

Healthcare infrastructures in the United States have been debated for long time and were recently reformed to provide more affordable service to the public. The government noted that one of the drivers of high healthcare premiums was high readmission rate in some hospitals. Thereby, the reform implemented scoring paradigms to evaluate hospital performances; namely, the readmission score. The high rate not only indicates that the initial treatment is inappropriate or insufficient, but also forces insurance companies to compensate for unexpected expenses. Consequently, the cost of healthcare premiums inevitably increases, making healthcare more expensive and unaffordable.

To investigate the readmission rate issue, I hypothesize that there may be a relationship among several financial factors and the readmission rate, i.e. low finance may correlate to low hospital performance. The factors are: 1) annual payroll by all industries per county, 2) total number of business establishments per county, 3) healthcare expenditures by a county government, 4) median income by a household, 5) municipal healthcare expenditures, 6) poverty level per county, 7) general revenue of a county government, and 8) state where a county locates. An outcome of interest is hospital performance per county (outcome = *better*, *national average*, or *worse*) evaluated by the aggregated score of within-30-day readmission and death rates in four categories: heart attack, heart failure, pneumonia, and overall-readmission rate per hospital.

The best model, based on financial factors, was able to predict hospital performance per county with an accuracy of 80%. However, this analysis has a serious shortcoming – imbalanced numbers of the outcome. Approximately, 75 % of outcome was *national average*, 16 % was *better*, and 9% was *worse*. To mitigate the imbalance, I trained my models with downsampled *national average* data, while the others were unmodified. However, it is inevitable to induce distortion of modeling and results due to probable inaccurate representation of *national average* data via downsampling.

As a conclusion, a simple model based on financial factors can give us some insights into ties between expected hospital performance and regional financial status. Predictions by the current model are certainly insufficient. Thus, to improve modeling and/or represent data with more authenticity, for example, ensemble prediction may be adequate, where predictions are made by several models and a final prediction is determined by averaging those predictions.