Group Members:

Junhua Deng, [Siddharth Saravanan](mailto:siddharth.saravanan@gwmail.gwu.edu), Amogh Ramagiri, Pramod Krishnachari

Research Topic

Our group is interested in finding lifestyles’ impacts on diabetes, or other chronics, in the United States. The data were collected by CDC. Dataset contains 21 different features and one indicator of diabetes. Features include basic physical status, lifestyles, gender, age, education, and income of each participant in the study. This dataset also includes high blood pressure/cholesterol level, heart attack, and stroke in binary form, so we can use them as indicators in our exploration.

SMART Questions

* Can eating vegetables reduce your chances of diabetes? What about fruits?
* Which habit causes more risk in high blood pressure? Consuming alcohol or smoking?

Source of Data

<https://archive.ics.uci.edu/dataset/891/cdc+diabetes+health+indicators>

Github Repository

Modeling Methods

* Logistic regression
* Random Forest
* LightGBM
* XGBoost
* etc.