Diabetes Patient's Hospital Readmission

INFO 6105-18687

Anish Sridhar
Anuja Thawali
Nikhil Prabhu Shankar
Sai Swaroop Reddy Pothireddy

Overview

A hospital readmission is when a patient who is discharged from the hospital, gets re-admitted again within a certain period of time. Hospital readmission rates for certain conditions are now considered an indicator of hospital quality, and also affect the cost of care adversely. For this reason, the Centers for Medicare & Medicaid Services established the Hospital Readmissions Reduction Program which aims to improve the quality of care for patients and reduce health care spending by applying payment penalties to hospitals that have more than expected readmission rates for certain conditions. Although diabetes is not yet included in the penalty measures, the program is regularly adding new disease conditions to the list, now totaling 6 for FY2018. In 2011, American hospitals spent over $41 billion on diabetic patients who got readmitted within 30 days of discharge. Being able to determine the factors that lead to higher readmission in such patients, and correspondingly being able to predict which patients will get readmitted can help hospitals save millions of dollars while improving quality of care.

Objectives

- Train the model to predict what factors are the strongest predictors of hospital readmission in diabetic patients
- Approaches to improve the accuracy of prediction using medical data with various machine learning algorithms and methods
- How well can we predict hospital readmission in this dataset with limited features?

Process

- ➔ Correct the dataset for any empty values and clean the data
- ➔ Identify the training and test dataset
- ➔ Train the data set to predict hospital readmission in diabetic patients based on certain parameters using various algorithms and see which one is giving the best result

Dataset

The data set represents 10 years (1999-2008) of clinical care at 130 US hospitals and integrated delivery networks. It includes over 50 features representing patient and hospital outcomes. Information was extracted from the database for encounters that satisfied the following criteria:

- It is an inpatient encounter (a hospital admission).
- It is a diabetic encounter, that is, one during which any kind of diabetes was entered to the system as a diagnosis.
- The length of stay was at least 1 day and at most 14 days.
- Laboratory tests were performed during the encounter.
- Medications were administered during the encounter.

The data contains such attributes as patient number, gender, age, admission type, time in hospital, medical specialty of admitting physician, number of lab tests performed, HbA1c test result, diagnosis, number of medications, diabetic medications, number of outpatient, inpatient, and emergency visits in the year before the hospitalization, etc.

Contains one csv file:

- ➔ Diabetes dataset: 102K records with 50 columns

Data Process

1. Clean the dataset to have complete, error-free records
2. Remove unnecessary data which has no impact on our results
3. Change data accordingly so that it runs on our algorithms

Resources

www.kaggle.com/datasets

Project Deliverables

Following is a complete list of all project deliverables:

- Python file
- PowerPoint presentation
- Project report
- Link to dataset

Outcomes

➔ Help understanding the major parameters causing diabetes
➔ Factors that are the strongest predictors of hospital readmission in diabetic patients
➔ Prediction of hospital readmission with limited features

Result

Using our model and process, we will be able to successfully predict whether a patient will be re-admitted or not. By being able to predict if a patient will need further care, proper measures can be used to care and be prepared for readmission or further health problems.