

2019

Research Experience for Undergraduates

**Predicting hospital readmission for patients
with multiple chronic conditions**

Ayzhamal Zhamangaraeva

Co-Advisors: Ioannis A. Kakadiaris and Dan Price

Motivation

Decreasing readmission rates will

- Decrease health care costs
- Help hospitals to avoid Medicare readmission penalties (since October 1, 2012)
- Improve patient care

Statistics

- In 2015, 2,592 US hospitals out of 5,627 registered hospitals received penalties from the CMS (losing a combined \$420 million)
- Historically, nearly 20% of all Medicare discharges had a readmission within 30 days.

Goal

To develop, implement and evaluate an algorithm to predict hospital readmission for patients with multiple chronic conditions

Objectives

1. Prepare the data to which the analysis will be based
2. Develop a prediction model
3. Evaluate the prediction model

Expected Impact

- To adjust the care of an individual with a high risk of readmission
 - Reduce costs
 - Improve quality of life

Deliverables

1. Dataset with computed features
2. Source code for models
3. Literature review XLS and report
4. Final report

Original Humana dataset

Includes 4 tables:

Med (57G) 716,464,506 rows x 35 columns
Lab (69G) x 13 columns
Rx (120G) 662,379,439 rows x 22 columns
Pat (1.3G) 12,913,657 rows x 42 columns

Deidentified records of three years

01/2013-12/2015

Challenge: 2 files out of 4 are corrupted.

Objective 1: Tasks

1. Define my cohort
2. Write R script to filter and analyze cohort
3. Compare positives and negatives
4. Derive new features from the comparison
5. Write R script to compute new features
6. Partition to training, testing, and validation datasets

Objective 1: Remaining Work

- ✓ 1. Define my cohort
2. Write R script to filter and analyze cohort
3. Compare positives and negatives
4. Derive new features from the comparison
5. Write R script to compute new features
6. Partition to training, testing, and validation datasets

Objective 2: Tasks

1. Conduct literature review on hospital readmission; highlight commonly used methods and features
2. Implement SVM
3. Implement RF

Objective 2: Results

Publication	Year	Methods	Features
Evaluating Patient Readmission Risk: A Predictive Analytics Approach	2018	SVM, RF, Gradient Boost	55 (HbA1c, Gender, Discharge disposition, Admission Source, Primary diagnosis, Race, Age, Time in hospital)
Assessment of Machine Learning vs Standard Prediction Rules for Predicting Hospital readmission	2019	CNN	382 including demographic data (sex, race, hospital service)
An integrated machine learning framework for hospital readmission	2018	DNN, SVM	demographic, social and economic status, treatment and clinical, health care utilization
Prediction modeling and pattern recognition for patient readmission	2016	FC NN, Regression	130 (patient data, claims data, drug count data, lab count data, outcome data)

Objective 2: Remaining Work

- ✓ 1. Conduct literature review on hospital readmission; highlight commonly used methods and features
- 2. Implement SVM
- 3. Implement RF

Objective 3: Tasks

1. Identify the failure cases
2. Improve features and iterate for a better accuracy and AUC

Objective 3: Remaining Work

1. Identify the failure cases
2. Improve features and iterate for a better accuracy and AUC

Self-reflection

1. Data pre-processing is a laborious task
2. Medical data is complex and hard to understand

Acknowledgements

The REU project is sponsored by NSF under award NSF-1659755. Special thanks to the following UH offices for providing financial support to the project: Department of Computer Science; College of Natural Sciences and Mathematics; Dean of Graduate and Professional Studies; VP for Research; and the Provost's Office. The views and conclusions contained in this presentation are those of the author and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the sponsors.