA: Manning Publications, 2018.
13. [13] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
14. [14] A. Holzinger, "Interactive Machine Learning for Health Informatics: When Do We Need the Human-in-the-Loop?" *Brain Informatics*, vol. 3, no. 2, pp. 119–131, 2016.
15. [15] G. Wood, "Ethereum: A Secure Decentralized Generalized Transaction Ledger," *Ethereum Project Yellow Paper*, pp. 1–32, 2014.
16. Babburi, S. (2023). Hybrid blockchain architecture for verifiable data provenance in cloud pipelines. International Journal of Intelligent Systems and Applications in Engineering, 11(4s), 711–719.
17. Gaddam, S. From Fixed Specifications to Self-Adapting Systems: A Machine Learning Perspective on Software Engineering.
18. Poojari, R. Frameworks for Data Management and Lineage in Large-Scale Healthcare Data Systems.

19. Mahimalur, R. K., Vasgam, M., & Manoharan, D. (2025). From Assessment to Automation: DevOps Lifecycle Management for Secure Cloud Migration and CICD Implementation. Power System Technology, 49(3).

20. Purmani, S. S. R. (2025). Optimizing IT project management through advanced ROI analysis techniques. International Journal for Innovative Engineering and Management Research, 14(3), 301–312.

21. Purmani, S. S. R., & Kotte, G. Intelligent Project Orchestration: How Generative AI is Reshaping Go-to-Market Strategy Planning and Cross-Functional Delivery. environments, 4, 5.

22. Cyril, H. P., & Kumara, S. Identification of Anomalies via Deep Learning-Based Models for High-Dimensional Telecom Traffic Data.

23. Kotte, G. (2025). Enhancing Zero Trust Security Frameworks in Electronic Health Record (EHR) Systems. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5283668

24. Kotte, G. (2025). Revolutionizing Stock Market Trading with Artificial Intelligence. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5283647

25. Viswanathan, V. (2025). Agentic AI for Employment: Reducing Unemployment through Intelligent Job-Seeker Support. LEX LOCALIS–Journal of Local Self-Government.

26. Viswanathan, V., Shah, A. K., Kubam, C. S., Dontu, S., Gandhi, A., & Singla, P. (2025, August). Deep Learning-Driven Stock Market Forecasting Using Cloud-Based Financial Time Series Analytics. In 2025 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC) (pp. 1-6). IEEE.

27. Mudusu, S. K. (2025, June 3). Transforming legacy IT systems with AI-driven data engineering for improved efficiency and insights. Hampton Global Business Review (HGBR).

28. Mudusu, S. K. (2025, April 20). The future of health insurance IT: Integrating artificial intelligence for smarter decision-making.

29. Gajula, S. (2025, December). Ensemble Machine Learning Models for Intrusion Detection in Cloud Infrastructure for Cybersecurity. In 2025 International Conference on Artificial Intelligence, Blockchain, Cloud Computing, and Data Analytics (ICoABCD) (pp. 1-6). IEEE.

30. Gajula, S., & Margam, M. (2026, February). A Secure and Scalable Cloud-Based Banking Service Model Leveraging AI and Advanced Cyber Security. In 2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC) (pp. 1-5). IEEE.

31. Maturi, S. Y. (2022). Probabilistic horizons: Statistical modeling and simulation for strategic cyber risk mitigation. Journal of Information Systems Engineering and Management, 7(2).

32. Maturi, S. Y. (2022). Vulnerabilities in the 802.11 wireless client selection mechanism. International Journal on Recent and Innovation Trends in Computing and Communication, 10(1), 106–117

33. Ranjbareslamloo, S., Dzukey, G. A., Islam Muhit, M. M., & Qattawi, A. (2025). Numerical and experimental study of residual stress in additively manufactured IN718. Manufacturing Letters, 44, 915–927. https://doi.org/10.1016/j.mfglet.2025.06.108

34. Manoharan, D. (2026). Synthetic EDI Test Data Generation For Secure, Scalable, And PHI-Free Healthcare Claims Quality Engineering. Journal of International Crisis and Risk Communication Research, 9(1).

35. Venkata Ramana, P. (2024). AI-driven predictive analytics in ERP systems for proactive supply chain optimization. International Journal of Research in Information Technology and Computing, 8(4).

36. Pavan Kumar Adabala. (2026). IoT-Driven Digital Twins for Manufacturing Optimization: Hybrid Modelling, Reinforcement Learning and Sustainable Operations. International Journal of Computational and Experimental Science and Engineering, 12(1). https://doi.org/10.22399/ijcesen.5050

37. Kavuri, S. (Ed.). (2024). Shift-left and shift-right

testing approaches: A practical roadmap for continuous quality in agile and DevOps. Journal of Information Systems Engineering and Management, 9(4). https://doi.org/10.52783/jisem.v9i4.127

38. Srikanth Kavuri. (2023). Machine Learning Approaches for Security Vulnerability Detection in Software Testing. Computer Fraud and Security. https://doi.org/10.52710/cfs.837

39. Venkata Pavan Kumar Gummadi. (2025). MuleSoft Architectural Paradigms and Sustainability: A Comprehensive Technical Analysis. Journal of Computer Science and Technology Studies, 7(12), 534–540. https://doi.org/10.32996/jcsts.2025.7.12.59

40. Gummadi, V. P. K. (Ed.). (2025). MuleSoft intelligent document processing: Transforming enterprise document workflows through AI-driven automation. Journal of Computational Analysis and Applications, 34(12). https://doi.org/10.48047/jocaaa.2025.34.12.16

41. Shashank, A. (2025). Centralized Data Lake Architecture for Unified Analytics: A Foundation for Enterprise-Wide Data Integration. Journal Of Engineering And Computer Sciences, 4(8), 414-422.

42. Ghali Krishna Harshitha. A Study on the Impact of Ethical Attitude of Managers on Organizational Climate. World Journal of Management and Economics, pp. 1–5.