A PROJECT REPORT ON

“Hospital Health Prediction System using Machine Learning”

Session: 2025-2026

DEPARTMENT OF INFORMATION TECHNOLOGY

HALDIA INSTITUTE OF TECHNOLOGY

Submitted To:- Debasish Sahoo

Submitted By :-

1. **Subhadip Dey – 10300223170**
2. **Srijani Dinda – 10300223169**
3. **Sathi Pramanik – 10300223137**

Hospital Health Prediction System using Machine Learning

1. Introduction

Modern healthcare systems generate a massive amount of patient data, including vital signs, laboratory values, diagnosis information, hospital admission details, and treatment costs. However, much of this data remains underutilized for predictive decision-making.

The Hospital Health Prediction System is a Machine Learning–based solution designed to analyse hospital patient records and generate predictive insights that can assist doctors and hospital administrators. The system uses structured medical and hospital data to forecast patient health outcomes, disease risks, readmission chances, and expected hospital stay duration.

The system is designed to:

- Predict patient recovery outcome
- Estimate patient survival status
- Identify disease probabilities
- Predict readmission risk
- Estimate treatment cost and recovery days
- Assess overall health risk level

This project integrates Machine Learning models with a Flask backend API and a React frontend to create a complete end-to-end intelligent healthcare prediction system.

2. Objectives of the Project

The primary objectives of this project are:

I. To analyse structured hospital patient data using Machine Learning techniques.
II. To predict multiple healthcare-related outcomes, including:
III. Outcome (Recovered / Not Recovered)
IV. Status (Alive / Deceased)
V. Readmission risk
VI. Disease prediction with top probabilities
VII. Expected recovery days and treatment cost
VIII. To implement an ensemble-based disease prediction model for higher accuracy.
IX. To develop a user-friendly web interface for hospital staff.
X. To demonstrate real-world healthcare decision support using ML.

3. Dataset Description

The dataset used in this project is a hospital dataset containing 10,000 patient records with demographic, clinical, lifestyle, and hospital-related attributes.

Patient Information

Age, Gender, City, Height_cm, Weight_kg

Medical & Vital Data

BP (Blood Pressure), Sugar level, Cholesterol (Chol), Pulse_Rate, Oxygen_Level

Lifestyle Factors

Smoking habit, Alcohol consumption, Stress_Level

Disease & Clinical Information

Disease type, Chronic_Disease indicator

Hospital Information

Days (hospital stay duration), Visits, ICU_Required, Admit_Type, Appointment_Type, Doctor_Specialization

Financial Information

Cost of treatment, Insurance availability

Target Variables

Outcome, Status, Readmitted

Additional derived features such as BMI and Health Risk Level were created using medical rules.

4. System Architecture

The system follows a three-layer architecture:

Frontend – React.js

- Collects patient input through forms
- Displays predictions, summaries, and risk indicators
- Provides interactive output visualization

Backend – Flask API

- Handles requests from frontend
- Loads trained ML models (.pkl files)
- Processes input data and returns predictions
- Machine Learning Layer
- Multiple ML models trained using scikit-learn
- Ensemble logic for disease prediction
- Encoders for categorical outputs

5. Machine Learning Models Used

5.1Outcome Prediction

Target: Outcome

Algorithm: Random Forest Classifier

Reason: Robust for non-linear medical data

5.2 Status Prediction

Target: Patient survival status

Algorithm: Random Forest Classifier

5.3 Readmission Prediction

Target: Readmitted

Algorithm: Logistic Regression

5.4 Disease Prediction (Ensemble Model)

Models combined:

Random Forest

Logistic Regression

K-Nearest Neighbours

Final prediction is based on majority voting and probability averaging. The system also provides top disease probabilities.

5.5 Recovery Days & Cost Prediction

Target: Days (hospital stay), Cost

Algorithm: Random Forest Regressor

Type: Regression

6. Feature Engineering & Risk Scoring

BMI Calculation

BMI = Weight (kg) / Height^2 (m^2)

Health Risk Score

A rule-based risk score was created using:

- Blood Pressure
- Sugar level
- Cholesterol
- Age
- Oxygen level
- Chronic disease

Risk categories:

- Low Risk
- Medium Risk
- High Risk

This improves clinical interpretability.

7. Results and Output

The system provides:

- BMI and BMI category
- Health Risk Level
- Outcome prediction with confidence
- Status prediction
- Readmission probability
- Predicted disease with probabilities
- Expected recovery days
- Estimated treatment cost

All outputs are shown through a clean and interactive user interface.

8. Applications and Use Cases

- Hospital decision support
- ICU resource planning
- Emergency triage
- Insurance risk assessment

- Healthcare analytics dashboards

9. Limitations

- Dataset is synthetic
- Predictions depend on input data quality
- System does not replace doctors
- Accuracy may reduce with highly complex cases

10. Future Enhancements

- Integration with real hospital databases
- Voice-enabled chatbot
- Cloud deployment

11. Conclusion

The Hospital Health Prediction System demonstrates the practical application of Machine Learning in healthcare. By combining multiple ML models, ensemble methods, and a full-stack architecture, the system provides predictive insights that support hospital decision-making and patient care planning.

This project highlights the potential of data-driven healthcare solutions and serves as a foundation for advanced medical AI systems