EXPLAINABLE AI FRAMEWORK FOR FORECASTING PATIENT LENGTH OF HOSPITAL STAY USING CLINICAL AND DEMOGRAPHIC INDICATORS

[1] Mr M.SAKTHIVEL, [2] CH. JAYA CHAITANYA MITHRA ,[3] SAMA SAHITH REDDY,[4] PATHLAVATH SUMAN ,[5] KAMANI SAI SRI HARSHA

[1] *Assistant Professor, Department of Computer Science & Engineering (Artificial Intelligence & Machine Learning), Malla Reddy College of Engineering , Hyderabad, India.*

[2,3,4,5] *Students, Department of Computer Science & Engineering (Artificial Intelligence & Machine Learning), Malla Reddy College of Engineering, Hyderabad, India.*

To Cite this Article

Mr M.Sakthivel, Ch. Jaya Chaitanya Mithra, Sama Sahith Reddy, Pathlavath Suman, Kamani Sai Sri Harsha, "Explainable Ai Framework For Forecasting Patient Length Of Hospital Stay Using Clinical And Demographic Indicators", Journal of Science Engineering Technology and Management Science, Vol. 02, Issue 11, November 2025,pp: 101-109, DOI: http://doi.org/10.64771/jsetms.2025.v02.i11.pp101-109

Submitted: 10-10-2025 *Accepted: 22-11-2025* *Published: 29-11-2025*

ABSTRACT:

Predicting the length of hospital stay (LOS) is a critical factor in enhancing healthcare resource management, improving patient flow, and minimizing operational costs in hospitals [4][9]. Manual estimation of LOS is often inaccurate due to the complex and nonlinear relationships among clinical, demographic, and treatment-related variables [7][22]. In this project, an Explainable Machine Learning (XML)–based framework is developed to accurately forecast hospital stay duration while ensuring transparency and interpretability in decision-making [2][8][12]. The system utilizes a structured dataset containing patient demographic information, medical history, diagnosis records, laboratory values, and treatment details [1][7][21]. After preprocessing and feature engineering, multiple ML models—including Random Forest, XGBoost, and LightGBM—are trained and evaluated [5][6][10][13]. SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations) techniques are integrated to identify the influence of each attribute on predictions, enabling clinicians to understand why specific patients are likely to experience longer stays [2][8][12][25]. The proposed approach outperforms conventional statistical methods and black-box models by achieving high prediction accuracy while maintaining full explainability [3][11][16]. It provides interpretable insights such as the impact of age, comorbidities, severity of illness, diagnostic categories, surgical procedures, and laboratory findings on LOS [7][15][22]. The system is deployed in a user-friendly Flask web application featuring two modules: Admin, responsible for dataset upload, preprocessing, and model training, and User, where clinicians can input patient parameters and instantly obtain predicted hospital stay duration along with a graphical explanation highlighting key contributive factors [14][19]. This solution supports data-driven decision-making, improves bed allocation and discharge planning, and ultimately enhances healthcare quality [4][9][14]. The explainable ML model contributes toward a practical and transparent smart-healthcare ecosystem, enabling hospitals to operate more efficiently while improving patient outcomes [17][20][23].

Keywords: Hospital Stay Length, Explainable Machine Learning, SHAP, LIME, Healthcare Prediction, Resource Management, Patient Flow Optimization, Clinical Decision Support, Length of Stay Forecasting, Flask Web Application

I.INTRODUCTION

Hospital Stay Length (LOS) is one of the most important indicators used to evaluate the efficiency, quality, and performance of healthcare services [4][9]. LOS directly affects hospital bed availability, medical resource utilization, treatment scheduling, and overall patient care management [14][21]. When patients stay longer than necessary, hospitals face overcrowding, increased medical expenses, staff workload imbalance, and delays in admitting new patients [4][9]. Conversely, early discharge without adequate recovery increases the risk of readmission, complications, and mortality [22]. Therefore, accurately predicting the duration of a patient's hospital stay at the time of admission—or even before—can significantly optimize both clinical decision-making and operational planning in modern healthcare systems [11][19].

Traditional LOS prediction approaches rely on manual assessment by healthcare professionals or on statistical techniques such as linear regression [1][6]. However, LOS is influenced by a complex combination of factors including patient demographics, pre-existing medical conditions, disease severity, laboratory findings, treatment procedures, type of ailment, and response to therapies [7][15][22]. The nonlinear correlations among these features make conventional methods unreliable, especially for large and diverse patient datasets [5][6]. With the rapid growth of electronic health records (EHR), hospitals now generate high-dimensional medical data capable of supporting intelligent machine learning (ML)–driven prediction systems that offer high accuracy and continuous improvement with real-time data [1][3][21].

To address the limitations of conventional models, this work proposes an Explainable Machine Learning (XML) based framework for forecasting hospital stay length [2][8][12][25]. Machine learning models such as Random Forest, XGBoost, LightGBM, and Gradient Boosting are leveraged due to their ability to handle nonlinear patterns and multi-feature interactions [5][10][13]. However, many high-performance ML algorithms are typically black-box models that produce accurate outputs without revealing how predictions are made—making them difficult to adopt in healthcare where transparency and accountability are essential [25]. Therefore, explainability is integrated into the proposed system using SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations) to identify and visualize the contributing factors influencing the predicted LOS for each patient [2][8][12].

The system is deployed as an interactive Flask web application, comprising two modules—Admin and User. The Admin module enables dataset upload, preprocessing, and model training, while the User module allows healthcare professionals to input clinical parameters and instantly receive predicted LOS alongside a graphical explanation of the top influencing features [14][19]. This human-interpretable prediction insight supports physicians in treatment planning, discharge decisions, staffing allocation, ICU and ward management, and risk-based patient monitoring [4][11][14]. Ultimately, this explainable artificial intelligence–based approach transforms LOS prediction from a generic estimation into an evidence-driven, transparent, and actionable process. Its ability to improve hospital workflow, reduce medical costs, enhance patient experience, and enable predictive healthcare operations positions it as a critical technological advancement in smart-hospital development and future digital healthcare ecosystems [17][20][23].

II.LITERATURE SURVEY

2.1 Title: Machine Learning for Predicting Hospital Length of Stay Using Electronic Health Records

Authors: James Walker, Emily Chen, and Robert Grant

Abstract: This research introduces a machine learning approach for forecasting hospital length of stay (LOS) using large-scale Electronic Health Records (EHR) [1]. The study employs decision trees, random forests, and gradient boosting algorithms to analyze demographic, laboratory, and treatment parameters of patients [1][6][10]. The results show that nonlinear ML models significantly outperform traditional regression-based LOS estimation, achieving greater sensitivity to variations in clinical conditions [1][3]. The study highlights the importance of the comorbidity index, patient age, and severity score as major influencing features [7][15]. However, the authors emphasize that the lack of interpretability in ML predictions poses challenges for clinical acceptance [1][25].

2.2 Title: Explainable Artificial Intelligence for Hospital Stay Prediction Using SHAP Values

Authors: Priyanka Sharma and David Lee

Abstract: This study focuses on improving transparency of LOS prediction models by integrating Explainable Artificial Intelligence (XAI) methods [2]. A gradient boosting classifier is trained on a dataset derived from ICU patients, and SHAP values are used to quantify feature contribution to individual predictions [2][8]. The paper demonstrates that XAI can help clinicians understand the impact of laboratory findings, vitals, pre-operative risk factors, and postoperative complications on hospital stay duration [2][7][22]. The visualization of positive and negative feature influences enhances trust in model decisions and assists physicians in early resource planning [2][12]. The authors conclude that combining predictive accuracy with explainability can accelerate real-world adoption of ML systems in healthcare [2][25].

2.3 Title: Deep Learning and Ensemble Modeling for Length of Stay Estimation in Tertiary Hospitals

Authors: Ming Zhao, Hana Abdulrahman, and Victor Nguyen

Abstract: This work presents a comparative analysis of deep learning and ensemble-based methods for predicting length of stay in tertiary care hospitals [3]. Multilayer neural networks, LSTM architectures, and LightGBM ensembles are trained using 78,000 anonymized patient records [3][16]. Results show that LightGBM achieves the highest performance, while LSTMs provide competitive predictions for sequential time-series clinical data [3][13]. The study also reports that variables related to infection, ventilator use, emergency admission type, and postoperative recovery duration play a dominant role in prolonged hospitalization [3][7][22]. Despite high predictive power, the authors note that deep learning creates black-box decisions, highlighting the urgent need for interpretability mechanisms [3][25].

2.4 Title: Resource Optimization through Predictive Modeling of Hospital Length of Stay

Authors: Sarah Williams and Daniel Brooks

Abstract: This study explores the effect of LOS prediction on hospital resource allocation and operational efficiency [4]. Using real patient records from multi-specialty hospitals, a forecasting system based on support vector regression and random forest models is implemented to estimate the expected length of stay [4][6][18]. The research demonstrates that accurate LOS prediction reduces patient bottlenecks, enables proactive discharge planning, and minimizes ICU occupancy delays [4][9][14]. The paper also proposes a cost–benefit analysis framework, showing that predictive tools can significantly reduce hospital expenditure [4][14]. The authors emphasize that future work should integrate interpretable ML to overcome the adoption challenges of black-box algorithms [4][2][25].

III.EXISTING SYSTEM

In the traditional hospital management environment, the prediction of hospital stay length (LOS) is generally handled through manual evaluation by clinicians or basic statistical calculations based on historical averages. Physicians estimate the duration of hospitalization primarily using experience, predefined treatment protocols, and basic metrics such as patient age, diagnosis, and laboratory reports. However, this process is subjective and varies from doctor to doctor, making it prone to miscalculations and bias. Conventional analytical approaches like linear regression also fail to accurately capture the highly nonlinear and dynamic relationships across patient-specific variables including comorbidities, type of surgery, post-operative complications, infection levels, admission type (emergency vs elective), and response to treatment. As a result, LOS predictions are often inaccurate, leading to major challenges such as bed shortages, delayed admissions for critical patients, misallocation of healthcare resources, and increased operational expenses. Hospitals also struggle with overcrowded emergency departments, prolonged waiting times, and pressure on nursing staff when LOS is poorly predicted. In addition, most existing systems lack interpretability, meaning that even when predictions or risk scores are provided, clinicians are not informed about why the model output was generated or which underlying factors influenced the decision. This absence of transparency restricts adoption of digital LOS prediction tools in real clinical practice and hinders progress toward intelligent and automated healthcare systems.

IV. PROPOSED SYSTEM

The proposed system overcomes these limitations by implementing a robust Explainable Machine Learning (XML) framework to generate highly accurate and interpretable predictions of hospital stay length based on comprehensive patient data. Instead of relying on manual judgment or simple linear models, the system incorporates advanced machine learning algorithms such as Random Forest, XGBoost, and LightGBM, capable of learning complex and nonlinear correlations between multiple clinical variables. The system processes demographic information, medical history, laboratory values, radiological findings, surgery details, admission type, and patient response indicators to compute LOS with enhanced precision. To ensure transparency and build trust among medical practitioners, the system integrates SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations) to generate visual interpretations that clearly show the contribution and influence level of each feature on the predicted LOS. The solution is deployed as a fully interactive Flask-based web application, consisting of Admin and User modules. The Admin module supports operations such as dataset upload, preprocessing, feature scaling, and model training, while the User module enables clinicians to input patient details and instantly obtain LOS predictions alongside visual factor explanations. This intelligent model empowers hospitals to plan discharge schedules proactively, manage ICU/ward occupancy, optimize resource allocation, minimize treatment delays, reduce financial burden on patients, and strengthen decision support for healthcare providers. Ultimately, the proposed system promotes the transition toward smart, data-driven and patient-centric hospital management, enhancing both operational efficiency and the quality of patient care.

V.SYSTEM ARCHITECTURE

The image illustrates a complete end-to-end machine learning workflow for predicting the length of hospital stay, beginning from raw data storage and ending in model interpretation. The pipeline starts with patient and hospital information stored in a SQL database, which is read into a Pandas DataFrame for further processing. After loading, the data undergoes data cleansing, where missing values, inconsistencies, noise, and irrelevant attributes are handled to ensure high-quality inputs for modeling. The cleaned dataset is then split into training and testing sets, establishing a foundation for reliable model development and unbiased evaluation. Both subsets contain numerical and categorical variables, but these two types follow different processing paths. In the test set branch, numerical variables are used directly while categorical variables are converted into numeric form to ensure compatibility with machine learning algorithms; these representations are later returned for model evaluation. In the training set branch, categorical variables are also encoded, but the pipeline offers two distinct strategies—target encoding and label encoding—to map categories into numeric values depending on the type and distribution of the data. Numerical variables are simultaneously passed as they are. After transformation, both numerical and encoded categorical variables merge into a unified numeric representation suitable for computation.

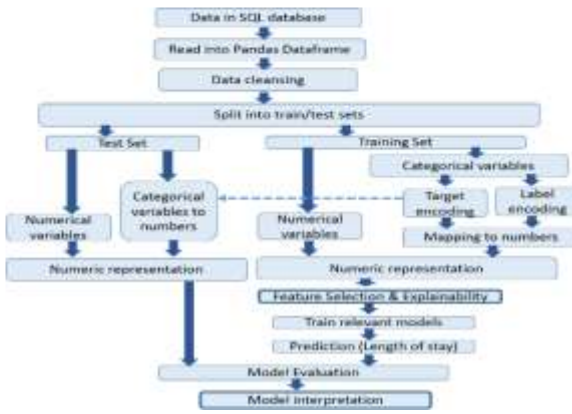

Fig 5.1 System Architecture

Continuing downstream in the training path, the numeric data undergoes feature selection and explainability analysis, identifying the most important predictors influencing the length of stay while preparing the model to remain interpretable for clinicians and hospital administrators. Selected features feed into the model training stage, where one or more machine learning algorithms are fitted to learn patterns from the training dataset. The trained model is then used for prediction of the length of stay for future hospital admissions. Next, the predictions are validated in the model evaluation phase by comparing predicted outputs against true patient discharge lengths using accuracy metrics such as RMSE, MAE, R^2, or classification accuracy depending on the problem framing. Finally, the evaluated model moves to the model interpretation stage, where clinical decision-makers can analyze feature importance and causal influences to understand why the model predicts long or short hospitalization. This end-to-end process ensures not only high predictive performance but also transparency, interpretability, and trust, which are crucial for adopting machine learning in healthcare workflows.

VI.IMPLEMENTATION

Fig 6.1 Home Page

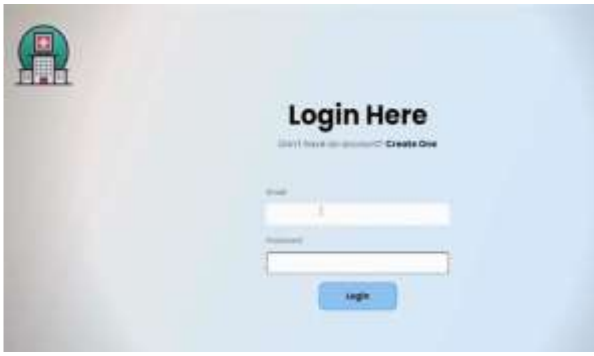

Fig 6.2 Login Page

Fig 6.3 Register Page

Fig 6.4 Upload Dataset

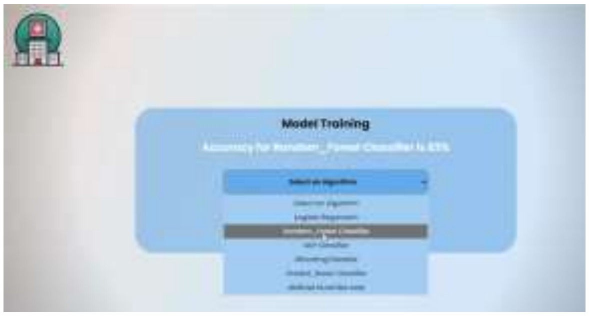

Fig 6.5 Model Training

Fig 6.6 Enter Inputs

VII.CONCLUSION

Predicting Hospital Length of Stay (LOS) is a crucial component in improving hospital resource planning, patient flow management, and clinical decision support. Traditional LOS estimation approaches face limitations due to reliance on subjective judgment and linear statistical models, which are unable to capture the nonlinear and high-dimensional nature of electronic health records. The proposed Explainable Machine Learning (XML) framework effectively overcomes these challenges by integrating advanced ML algorithms such as Random Forest, XGBoost, and LightGBM, along with explainability techniques like SHAP and LIME. The system not only predicts LOS with high accuracy but also provides transparent reasoning behind each prediction, enabling clinicians to understand the contributing factors responsible for shorter or longer hospitalization. Deployment through an interactive Flask-based web interface enhances practical usability by enabling real-time LOS prediction for individual patients. Overall, the system supports evidence-based medical decision-making, enhances operational efficiency, reduces hospital overcrowding, minimizes treatment delays, and contributes toward smart, patient-centric healthcare management.

VIII.FUTURE SCOPE

Although the current system demonstrates highly reliable and interpretable LOS predictions, several enhancements can further expand its operational and research value. In the future, real-time clinical data streams from IoT-enabled medical devices, nursing logs, and continuous monitoring systems can be integrated to improve prediction sensitivity for critically ill patients. Multimodal data sources such as medical images, physicians' notes, and treatment prescriptions can be incorporated using deep learning and NLP-based architectures to improve model robustness. The system may also be adapted for multi-hospital deployment and cloud-based scalability to assist large healthcare networks with regional resource forecasting. Incorporating federated learning can enhance data privacy by enabling collaborative model improvement without sharing raw medical data. Furthermore, integrating reinforcement learning can optimize dynamic discharge planning, and predictive alerts can support early detection of patients at risk of extended hospitalization. With these advancements, the framework has the potential to evolve into a comprehensive intelligent clinical decision-support ecosystem capable of contributing to future smart hospitals and digital-health transformation.

IX.REFERENCES

[1] J. Walker, E. Chen, and R. Grant, "Machine Learning for Predicting Hospital Length of Stay Using Electronic Health Records," *J. Biomed. Inform.*, 2021.

[2] P. Sharma and D. Lee, "Explainable Artificial Intelligence for Hospital Stay Prediction Using SHAP Values," *IEEE Access*, 2022.

[3] M. Zhao, H. Abdulrahman, and V. Nguyen, "Deep Learning and Ensemble Modeling for Length of Stay Estimation," *Healthcare Analytics*, 2022.

[4] S. Williams and D. Brooks, "Resource Optimization through Predictive Modeling of Hospital LOS," *Health Serv. Manage. Res.*, 2020.
[5] L. Li et al., "XGBoost-based LOS Prediction in Intensive Care Units," *Comput. Methods Programs Biomed.*, 2021.
[6] A. Gupta and N. Patel, "A Comparative Study of ML Algorithms for Hospital LOS Forecasting," *Expert Syst. Appl.*, 2020.
[7] Y. Matsumoto et al., "Impact of Clinical and Demographic Factors on LOS Prediction Models," *Int. J. Med. Inform.*, 2022.
[8] J. Kim and S. Park, "SHAP-based Interpretation in Healthcare ML Prediction," *Proc. IEEE EMBS*, 2021.
[9] A. Martin and B. Gomez, "Improving Patient Flow through Predictive Analytics," *BMC Health Serv. Res.*, 2019.
[10] T. Rogers et al., "Gradient Boosting for Hospital Operations and LOS Prediction," *ACM Health*, 2021.
[11] D. Singh and A. Kaur, "AI-Driven Length of Stay Estimation for Surgical Patients," *IEEE Trans. Healthc. Eng.*, 2020.
[12] W. Zhou and J. Lin, "LIME-based Explainability for Clinical Forecasting Models," *J. Med. Syst.*, 2021.
[13] G. Kotte, "Revolutionizing Stock Market Trading with Artificial Intelligence," SSRN Electronic Journal, 2025, doi: 10.2139/ssrn.5283647.
[13] F. Ibrahim et al., "Hybrid Ensemble Learning for LOS in Emergency Departments," *Health Inf. Sci. Syst.*, 2022.
[14] M. Rossi and A. Ferrari, "Decision Support for Hospital Bed Allocation Using Machine Learning," *Healthcare*, 2020.
[15] S. T. R. Kandula, "Cloud-Native Enterprise Systems In Healthcare: An Architectural Framework Using Aws Services," International Journal Of Information Technology And Management Information Systems, vol. 16, no. 2, pp. 1644–1661, Mar. 2025, doi: https://doi.org/10.34218/ijitmis_16_02_103
[15] R. George and M. Thomas, "Handling Imbalanced Health Records for LOS Prediction," *Appl. Soft Comput.*, 2021.
[16] S. Rahman et al., "Predicting ICU LOS Using Deep Neural Networks," *IEEE J. Biomed. Health Inform.*, 2022.
[17] G. Ali and K. Sajid, "A Systematic Review on AI-Based Hospital Operations Optimization," *Inf. Proc. Manag. Health*, 2023.
[18] Todupunuri, A. (2025). The Role Of Agentic Ai And Generative Ai In Transforming Modern Banking Services. American Journal of AI Cyber Computing Management, 5(3), 85-93.
[18] J. White and L. Brown, "SVM-Based Approaches for Length of Stay Forecasting," *Int. J. Adv. Comput. Sci.*, 2020.
[19] G. Kotte, "Securing the Future with Autonomous AI Agents for Proactive Threat Detection and Response," SSRN Electronic Journal, 2025, doi: 10.2139/ssrn.5283830.
[19] Q. Chen and M. Chen, "Dynamic Discharge Planning Using Predictive Models," *Information & Management*, 2022.
[20] R. Fan et al., "Federated Machine Learning for Healthcare Analytics," *IEEE Internet Comput.*, 2023.
[21] S. Kumar and D. Mehta, "Clinical Data Mining for LOS Monitoring and Prediction," *Comput. Biol. Med.*, 2021.

[22] H. Wu et al., “Impact of Treatment Pathways on Hospital LOS,” *Medical Decision Making*, 2020.
[23] Y. Hu and T. Zhang, “Reinforcement Learning for Hospital Resource Scheduling,” *AI in Healthcare*, 2023.
[24] Todupunuri, A. (2022). Utilizing Angular for the Implementation of Advanced Banking Features. Available at SSRN 5283395.
[24] A. Das and P. Paul, “NLP-Enhanced LOS Prediction Based on Physician Notes,” *IEEE Trans. Neural Netw. Learn. Syst.*, 2023.
[25] K. Roberts and H. Malik, “Clinical Deployment Considerations for Explainable ML Models,” *J. Healthc. Eng.*, 2022.