



# Hybrid Blockchain and Big Data Framework for Privacy-Preserving Medical Data Sharing

<sup>1</sup> Mr.P.Selvaprasanth

<sup>1</sup>Assistant Professor, Electronics and Communication Engineering, Sethu Institute of Technology, Virudhunagar, India.

<sup>1</sup>Corresponding Author's email: [selvaprasanthapece@sethu.ac.in](mailto:selvaprasanthapece@sethu.ac.in)

**Abstract.** In the healthcare sector, the need for privacy-preserving and secure data sharing is paramount, especially as the volume of medical data continues to grow due to advancements in big data and digital health technologies. To address these challenges, a hybrid blockchain and big data framework offers a promising solution for secure medical data sharing. Blockchain technology provides a decentralized and immutable ledger that ensures the security, transparency, and privacy of medical data. However, traditional blockchain systems face scalability and efficiency issues when handling large-scale datasets, such as electronic health records (EHRs), medical imaging data, and real-time health monitoring data. By integrating big data technologies into the blockchain infrastructure, the hybrid framework can leverage the strengths of both systems to enable efficient and privacy-preserving medical data sharing. This paper proposes a hybrid framework that combines the immutability and security of blockchain with the scalability and processing power of big data analytics. Through the use of smart contracts, healthcare institutions can securely share medical data while maintaining compliance with regulations like HIPAA and GDPR. The framework also incorporates privacy-preserving mechanisms such as zero-knowledge proofs (ZKP) and differential privacy to protect patient data during data sharing. This paper provides a detailed analysis of current challenges in medical data sharing and presents a novel hybrid blockchain and big data solution that enhances the security, privacy, and scalability of healthcare systems.

**Keywords.** Hybrid Blockchain, Big Data, Privacy-Preserving Medical Data Sharing, Smart Contracts, Electronic Health Records (EHR), Zero-Knowledge Proofs, Differential Privacy, Healthcare Data Security, Decentralized Systems, GDPR Compliance.

## 1. INTRODUCTION

The rapid digital transformation of healthcare has led to an unprecedented increase in the volume of medical data generated by patients, healthcare providers, and medical devices. The digitization of electronic health records (EHRs), medical imaging, and real-time health monitoring devices has opened up new opportunities for data-driven healthcare, personalized medicine, and predictive diagnostics. However, the sheer volume of healthcare data, coupled with its sensitive nature, presents significant challenges in terms of privacy, security, and data sharing.

Healthcare organizations are increasingly required to share medical data for research, diagnostics, and collaboration. However, ensuring the privacy and security of patient data during these exchanges is a major concern, especially with the stringent requirements of data privacy regulations such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in Europe. Traditional centralized data storage systems are vulnerable to data breaches, unauthorized access, and cyberattacks, which can compromise patient privacy and lead to legal and financial repercussions.

Blockchain technology has emerged as a potential solution for addressing these privacy and security challenges. Blockchain's decentralized and immutable architecture ensures that medical data is securely stored and shared across multiple entities, providing transparency and traceability while reducing the risk of data



breaches. However, blockchain's limitations in handling large datasets and high throughput, particularly in healthcare settings, have raised concerns about its scalability and efficiency.

To address these challenges, this paper proposes a hybrid blockchain and big data framework for privacy-preserving medical data sharing. By integrating big data technologies with blockchain, the framework can effectively manage the storage, processing, and sharing of large-scale medical datasets. Smart contracts enable automated data sharing agreements, while privacy-preserving techniques such as zero-knowledge proofs (ZKP) and differential privacy ensure that patient data remains secure and private during data exchanges. This framework offers a novel solution for enhancing the security, privacy, and scalability of healthcare data sharing while complying with data privacy regulations.

## 2. LITERATURE SURVEY

The application of blockchain technology in healthcare has been widely explored due to its potential to improve data security, privacy, and traceability. Blockchain's decentralized nature allows healthcare providers, patients, and researchers to share data without relying on a centralized authority, thus reducing the risk of data breaches and ensuring data integrity. In their study, Azaria et al. (2016) introduced MedRec, a blockchain-based system for managing electronic health records (EHRs). MedRec demonstrated how blockchain could provide a secure and transparent infrastructure for EHR sharing while allowing patients to retain control over their data. However, the study also highlighted scalability concerns, as blockchain networks can become slow and inefficient when dealing with large datasets.

Xia et al. (2017) proposed a blockchain-based privacy-preserving healthcare system that leveraged smart contracts to facilitate secure data sharing between healthcare providers. Their study showed that smart contracts could automate the process of granting and revoking access to medical data, reducing the administrative burden on healthcare organizations. However, the study noted that blockchain alone could not efficiently handle the storage and processing of large medical datasets, such as medical images and real-time monitoring data.

The advent of big data technologies has revolutionized the way healthcare organizations store, process, and analyze large datasets. Wang et al. (2018) demonstrated the potential of big data analytics to improve healthcare outcomes by enabling real-time monitoring, predictive analytics, and personalized treatment plans. Big data technologies, such as Apache Hadoop and Apache Spark, have the ability to process massive amounts of data in parallel, making them ideal for handling large-scale healthcare datasets. However, big data systems are typically centralized, which raises concerns about data privacy and security.

To address these challenges, researchers have explored the integration of blockchain with big data technologies. Zhang et al. (2020) proposed a hybrid blockchain-big data architecture for managing healthcare data. The hybrid approach leveraged blockchain for secure data sharing while using big data technologies to handle the processing and storage of large datasets. The study demonstrated the feasibility of combining these two technologies but did not address specific privacy-preserving techniques that could enhance patient data protection.

This paper builds on the existing literature by proposing a hybrid blockchain and big data framework that incorporates privacy-preserving mechanisms such as zero-knowledge proofs (ZKP) and differential privacy. These techniques ensure that patient data remains confidential during data sharing and analysis, while the integration of blockchain and big data enhances the system's scalability and efficiency.

## 3. PROPOSED METHODOLOGY

The proposed Hybrid Blockchain and Big Data Framework is designed to enable secure and privacy-preserving medical data sharing while addressing the scalability challenges of blockchain in handling large datasets. The system integrates blockchain technology with big data analytics and incorporates privacy-preserving techniques such as zero-knowledge proofs (ZKP) and differential privacy.

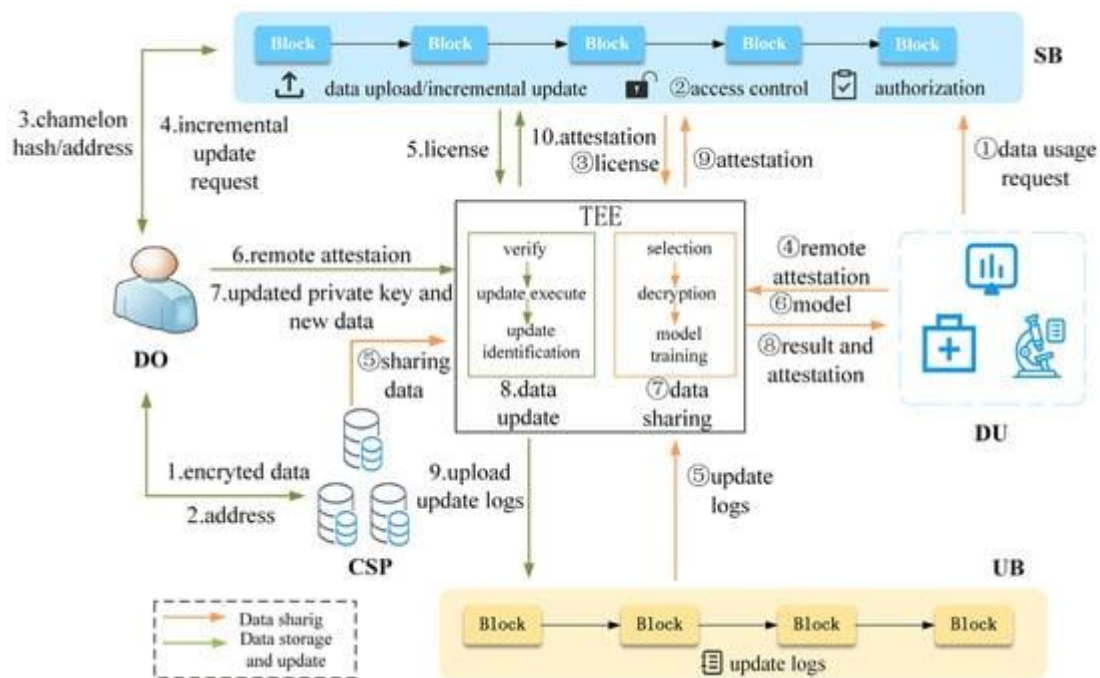


FIGURE 1. A Blockchain-Based Privacy-Preserving Healthcare Data Sharing Scheme

## Blockchain Layer for Data Security and Sharing

The blockchain layer serves as the foundation for securing medical data and enabling decentralized data sharing between healthcare providers, patients, and researchers. Medical data is hashed and stored on the blockchain, ensuring data integrity and immutability. Smart contracts are used to automate data-sharing agreements and enforce access control policies. Each healthcare provider, patient, or researcher has access to the blockchain, allowing them to share data securely and transparently.

## Big Data Layer for Scalable Data Processing

The big data layer handles the storage and processing of large medical datasets, such as EHRs, medical imaging, and real-time health monitoring data. The system utilizes big data technologies such as Apache Hadoop and Apache Spark to process and analyze these large datasets in parallel. This layer ensures that the system can handle high-throughput data processing while maintaining efficiency and scalability.

## Privacy-Preserving Mechanisms

To ensure patient privacy during data sharing, the system incorporates zero-knowledge proofs (ZKP) and differential privacy. ZKP allows one party to prove to another party that they possess certain data without revealing the actual data. This ensures that medical data can be verified without exposing sensitive information. Differential privacy is used to add noise to datasets before sharing them with researchers, ensuring that individual patient data cannot be re-identified.



## Smart Contracts for Data Access Control

Smart contracts are deployed on the blockchain to automate the process of granting and revoking access to medical data. These contracts define the terms and conditions under which data can be accessed, ensuring compliance with privacy regulations such as HIPAA and GDPR. The use of smart contracts reduces administrative overhead and ensures that only authorized parties can access the data.

## 4. CONCLUSION

The proposed Hybrid Blockchain and Big Data Framework offers a secure, scalable, and privacy-preserving solution for medical data sharing in healthcare. By integrating the immutability and security of blockchain with the processing power of big data technologies, the framework addresses the challenges of handling large medical datasets while ensuring compliance with data privacy regulations. The use of privacy-preserving techniques such as zero-knowledge proofs (ZKP) and differential privacy enhances patient data protection during data exchanges, allowing healthcare organizations to securely collaborate and share medical information. This hybrid framework not only improves data security and privacy but also enables real-time data analysis, which can enhance healthcare outcomes and drive advancements in medical research. Future work should focus on optimizing the system for real-time health monitoring and exploring additional privacy-enhancing technologies.

## REFERENCES

1. Lopez, S., Sarada, V., Praveen, R. V. S., Pandey, A., Khuntia, M., & Haralayya, D. B. (2024). Artificial intelligence challenges and role for sustainable education in india: Problems and prospects. *Sandeep Lopez, Vani Sarada, RVS Praveen, Anita Pandey, Monalisa Khuntia, BhadrappaHaralayya (2024) Artificial Intelligence Challenges and Role for Sustainable Education in India: Problems and Prospects. Library Progress International*, 44(3), 18261-18271.
2. Kumar, N., Kurkute, S. L., Kalpana, V., Karuppannan, A., Praveen, R. V. S., & Mishra, S. (2024, August). Modelling and Evaluation of Li-ion Battery Performance Based on the Electric Vehicle Tiled Tests using Kalman Filter-GBDT Approach. In *2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS)* (pp. 1-6). IEEE.
3. Sharma, S., Vij, S., Praveen, R. V. S., Srinivasan, S., Yadav, D. K., & VS, R. K. (2024, October). Stress Prediction in Higher Education Students Using Psychometric Assessments and AOA-CNN-XGBoost Models. In *2024 4th International Conference on Sustainable Expert Systems (ICSES)* (pp. 1631-1636). IEEE.
4. Yamuna, V., Praveen, R. V. S., Sathya, R., Dhivva, M., Lidiya, R., & Sowmiya, P. (2024, October). Integrating AI for Improved Brain Tumor Detection and Classification. In *2024 4th International Conference on Sustainable Expert Systems (ICSES)* (pp. 1603-1609). IEEE.
5. Anuprathibha, T., Praveen, R. V. S., Jayanth, H., Sukumar, P., Suganthi, G., & Ravichandran, T. (2024, October). Enhancing Fake Review Detection: A Hierarchical Graph Attention Network Approach Using Text and Ratings. In *2024 Global Conference on Communications and Information Technologies (GCCIT)* (pp. 1-5). IEEE.
6. Praveen, R. V. S. (2024). *Data Engineering for Modern Applications*. Addition Publishing House.
7. Dhivya, R., Sagili, S. R., Praveen, R. V. S., VamsiLala, P. N. V., Sangeetha, A., & Suchithra, B. (2024, December). Predictive Modelling of Osteoporosis using Machine Learning Algorithms. In *2024 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS)* (pp. 997-1002). IEEE.
8. Kemmannu, P. K., Praveen, R. V. S., Saravanan, B., Amshavalli, M., & Banupriya, V. (2024, December). Enhancing Sustainable Agriculture Through Smart Architecture: An Adaptive Neuro-Fuzzy Inference System with XGBoost Model. In *2024 International Conference on Sustainable Communication Networks and Application (ICSCNA)* (pp. 724-730). IEEE.



9. Praveen, R. V. S., Raju, A., Anjana, P., & Shibi, B. (2024, October). IoT and ML for Real-Time Vehicle Accident Detection Using Adaptive Random Forest. In *2024 Global Conference on Communications and Information Technologies (GCCIT)* (pp. 1-5). IEEE.
10. Praveen, R. V. S., Hemavathi, U., Sathya, R., Siddiq, A. A., Sanjay, M. G., & Gowdish, S. (2024, October). AI Powered Plant Identification and Plant Disease Classification System. In *2024 4th International Conference on Sustainable Expert Systems (ICSES)* (pp. 1610-1616). IEEE.
11. Thamilarasi, V., & Roselin, R. (2021, February). Automatic classification and accuracy by deep learning using cnn methods in lung chest X-ray images. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1055, No. 1, p. 012099). IOP Publishing.
12. Thamilarasi, V., & Roselin, R. (2019). Lung segmentation in chest X-ray images using Canny with morphology and thresholding techniques. *Int. j. adv. innov. res.*, 6(1), 1-7.
13. Thamilarasi, V., & Roselin, R. (2019). Automatic thresholding for segmentation in chest X-ray images based on green channel using mean and standard deviation. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 8(8), 695-699.
14. Thamilarasi, V., & Roselin, R. (2021). U-NET: convolution neural network for lung image segmentation and classification in chest X-ray images. *INFOCOMP: Journal of Computer Science*, 20(1), 101-108.
15. Asaithambi, A., & Thamilarasi, V. (2023, March). Classification of Lung Chest X-Ray Images Using Deep Learning with Efficient Optimizers. In *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 0465-0469). IEEE.
16. Jadhav, S., Machale, A., Mharnur, P., Munot, P., & Math, S. (2019, September). Text based stress detection techniques analysis using social media. In *2019 5th International Conference On Computing, Communication, Control And Automation (ICCUBE)* (pp. 1-5). IEEE.
17. Anitha, C., Tellur, A., Rao, K. B., Kumbhar, V., Gopi, T., Jadhav, S., & Vidhya, R. G. (2024). Enhancing Cyber-Physical Systems Dependability through Integrated CPS-IoT Monitoring. *International Research Journal of Multidisciplinary Scope*, 5(2), 706-713.
18. Kiran, A., Sonker, A., Jadhav, S., Jadhav, M. M., Naga Ramesh, J. V., & Muniyandy, E. (2024). Secure Communications with THz Reconfigurable Intelligent Surfaces and Deep Learning in 6G Systems. *Wireless Personal Communications*, 1-17.
19. Thepade, D. S., Mandal, P. R., & Jadhav, S. (2015). Performance Comparison of Novel Iris Recognition Techniques Using Partial Energies of Transformed Iris Images and Energy Compaction With Hybrid Wavelet Transforms. In *Annual IEEE India Conference (INDICON)*.
20. Vandana, C. P., Basha, S. A., Madijagan, M., Jadhav, S., Matheen, M. A., & Maguluri, L. P. (2024). IoT resource discovery based on multi faected attribute enriched CoAP: smart office seating discovery. *Wireless Personal Communications*, 1-18.
21. Jadhav, S., Durairaj, M., Reenadevi, R., Subbulakshmi, R., Gupta, V., & Ramesh, J. V. N. (2024). Spatiotemporal data fusion and deep learning for remote sensing-based sustainable urban planning. *International Journal of System Assurance Engineering and Management*, 1-9.
22. Jadhav, S., Chaudhari, V., Barhate, P., Deshmukh, K., & Agrawal, T. (2021). Extreme Gradient Boosting for Predicting Stock Price Direction in Context of Indian Equity Markets. In *Intelligent Sustainable Systems: Selected Papers of WorldS4 2021, Volume 2* (pp. 321-330). Singapore: Springer Nature Singapore.
23. Jadhav, S., Chaudhari, V., Barhate, P., Deshmukh, K., & Agrawal, T. (2021). REVIEW PAPER ON: ALGORITHMIC TRADING USING ARTIFICIAL INTELLIGENCE.
24. Jadhav, S., Chaudhari, V., Barhate, P., Deshmukh, K., & Agrawal, T. (2021). in Context of Indian Equity Markets. *Intelligent Sustainable Systems: Selected Papers of WorldS4 2021, Volume 2*, 334, 321.
25. Jadhav, S. R., Bishnoi, A., Safarova, N., Khan, F., Aurangzeb, K., & Alhussein, M. (2024). Dual-Attention Based Multi-Path Approach for Intensifying Stock Market Forecasting. *Fluctuation and Noise Letters*, 23(02), 2440009.
26. Vishwanath, B., & Vaddepalli, S. (2023). The future of work: Implications of artificial intelligence on hr practices. *Tuijin Jishu/Journal of Propulsion Technology*, 44(3), 1711-1724.
27. Surendar Vaddepalli, D. B. V. (2025). ENTREPRENEURIAL ECOSYSTEMS IN THE GCC- ASSESSING SUPPORT SYSTEMS FOR WOMEN AND DISABLED ENTREPRENEURS IN OMAN. *Machine Intelligence Research*, 19(1), 126-143.





28. Vaddepalli, S., & Vishwanath, B. (2024). MERGERS AND ACQUISITIONS: DRIVERS, CHALLENGES, AND PERFORMANCE OUTCOMES IN GCC NATIONS. *International Journal of Central Banking*, 20(1), 298-310.
29. Sangam, V. G., Priyadarshini, S. H., Anand, N., Prathibha, P., Purohit, P., & Nalamitha, R. (2021, June). Early Detection of Diabetic Foot Ulcer. In *Journal of Physics: Conference Series* (Vol. 1937, No. 1, p. 012049). IOP Publishing.
30. Kumar, C. R., Vijayalakshmi, B., Priyadarshini, S. H., Sikdar, S., Bhat, S. N., & Neelam, M. (2020). Standing wheelchair with voice recognition system. *J. Crit. Rev*, 7, 2042-2047.
31. Priyadarshini, S. H., Dutt, D. N., & Rajan, A. P. (2019). Nonlinear Processing of Wrist Pulse Signals to Distinguish Diabetic and Non-Diabetic Subjects. *Int. J. Eng. Adv. Technol.*, 9(1), 7105-7110.
32. Priyadarshini, S. H., Poojitha, S., Vinay, K. V., & VA, A. D. (2023, October). AQUASENSE: Sensor Based Water Quality Monitoring Device. In *2023 International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)* (pp. 1786-1789). IEEE.
33. Padma, C. R., Priyadarshini, S. H., Nanditha, H. G., Pavithra, G., & Manjunath, T. C. (2022, August). Design & Development of micro-controlled system using VHDL with the help of UART Tx & Rx. In *2022 2nd Asian Conference on Innovation in Technology (ASIANCON)* (pp. 1-11). IEEE.
34. Rao, M. R., Mangu, B., & Kanth, K. S. (2007, December). Space vector pulse width modulation control of induction motor. In *IET-UK International Conference on Information and Communication Technology in Electrical Sciences (ICTES 2007)* (pp. 349-354). Stevenage UK: IET.
35. Rao, M. R., & Prasad, P. V. N. (2014). Modelling and Implementation of Sliding Mode Controller for PMBDC Motor Drive. *International journal of advanced research in electrical, electronics and instrumentation engineering*, 3(6).
36. Sameera, K., & MVR, S. A. R. (2014). Improved power factor and reduction of harmonics by using dual boost converter for PMBLDC motor drive. *Int J Electr Electron Eng Res*, 4(5), 43-51.
37. Srinivasu, B., Prasad, P. V. N., & Rao, M. R. (2006, December). Adaptive controller design for permanent magnet linear synchronous motor control system. In *2006 International Conference on Power Electronic, Drives and Energy Systems* (pp. 1-6). IEEE.
38. Al-Ghanimi, M. G., Hanif, O., Jain, M. V., Kumar, A. S., Rao, R., Kavim, R., ... & Hossain, M. A. (2022, December). Two TS-Fuzzy Controllers based Direct Torque Control of 5-Phase Induction Motor. In *2022 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)* (pp. 1-6). IEEE.
39. Prathap, P. B., & Saara, K. (2024). Quantifying efficacy of the fiberbragg grating sensors in medical applications: a survey. *Journal of Optics*, 53(5), 4180-4201.
40. Kumar, T. V. (2024). A Comprehensive Empirical Study Determining Practitioners' Views on Docker Development Difficulties: Stack Overflow Analysis.
41. Kumar, T. V. (2024). A New Framework and Performance Assessment Method for Distributed Deep Neural NetworkBased Middleware for Cyberattack Detection in the Smart IoT Ecosystem.
42. Arora, P., & Bhardwaj, S. (2017). A Very Safe and Effective Way to Protect Privacy in Cloud Data Storage Configurations.
43. Arora, P., & Bhardwaj, S. (2017). Combining Internet of Things and Wireless Sensor Networks: A Security-based and Hierarchical Approach.
44. Arora, P., & Bhardwaj, S. (2017). Enhancing Security using Knowledge Discovery and Data Mining Methods in Cloud Computing.
45. Yendluri, D. K., Ponnala, J., Tatikonda, R., Kempanna, M., Thatikonda, R., & Bhuvanesh, A. (2023, November). Role of rpa& ai in optimizing network field services. In *2023 7th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)* (pp. 1-6). IEEE.
46. Yendluri, D. K., Ponnala, J., Thatikonda, R., Kempanna, M., Tatikonda, R., & Bhuvanesh, A. (2023, November). Impact of Robotic Process Automation on Enterprise Resource Planning Systems. In *2023 International Conference on the Confluence of Advancements in Robotics, Vision and Interdisciplinary Technology Management (IC-RVITM)* (pp. 1-6). IEEE.
47. Sidharth, S. (2021). MULTI-CLOUD ENVIRONMENTS: MITIGATING SECURITY RISKS IN DISTRIBUTED ARCHITECTURES.
48. Sidharth, S. (2020). THE GROWING THREAT OF DEEPAKES: IMPLICATIONS FOR SECURITY AND PRIVACY.



49. Sidharth, S. (2019). QUANTUM-ENHANCED ENCRYPTION TECHNIQUES FOR CLOUD DATA PROTECTION.
50. Sidharth, S. (2019). SECURING CLOUD-NATIVE MICROSERVICES WITH SERVICE MESH TECHNOLOGIES.
51. Sidharth, S. (2019). DATA LOSS PREVENTION (DLP) STRATEGIES IN CLOUD-HOSTED APPLICATIONS.
52. Sidharth, S. (2018). RANSOMWARE TRENDS AND EFFECTIVE MITIGATION TECHNIQUES IN 2018.
53. Sidharth, S. (2018). POST-QUANTUM CRYPTOGRAPHY: PREPARING FOR A QUANTUM COMPUTING ERA.
54. Sidharth, S. (2017). CYBERSECURITY STRATEGIES FOR IOT DEVICES IN SMART CITIES.
55. Sidharth, S. (2017). ACCESS CONTROL MODELS FOR SECURE HYBRID CLOUD DEPLOYMENT.
56. Sidharth, S. (2017). MACHINE LEARNING ALGORITHMS FOR REALTIME MALWARE DETECTION.
57. Ara, T., Ambareen, J., Venkatesan, S., Geetha, M., & Bhuvanesh, A. (2024). An energy efficient selection of cluster head and disease prediction in IoT based smart agriculture using a hybrid artificial neural network model. *Measurement: Sensors*, 32, 101074.
58. Divyashree, H. S., Avinash, N., Manjunatha, B. N., Vishesh, J., & Mamatha, M. (2024). Enhancing secrecy using hybrid elliptic curve cryptography and Diffie Hellman key exchange approach and Young's double slit experiment optimizer based optimized cross layer in multihop wireless network. *Measurement: Sensors*, 31, 100967.
59. NR, D., GK, D. S., & Kumar Pareek, D. P. (2022, February). A Framework for Food recognition and predicting its Nutritional value through Convolution neural network. In *Proceedings of the International Conference on Innovative Computing & Communication (ICICC)*.
60. Prasath, D. S., & Selvakumar, A. (2015). A Novel Iris Image Retrieval with Boundary Based Feature Using Manhattan Distance Classifier. *International Journal Of Innovative Technology And Creative Engineering (Issn: 2045-8711) Vol, 5*.
61. Nirmala, K., & Prasath, S. (2020). Probabilistic mceliece public-key cryptography based identity authentication for secured communication in VANET. *Solid State Technology*, 63(6), 10167-10182.
62. Sivasankaran, P., & Dhanaraj, K. R. (2024). Lung Cancer Detection Using Image Processing Technique Through Deep Learning Algorithm. *Revue d'IntelligenceArtificielle*, 38(1).
63. Pannirselvam, S., & Prasath, S. (2015). A Novel Technique for Face Recognition and Retrieval using Fiducial Point Features. *Procedia Computer Science*, 47, 301-310.
64. Tamilselvi, R., Mohanasathiya, K. S., & Prasath, S. (2024). Developed a Smooth Support Vector Machine to Predict the Crop Production in Alluvial Soil and Red Soil Regions of Tamil Nadu India [J]. *Naturalista Campano*, 28(1), 279-297.