**Table 6 Explanation of features**

|  |  |  |
| --- | --- | --- |
| Feature name | Type | Description and values |
| encounter\_id | Numeric | unique identifier of a patient visit |
| patient\_nbr | Numeric | unique identifier of a patient |
| race | Categorical | values: Caucasian, Asian, African American, Hispanic, and other |
| gender | Categorical | values: male, female, and unknown/invalid |
| age | Categorical | grouped in 10-year intervals up to 100 |
| weight | Numeric | patient weight in pounds |
| admission\_type\_id | Categorical | integer indicating emergency, urgent, elective, newborn, and not available |
| discharge\_disposition\_id | Categorical | integer indicating discharged to home, expired, and not available |
| admission\_source\_id | Categorical | integer indicating physician referral, emergency room, and transfer from a hospital |
| time\_in\_hospital | Numeric | number of days between admission and discharge |
| payer\_code | Categorical |  |
| medical\_specialty | Categorical |  |
| num\_lab\_procedures | Numeric | number of procedures (other than lab tests) performed |
| num\_procedures | Numeric | number of medications administered during the encounter |
| num\_medications | Numeric | number of medications administered during the encounter |
| number\_outpatient | Numeric | for year preceding the encounter |
| number\_emergency | Numeric | for year preceding the encounter |
| number\_inpatient | Numeric | for year preceding the encounter |
| diag\_1 | Categorical | primary diagnosis, 848 distinct values/codes |
| diag\_2 | Categorical | secondary diagnosis, 923 distinct values/codes |
| diag\_3 | Categorical | additional secondary diagnosis, 954 distinct values/codes |
| number\_diagnoses | Numeric |  |
| max\_glu\_serum | Categorical | indicates the range of the result or if the test was not taken |
| A1Cresult | Categorical | indicates the range of the result or if the test was not taken |
| change | Categorical | dichotomous indicator of change in diabetic medications |
| diabetesMed | Categorical | dichotomous indicator of diabetic medication being prescribed |
| 18 features for medications | Categorical | metformin, repaglinide, nateglinide, chlorpropamide, glimepiride, glipizide, glyburide, tolbutamide, pioglitazone, rosiglitazone, acarbose, miglitol, troglitazone, tolazamide, insulin, glyburide-metformin, metforminrosiglitazone, metformin-pioglitazone  one variable exists for each medication; values indicate increased, decreased, held or no usage |