

Capstone Project Ideas:

1) Recommend a Hospital Based on Yelp Overall Reviews in State of California

Dataset: Collect hospital reviews from Yelp.com.

Description of dataset: Reviews are textual data.

Question: Can we recommend a hospital to patients based on overall hospital ratings in state of California?

Method: Use sentence sentiment classifier to calculate overall hospital ratings from textual reviews.

2) Predict California Hospital Profitability

Dataset: Hospital Profitability in 2009-2013 from <http://www.healthdata.gov/dataset/hospital-profitability-2009-2013>.

Description of dataset: The dataset contains income statement information for all licensed, comparable hospitals in the state of California. Kaiser hospitals, state mental hospitals, psychiatric health facilities, and hospitals with mainly long-term care patients are excluded. Deductions from Revenue, Net Patient Revenue, Net from Operations (Operating Revenue less Operating Expense), and Net Income for public hospitals has been adjusted for Disproportionate Share intergovernmental transfers for funding the Disproportionate Share Hospital Program. The program gets federal matching funds to pay supplemental payments to hospitals with a disproportionate share of uninsured, underinsured, and Medi-Cal patients.

Question: Can we predict future hospital profitability using the data from 2009-2013?

Method: Use classification to predict what funds contribute to hospital profitability.

3) Prediction of Future Currency Exchange Rate

Dataset: Collect data from Yahoo Finance for currency exchange.

Description of dataset: Time series data.

Method: Time series analysis

Question: Can we predict future currency exchange rate, does it depend on overall market performance?