Predicting Hospital Readmissions Using Machine Learning

1. Introduction

1.1 Problem Statement

The primary goal of this project is to build a predictive model that can identify patients who are at high risk of hospital readmission within 30 days after their initial discharge.

1.2 Business Impact

- Reducing hospital readmissions can improve patient care and hospital efficiency.
- Enables proactive interventions for high-risk patients.
- Helps optimize resource allocation and reduce healthcare costs.

1.3 Dataset Overview

The dataset consists of patient demographics, medical history, hospitalization details, and lab test results. The key columns include:

- **Demographics**: Age, Gender
- **Hospitalization Data**: Admission Type, Number of Outpatient/Inpatient/Emergency Visits
- **Medical History**: Diagnoses, A1C Result, Number of Medications
- **Target Variable**: Readmission (Yes/No)

2. Methodology

2.1 Data Preprocessing

- Handled missing values.
- Encoded categorical features.

- Removed outliers from the feature 'Age'.

2.2 Feature Engineering

- **Age Groups**: Patients were categorized into Child, Young Adult, Adult, Senior, and Elderly.
- **Total Visits**: Combined outpatient, inpatient, and emergency visits.
- **High Diagnosis Count**: Created a binary flag for patients with more than 5 diagnoses.

2.3 Model Selection & Training

- Split data into input and output.
- Split data into training (70%) and testing (30%) sets.
- Used **Logistic Regression** for baseline modeling.

3. Results & Evaluation

3.1 Model Performance Metrics

- **Accuracy**: 51%
- **Precision**: 50%
- **Recall**: 42%
- **F1-Score**: 46%
- **ROC-AUC Score**: 52%

3.2 Feature Importance Analysis

- The most important factors for predicting readmission were:
  - Number of inpatient visits
  - Total number of diagnoses
  - Admission type (emergency admissions had higher readmission rates)

4. Insights & Recommendations

- **Proactive Patient Monitoring**: Patients with frequent inpatient visits and high diagnosis counts should receive extra post-discharge care.
- **Targeted Interventions**: Special follow-up programs for patients with high-risk conditions can help reduce readmissions.
- **Improved Discharge Planning**: Hospitals should ensure proper post-discharge instructions and follow-ups for high-risk patients.