

ORIE 4741 Project Proposal

Anusha Avyukt (aa2686), Fei Xia (fx43), Siyang Liu (sl2687)

September 29th, 2019

1 Problem

When seeking medical treatment, there are constraints posed by health insurance provider to visit certain hospitals covered by the insurance policy. Along with the complex healthcare system, this makes it difficult for the patient to select the hospital for required treatment. By making this selection process easier, we would improve the decision at critical times. We plan to improve the process of hospital selection using analysis of the data provided by NY state, for in-patient visits across all hospitals in the state.

Our focus is on hospital selection for specific situations, like emergency cases, severe chronic diseases, small diseases. Our feature space considers severity of illness, cost, number of procedures, distance etc, combined with other metrics such as re-admissions rate and hospital quality. We will also explore the possibility of selecting the hospital based on pre-diagnosis of the disease from the patient's health reports.

2 Dataset

The Statewide Planning and Research Cooperative System (SPARCS) inpatient de-Identified dataset provides details on patient characteristics, diagnoses, treatments, services, and charges. This data contains basic record level detail regarding the discharge; however, the data does not contain protected health information (PHI) under the Health Insurance Portability and Accountability Act (HIPAA). Till now, there are 2.54 million rows and 34 columns entries in the dataset. Link to the datasets:

<https://health.data.ny.gov/Health/Hospital-Inpatient-Discharges-SPARCS-De-Identified/u4ud-w55t>

<https://www.kaggle.com/c/pf2012>

MEPS is a set of large-scale surveys of families and individuals, their medical providers, and employers across the United States. MEPS is the most complete source of data on the cost and use of health care and health insurance coverage. Link to the datasets:

<https://meps.ahrq.gov/mepsweb/>

3 Significance

Finding a hospital for a given condition can be particularly challenging when looking for affordable care. With the constraints imposed by health insurance, it is important to identify suitable options for treatment. We hope to gain some insights by analyzing the SPARCS dataset and have a data driven way of choosing the hospital to visit.