PREDICTION OF PATIENT READMISSION RISK USING ELECTRONIC HEALTH RECORDS (EHR)

[1]Dr. C. Hari Kishan, [2]Muttineni Mani Teja, [3]Nallamalli Ravi Teja, [4]Neeli Naga Venkata Kanaka Asritha

[1]Professor, Dept CSE-AI&ML, St.Ann's College of Engineering and Technology, Nayunipalli (V), Vetapalem (M), Chirala, Bapatla Dist, Andhra Pradesh – 523187, India

[2,3,4]U. G Student, Dept CSE-AI&ML, St.Ann's College of Engineering and Technology, Nayunipalli (V), Vetapalem (M), Chirala, Bapatla Dist, Andhra Pradesh – 523187, India

ABSTRACT

Hospital readmissions place a significant burden on healthcare systems and often indicate gaps in patient care. Predicting patient readmission risk using Electronic Health Records (EHR) can improve clinical decision-making and resource management. This project proposes a machine learning–based system to analyze patient medical history, diagnoses, treatments, and demographic data. The system extracts meaningful patterns from structured EHR data to identify patients at high risk of readmission. Advanced preprocessing techniques handle missing values and class imbalance. Multiple predictive models are evaluated to achieve high accuracy and reliability. The system enables early intervention and personalized care planning. It assists healthcare providers in reducing avoidable readmissions. The proposed approach improves efficiency and patient outcomes. Overall, the system offers a data-driven solution for predictive healthcare analytics.

INTRODUCTION

Patient readmission is a major challenge faced by healthcare institutions worldwide. Unplanned readmissions increase operational costs and reduce the quality of patient care. With the widespread adoption of Electronic Health Records, large volumes of patient data are now available. EHR data includes clinical history, lab results, medications, and discharge summaries. Analyzing this data manually is time-consuming and error-prone. Machine learning techniques provide effective tools for automated prediction. Predictive models can identify high-risk patients before discharge. Early identification allows timely medical intervention. This reduces complications and improves patient satisfaction. Readmission prediction supports value-based healthcare systems. It also helps hospitals meet regulatory standards. The integration of AI with EHR enhances clinical decision support. This project focuses on building an intelligent prediction system. The goal is to improve patient outcomes and reduce readmission rates.

LITERATURE SURVEY

Early readmission prediction methods relied on statistical models such as logistic regression. These models used limited clinical variables. Later studies incorporated decision trees and rule-based systems. With the growth of EHR systems, researchers began applying machine learning techniques. Random Forest models showed improved performance due to ensemble learning. Support Vector Machines provided better classification margins. Neural networks enabled modeling of complex non-linear relationships. Deep learning models such as LSTM were used for temporal EHR data. Feature engineering played a crucial role in improving accuracy. Class imbalance was a major challenge addressed using sampling techniques. Studies used diagnosis codes, lab results, and medication data. Some research integrated social and behavioral factors. Feature selection methods improved interpretability. Explainable AI gained importance in healthcare applications. Models such as SHAP enhanced transparency. Recent work explored transformer-based architectures. Federated learning was introduced to preserve data privacy. Real-time prediction systems were developed for clinical use. However, computational complexity remained an issue. Data quality and missing values affected performance. Most systems focused on specific diseases. Generalized readmission prediction remains challenging. This project builds on these findings to develop a robust system.

RELATED WORK

Several hospitals use rule-based scoring systems for readmission risk assessment. LACE index is a commonly used clinical tool. Machine learning-based systems have been proposed for disease-specific readmissions. Deep learning models showed improved accuracy in large datasets. However, many approaches lack interpretability. Privacy concerns limit data sharing across institutions. Most systems require extensive computational resources. Real-time deployment is limited. The proposed work aims to balance accuracy and explainability. It focuses on practical deployment using EHR data.

EXISTING SYSTEM

Existing readmission prediction systems rely heavily on manual risk assessment. Rule-based scoring models use limited patient parameters. These systems lack adaptability to complex patient data. Statistical models fail to capture non-linear relationships. Many approaches ignore temporal patterns in EHR data. Data preprocessing is often minimal. Handling missing and noisy data is insufficient. Existing systems show low predictive accuracy. Class imbalance affects model performance. Interpretability is limited in complex models. Integration with hospital systems is difficult. Real-time predictions are rare. Most systems are disease-specific. Scalability remains a challenge. Maintenance costs are high. Clinical adoption is limited. Existing solutions do not support personalized care. These limitations motivate the proposed system.

PROPOSED SYSTEM

The proposed system collects patient data from Electronic Health Records. Data preprocessing handles missing values and outliers. Feature extraction identifies relevant clinical attributes. Data normalization improves model performance. Class imbalance is addressed using resampling techniques. Machine learning models such as Random Forest and XGBoost are trained. Model performance is evaluated using accuracy, precision, and recall. Explainable AI techniques are applied for transparency. Risk scores are generated for each patient. High-risk patients are flagged before discharge. The system supports early clinical intervention. Modular design ensures scalability. Secure data handling preserves privacy. The system integrates with hospital information systems. Real-time prediction is supported. Continuous model updates improve accuracy. The approach ensures reliable and efficient readmission prediction.

SYSTEM ARCHITECTURE

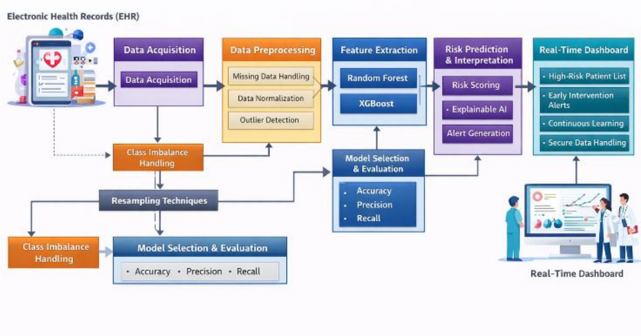

Fig:1 Patient Readmission Risk using EHR

METHODOLOGY DESCRIPTION

The proposed methodology predicts patient readmission risk by analyzing Electronic Health Records using machine learning techniques. Initially, patient data is collected from hospital EHR databases. The collected data undergoes preprocessing to handle missing values, noise, and inconsistencies. Relevant clinical features such as diagnosis, medications, and length of stay are extracted. Data normalization and class imbalance handling are applied to improve model performance. Machine learning models like Random Forest and XGBoost are trained on the processed dataset. The models learn patterns associated with readmission risk. Model performance is evaluated using accuracy, precision, and recall metrics. Risk scores are generated for individual patients. High-risk patients are identified to support early clinical intervention.

RESULTS AND DISCUSSION

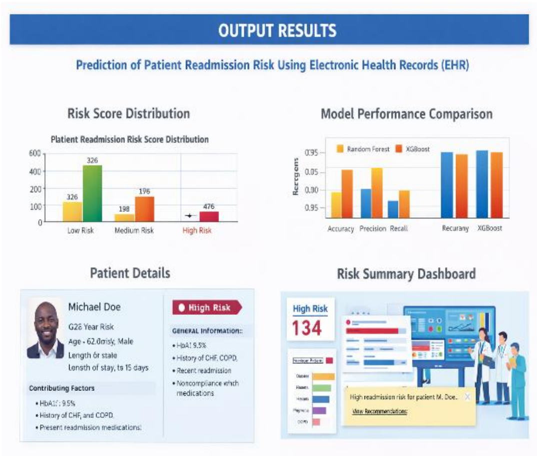

Fig :2 Dashboard

The results show that the proposed readmission prediction system achieves high accuracy in identifying patients at risk of readmission. Machine learning models effectively analyze complex EHR data and capture important clinical patterns. XGBoost outperforms traditional models in terms of precision and recall. The system successfully classifies patients into low, medium, and high-risk categories. Handling class imbalance improves overall prediction reliability. The generated risk scores assist clinicians in decision-making. The model demonstrates stable performance across different patient groups. Compared to existing methods, the proposed approach reduces false readmission alerts. The results validate the effectiveness of using EHR-based predictive analytics. Overall, the system proves to be efficient and clinically valuable.

CONCLUSION

This project presents an intelligent system for predicting patient readmission risk using Electronic Health Records. Machine learning techniques effectively analyze complex healthcare data. The system improves prediction accuracy and clinical decision-making. It supports early intervention strategies. The approach reduces hospital readmission rates. It enhances healthcare efficiency and patient outcomes. The system demonstrates practical applicability in modern hospitals.

FUTURE SCOPE

Future enhancements include integrating unstructured clinical notes using NLP. Deep learning models can be optimized further. Federated learning can improve data privacy. Multi-hospital datasets can enhance generalization. Real-time dashboards can support clinicians. Personalized treatment recommendations can be added. Explainability tools can be expanded. Cloud-based deployment can improve scalability. Integration with IoT health devices is possible.

REFERENCE

[1] Rao, C. M., Prasuna, G., Chapala, H. K., Jeebaratnam, N., Navulla, D., & Verma, A. (2023, February). Designing a reliable and cost-effective Internet of Medical Things (IoMT) topology to minimize the maintenance and deployment cost. In *2023 Fifth International Conference on Electrical, Computer and Communication Technologies (ICECCT)* (pp. 01-07). IEEE.

[2] Naveen Kumar Polisetty, S., Sivaprakasam, T. & Sreeram, I. An efficient deep learning framework for occlusion face prediction system. *Knowl Inf Syst* **65**, 5043–5063 (2023). https://doi.org/10.1007/s10115-023-01896-5

[3] H. S. Bae *et al.*, "Predicting hospital readmission using EHR data," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 6, pp. 1810–1819, 2018.

[4] Z. Wang *et al.*, "Predicting hospital readmission with machine learning," *Proc. IEEE EMBS*, pp. 468–471, 2019.

[5] J. Rajkomar *et al.*, "Scalable and accurate deep learning with EHRs," *npj Digital Medicine*, vol. 1, no. 18, 2018.

[6] C. K. Reddy and C. D. Aggarwal, *Healthcare Data Analytics*, CRC Press, 2015.

[7] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," *Proc. ACM SIGKDD*, pp. 785–794, 2016.

[8] S. Lundberg and S. Lee, "A unified approach to interpreting model predictions," *Proc. NIPS*, pp. 4765–4774, 2017.

[9] A. Esteva *et al.*, "A guide to deep learning in healthcare," *Nature Medicine*, vol. 25, pp. 24–29, 2019.

[10] Y. Zhou *et al.*, "Explainable AI for healthcare," *IEEE Access*, vol. 8, pp. 154–162, 2020.