HDAT9500 Machine Learning I

Data Dictionary: Diabetes inpatient encounters 130 USA hospitals (1999-2008)

Feature Name	Description and values	Format	% Missing
Admission id	Unique identifier of an admission	Numeric	0%
Patient id	Unique identifier of a patient	Numeric	0%
Sex	Female, Male, Unknown/Invalid	Nominal	0%
Age	Age of the patients	Numeric	0%
Weight	Weight in pounds	Numeric	97%
Admission type	Integer identifier corresponding to 9 distinct values (Appendix A)	Nominal	0%
Discharge disposition	Integer identifier corresponding to 29 distinct values (Appendix B)	Nominal	0%
Admission source	Integer identifier corresponding to 21 distinct values (Appendix C)	Nominal	0%
Payer code	Integer identifier corresponding to 23 distinct values, for example, Blue Cross/Blue Shield, Medicare, and self-pay	Nominal	52%
Medical specialty	Including: cardiology, internal medicine and surgeon	Nominal	53%
Glucose serum test result	Indicates the range of the result or if the test was not taken	Nominal	0%
A1C test result	Indicates the range of the results or if the test was not taken. Values: >8 if the result was greater than 8%, >7 if the results was greater than 7% but less than 8%, 'normal' if the results was less than 7% and 'none' if not measured	Nominal	0%
Diagnosis 1	The primary diagnosis (coded as first three digits of ICD-10)	Nominal	0%
Diagnosis 2	Secondary diagnosis (coded as first three digits of ICD-10)	Nominal	0%
Diagnosis 3	Additional diagnosis (coded as first three digits of ICD-10)	Nominal	1%
Length of stay	Integer number of days between admission and discharge	Numeric	0%
Number of diagnoses	Number of diagnoses entered to the system	Numeric	0%
Number of lab procedures	Number of lab test performed during encounter	Numeric	0%
Number of procedures	Number of procedures (other than lab test) performed during the encounter	Numeric	0%
Number of medications	Number of distinct generic names administered during the encounter	Numeric	0%
Number emergency	Number of emergency visits of the patient in the year preceding the encounter	Numeric	0%
Number inpatient	Number of inpatient visits of the patient in the year preceding the encounter	Numeric	0%
Number outpatient	Number of outpatient visits of the patient in the year preceding the encounter	Numeric	0%
Group name 1	Broad categories of Diagnosis 1 code (Appendix E)	Nominal	0%
Group name 2	Broad categories of Diagnosis 2 code (Appendix E)	Nominal	0%
Group name 3	Broad categories of Diagnosis 3 code (Appendix E)	Nominal	0%
Readmission	Inpatient Readmission: 'Yes' if the patient was readmitted in less than 30 days and 'no' if the patient was readmitted in more than 30 days or was not readmitted.	Nominal	0%

Appendix A: Admission type categories

admission_type_id	Original Description	Grouped Description
1	Emergency	Emergency
2	Urgent	Urgent
3	Elective	Elective
4	Not Available	Not available/Null
5	NULL	Not available/Null
6	Trauma Centre	Trauma Centre
7	Not Mapped	Not available/Null

Appendix B: Discharge disposition categories

discharge_disposition_id	Original Description	Grouped Description
1	Discharged to home	Discharged to home
2	Discharged/transferred to another short-term hospital	Short term hospital
3	Discharged/transferred to SNF	Transferred to SNF
4	Discharged/transferred to ICF	Other
5	Discharged/transferred to another type of inpatient care institution	Short term hospital
6	Discharged/transferred to home with home health service	Home health service
7	Left AMA	Other
8	Discharged/transferred to home under care of Home IV provider	Other
9	Admitted as an inpatient to this hospital	Other
10	Neonate discharged to another hospital for neonatal aftercare	Other
11	Expired	Other
12	Still patient or expected to return for outpatient services	Other
13	Hospice / home	Other
14	Hospice / medical facility	Other
15	Discharged/transferred within this institution to Medicare approved swing bed	Other
16	Discharged/transferred/referred another institution for outpatient services	Other
17	Discharged/transferred/referred to this institution for outpatient services	Other
18	NULL	Other
19	Expired at home. Medicaid only, hospice	Other
20	Expired in a medical facility. Medicaid only, hospice	Other
21	Expired, place unknown. Medicaid only, hospice	Other
22	Discharged/transferred to another rehab fac including rehab units of a hospital	Short term hospital
23	Discharged/transferred to a long-term care hospital	Other
24	Discharged/transferred to a nursing facility certified under Medicaid but not certified under Medicare	Other
25	Not Mapped	Other
27	Discharged/transferred to a federal health care facility	Other
28	Discharged/transferred/referred to a psychiatric hospital of psychiatric distinct part unit of a hospital	Other
29	Discharged/transferred to a Critical Access Hospital	Other
30	Elsewhere	Other

Appendix C: Admission source categories

admission_source_id	Original Description	Grouped Description
1	Physician Referral	Physician Referral
2	Clinic Referral	Other
3	HMO Referral	Other
4	Transfer from a hospital	Transfer from another health care facility
5	Transfer from a Skilled Nursing Facility (SNF)	Other
6	Transfer from another health care facility	Transfer from another health care facility
7	Emergency Room	Emergency Room
8	Court/Law Enforcement	Other
9	Not Available	Other
10	Transfer from critical access hospital	Other
11	NULL	Other
12	Not Mapped	Other
13	Transfer from hospital inpt/same fac reslt in a sep claim	Other
14	Transfer from Ambulatory Surgery Center	Other

Appendix E: Group name 1, 2 and 3 categories

Diagnoses ICD10 (1, 2, 3)*	Group Names
A00 – B99	Infectious
C00 – D48	Cancer
D50 – D89	Blood and Immune
E00 – E90	Endocrine
F00 – F99	Mental and Substances
G00 – G99	CNS
I00 – I99	Cardiac and Circulatory
J00 – J99	Respiratory
K00 – K93	Digestive
X60 – X84	Suicide
X85 – Y09	Homicide
ICD-10 codes not mentioned above	Other

Original source of the dataset: https://archive.ics.uci.edu/ml/datasets/Diabetes+130-US+hospitals+for+years+1999-2008

*The original dataset contained ICD9. We converted the 9tn version to ICD10.