APPLICATION OF DATA MINING IN HEALTH CARE

Yabin Paul Thomas[*1], Sobin Thomas Sony[*2], Alan Sabu[*3], Abhijith G Nair[*4]

[*1,2,3,4]Third Year BCA Student, PG Department Of Computer Applications And Artificial Intelligence, Saintgits College Of Applied Sciences Kottayam, Kerala, India.

ABSTRACT

In In today's rapidly evolving technological landscape, organizations across various fields are increasingly relying on data mining technology. Notably, the healthcare sector, vital for everyone's well-being, has witnessed significant advancements. Technologies for examining physical conditions and detecting diseases have become instrumental in preventive healthcare. Awareness about symptoms empowers individuals to take proactive measures, preventing the onset of certain diseases. Given the vast amount of medical data available, efficient methods for data collection from diverse databases are essential. Modern data mining technologies play a pivotal role in this context. This study explores the specific applications of data mining in the medical field, emphasizing its relevance and success in healthcare due to its high accuracy in prediction and results.

Keywords: Data Mining Technology, Healthcare Sector, Examining Physical Conditions, Detecting Diseases, Preventive Healthcare, Symptoms Awareness, Proactive Measures, Medical Data, Modern Data Mining Technologies, Applications.

I.   INTRODUCTION

Data mining, a process of uncovering hidden data, relationships, and patterns from extensive datasets, stands as a transformative method compared to traditional approaches. Organizations harness data mining technology, extracting valuable insights from their data warehouses. In the medical industry, diverse data mining tools help identify contemporary patterns, making it a quantitative and user friendly approach. Readability and error-correction capabilities enhance its utility. Data Mining techniques are used for various industrial purposes like automation, accident detection especially in cloud environment, health care and business sector. The healthcare sector in India, comprising public and private organizations, faces the challenge of managing substantial data related to patients, medicines, and diseases. With a focus on low-cost medical treatment for the economically disadvantaged, public sector healthcare organizations play a crucial role. Given the enormity of healthcare data, the demand for new technologies and techniques for efficient data management is evident. This paper explores the applications of data mining in healthcare, emphasizing its indispensability in addressing contemporary challenges and enhancing healthcare practices.

II.   LITERATURE SURVEY

1. **Data Mining Techniques in Healthcare: A Review (2005) - Han, J., & Kamber, M.**

- This foundational work provides an overview of various data mining techniques and their applications in healthcare.
- Discusses the challenges and opportunities of applying data mining in healthcare, emphasizing the potential for improved decision-making and patient care.

2. **Applications of Data Mining Techniques in Healthcare and Prediction of Heart Attacks (2012) - Rana, S., & Sharma, A.**

- Focuses on the application of data mining in predicting heart attacks, showcasing how predictive modeling can assist in identifying high-risk patients.
- Discusses the use of decision trees, neural networks, and other techniques for predicting cardiovascular events.

3. **A Survey of Data Mining Techniques in the Field of Healthcare (2014) - Bellazzi, R., & Zupan, B.**

- Provides a comprehensive survey of data mining techniques used in healthcare, covering areas such as disease prediction, diagnosis, and treatment planning.
- Highlights the challenges and future directions in the integration of data mining in healthcare systems.

4. **Data Mining in Healthcare - A Review (2015) - Kumar, P., & Deep, A.**

- Discusses the role of data mining in healthcare management, patient care, and disease prediction.
- Emphasizes the importance of efficient data preprocessing and feature selection in improving the accuracy of predictive models.

5. Big Data Analytics in Healthcare: Promise and Potential (2016) - Dumbill, E.

- Explores the impact of big data analytics, including data mining, in healthcare, emphasizing the potential for personalized medicine and improved patient outcomes.
- Discusses challenges related to data privacy, security, and interoperability.

III. METHODOLOGY

Method and analysis which is performed in your research work should be written in this section. A simple strategy to follow is to use keywords from your title in first few sentences.

Data Mining

Data mining is a dynamic process that involves the extraction and sorting of valuable information from extensive datasets. Essentially, data mining tools serve as predictive instruments, utilizing existing data or conducting analyses to foresee future trends. Within the realm of data mining, various parameters play crucial roles, each serving distinct purposes. Path analysis involves uncovering patterns within interconnected data points, revealing relationships that exist among them.

Data Mining Application Areas

Nowadays data mining is broadly used in many fields. Data mining is used to automate the predication from the large data. Some of the data mining applications in various fields are;

- Business
- Healthcare
- Telecommunication
- Marketing and Sales
- Road Safety
- Education

Data Mining Application of HealthCare

Data mining has been used intensively and widely by numerous industries. In healthcare, data mining is becoming more popular nowadays. Data mining applications can incredibly benefit all parties involved in the healthcare industry. For example, data mining can help the healthcare industry in fraud detection and abuse, customer relationship management, effective patient care, and best practices for affordable healthcare services. The large amounts of data generated by healthcare transactions are too complex and huge to be processed and analysed by conventional method.

Treatment Effectiveness: Data mining facilitates the identification of specific patient characteristics or demographic factors that may influence the success of certain treatments. By analysing large datasets, healthcare professionals can gain insights into the nuances of treatment responses among diverse patient populations. This knowledge can inform personalized and more effective treatment plans, ultimately improving overall healthcare outcomes.

Healthcare Management: Data mining involves using advanced analytical techniques to extract valuable insights from medical data. It aids in patient management, disease prediction, resource allocation, fraud detection, decision support, quality improvement, personalized medicine, and epidemiological studies. This leads to better decision-making, improved resource utilization, and enhanced patient care.

Customer Relationship Management: Data mining involves using analytical techniques to understand and improve interactions with customers. It includes segmentation for targeted strategies, predictive analytics for future behaviour, churn prediction, personalized marketing, and analysis of customer satisfaction. Additionally, data mining helps in detecting fraud, predicting customer lifetime value, and leveraging social media data for better insights. Overall, CRM in data mining enhances customer engagement and supports informed decision-making for businesses.

Fraud and Abuse: In data mining, fraud and abuse detection are critical applications. By analysing patterns and anomalies within large datasets in cloud [5], data mining helps identify irregularities or fraudulent activities. It plays a key role in safeguarding systems, preventing financial losses, and ensuring the integrity of data. The technology enables early detection of suspicious behaviour, making it an indispensable tool for maintaining security and trust in various domains, including finance, healthcare, and online platforms. This paper explores the synergy between data mining and healthcare by highlighting how data mining techniques contribute to the analysis, storage, and processing of substantial healthcare datasets. The integration of data mining and healthcare forms a cohesive model within the realm of distributed network technology. In this context, the cloud plays a pivotal role, providing support for big data in terms of data security, encryption, integrity, transformation, heterogeneity, and quality.

Despite encountering challenges such as scalability, availability, and bandwidth limitations for data transfer in the integration of data mining with cloud technology, ongoing efforts by cloud providers aim to address these issues. Solutions are being continuously developed to enhance the efficiency of utilizing data mining in healthcare on the cloud platform. The integration and application of data mining in healthcare are anticipated to have a profound and expanding impact in the years to come.

IV.   RESULTS

Table 1. Results of comparative analysis of various disease in Healthcare:

| S. No	Types of Disease	Data Mining Tool	Technique	Algorithm	Traditional Method	Accuracy Level for DM Applications
1	Tuberculosis	WEKA	Naïve Bayes classifier	KNN	Probability statistics	78 Percent
2	Heart Disease	ODND,NCC2	Classification	Naive	Probability	60 Percent
3	Kidney	RST	Classification	Decision making	Statistics	76 Percent
4	Diabetes Mellitus	ANN	Classification	C4.5 Algorithm	Neural Network	82 Percent
5	Blood bank Sector	WEKA	Classification	J48		90 Percent
6	Dengue	SPSS Modeler		C5.0	Statistics	80 Percent
7	Hepatitis C	SNP	INFORMATION	Gain	Decision rule	74 Percent
8	Heart disease	Python	Classification	MLP	Neural Network	84.9 Percent
9	Diabetes	WEKA	Classification	SMO	Machine learning	79.8 Percent
10	Heart Disease	Python	Classification	ANN	Neural Network	95.025 Percent
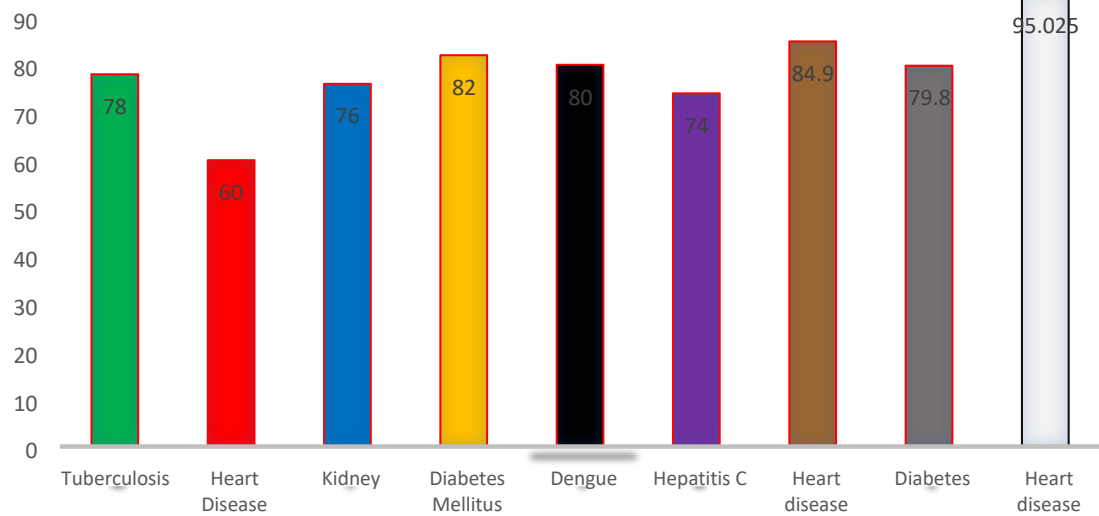
Figure 2: Accuracy Level Chart

V.   CONCLUSION

This paper explores the synergy between data mining and healthcare by highlighting how data mining techniques contribute to the analysis, storage, and processing of substantial healthcare datasets. The integration of data mining and healthcare forms a cohesive model within the realm of distributed network technology. In this context, the cloud plays a pivotal role, providing support for big data in terms of data security, encryption, integrity, transformation, heterogeneity, and quality.

Despite encountering challenges such as scalability, availability, and bandwidth limitations for data transfer in the integration of data mining with cloud technology, ongoing efforts by cloud providers aim to address these issues. Solutions are being continuously developed to enhance the efficiency of utilizing data mining in healthcare on the cloud platform. The integration and application of data mining in healthcare are anticipated to have a profound and expanding impact in the years to come.

VI.   REFERENCES

[1]   NV Balaji, 2021, Heart disease prediction system using Correlation Based Feature Selection with Multilayer Perceptron approach, IOP conference series: materials science and engineering

[2]   Ambily Merlin Kuruvilla, 2019, Predicting Diabetes Mellitus Using Feature Selection And Classification Techniques In Machine Learning Algorithms, Karpagam JCS

[3]   Ambily Merlin Kuruvilla, 2021, Improved Artificial Neural Network Through Metaheuristic Methods And Rough Set Theory For Modern Medical Diagnosis, Indian Journal of Computer Science and Engineering.

[4]   PC Sherimon, 2023, Customized Hybrid Deep Learning Model for Road Accident Detection Based on CCTV Images, 2023 IEEE International Performance, Computing, and Communications Conference (IPCCC).

[5]   RB Ansaba, 2022, Big Data in Cloud Computing Environment, International Journal Of Recent Trends In Multidisciplinary Research.

[6]   Dr. John T Abraham, 2018 Volume 3, Issue 11, Application of artificial intelligence in educational scenario: computer assisted instructions (CAI) - a tool for improving teaching learning process , International Journal of Scientific Development and Research (IJSDR)

[7]   V.V. Ramalingam, 2018," Prediction of Heart Diseases Using Machine Learning", International Journal of Engineering and Technology.

[8]   Amin Ul Haq 2019 A Hybrid Intelligence System Frame Work for Prediction of Heart Disease Using ML, Mobile Information Systems.

[9]   Shadman Nashif Heart Disease Detection by Using Machine Learning Algorithm and a Real time Cardiovascular Health Monitoring System World journal of Engineering and Technology,vol.06 no.04,2018,article id 88650.

[10]   Poornima Singh 2018 Effective heart disease prediction system using data mining techniques International journal of Nano medicine.

[11]   M. Gunay and T. Ensarı 2019 Predictive churn analysis with machine learning methods IEEE 26th Signal Processing and Communications Applications Conference (SIU), Izmir, 2018, pp. 1- 4.

[12]   R. Suguna, M. Shyamala Devi and Rincy Merlin Mathew 2019 Customer Churn Predictive Analysis by Component Minimization using Machine Learning International Journal of Innovative Technology and Exploring Engineering.

[13]   Devansh Shah, Samir Patel and Santosh Kumar Bharti, 2020, Heart Disease Prediction using Machine Learning Techniques, Springer.

[14]   Hamidreza and Ashrafi Esfahani, 2017, Cardiovascular disease detection using a new ensemble classifier, IEEE International Conference on Knowledge-Based Engineering and Innovation (KBEI).