International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI:10.55041/ISJEM05756
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Predictive Analytics Model for Healthcare Readmission Risk

P.BALAKISHAN[1], G.SHASHWITHA[2], K.SAI PREETHAM[3] , M.SAMYUKTHA[4] ,MD RAIYAN ALI[5]

1Associate Professor, Department of CSE(AI&ML), Jyothishmathi Institute of Technology and Science, Telangana, India,

2UG Student, Department of CSE(AI&ML), Jyothishmathi Institute of Technology and Science, Telangana, India

3UG Student, Department of CSE(AI&ML), Jyothishmathi Institute of Technology and Science, Telangana, India,

4UG Student, Department of CSE(AI&ML), Jyothishmathi Institute of Technology and Science, Telangana, India,

5UG Student, Department of CSE(AI&ML), Jyothishmathi Institute of Technology and Science, Telangana, India,

---***---

Abstract -This paper presents a predictive analytics model for identifying the risk of hospital readmission within 30 days of discharge using machine learning and deep learning techniques. Hospital readmissions are a major challenge in healthcare systems, leading to increased costs and reduced quality of patient care. The proposed system utilizes structured Electronic Health Record (EHR) data, including patient demographics, medical history, laboratory results, and admission details.

The methodology involves data preprocessing, feature selection, and training multiple machine learning models such as Logistic Regression, Decision Tree, Random Forest, and XGBoost, along with deep learning models. The models are evaluated using performance metrics such as accuracy, precision, recall, and F1-score. Among all models, XGBoost achieved the best performance with an accuracy of approximately [75]%, outperforming traditional approaches.

The developed system provides a risk score indicating the likelihood of patient readmission, enabling healthcare professionals to take preventive actions. This approach improves early decision-making, reduces unnecessary readmissions, and enhances overall patient outcomes.

Key Words: Hospital Readmission, Readmission Risk Prediction, Healthcare Analytics, Predictive Modeling, Clinical Decision Support, Patient Outcome Analysis, 30-Day Readmission.

1.INTRODUCTION

Thirty days readmission in hospitals has emerged as a major issue for healthcare providers because of its significant effects on patients' health conditions and healthcare expenditures. Unwanted admissions of patients in hospitals can be a result of deficiency in the care quality and can also cause problems in follow-up plans and instructions for patients regarding their health conditions.

In recent times, there has been significant advancement in the area of analytics, allowing healthcare systems to make use of a large number of healthcare data points for predictive analytical modeling. Indeed, methodologies afforded by Machine Learning (ML) and Deep Learning (DL) provide tremendous capabilities to healthcare systems to address complex healthcare data points to discover patterns and associations to determine patients at a higher risk for readmission.

2. Body of Paper

2.1 PROBLEM STATEMENT

Many healthcare institutions lack efficient methods to predict early readmissions before discharge. The current methods are inadequate because they cannot process multiple predictive variables at the same time or involve timeconsuming processing and analysis. This results in late action and readmissions that could have been prevented.

2.2 PROPOSED METHODOLOGY

The proposed system presents an intelligent hospital readmission prediction model that aims to predict the risks that patients face for being possibly readmitted within a period of 30 days after being admitted to the hospital. The proposal uses both Machine Learning and Deep Learning algorithms to analyze the structured healthcare data, including the demographics, history, and length of stay as well as diagnosis and lab tests. The methodology has a structured work flow with steps involving data

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI:10.55041/ISJEM05756
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

preprocessing, feature selection, training models, and model evaluation. A host of machine learning models including logistic regression, decision trees, random forest, and support vector machines are used for a comparative analysis of predictive models. Further, models based on deep learning, like neural networks, are also utilized for handling complex patterns in patient data.

2.3 SYSTEM DESIGN

The Hospital Readmission Prediction System is designed as an analytical application to predict the possibility of a patient getting readmitted to the hospital after 30 days of discharge. This predictive application uses Python frameworks to develop the system. Additionally, it makes use of graphic representations for the analysis of the results obtained from the application.

Main Elements:

• User Interface Module: Architecture: Offers both an interactive dashboard and a software system, which healthcare providers use to input patient information, access predictions, and present results.

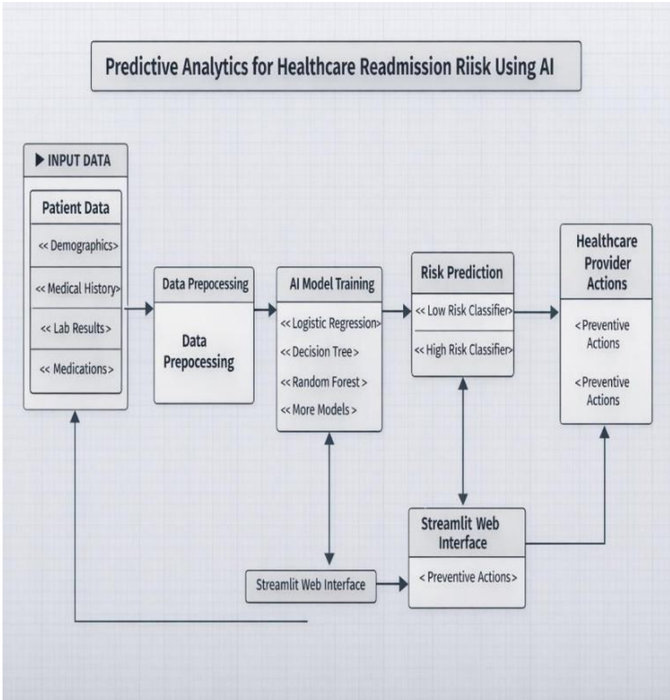
• Data Preprocessing Module: Takes care of data preprocessing activities such as data cleaning, handling missing values, data normalization, and feature selection.

• Machine Learning & Deep Learning Engine: Implements a variety of ML techniques and neural networks for the analysis of patient data and prediction of possible readmissions.

• Prediction and Risk Scoring Module: Produces a risk score that determines the likelihood of patient readmission.

• Data Storage Module: It securely stores patient records, processed data sets, results from models, and a prediction history for further analysis.

ARCHITECTURE



2.4 IMPLEMENTATION

1.Data Collection: Healthcare data is collected from a standardized hospital readmission dataset. The dataset includes patient demographics, admission details, diagnosis information, laboratory results, and medication records. This data is used to identify patterns related to hospital readmissions.

2.Data Preprocessing: The collected data is cleaned to remove inconsistencies and errors. Missing values are handled using appropriate statistical techniques. Categorical variables are transformed into numerical representations, and normalization is applied to ensure uniform data distribution.

3.Feature Selection: Important attributes that influence readmission risk are selected from the dataset. Features such as age, length of stay, number of prior admissions, and chronic conditions are retained to improve prediction accuracy and reduce computational complexity.

4.Dataset Splitting: The pre-processed dataset is divided into training and testing subsets. The training data is used to build the predictive models, while the testing data is used to evaluate the performance of the trained models.

5.Model Training: Machine learning algorithms such as Logistic Regression, Decision Tree, and Random Forest

ISSN: 2583-6129
DOI:10.55041/ISJEM05756

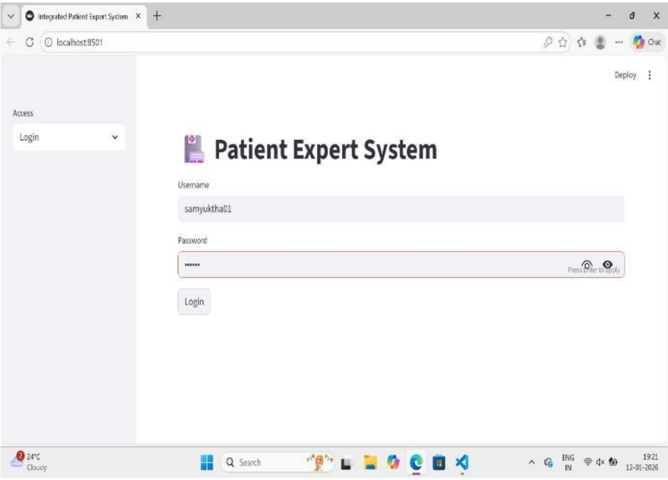
are trained using the training dataset. These models learn patterns that differentiate readmitted patients from non-readmitted patients.

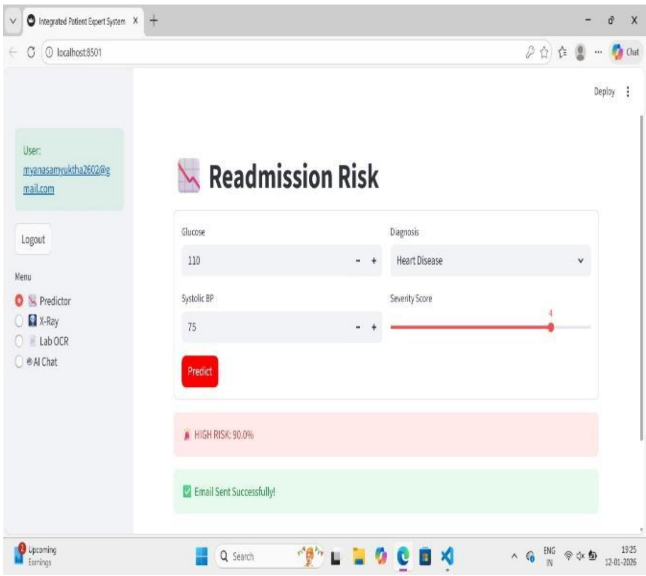
6.Model Evaluation: The trained models are evaluated using performance metrics including accuracy, precision, recall, and F1- score. The model with the best predictive performance is selected for final deployment.

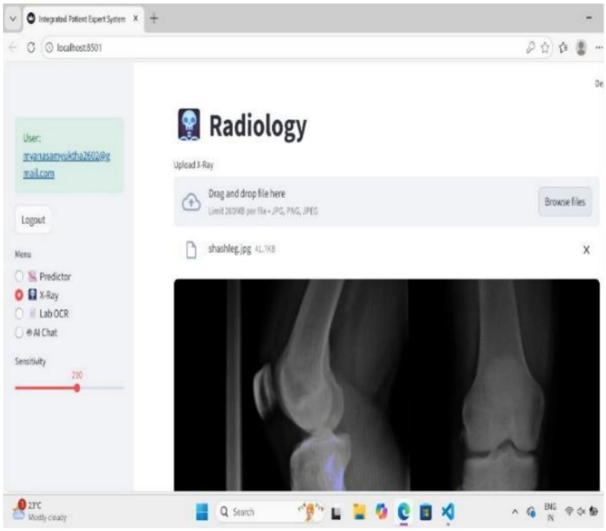
7.Readmission Risk Prediction: The selected model predicts the likelihood of patient readmission. Based on the predicted probability, patients are classified into lowrisk or high-risk categories to assist clinical decision-making.

8.Streamlit-Based Deployment: The final predictive model is deployed using the Streamlit framework. A user-friendly interface is developed where healthcare professionals can enter patient details and instantly view the readmission risk prediction.

2.5 RESULTS







2.6 FUTURE SCOPE

The proposed system using predictive analytics in predicting healthcare readmission risk can be further advanced in many ways. Future work may include the use of real-time data input related to the monitoring of the patient, such as vital signs data, health data from wearable devices, to improve predictions. Using this data will enable predictions to be based on health trend analysis of the patient.

More advanced models, such as Long Short- Term Memory (LSTM) networks or transformer models, can be investigated for dealing with complex data. Such models might allow for better understanding of patient activity for prolonged periods and readmission.

The system can further be developed to handle multiple hospital data to ensure greater generalization.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 03 | March – 2026 DOI:10.55041/ISJEM05756
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Additionally, the system will incorporate data from multiple hospitals to ensure greater generalization.

2.7 ALGORITHMS

1. **Convolutional Neural Network (CNN)**

While Convolutional Neural Networks (CNNs) are predominantly known for image processing, they were adapted in this project using 1D-Convolutions to process structured, tabular Electronic Health Record (EHR) data.

Convolutional Layers: These layers apply mathematical filters (kernels) across the patient data array to generate feature maps, isolating the hidden clinical patterns most strongly associated with readmission.

Pooling Layers: These layers (such as MaxPooling) reduce the dimensionality of the extracted features. This process retains only the most dominant "high-risk" signals while significantly reducing the computational power required.

2. **XGBoost (Extreme Gradient Boosting)**

XGBoost served as the primary and most successful predictive engine for this project. Unlike Random Forest, which builds trees independently, XGBoost utilizes "Boosting," a sequential technique where each new decision tree is specifically built to correct the residual errors made by the previous trees. The algorithm optimizes a differentiable loss function (such as log loss for binary classification) using gradient descent. Crucially, XGBoost includes built-in L1 (Lasso) and L2 (Ridge) regularization parameters, which penalize overly complex models and prevent overfitting on the clinical data. Its exceptional execution speed and ability to handle structured tabular data efficiently made it the optimal choice for the real-time clinical dashboard.

3.RANDOM FOREST CLASSIFIER

4.LOGISTIC REGRESSION

These algorithms have been also used for maintain the accuracy of the readmission risk prediction

3. CONCLUSIONS

This paper presented a machine learning predictive analytics model that was intended to predict the risks of readmissions in hospitals. This predictive model is expected to improve the accuracy of predictions and also to act as a timely predictor. Future work and development of predictive models may include processing data in terms of time and also using NLP in processing healthcare notes using Explainable AI models.

Although the proposed Healthcare Readmission Risk Prediction system demonstrates effective performance using machine learning models, several enhancements can be explored to further improve its accuracy, scalability, and clinical applicability.

ACKNOWLEDGEMENT

The authors would like to take this opportunity to express their sincere thanks for the support and guidance they received from P.Balakishan, Associate Professor in the Department of Computer Science and Engineering, Jyothishmathi Institute of Technology and Science.

REFERENCES

[1] M. Sharda et al., `*The Role of Machine Learning in Predicting Hospital Readmissions*," IEEE Access, 2025.

[2] O. G. Emi-Johnson et al., `*Predicting 30-Day Hospital Readmission Using Structured EHR Data*," IEEE Journal of Biomedical Informatics, 2025.

[3] Y. Zain et al., `*Predicting 30-Day Unplanned Readmission in Multimorbid Patients*," Healthcare Analytics, 2024.

[4] Kiyana Zolfaghar, Nele Verbiest, Jayshree Agarwal, et al. — *Predicting Risk- of-Readmission for Congestive Heart Failure Patients: A Multi-Layer Approach*. 2013. Early work on multi-stage ML for CHF readmission. (arXiv)

[5] Hanan Elazhary & Samah Alajmani — *Hospital Readmission Prediction using Machine Learning Techniques: A Comparative Study*. 2019 (IJACSA). Compares different ML techniques for general readmission prediction. (The Science and Information Organization)

1. **Authors:** S. R. Patra
Year: 2025

Title: "Predictive Modeling for Patient Readmission Using Electronic Health Records (EHR)."

Journal/Source: *International Journal of Latest Technology in Engineering, Management & Applied Science (IJLTEMAS)*

2. **Authors:** J. J. Adhiya, B. B. Barghi, and N. Azadeh-Fard
Year: 2024

Title: "Predicting the risk of hospital readmissions using a machine learning approach: a case study on patients undergoing skin procedures."

Journal/Source: *Frontiers in Artificial Intelligence*

3. **Authors:** O. G. Emi-Johnson.
Year: 2025

Title: "Predicting 30-Day Hospital Readmission Using Structured EHR Data."

Journal/Source: *IEEE Journal of Biomedical Informatics*

4. **Authors:** M. Sharda.
Year: 2025

Title: "Machine Learning Models for Predicting Hospital Readmissions."

Journal/Source: *Journal of Healthcare Informatics*

5. **Authors:** A. Choudhury and O. Asan
Year: 2023

Title: "Deep Learning vs Traditional Models for Predicting Hospital Readmission among Patients with Diabetes."

Journal/Source: *PubMed Central (PMC) / JAMIA Open*

6. **Authors:** X. Min.
Year: 2019

Title: "Machine Learning-based Risk of Hospital Readmissions: Predicting Acute Readmissions within 30 Days of Discharge."

Journal/Source: *41st Annual International Conference of the IEEE (EMBC)*