Impact of Medication for Lifestyle Diseases on Hospital Readmission

Table of Contents

Problem Statement

Management of lifestyle diseases in hospitalized patients has a significant bearing on outcome, in terms of both morbidity and mortality. The main object for this problem is to predict whether a patient is likely to be readmitted to hospital based on the previous details of the patient.

Dataset Description

Target attribute: "readmitted" (discrete variable: 2 classes)

The dataset contains 48 features and 66587 rows in train dataset and 16647 data in test dataset. The Column readmission is the target dataset. The description of each feature is as below:

Colum Name	Description	Description and values
encounter_id	Encounter ID	Unique identifier of an encounter
patient_id	Patient number	Unique identifier of a patient
race	Race	Values: Caucasian, Asian, African American, Hispanic, and other

gender	Gender	Values: male, female, and unknown/invalid
age	Age	Grouped in 10-year intervals: 0, 10), 10, 20), …, 90, 100)
weight	Weight	Weight in pounds.
time_in_hospital	Time in hospital	Integer number of days between admission and discharge
medical_specialty	Medical specialty	Integer identifier of a specialty of the admitting physician, corresponding to 84 distinct values, for example, cardiology, internal medicine, family⁄general practice, and surgeon
num_lab_procedures	Number of lab procedures	Number of lab tests performed during the encounter
num_procedures	Number of procedures	Number of procedures (other than lab tests) performed during the encounter
num_medications	Number of medications	Number of distinct generic names administered during the encounter
number_outpatient	Number of outpatient visits	Number of outpatient visits of the patient in the year preceding the encounter
number_emergency	Number of emergency visits	Number of emergency visits of the patient in the year preceding the encounter
number_inpatient	Number of inpatient visits	Number of inpatient visits of the patient in the year preceding the encounter
diag-1 to diag_5	Diagnosis 1	The primary diagnosis values
number_diagnoses	Number of diagnoses	Number of diagnoses entered to the system
change	Change of medications	Indicates if there was a change in diabetic medications (either dosage or generic name). Values: “change” and “no change”
diabetesMed	Diabetes medications	Indicates if there was any diabetic medication prescribed. Values: “yes” and “no”

X1 - X25	25 features for medications	Different Medications and dosages given to the patient X1 and X2 Values: "None" if no administered else the dosages X-3 to X25 Values: “up” if the dosage was increased during the encounter, “down” if the dosage was decreased, “steady” if the dosage did not change, and “no” if the drug was not prescribed
readmitted	Readmitted	"0" - Not Admitted "1" - Admitted

Hackathon Tasks

As part of this hackathon, you are expected to complete three tasks: i) Exploratory Analysis ii) ML Modelling iii) Recommendations

Exploratory analysis

Exploratory Data Analysis using visualizations, numerical analysis, and describing the findings.

- List down the insights/patterns observed from the visualizations
- Explain the impact of the most important attributes on the target attribute observed from data visualizations.

ML Modelling

You are expected to create a robust framework by feature-engineering and improving the baseline ML model performance.

Recommendations

- Any final visualizations you would use to convey your results?
- Can you explain your ML model using non-technical terms?

Evaluation Metric

- The evaluation metric for this hackathon is the **F1 Score**