




DAV-Project-Dataset

Ayushi 19/19018

Rachit Mittal 19/19015

Overview

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.

Specifications

Attributes

Total no of attributes are 55.

Attributes are of type -- integers, strings and float

The data contains attributes such 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.

Queries can be formed

- 1- Total number of Females between the age group 20-30, who all are diagnosed with high insulin levels.
- 2- Total number of Males between the age group 30-50, who all are diagnosed with low insulin levels.
- 3- No. of patients expired due to diabetes.
- 4- Patients admissioned to trauma centers between the age group 20-40.
- 5- Distribution of race of admitting physician with respect to speciality
- 6- Females with number of lab procedures > 50
- 7- Number of emergency patients discharged without being inpatient.
- 8- No of new-born babies admissioned due to premature delivery.
- 9- Patients suffering from high levels of glipizide being discharged to home in < 8 days from hospital.

10- Number of expired outpatients vc Number of expired inpatients. ...etc...

Diabetes 130-US hospitals for years 1999-2008 Data Set

Download: [Data Folder](#), [Data Set Description](#)

Abstract: This data has been prepared to analyze factors related to readmission as well as other outcomes pertaining to patients with diabetes.

Data Set Characteristics:	Multivariate	Number of Instances:	100000	Area:	Life
Attribute Characteristics:	Integer	Number of Attributes:	55	Date Donated	2014-05-03
Associated Tasks:	Classification, Clustering	Missing Values?	Yes	Number of Web Hits:	362955

Source:

The data are submitted on behalf of the Center for Clinical and Translational Research, Virginia Commonwealth University, a recipient of NIH CTSA grant UL1 TR00058 and a recipient of the CERNER data. John Clore (jclore@vcu.edu), Krzysztof J. Cios (kcios@vcu.edu), Jon DeShazo (jpdeshazo@vcu.edu), and Beata Strack (strackb@vcu.edu). This data is a de-identified abstract of the Health Facts database (Cerner Corporation, Kansas City, MO).

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.

Link for dataset

<https://archive.ics.uci.edu/ml/machine-learning-databases/00296/>