Hospital Automation: A Data Driven Approach for Predicting Patient Readmission

Introduction

A major concern for modern hospitals is counting how many patients are readmitted to their hospital within 30 days after an original admission. As the healthcare system moves toward value-based care, many programs to improve the quality of care of patients were developed. One of these programs is called the Hospital Readmission Reduction Program (HRRP), which reduces reimbursement to hospitals with above average readmissions. For those hospitals which are currently penalized under this program, one solution is to create interventions to provide additional assistance to patients with increased risk of readmission. Such an early readmission may be planned or unplanned, but the Medicare Payment Advisory Committee reported that 17.6% of U.S. hospital admissions resulted in readmissions within 30 days of discharge.

Medicare's definition of an early readmission is a readmission that occurred within 30 days of discharge. One patient population that is at increased risk of hospitalization and readmission is that of diabetes. Diabetes is a medical condition that affects approximately 1 in 10 patients in the United States. According to Ostling et al, patients with diabetes have almost double the chance of being hospitalized than the general population.

Problem Statement

Predict if a patient with diabetes will be readmitted to the hospital within 30 days.

Objectives

1. To identify the patients with increased risk of readmission.
2. Build a predictive model that focuses on predicting hospital readmission for patients with diabetes.

Methodology

The methodology for the project must essentially highlight the procedures involved in building a model for predicting patient readmission which include:

- Data Exploration
- Feature Engineering
- Building training and test samples
- Model selection
- Model evaluation

Data Set Information

The dataset 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:

(1) It is an inpatient encounter (a hospital admission).
(2) It is a diabetic encounter, that is, one during which any kind of diabetes was entered to the system as a diagnosis.

(3) The length of stay was at least 1 day and at most 14 days.
(4) Laboratory tests were performed during the encounter.
(5) Medications were administered during the encounter.
The data contains such attributes as patient number, race, gender, age, admission type, time in hospital, medical specialty of admitting physician, number of lab test performed, HbA1c test result, diagnosis, number of medication, diabetic medications, number of outpatient, inpatient, and emergency visits in the year before the hospitalization, etc.