ORIE 5741/4741 Project Proposal

Ruobing Shui, Hehong Li, Ruize Ren

Spend one minute, and think which hospital you'd like to visit when you feel uncomfortable. Are there pounds of information rushing into your brain? There are 6,090 hospitals in US.[1] If we narrow down to New York city, there are still 214 hospitals.[2] The question of how to make right decisions always annoys people, even when there are only three choices.

The process of making decisions of the hospitals normally costs time and money, sometimes even lives. There are several barriers for people to pick out the best option. Not only the tons of hospitals, but also information barriers stop people from making decisions properly. Besides, due to the different education level and poor health announcements, many people even have no idea what the ill they get is and how serious it is, which might result in a tragedy.

Our project goal is to create a predictable model to help people make smart decisions.

The question of choosing the most suitable hospitals for the patients is raised based on the considerations from different aspects, especially their health conditions. After the analysis in our project, the user will get a customized plan and some suggestions on which hospitals are appropriate to visit from a huge amount of them. To make a comprehensive decision, the question will be not only limited to the medical problems of the patients, but also some of their personal conditions, such as insurance, locations and their outcomes.

Based on these questions, we specifically choose the dataset from the Hospital Inpatient Discharges (SPARCS De-Identified): 2012.[3] This dataset could provide various areas in the New York State. Since our goal is to help people living in New York State make smart decisions. Starting from the individuals, this data set contains their final choice of the hospital, their specific health issues, other factors such as estimated cost, payment type, APR information that might affect their final decision.

The dataset has already supported all the features we considered for our model. If there are several other features not included we'd like to put into our model in the process, we'll explore extra datasets in relevant hospital associations. If the classification is too specific and data distribution is sparse, then we can merge these specific classifications into a more general feature

1. *Fast Facts on U.S. Hospitals, 2021, The American Hospital Association, Retrieved Sep 30th, 2021 from https://www.aha.org/statistics/fast-facts-us-hospitals*
2. *NYS Health Profiles. Retrieved Sep 30th, 2021, from https://profiles.health.ny.gov/hospital/bed_type/Total+Beds*
3. *Hospital Inpatient Discharges (SPARCS De-Identified): 2012, Retrieved Sep 30th, 2021, from https://health.data.ny.gov/Health/Hospital-Inpatient-Discharges-SPARCS-De-Identified/u4ud-w55t*