Medical Insurance Costs

Team Members: Karina Hutula, Ling Lv, Kyle Shimberg, Ranjani Anjur Venkatraman, Dawn Washington

Business Objective: Predicting medical insurance costs for potential employees.

Overview: As a new data startup, we have investigated many avenues to begin making money. The medical industry is one of the most lucrative, so we decided to go that route. After initial data exploration, we have discovered a huge gap for employers – predicting medical insurance costs. This is a large expense for all businesses that provide healthcare coverage. We will be using existing basic personal factors and insurance cost data to generate a machine learning model to help companies predict a potential employee’s insurance impact. This is a huge opportunity for our startup to make a big impact in the data science world.

Dataset Used: Medical Cost Personal Datasets from Kaggle (https://www.kaggle.com/mirichoi0218/insurance)

Dataset Limitations: The data is not recently collected and not upkept. The dataset is small (less than 1400 rows and only 7 columns).

Tasks:

* Read in the data
* Exploratory data analysis
* Data preprocessing
* Machine learning models
* Model deployment
* Web page design
* Heroku deployment
* README
* Slide deck