ChainMedIQ: Diagnostics Powered by ML and Blockchain

[1]Sanket Koli, [2]Pratik Nagtilak, [3]Shridhar Marda,[4]Mohammed Sahbi Inamdar

[1,2,3,4]UG Students, Department of Computer Science and Engineering,
SVERI's College of Engineering Pandharpur,
Maharashtra India

[5]S. M. Shinde

*[5]Assistant Professor, Department of Computer Science and Engineering,
SVERI's College of Engineering Pandharpur,
Maharashtra India*

ABSTRACT

ChainMedIQ is an innovative web-based application that merges machine learning (ML) with blockchain technology to revolutionise diagnostics in healthcare. Focused on predicting heart disease, the ML model is trained on extensive patient records, providing accurate and reliable predictions. Blockchain technology ensures that patient data is stored securely, transparently, and immutably, allowing patients to retain control over their health information. By decentralising data storage, ChainMedIQ eliminates single points of failure, enhancing the reliability and availability of diagnostic services. The platform fosters trust between patients and healthcare providers by ensuring data integrity and security. ChainMedIQ not only aids clinicians in making data-driven decisions but also empowers patients with secure, decentralised data management, leading to improved patient outcomes and a more resilient healthcare system.

Keywords:

Machine Learning (ML), Blockchain Technology, Heart Disease Prediction, Healthcare Diagnostics, Data Security, Decentralised Data Management, Patient Empowerment Transparency , Predictive Analytics Scalability .

I. INTRODUCTION

ChainMedIQ is an innovative platform that merges the strengths of machine learning (ML) and blockchain technology to advance healthcare diagnostics, with a particular focus on heart disease. As heart disease remains a leading cause of death worldwide, accurate and timely diagnosis is crucial. ChainMedIQ addresses this need by employing a sophisticated ML model trained on a vast dataset of patient records to predict the likelihood of heart disease with high accuracy. Blockchain technology is integral to ChainMedIQ's approach, ensuring that patient data is stored in a secure, decentralised, and immutable manner. This not only protects sensitive health information from unauthorized access but also gives patients control over their data, fostering a more transparent and trust-based relationship between patients and healthcare providers. The decentralized nature of blockchain also eliminates the risk of single points of failure, ensuring the continuous availability and reliability of diagnostic services.[4] ChainMedIQ represents a significant step forward in healthcare, offering a powerful tool for clinicians to make data-driven decisions while empowering patients with secure, transparent data management. By combining ML's predictive capabilities with blockchain's security and transparency, ChainMedIQ aims to improve patient outcomes and contribute to a more resilient and trustworthy healthcare system.

II. LITERATURE SURVEY

2.1 Existing model The first literature source, Rajkomar, A., Dean, J., and Kohane, I. (2018). "Machine Learning in Medicine," highlights key aspects of machine learning applications in healthcare. Here's a breakdown of the existing model in key points:

1. Promise of ML in Healthcare: Machine learning has the potential to revolutionize healthcare by analyzing vast datasets to improve diagnosis, treatment planning, and patient outcomes. [2]

2. Data Utilization: Existing models focus on processing massive amounts of pa tient data, including medical records, imaging, and genomic data, for more precise decision-making. [3]

3. Predictive Analytics: ML models in healthcare are often used for predictive analytics, helping clinicians forecast patient risks, potential diseases, and treatment responses. [1]

4. Risk Stratification: These models assist in identifying high-risk patients, particularlchroni

5.Algorithm Types: The study highlights algorithms such as Support Vector Machines (SVMs), Neural Networks, and Decision Trees as commonly used techniques in healthcare ML models. [7]

6. Clinical Validation Needs: One of the challenges in existing models is the need for clinical validation to ensure that predictions translate effectively into real-world medical practice.

7. Data Quality: The performance of ML models heavily depends on the quality of the data, requiring clean, standardized, and comprehensive datasets.

III. PROBLEM DEFINITION

The healthcare system faces significant challenges, particularly in diagnosing and managing heart disease. Traditional diagnostic methods are often inaccurate, leading to delayed interventions. Centralized patient data storage raises concerns about privacy, security, and inefficiency, while patients lack control over their medical records, limiting transparency and trust. Interoperability issues between healthcare systems hinder effective collaboration. Furthermore, the limited adoption of advanced technologies like Machine Learning (ML) and Blockchain restricts diagnostic accuracy and data security. Delays in accessing real-time data and regulatory hurdles, including ethical biases and compliance with privacy laws, exacerbate these problems.

IV. OBJECTIVE

1. To develop a machine learning-based diagnostic system.

2. To integrate blockchain technology.

3. To create a synergistic framework.

4. To design a user-friendly mobile application.

5. To provide real-time access to medical services.

6. To ensure data interoperability and security.

7. To optimize the performance of machine learning models.

8. To explore the potential of personalized medicine.

V. ALGORITHMS

1.Supervised Learning Algorithms (for Disease Prediction):

Random Forest: Used for classification of patients based on their medical history and test results. It builds multiple decision trees and outputs the most accurate prediction for heart disease.

Support Vector Machine (SVM): Utilized to create hyperplanes that differentiate between healthy and at-risk patients based on input data features.

Neural Networks (NN): A deep learning approach that processes complex medical data, identifying patterns and making diagnostic predictions. The network is trained with labeled data to ensure high accuracy in heart disease detection.

2.Blockchain (for Security and Transparency):

Distributed Ledger Technology (DLT): Blockchain is used to maintain a de centralized ledger of patient records, ensuring that no single party controls the data. All transactions (data access, modifications, etc.) are recorded and cannot be altered.

Smart Contracts: Automated contracts are coded to enforce the secure sharing of patient data between doctors and healthcare providers. For example, a smart contract automatically allows a doctor to access patient data once consent is provided by the patient.

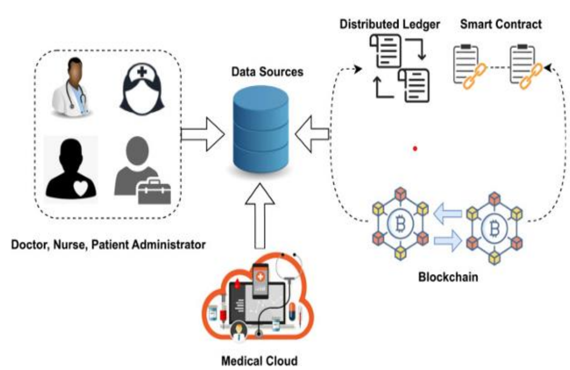

Fig1: Block Diagram

VI. PROPOSED METHODOLOGY

The development of the ChainMedIQ project follows a structured approach that con sists of multiple phases to ensure the successful design, development, and deploy ment of the system. The project planning involved the following key steps:

1. Requirement Gathering and Analysis: This phase involved a detailed analysis of the healthcare domain to understand the need for a predictive diagnostic system. Meetings with stakeholders, including doc tors, healthcare administrators, and patients, were held to gather comprehensive data on the current challenges in heart disease diagnosis and data security. Based on the analysis, system requirements were defined.[2][5]

2. Feasibility Study: A technical and economic feasibility study was conducted to evaluate the practicality of using Machine Learning (ML) and Blockchain technologies in healthcare diagnostics. This study ensured that the solution would be viable, scalable, and cost effective for both hospitals and individual users.[7]

3. System Design: In this phase, both high-level and detailed system architecture designs were created. The design focused on the integration of ML models for disease prediction and blockchain for secure data management[3].Various modules, such as the user interface, blockchain, and cloud storage, were designed to interact efficiently within the system.

4. Prototype Development: A basic prototype was developed to test the core functionalities of the system. This included testing the blockchain's ability to securely store and manage medical records and ML's ability to predict heart disease based on historical data[4].

5. ML Model Training and Testing: In this phase, machine learning models were trained using large datasets related to heart disease. The models were fine-tuned to improve diagnostic accuracy. Multiple algorithms were tested to identify the most accurate one for predicting heart conditions.

6.Blockchain Integration: The blockchain infrastructure was integrated with the cloud storage and the ML di agnostic system. This phase involved developing smart contracts for automating data access, patient consent management, and securing medical transactions in a decetralized manner[6].

7. System Implementation: After successful integration, the full system was implemented, including the user interface for doctors, patients, and administrators. The system was deployed in a controlled environment for initial testing and evaluation.[3]

8. Testing and Validation: Extensive testing was performed to ensure the system met the required standards for accuracy, security, and usability. Multiple test cases were developed to validate both the diagnostic performance and the blockchain security mechanisms.

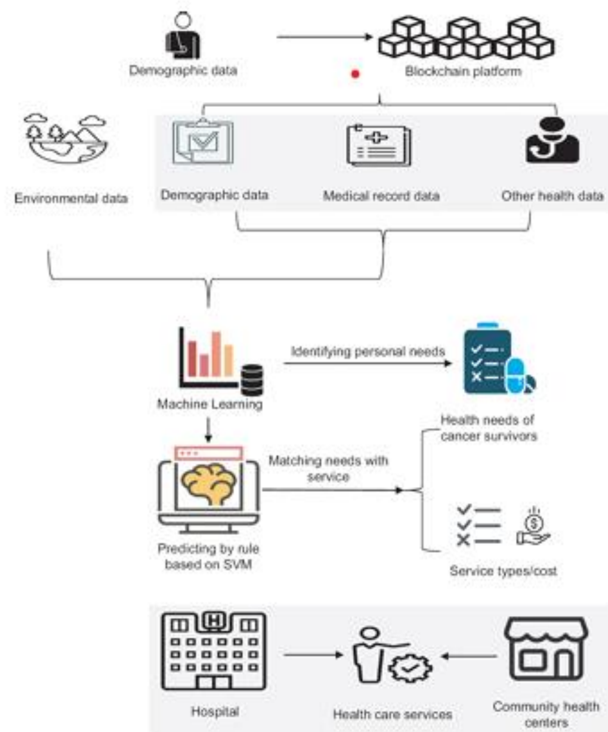

Fig2: WorkFlow

VII. RESULTS AND DISCUSSION

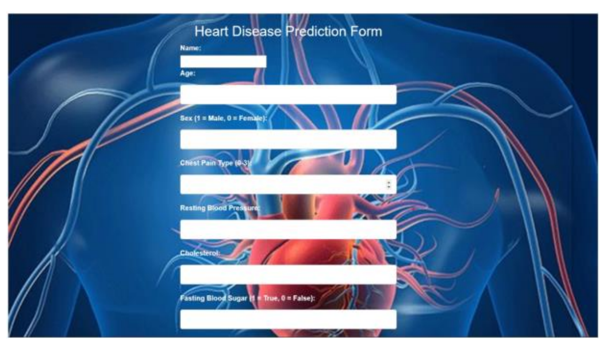

Fig3: Form

1. age- age in years

2. sex- (1 = male; 0 = female)

3. cp- chest pain type

4. trestbps-resting blood pressure (in mm Hg on admission to the hospital)

5. chol- serum cholestoral in mg/dl

6. fbs- (fasting blood sugar ¿ 120 mg/dl) (1 = true; 0 = false)

7. restecg- resting electrocardiographic results

8. thalach- maximum heart rate achieved

9. exang- exercise induced angina (1 = yes; 0 = no)

10. oldpeak- ST depression induced by exercise relative to rest looks at stress of heart during excercise unhealthy heart will stress more

11. slope- the slope of the peak exercise ST segment

12. ca- number of major vessels (0-3) colored by flourosopy colored vessel means the doctor can see the blood passing through the more blood movement the better (no clots)

13. thal- thalium stress result

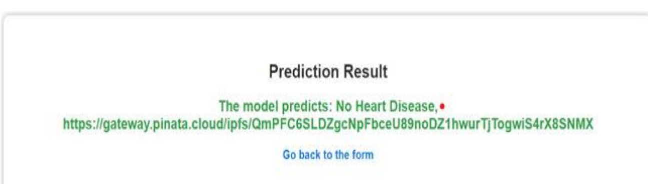

Fig4: Output

It predicts about the presence or absense of heart disease and also gives the connected link of blockchain where all the data of patient is stored and which is immutable.

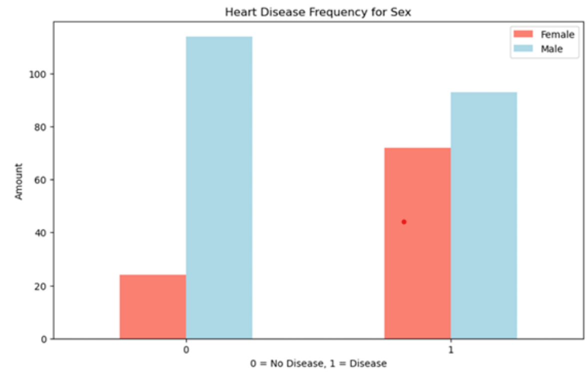

Fig5: Number of Males and Females

The above graph indicates the frequency of the males and females having disease and not having disease respectively

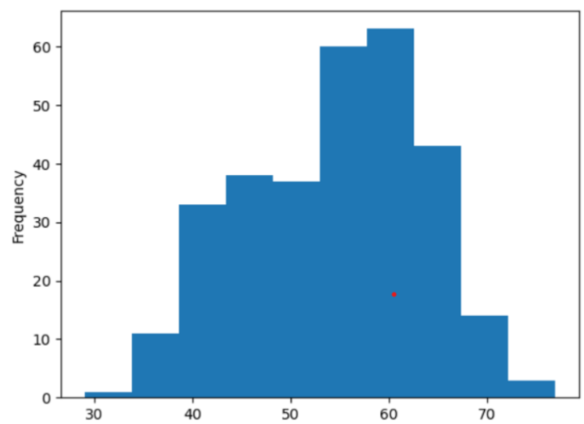

Fig6: Histogram of frequency

The histogram depicts the frequency distribution of patient ages in the dataset. The majority of patients fall within the age range of 50 to 60 years, indicating that middle-aged and older individuals are more common in this heart disease dataset. There is a gradual increase in frequency from age 40, peaking around the 55-60 age group, and then tapering off after age 65. This distribution helps in identifying the target age group most affected by heart conditions, which can assist in focusing diagnostic and preventive efforts.

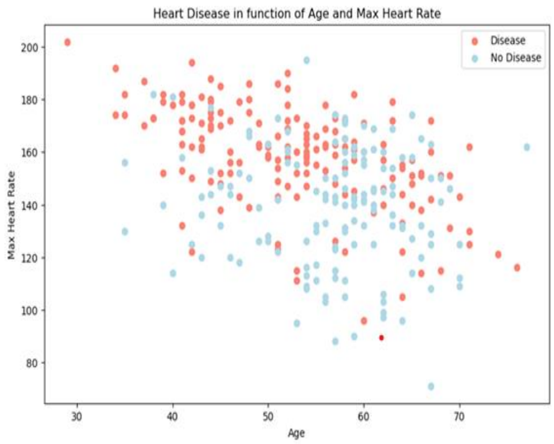

Fig7: Heart Rates

The scatter plot illustrates the relationship between age and maximum heart rate in individuals with and without heart disease. Red dots represent patients diagnosed with heart disease, while blue dots indicate those without. The plot shows that people with heart disease generally have a lower maximum heart rate across different age groups, compared to those without heart disease. The trend suggests that as age increases, the maximum heart rate tends to decrease, with heart disease patients having more restricted heart rates. This visualization helps in understanding how heart rate metrics correlate with heart disease diagnosis.

The bar chart represents the accuracy of three machine learning algorithms used in the project: K-Nearest Neighbors (KNN), Logistic Regression, and Random For est. Among the models, Logistic Regression shows the highest accuracy, followed by Random Forest, while KNN has the lowest. This comparison highlights the su perior performance of Logistic Regression for the dataset used in this study, making it the most reliable model for disease prediction. The difference in accuracy show cases the effectiveness of different algorithms in handling medical data and making diagnostic predictions.

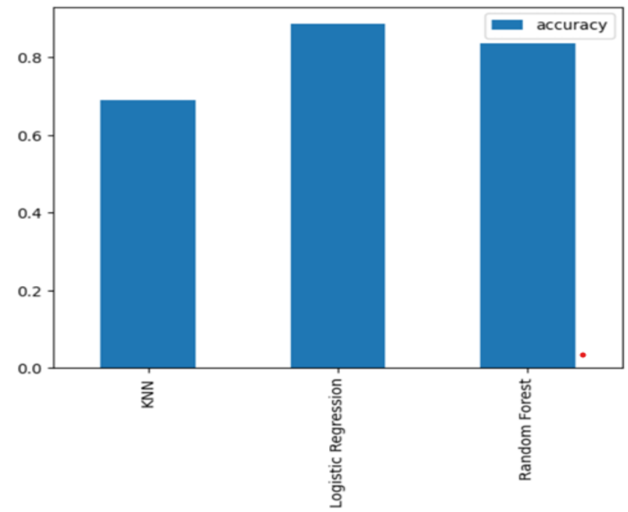

Fig8: Model Accuracy

VIII . CONCLUSION

ChainMedIQ represents a significant advancement in the realm of healthcare diagnostics by combining the strengths of machine learning (ML) and blockchain technology. By leveraging ML, the application aims to deliver accurate and reliable pre dictions of heart disease, enhancing clinicians' ability to make informed decisions and improving patient outcomes. Blockchain technology complements this by providing a secure, decentralized, and transparent framework for managing patient data, ensuring its integrity and confidentiality. The integration of these technologies not only addresses critical challenges in health care, such as data security and diagnostic accuracy but also empowers patients by giving them control over their health information. Chain MedIQ's approach fosters greater trust and collaboration between patients and healthcare providers, ultimately leading to a more efficient and resilient healthcare system. While the project faces challenges such as integration complexity, regulatory compliance, and scalability, it holds the promise of transforming diagnostic processes and setting a precedent for future technological innovations in healthcare. Through continued development and refinement, Chain MedIQ has the potential to significantly improve how heart disease is diagnosed and managed, paving the way for more se cure and accurate healthcare solutions.

XI.FUTURE SCOPE

1. Advanced AI for Multi-Disease Prediction: Expanding the system to predict a wider range of diseases beyond heart conditions, using more advanced machine learning models

and incorporating more diverse med ical data for comprehensive diagnostics.

2. Integration with IoT Devices: Connecting the system to wearable devices and IoT sensors for real-time health mon itoring, allowing continuous data collection and more accurate, personalized predic tions for patients.

3. Interoperability with Other Healthcare Systems: Enhancing the system to be interoperable with existing healthcare management plat forms, electronic health records (EHR), and telemedicine services to provide seam less data sharing and improved healthcare workflows across organizations.

X.REFERENCES

1.Rajkomar, A., Dean, J., and Kohane , I. (2018). *Machine Learning in Medicine.* New England Journal of Medicine, 380(14), 1347-1358. DOI: 10.1056/NEJMra1614398

2.Deo, R. C. (2015). *Machine Learning in Medicine.* Circulation, 132(20), 1920-1930. DOI:10.1161/CIRCULATIONAHA.115.015694

3.Agbo, C. C., Mahmoudi, M., and Eklund, J. M. (2019). *Blockchain Technology for Healthcare: A Comprehensive Review and Directions for Future Research.* Applied Sciences, 9(13), 2731. DOI: 10.3390/app9132731

4.Mamoshina, P., Kochetov, I., and Putin, E. (2018). *The Role of Blockchain Technology in Healthcare: A Systematic Review.* Journal of Biomedical Informatics, 87, 69-79. DOI: 10.1016/j.jbi.2018.09.014

5.Xia, Q., Wu, Y., and Zhang, L. (2017). *Blockchain-Based Data Management for Secure and Reliable Healthcare Services.* IEEE Access, 5, 13715-13724. DOI: 10.1109/ACCESS.2017.2723798

6.Gaur, P., and Kaur, N. (2019). *Blockchain and Machine Learning: A Synergistic Approach for Healthcare Services.* Procedia Computer Science, 152, 45-52. DOI: 10.1016/j.procs.2019.05.001

7.Dey, N., Ashour, A. S., and Tiwari, P. (2016). *Machine Learning Approaches for Heart Disease Prediction.* Springer. ISBN: 978-3-319-33119-8

8.Coudray, N., Ocampo, P. S., and Sakellaropoulos, T. (2018). *Classification and Prediction of Heart Disease Using Machine Learning Techniques.* Journal of Healthcare Engineering, 2018. DOI: 10.1155/2018/5706573