

PREDICTING HOSPITAL STAY LENGTH USING EXPLAINABLE MACHINE LEARNING



Mini project submitted in partial fulfilment of the requirement for the award of the degree

of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING (DATA SCIENCE)

Under the esteemed guidance of

Dr. A. ANIL KUMAR REDDY

Assoc. Professor-HoD (DS)

By

K PRANATHI

22U11A6724



Department of Computer Science and Engineering (Data Science)

Samskruti College of Engineering and Technology

Approved by AICTE & Affiliated to JNTUH, Accredited by NBA and NAAC with A Grade

Kondapur(V), Ghatkesar (M), Medchal District, Telangana – 501301

MAY – 2025

Samskruti College of Engineering and Technology
Approved by AICTE & Affiliated to JNTUH, Accredited by NBA and NAAC with A Grade
Kondapur(V), Ghatkesar (M), Medchal District, Telangana – 501301

Department of Computer Science and Engineering (Data Science)



CERTIFICATE

This is to certify that the B.Tech Mini Project report entitled “**Predicting Hospital Stay Length Using Explainable Machine Learning**” is a bonafide work done by **K Pranathi (22U11A6724)** in partial fulfillment of the requirement of the award for the degree of Bachelor of Technology in “**Computer Science and Engineering (Data Science)**” from Jawaharlal Nehru Technological University, Hyderabad during the year 2024-2025.

Guide Signature

Dr. A. Anil Kumar Reddy, M.Tech, Ph.D,
Assoc. Professor

HoD-CSE(DS)

Dr. A. Anil Kumar Reddy, M.Tech, Ph.D,
Assoc. Professor

External Examiner

Department of Computer Science and Engineering (Data Science)



DECLARATION BY THE CANDIDATE

I, **K Pranathi**, bearing Roll Number **(22U11A6724)** hereby declare that the project report entitled **“Predicting Hospital Stay Length Using Explainable Machine Learning”** is done under the guidance of **Dr. A. Anil Kumar Reddy, M.Tech, PHD, Associate Professor**, Department of Computer Science and Engineering (Data Science), Samskruti College of Engineering and Technology, is submitted in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology in Computer Science and Engineering (Data Science)**.

This is a record of bonafide work carried out by me and the results embodied in this project have not been reproduced or copied from any source. The results embodied in this project report have not been submitted to any other University or Institute for the award of any other degree or diploma.

K Pranathi (22U11A6724)
Department of CSE (Data Science)
Samskruti College of Engineering and Technology

ACKNOWLEDGEMENT

It is with profound gratitude and sincere appreciation that I extend my heartfelt thanks to all those who have played a significant role in the successful completion of this undergraduate project.

First and foremost, I express my deep sense of respect and gratitude to our **Honourable Chairman, Mr. A V Ramana Reddy**, for his constant encouragement and for nurturing a culture of academic excellence and innovation within the institution.

I would also like to express my sincere thanks to **Dr. M. Ramakanth Reddy, Director**, for his visionary leadership and continued support, which have provided the ideal environment and motivation for academic pursuits such as this project.

My heartfelt appreciation goes to **Dr. P Janaki Ramulu, Principal**, for his steadfast guidance, infrastructural support, and encouragement that have helped bring this project to successful fruition.

I am deeply grateful to **Dr. Anil Kumar Reddy, Head of the Department of Computer Science and Engineering (Data Science)**, for his academic leadership, valuable feedback, and continuous support throughout the duration of this project.

A special note of gratitude is reserved for my project guide **Dr. A. Anil Kumar Reddy, M.Tech, Ph.D, Assoc. Professor, Department of Computer Science and Engineering (Data Science)**, whose expert supervision, insightful suggestions, and dedicated mentorship have been instrumental in shaping the direction and outcome of this work.

Lastly, I am ever grateful to my parents and family for their unconditional love, encouragement, and moral support. Their unwavering faith in me has been my greatest strength throughout this journey.

With genuine appreciation, my acknowledge every individual who has, in one way or another, contributed to the successful completion of this project.

With warm regards,

K Pranathi (22U11A6724)

Department of Computer Science and Engineering (Data Science)

Samskruti College of Engineering and Technology Kondapur,

Ghatkesar.

ABSTRACT

Accurate prediction of hospital length of stay (LOS) is critical for efficient resource allocation, improved patient care, and cost management in healthcare systems. This project presents a machine learning-based approach to predict hospital stay duration using patient demographic, clinical, and treatment-related data. Traditional models often act as "black boxes," making it difficult for clinicians to understand the reasoning behind predictions. To address this, we incorporate explainable machine learning techniques, such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations), to enhance transparency and trust. Our system not only predicts LOS with high accuracy using algorithms like Random Forest, XGBoost, and Support Vector Machines but also provides interpretable insights into which features most influence each prediction. The integration of explainability bridges the gap between complex models and clinical decision-making, enabling healthcare professionals to make informed and data-driven interventions.

KEYWORDS:

- Hospital Length of Stay (LOS)
- Machine Learning
- Explainable AI (XAI)
- SHAP
- LIME
- Healthcare Analytics
- Predictive Modeling
- Clinical Decision Support
- Random Forest
- XGBoost
- Patient Data Analysis
- Model Interpretability

LIST OF FIGURES

S. NO.	FIGURE NAME
1	SYSTEM ARCHITECTURE DIAGRAM OF PREDICTING HOSPITAL STAY LENGTH
2	SAMPLE DATA FLOW DIAGRAM (ER)
3	UML CLASS DIAGRAM
4	UML USE CASE DIAGRAM
5	UML SEQUENCE DIAGRAM
6	ACTIVITY DIAGRAM

LIST OF TABLES

S. NO.	TABLE NAME	PAGE NUMBER
1	COMPARATIVE STUDY OF EXISTING VS PROPOSED SYSTEM	8-9
2	TEST CASES AND RESULTS	43
3	SDLC FORMS	55-56

LIST OF ABBREVIATIONS

S. NO.	ABBREVIATION	FULL FORM
1	DLD	Data Flow Diagram
2	UML	Unified Modeling Language
3	SRS	Software Requirements Specification
4	SDLC	Software Development Life Cycle
5	ML	Machine Learning

Table of Contents

S. NO.	CONTENT	PAGE NO.
1	Title Page	i
2	Certificate	ii
3	Declaration	iii
4	Acknowledgement	iv
5	Abstract	v
6	List of Figures	vi
7	List of Tables	vii
8	List of Abbreviations	viii
9	Table of Contents	x

Chapter 1: Introduction

1.1	Overview of the Project	2
1.2	Problem Statement	2
1.3	Objectives of the Project	3
1.4	Scope of the Project	3
1.5	Methodology	4-5
1.6	Organization of the Project	5

Chapter 2: Literature Survey

2.1	Review of Existing System	7
2.2	Limitations of Existing Approaches	7-8
2.3	Need for the Proposed System	8
2.4	Comparative Study	8-9
2.5	Summary	9

Chapter 3: System Analysis

3.1	Feasibility Study	11
3.1.1	Technical Feasibility	11
3.1.2	Economic Feasibility	11
3.1.3	Operational Feasibility	11-12
3.1.4	Time Estimation & Cost Estimation	12
3.2	Software Requirements Specification	12
3.2.1	Software Requirements	12
3.2.2	Hardware Requirements	13
3.3	Functional and non-functional Requirements	13

Chapter 4: System Design

4.1	System Architecture	15
4.2	Database Design	16
4.3	UML Diagrams	18-20
4.4	User Interface Design	21
4.5	Design Standards Followed	22-23
4.6	Safety & Risk Mitigation Measures	23-25

Chapter 5: Implementation

5.1	Technology Stack	26-30
5.2	Module-wise Implementation	30-33
5.3	Code Integration Strategy	33-35
5.4	Sample Code Snippets	35-40

Chapter 6: Testing

6.1	Testing Strategy	42
6.2	Unit Testing	42
6.3	Integration Testing	42
6.4	System Testing	42-43
6.5	Cases and Results	43
6.6	Bug Reporting and Tracking	44

Chapter 7: Results and Discussion

7.1	Output Screenshots	46-48
-----	--------------------	-------

Chapter 8: Conclusion and Future Scope

8.1	Summary of Work Done	50
8.2	Limitations	50
8.3	Challenges Faced	51
8.4	Future Enhancements	51-52

Chapter 9: References

9.1	References	54
-----	------------	----

Chapter 10: Appendices

10.1	Sdlc Forms	56
10.2	Project Timeline	56
10.3	Ethical Considerations	56-57
10.4	Plagiarism Report	57
10.5	Journal Paper Published	57

