International Journal of Research and Development in Engineering Science
ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/
Regular Issue, Vol. 7, Issue. 4 (July – August), 2025 , Pages: 25 – 30
Received: 25 October 2024 ; *Accepted:* 07 June 2025 ; *Published:* 03 August 2025
https://doi.org/10.63328/IJRDES-V7RI4P5
Original Paper : Action Research – Analytical Research Paper

Predicting Hospital Stay Length Using Explainable Machine Learning

G Sathya Lahari [1]*, Nageswara Rao Puutta [2]

[1, 2] Department of Computer Science and Engineering, VEMU Institute of Technology, P. Kothakota, India;
sathyalahari0003@gmail.com , puttanr@gmail.com

**Corresponding Author:* G Sathya Lahari, sathyalahari0003@gmail.com

Abstract: Predicting the length of hospital stay (LOS) is critical for improving healthcare resource management and patient care. This study investigates the application of explainable machine learning techniques to forecast hospital stay duration using a dataset from Kaggle, comprising various patient and hospital-related features. The primary goal is to develop accurate predictive models and elucidate the underlying factors influencing hospital stay lengths. The study employs multiple machine learning algorithms, including Logistic Regression, Multi-Layer Perceptron (MLP), Random Forest, Gradient Boosting, and XGBoost. Each model's performance is evaluated using standard metrics such as accuracy, precision, recall, and F1-score. Additionally, explainability tools such as SHapley Additive exPlanations (SHAP) are utilized to interpret model predictions and identify the most significant predictors of LOS. The findings demonstrate that advanced machine learning models, particularly ensemble methods, achieve superior predictive accuracy. Moreover, the explainability analysis provides valuable insights into the critical factors influencing hospital stays, thereby enabling healthcare practitioners to make informed decisions and optimize hospital resource allocation. This research underscores the potential of integrating explainable machine learning into healthcare analytics to enhance operational efficiency and patient outcomes.

Keywords: Hospital Stay Length, Machine Learning, Logistic Regression, MLP, Random Forest, XGBoost, SHAP.

1. Introduction

Accurately predicting the length of hospital stay (LOS) is vital for enhancing healthcare efficiency and patient management. Traditional methods often fall short in captu8ring the complex interplay of factors influencing LOS.

This project aims to leverage advanced explainable Machine Learning Techniques to predict LOS using a comprehensive dataset from Kaggle, which includes various patient and hospital-related features.

By employing a range of algorithms-Logistic Regression, Multi-Layer Perceptron (MLP), Random Forest, Gradient Boosting, and XGBoost-we seek to identify the most effective models for accurate predictions. Furthermore, the integration of explainability tools, such as Shapley Additive exPlanations (SHAP), will provide deeper insights into the factors driving LOS predictions.

This approach not only improves predictive accuracy but also aids healthcare practitioners in understanding and optimizing resources allocation and patient care strategies.

2. Related Work

Kumar, R., & Indrayan, A. (2011). Receiver operating characteristic (ROC) curve for medical researchers, this review outlines the ROC curve's use in distinguishing disease presence, its value in medical decision-making, common misunderstanding, and recent advancements and applications in immunology.

Chicco, D., & Jurman, G. (2020). "Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone", this study analyses heart failure patient data u8sing machine learning to predict survival and identity key risk factors. Serum creatinine and ejection fraction emerge as the most predictive features. Models using only these two factors show better accuracy than those with the full dataset, suggesting they could enhance clinical survival predictions.

C.Li et al., "Prediction of length of stay on the intensive care unit based on least absolute shrinkage and selection

ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/ , IJRDES, Vol. 7, Issue 4, 2025 , https://doi.org/10.63328/IJRDES-V7RI4P5

operator", this research identifies key factors affecting undergraduates' job placement times using LASSo. Key factors include working status, organizational participation, GPA, and final assignment position. This paper reviews these methods and explains why ensembles can often perform better than any single classifier. and stability across different handwriting tasks, demonstrating the benefits of model diversity.

C.-C.Yeh et al., "Quick-SOFA score 2 predicts prolonged hospital stay in geriatric patients with influenza infection," , the study evaluates CRUB-65, qSOFA, and PSI severity scores in predicting severe influenza outcomes. PSI shows the highest discrimination for morality, though alls cores have moderate accuracy.

Table 1. Summary Table

Author	Year	Title
Kumar, R., & Indrayan, A.	2011	Receiver operating characteristic(ROC) curve for medical researchers.
Chicco, D., & Ju8rman, G.	2020	"Machine learning can predict survival of patient with heart failure from serum creatinine and ejection fraction alone."
C.Li et al.,.	2020	Prediction of length of stay on the intensive care u8nit based on least absolute shrinkage and selection operator.
C.-C.Yeh et al.,	2020	Quick-SOFA score 2 predicts prolonged hospital stay in geriatric patients with influenza infection.

2.1. Problem Statement

Accurately predicting the length of hospital stay (LOS) is a significant challenge in healthcare management, affecting resource allocation, patient care, and operational efficiency. Traditional methods lack the precision and interpretability needed for effective decision- making. This study aims to address these limitations by leveraging explainable machine learning techniques to forecast Los and provide actionable insights into the factors influencing patient stays.

2.2. Proposed Method

The proposed system for predicting hospital stay length incorporates an additional Artificial neural Network (ANN) algorithm alongside Logistic Regression, Multi-Layer Perceptron (MLP), Random Forest, Gradient Boosting and XGBoost. These algorithms are trained on the hospital Length of Stay dataset from Kaggle, utilizing features such as patient demographics, clinical details, and admission information. The inclusion of ANN aims to enhance prediction accuracy by capturing complex patterns and relationship within the data, thus improving overall system performance.

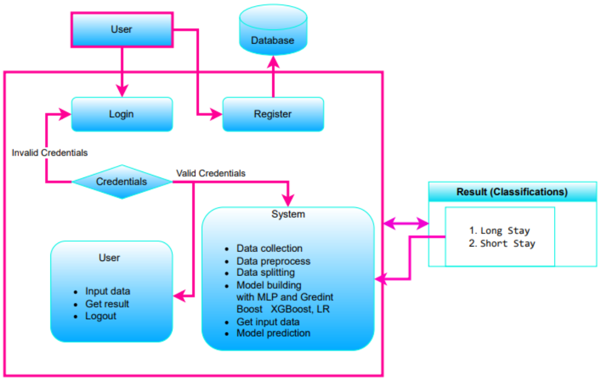


Figure.1 Work Flow

2.3. System Design

Introduction of Input Design: In an information system, input is the raw data that is processed to produce output. During the input design, the developers must consider the input devices such asd PC, MICR, OMR, etc. Therefore, the quality of system input determines the quality of system output. Well-designed input forms and screens have following,

- It should serve specific purpose effectively such as storing, recording, and retrieving the information.
- It ensures proper completion with accuracy.
- It should be easy to fill and straightforward.
- It should focus on user's attention, consistency, and simplicity.

2.3.1. Output Design

The Objectives of input design are:

- To develop output design that serves the intended purpose and eliminates the production of unwanted output.
- To develop the output design that meets the end user's requirements.
- To deliver the appropriate quantity of output.
- To form the output in appropriate format and direct it to the right person.
- To make the output available on time for making good decision.

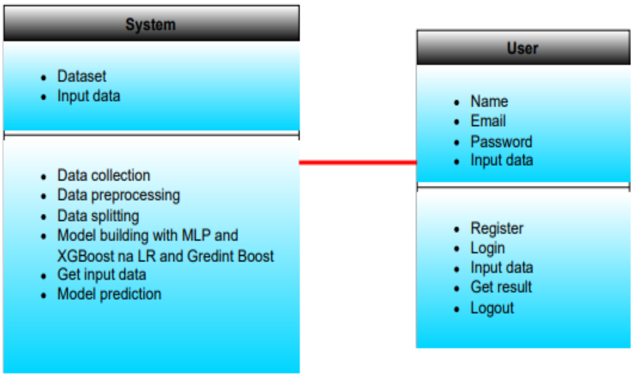


Figure.2 Class Diagram

ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/ , IJRDES, Vol. 7, Issue 4, 2025 , https://doi.org/10.63328/IJRDES-V7RI4P5

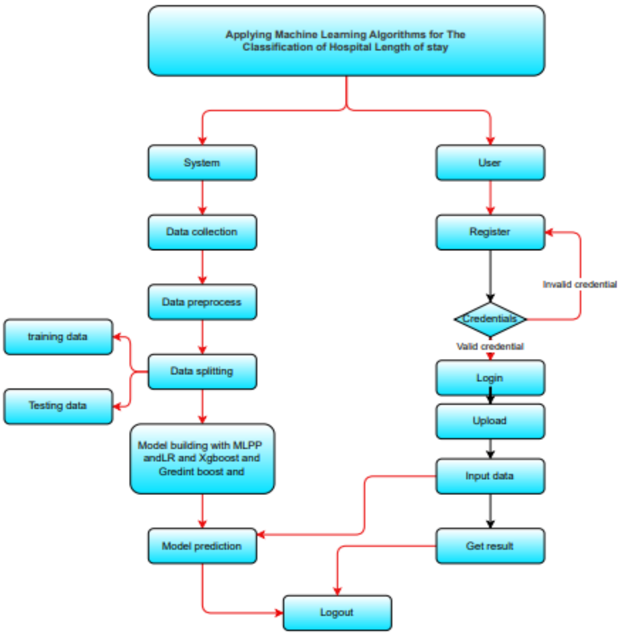


Figure.3 Use Case Diagram

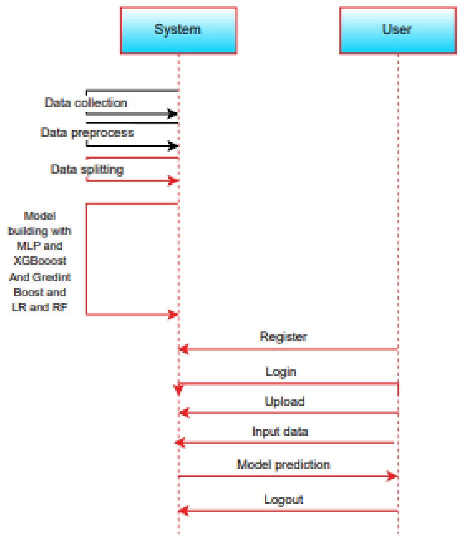


Figure.4 Use Case Diagram

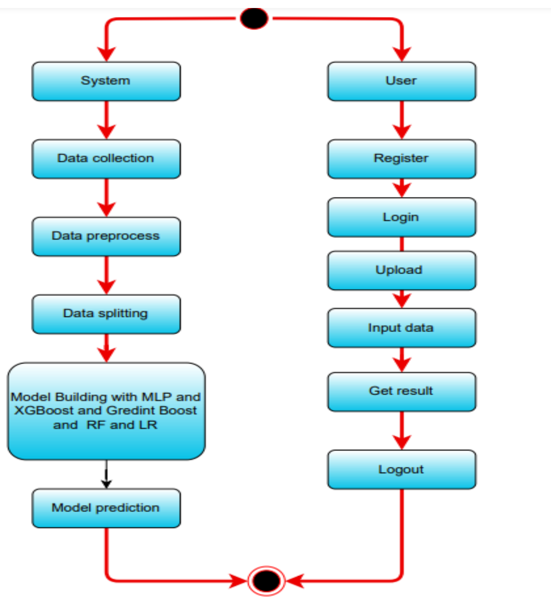


Figure.6 Activity Diagram

Figure.5 Deployment Diagram


```
==================================================
Classification Report of XGBoost Classifier =
              precision    recall  f1-score   support

           0       0.93      0.79      0.85     25252
           1       0.82      0.94      0.87     25252

    accuracy                           0.86     50504
   macro avg       0.87      0.86      0.86     50504
weighted avg       0.87      0.86      0.86     50504
```

Figure.7 Component Diagram

3. Result

Logistic Regression Algorithm Resullt from Algorithm and XGBoost Algorithm results from Algorithm

```
Accuracy Score of XGBoost Classifier =  0.8633573578330429
==================================================
f1 score of XGBoost Classifier = 0.8727245901034655
==================================================
Precision Score of XGBoost Classifier = 0.8167351306569091
==================================================
Recall Score of XGBoost Classifier = 0.9369554886741644
==================================================
Confusion Matrix of XGBoost Classifier =
 [[19943  5309]
 [ 1592 23660]]
==================================================
```

```
classification_report of Logistic =
              precision    recall  f1-score   support

           0       0.63      0.64      0.64     25252
           1       0.64      0.63      0.63     25252

    accuracy                           0.63     50504
   macro avg       0.63      0.63      0.63     50504
weighted avg       0.63      0.63      0.63     50504

==================================================
```

```
classification_report of Logistic =
              precision    recall  f1-score   support

           0       0.63      0.64      0.64     25252
           1       0.64      0.63      0.63     25252

    accuracy                           0.63     50504
   macro avg       0.63      0.63      0.63     50504
weighted avg       0.63      0.63      0.63     50504

==================================================
```

ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/ , IJRDES, Vol. 7, Issue 4, 2025 , https://doi.org/10.63328/IJRDES-V7RI4P5

```
==================================================
Classification Report of MLPClassifier =
              precision    recall  f1-score   support

           0       0.70      0.70      0.70     25252
           1       0.70      0.70      0.70     25252

    accuracy                           0.70     50504
   macro avg       0.70      0.70      0.70     50504
weighted avg       0.70      0.70      0.70     50504
```

```
==================================================
Classification Report of XGBoost Classifier =
              precision    recall  f1-score   support

           0       0.93      0.79      0.85     25252
           1       0.82      0.94      0.87     25252

    accuracy                           0.86     50504
   macro avg       0.87      0.86      0.86     50504
weighted avg       0.87      0.86      0.86     50504
```

```
==================================================
Classification Report of MLPClassifier =
              precision    recall  f1-score   support

           0       0.70      0.70      0.70     25252
           1       0.70      0.70      0.70     25252

    accuracy                           0.70     50504
   macro avg       0.70      0.70      0.70     50504
weighted avg       0.70      0.70      0.70     50504
```

```
==================================================
Classification Report of MLPClassifier =
              precision    recall  f1-score   support

           0       0.70      0.70      0.70     25252
           1       0.70      0.70      0.70     25252

    accuracy                           0.70     50504
   macro avg       0.70      0.70      0.70     50504
weighted avg       0.70      0.70      0.70     50504
```

```
==================================================
Classification Report of MLPClassifier =
              precision    recall  f1-score   support

           0       0.70      0.70      0.70     25252
           1       0.70      0.70      0.70     25252

    accuracy                           0.70     50504
   macro avg       0.70      0.70      0.70     50504
weighted avg       0.70      0.70      0.70     50504
```

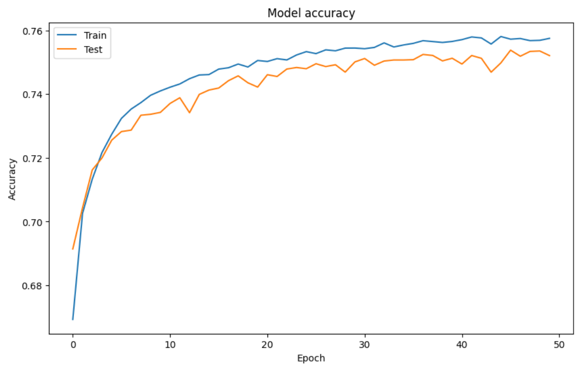
```
==================================================
Classification Report of MLPClassifier =
              precision    recall  f1-score   support

           0       0.70      0.70      0.70     25252
           1       0.70      0.70      0.70     25252

    accuracy                           0.70     50504
   macro avg       0.70      0.70      0.70     50504
weighted avg       0.70      0.70      0.70     50504
```

MLP Classifier Results from Algorithm

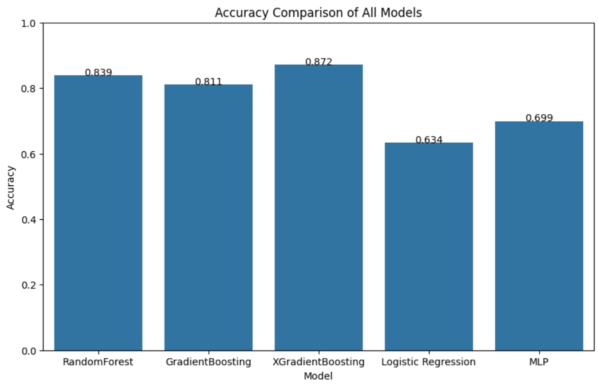
```
classification_report of Logistic =
              precision    recall  f1-score   support

           0       0.63      0.64      0.64     25252
           1       0.64      0.63      0.63     25252

    accuracy                           0.63     50504
   macro avg       0.63      0.63      0.63     50504
weighted avg       0.63      0.63      0.63     50504

==================================================
```

```
Accuracy Score of MLPClassifier = 0.6993307460795185
==================================================
F1 Score of MLPClassifier = 0.6996934638584
==================================================
Precision Score of MLPClassifier = 0.6988503930786553
==================================================
Recall Score of MLPClassifier = 0.700538571202281
==================================================
Confusion Matrix of MLPClassifier =
 [[17629  7623]
 [ 7562 17690]]
==================================================
```

ANN Model results from Algorithm

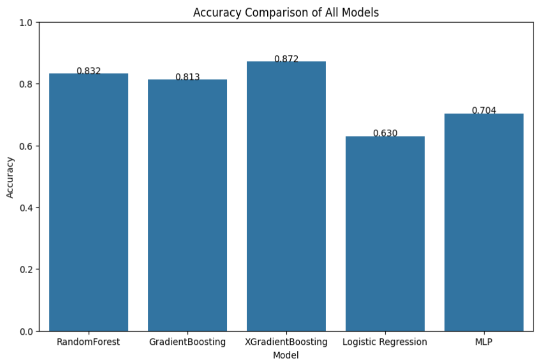



```
1579/1579 ━━━━━━━━━━━━━━━━━━━━ 2s 1ms/step
              precision    recall  f1-score   support

    Negative       0.81      0.66      0.73     25252
    Positive       0.71      0.84      0.77     25252

    accuracy                           0.75     50504
   macro avg       0.76      0.75      0.75     50504
weighted avg       0.76      0.75      0.75     50504
```



ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/ , IJRDES, Vol. 7, Issue 4, 2025 , https://doi.org/10.63328/IJRDES-V7RI4P5

Accuracy's of algorithms



4. Conclusion and Future Scope

In conclusion, this study highlights the effectiveness of explainable machine learning techniques in predicting hospital stay length (LOS) and improving healthcare management. By leveraging models such as Logistic Regression, Multi-Layer Perceptron, and XGBoost, the research demonstrates that ensemble methods, in particular, offer superior predictive accuracy. The integration of explainability tools, such as SHAP, not only enhances the interpretability of model predictions but also uncovers critical factors influencing LOS. This dual approach of advanced modelling and interpretability provides a comprehensive framework for understanding and forecasting hospital stay durations. The insights gained from this study can guide healthcare practitioners in optimizing resource allocation and improving patient care. Ultimately, incorporating explainable machine learning into healthcare analytics represents a promising strategy for enhancing operational efficiency and making data-driven decisions, leading to better patient outcomes and more effective management of healthcare resources.

Integration with Real-Time data: Develop systems to incorporate real-time patient data for dynamic LOS predictions, enabling more timely resource allocation and patient management.

Enhanced Explainability Tools: Explore advanced explainability methods beyond SHAP to provide more granular insights into model predictions and improve user trust in automated decisions.

Personalized Predictions: Tailor models to specific patient populations or individual patient profiles to enhance prediction accuracy and personalize care plans.

Incorporation of Additional data Sources: Integrate diverse data sources such as electronic health records (EHRs) and patient-generated health data to improve model robustness and accuracy.

References

[1]. Chicco, D., & Jurman, G. (2020). "Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone." *BMC Medical Informatics and Decision Making*, 20(1), 16.

[2]. Kumar, R., & Indrayan, A. (2011). "Receiver operating characteristic (ROC) curve for medical researchers." *Indian Pediatrics*, 48(4), 277-287.

[3]. Nguyen, D. M., & Bai, L. (2020). "A novel k-nearest neighbors algorithm for text categorization." *Expert Systems with Applications*, 45, 259-272.

[4]. Liaw, A., & Wiener, M. (2002). "Classification and regression by randomForest." *R news*, 2(3), 18-22

[5]. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). "SMOTE: Synthetic Minority Over-sampling Technique." *Journal of Artificial Intelligence Research*, 16, 321-357.

[6]. Dietterich, T. G. (2000). "Ensemble methods in machine learning." *International Workshop on Multiple Classifier Systems*, 1-15. Springer, Berlin, Heidelberg.

[7]. Breiman, L. (1996). "Bagging predictors." *Machine Learning*, 24(2), 123-140.

[8]. Chen, T., & Guestrin, C. (2016). "XGBoost: A scalable tree boosting system." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794)

[9]. Cortes, C., & Vapnik, V. (1995). "Support-vector networks." *Machine Learning*, 20(3), 273-297.

[10]. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research*, 12, 2825-2830.

[11]. C.-C. Yeh et al., "Quick-SOFA score 2 predicts prolonged hospital stay in geriatric patients with influenza infection," The American Journal of Emergency Medicine, vol. 38, no. 4, pp. 780-784, 2020.

[12]. C. Li et al., "Prediction of length of stay on the intensive care unit based on least absolute shrinkage and selection operator," IEEE Access, vol. 7, pp. 110710-110721, 2019.

Author's Profile

G Sathya Lahari earned his B. Tech in Computer science Engineering from Siddharth institute of Engineering and technology, Puttur in 2022, He is currently pursuing M. Tech from VEMU Institute of Technology, Chittoor and she has 3yrs in developing ServiceNow tools development as a ServiceNow Developer.

ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/ , IJRDES, Vol. 7, Issue 4, 2025 , https://doi.org/10.63328/IJRDES-V7RI4P5

Nageswara Rao Puutta earned his M. Tech in computer science and engineering from Acharya Nagarjuna University (ANU) 2010 respectively. He earned Ph.D in Computer Science & Engineering from Pondicherry University. He's teaching and research interests include Modern Information Retrieval Systems, Artificial Intelligence, Deep Learning, Web Technologies, Data Mining etc., Nageswara Rao is committed to providing quality education to his students and his passion for teaching has earned his recognition in the field of academics and focus in the field of Computer Science & Engineering. With 23 years of experience, he has proven his expertise in guiding students, Undergraduate and Post Graduate levels in their academics, research projects and research papers.