Predict whether a patient will be readmitted or not?

Presenter :
Anushi Shah

Hospital readmission : A hospitalization that occurs within 30 days after a discharge.


(Within 30 days)
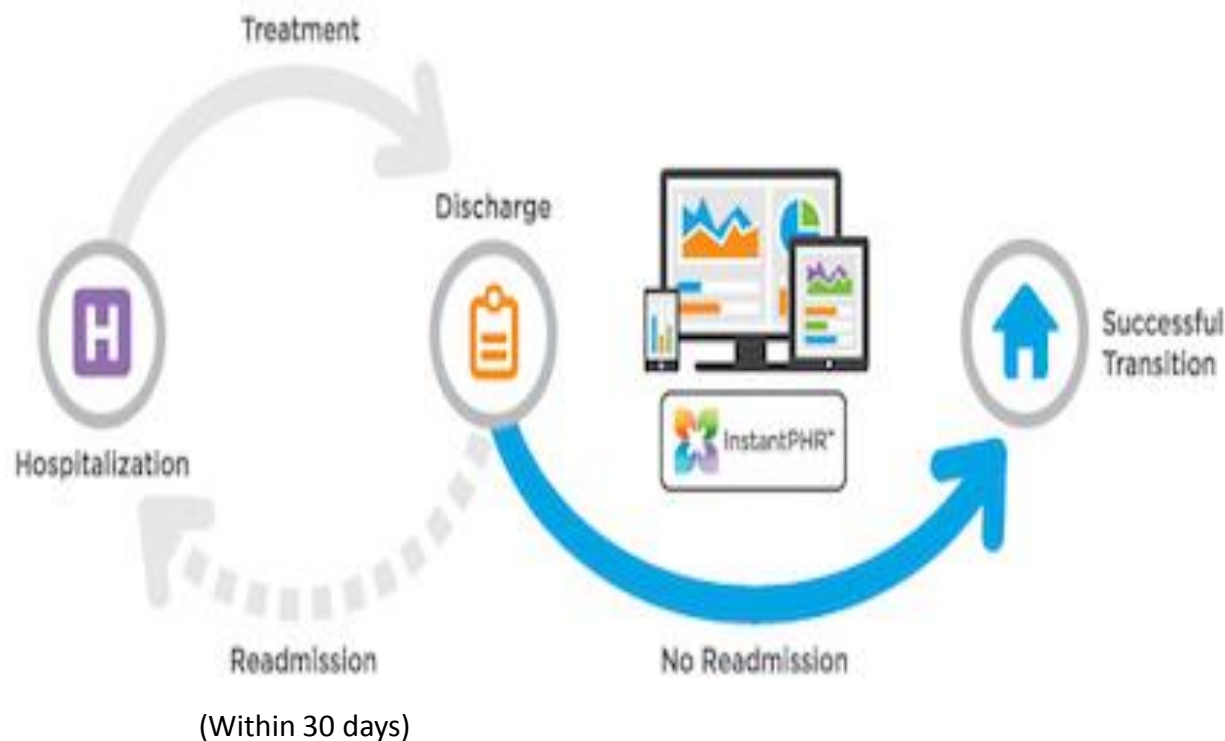
Why are readmissions a problem?
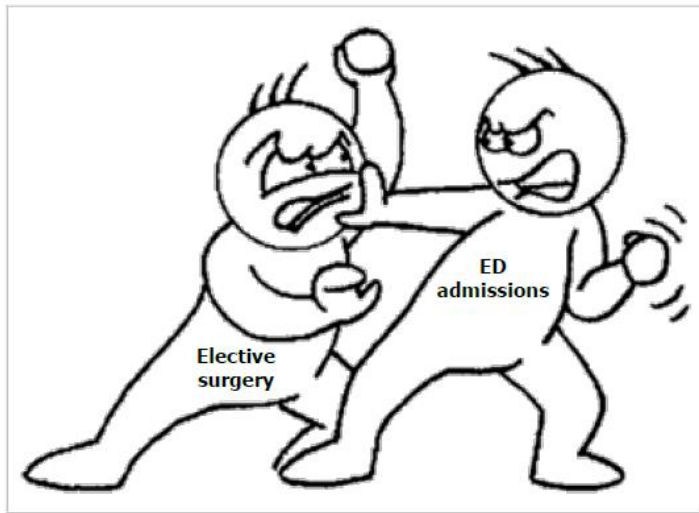
- Increased wait times
- Increased medical erros & negligence
- Patient health at risk

- Huge cost burden
- 2011(US) – Readmissions cost $41 billion
- High readmission rate – Poor quality of care

Data set

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

- Center for Clinical and Translational Research, Virginia Commonwealth University

- UCI Machine learning repository

- 101766 observations, 55 features

Data exploration
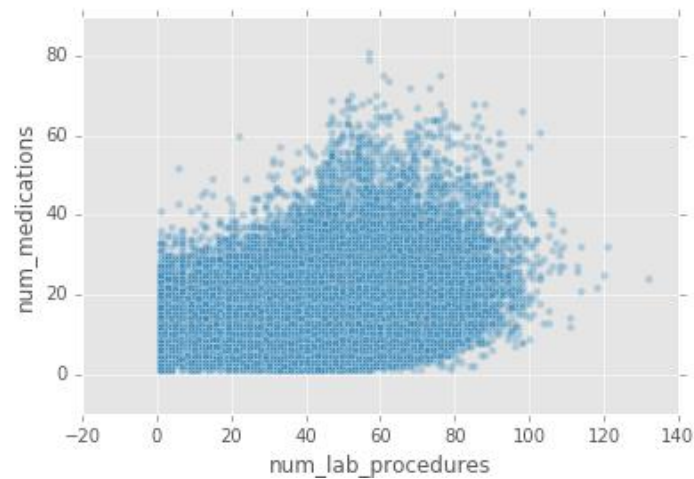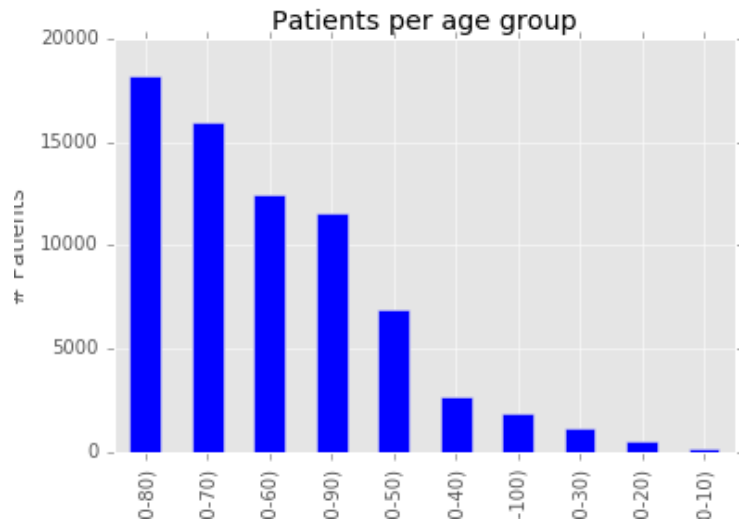
Data pre-processing

- Drop :

    - irrelevant columns (Payer_code)

    - features with high % of missing values : weight (97%), medical_speciality (50%)

    - rows with missing values.

- Transformation of categorical features

- Use only first hospitalization per patient

Modelling Techniques

Total sample size : 71518

Split data set : Training(80%), Test(20%)

Target/Outcome : Readmitted (1/0)

Models :

- Logistic regression

- Decision trees

- Random Forest

- Xgboost

Results

Classifier	Accuracy
Logistic regression	0.624
Decision Trees	0.625
Random Forest	0.62

index	Feature_name	Rank
2	num_lab_procedures	0.133361
1	num_medications	0.118097
0	time_in_hospital	0.083822
6	number_diagnoses	0.058923
3	num_procedures	0.05654
5	number_inpatient	0.02407
4	number_outpatient	0.02059
38	diag_1_map_1.0	0.017792
53	Age_new_7	0.016179
10	race_new_2.0	0.014649
54	Age_new_8	0.014135
52	Age_new_6	0.014035
119	insulin_new_1	0.013937
46	diag_1_map_9.0	0.013926
39	diag_1_map_2.0	0.013565
8	race_new_0.0	0.012527
55	Age_new_9	0.012487
118	insulin_new_0	0.012196
61	A1Cresult_new_0	0.012081

Learning

- Models

- Statistical & theoretical concepts

- Detailed level features may be helpful:

    - procedures (which ones, when performed)

    - medications (dose, duration)

    - BMI

    - medical speciality

    - Blood pressure

    - Electronic medical records with past history, family history

Future steps

- Feature engineering & more investigation

- Try other models

- Try similar datasets

- Dashboard

Thank You !!

☺

Life is not a piece of cake and so is machine learning.

Keep persisting!!