"A Medical Decision Support System to Identify the Mortality Rates using ML"

Deepthi N[1], Sushmitha V Dundaraddi[2], Sandeep Kumar L[3], Jyothsna A[4], Subhash K N[5]

[1]*Assistant Professor, Dept. of Information Science & Engineering, VVIET, Karnataka, India*

[2]*Student, Dept. of Information Science & Engineering, VVIET, Karnataka, India*

[3]*Student, Dept. of Information Science & Engineering, VVIET, Karnataka, India*

[4]*Student, Dept. of Information Science & Engineering, VVIET, Karnataka, India*

[5]*Student, Dept. of Information Science & Engineering, VVIET, Karnataka, India*

Abstract

Patient mortality is a common occurrence in hospitals due to various factors, including the availability of resources, medical staff, and facilities. The rising rate of patient deaths is a significant concern, influenced by diseases, insufficient medical resources, and inadequate healthcare services. To address this challenge, we propose a system that automatically identifies the factors contributing to mortality rates. Our project aims to demonstrate the association between mortality and healthcare services using the ECLAT algorithm. The system will analyze the relationship between hospital resources and mortality rates using Microsoft technologies, considering parameters such as specialists, beds, ICU facilities, and nursing staff. Efficient machine learning algorithms will be employed to identify critical factors influencing mortality rates, utilizing "Visual Studio" for the front end and "SQL Server" for the backend due to their robust library and tool support for real time applications.

Keywords: ***Data Science, Machine Learning, Association Learning, Mortality Rates, Visual Studio, SQL Server***

1. Introduction

With the advancement of big data and AI, data analysis and mining have become essential in various fields, including healthcare. Our system collects multi-source electronic medical record data and utilizes data analysis and mining technologies for intelligent mortality prediction. Manual identification of mortality rate factors is complex, time-consuming, and expensive. Existing systems primarily focus on data collection, storage, and retrieval, lacking the capability to extract actionable insights for medical practitioners. Our proposed system employs the ECLAT algorithm to find relationships between hospital resources and mortality rates, aiming to provide valuable information to reduce patient deaths.

2. LITERATURE SURVEY

2.1 Paper 1: Machine Learning Algorithms in Healthcare Systems: A Review

Authors: Pradeep Kumar Kushwaha, M. Kumaresan

Year of Publication: 2021

Methods Used: Naïve Bayes, Decision Tree, SVM, Regression, and KNN algorithms

Result: This paper reviews the application of various machine learning algorithms in healthcare, specifically focusing on areas such as diabetes management, cancer detection, brain tumor identification, and bioinformatics.

Remarks: The algorithms are applied in diverse healthcare fields, but they do not address the prediction of mortality rates.

2.2 Paper 2: Artificial Intelligence-Based Comparative Study of Mortality Prediction

Authors: Satyam Prasad Tiwari, Ashutosh Upadhyay, Karthikeyan S

Year of Publication: 2020

Methods Used: Logistic Regression, Random Forest, and Support Vector Machine

Result: The study found that logistic regression outperformed other models in predicting patient mortality, leading to its selection for further use.

Remarks: Predicting mortality can aid in critical decision-making and optimize hospital resources, providing valuable insights for doctors and family members. However, the study does not explore the relationship between resource allocation and mortality rates.

2.3 Paper 3: Prediction of Mortality in Patients with Cardiovascular Disease Using Data Mining Methods

Authors: Damir Imamovic, Elmir Babovic, Nina Bijedic

Year of Publication: 2020

Methods Used: Decision Tree Mining, Neural Network, and Logistic Regression

Result: The research aims to compare the effectiveness of different data mining methods in predicting mortality in cardiovascular disease patients.

Remarks: The focus is exclusively on cardiovascular diseases. The use of Convolutional Neural Networks (CNNs) for data processing may require significant time and large datasets, potentially leading to less accurate results.

2.4 Paper 4: Prediction of Mortality in Patients with Cardiovascular Disease Using Data Mining Methods **Authors:** Damir Imamovic, Elmir Babovic, Nina Bijedic

Year of Publication: 2020

Methods Used: Decision Tree Mining, Neural Network, and Logistic Regression

Result: This study evaluates the effectiveness of various data mining techniques in predicting mortality among cardiovascular disease patients.

Remarks: Similar to the previous paper, this study is limited to cardiovascular disease. The application of CNNs necessitates substantial processing time and large data volumes, which may impact the accuracy of the results.

3. Comparative Analysis of LS

Constraints	**P 1**	**P 2**	**P 3**	**P 4**	**Our System**
Unsupervised Learning techniques used	No	No	No	No	Yes
More Parameters Used	No	No	Yes	No	Yes
Real Time Implementations	No	No	No	No	Yes
Dynamic Data	No	No	No	No	Yes
Identification of Most important resource for death rates	No	No	No	No	Yes
GUI based	No	No	No	No	Yes

4. Proposed Work

Analyzing the relationship between hospital resources and mortality rates is essential for effective public health policy deployment. High-quality health services play a crucial role in reducing mortality rates. By using data mining techniques, systems can uncover correlations between health services and mortality rates. The proposed system aims to assist medical departments in lowering mortality rates by identifying hidden correlations between various hospital resources, such as doctors, dentists, pharmacists, nurses, technical nurses, and scanning departments, and mortality rates. Identifying hospital resources that are associated with higher patient death rates will enable hospitals to implement strategies to mitigate these issues. This system allows organizations to operate effectively and efficiently, ensuring high standards and delivering key results. Consequently, patient mortality rates will decrease, hospital services will improve, and the hospital's reputation will be enhanced. Machine learning techniques are employed to achieve accurate results, using appropriate disease parameters for prediction and enabling faster decisionmaking. The system is designed to work with dynamic data, ensuring adaptability and continuous improvement.

4.1 Outcome

Input: Various hospital resources, including doctors, dentists, pharmacies, nurses, technical nurses, scanning departments, and mortality rates from previous years.

Output: Identifies the relationship between healthcare resources and mortality rates.

5. Methodology

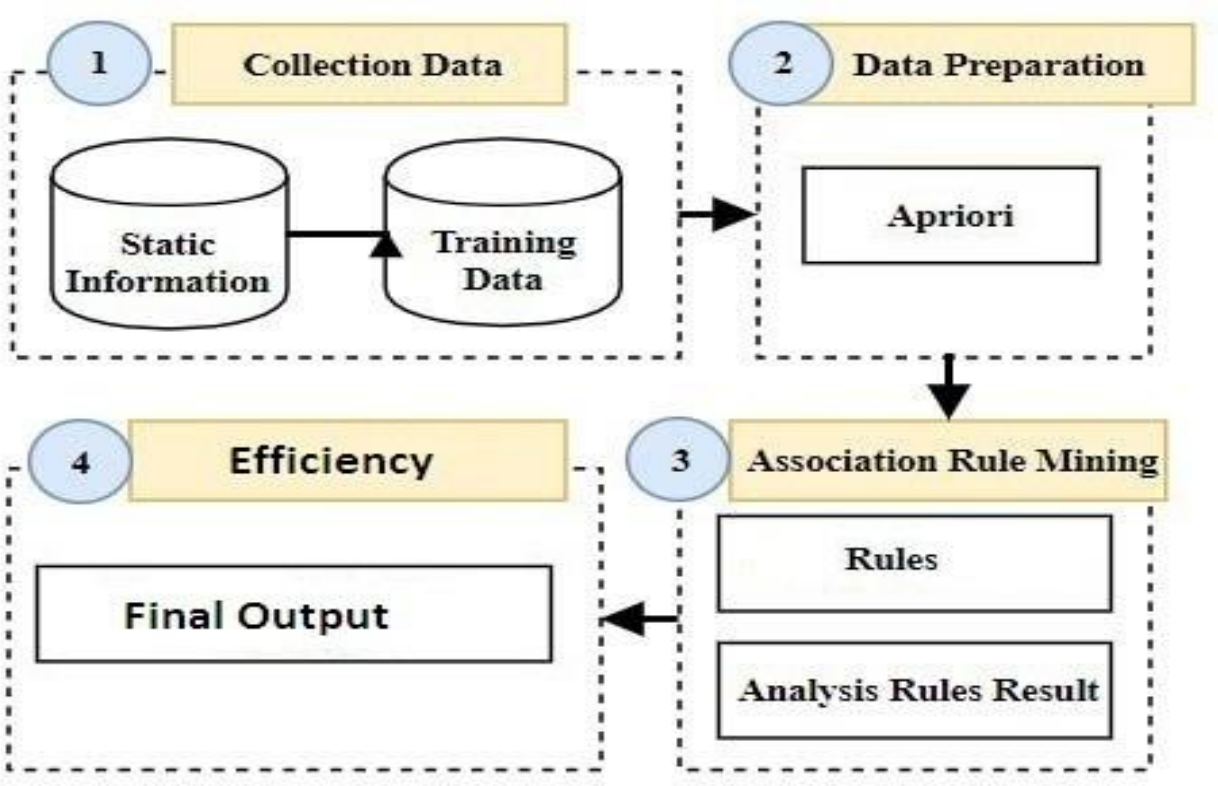

Fig: 1 Methodology Diagram

Step 1: Data Collection

We are developing a real-time application with dedicated data servers for storage. Data collection involves gathering information from various sources, including hospital resources and death rates.

Step 2: Data Preparation

Data from the servers is extracted and analyzed. This step involves removing irrelevant data and retaining only the necessary information related to hospital resources and mortality rates for pattern prediction.

Step 3: Specify Constraints

Constraints refer to the factors considered in the system, such as the number of heart specialists, neurologists, nurses, pharmacy units, physicians, beds, and death rates.

Step 4: Association Learning (Apriori Algorithm/Apriori TID Algorithm)

Efficient algorithms are applied to the training datasets to find patterns. The selected algorithms offer several advantages:

1. Faster results with quick predictions.

2. Effective for both small and large datasets.

3. Requires only one database scan.

4. Handles multiple constraints efficiently.

Step 5: Pattern Prediction

The system predicts the correlation between hospital resources and mortality rates using historical data processed by the Apriori and Apriori TID algorithms. Step 6: Results The generated patterns indicate relationships such as: Low number of heart specialists correlates with high death rates.High number of neurologists correlates with average death rates.Low number of physicians correlates with high death rates. Low number of physicians correlates with high death rates. Additional patterns continue in a similar manner.

Step 7: Visual Representation

The final patterns are displayed to users through a graphical user interface (GUI). When users log into the application, the system presents the outputs on the GUI.

6. Experiment Results

Result Analysis

6.1 APRIORI Algorithm

Performance Factor

Data Structure: Array-based.

Memory Utilization: Varies with the dataset size (less for smaller datasets).

Number of Scans: Only one scan required.

Execution Time: Depends on the candidate generation process.

No of Instances (records)	Execution Time (milli Secs)
Around 2k	6557
Around 1k	6495
Around 500 records	6465
100 records	6445

6.2 Apriori TID Algorithm

Performance Factor

Data Structure – array based

Memory Utilization – depends on the data set [less for small datasets]

No.of.scans - multiple scan to generate candidate set

Execution time - execution time depends on producing candidates

No of Instances (records)	Execution Time (milli Secs)
Around 2k	2572
Around 1k	2505
Around 500 records	2475
100 records	2450

7. Conclusion

Predicting patient mortality and understanding the reasons behind it are crucial in the current medical sector. Identifying the factors that contribute to increased patient mortality is essential. The proposed system utilizes efficient unsupervised learning algorithms, such as Apriori and Apriori TID, to process medical datasets and identify the relationship between hospital resources and patient mortality rates. Additionally, the system pinpoints the most significant factors contributing to increased patient mortality. This system is beneficial for hospitals, enabling them to address these critical issues effectively.

8. Future Enhancements

The proposed system employs the Apriori algorithm to identify the relationship between hospital resources and mortality rates. In data science, various algorithms such as the Eclat algorithm, SFIT algorithm, and AIT algorithm can also be used to discover patterns. In the future, additional algorithms can be incorporated to compare and determine the most efficient one. Moreover, a visitor query module can be added, allowing visitors to post queries to the administrator, who can then respond to them. Additional parameters related to hospital resources can also be included to enhance pattern discovery. **9.**

9. References

[1] Machine learning algorithm in healthcare system: A Review, Pradeep Kumar Kushwaha, M. Kumaresan, 2021

[2] Artificial Intelligence based Comparative Study of Mortality Prediction, Satyam Prasad Tiwari, Ashutosh Upadhyay, Karthikeyan S, 2020

[3] Prediction of mortality in patients with cardiovascular disease using data mining methods, Damir Imamovic, Elmir Babovic, Nina Bijedic , 2020

[4] Prediction of mortality in patients with cardiovascular disease using data mining methods, Damir Imamovic, Elmir Babovic, Nina Bijedic ,2020

[5] Intensive Care Unit Mortality Prediction: An Improved Patient-Specific Stacking Ensemble Model, NORA EL-RASHIDY, SHAKER EL-SAPPAGH, TAMER ABUHMED, SAMIR ABDELRAZEK, AND HAZEM M. EL-BAKRY , 2020