New York University Shanghai

Suggested Final Project List

BUSF-SHU 210: Business Analytics (Spring 2019)

The final project counts for 40% of the course grade for Business Analytics. This document provides a suggested list of project topics. If you want to choose a topic beyond this list, please talk to me and get my approval. For each of the project listed below (and a project topic of your own choice), you may need to collect your own data set to address the questions of interest.

You should email me your choice of final project topic by **10:00pm, Sunday, April 14**.

Local/Heterogeneous Treatment Effect

Most causal inference method assumes constant causal effect, i.e., the treatment effect is independent of covariates. In practice, however, the causal effect of treatment may be different for different subjects. Applying the idea of decision trees, Athey and Imbens (2015) propose a method that could efficiently estimate heterogeneous treatment effects for subjects with differentiated covariates. Your task is to implement the method proposed by Athey and Imbens (2015). Furthermore, you need to use real or simulated data to validate (or invalidate) their method. You may use the simple linear regression with cross terms of the treatment assignment and some covariates as the benchmark.

Multi-Armed Bandit Experiments

The multi-armed bandit framework has been widely adopted by tech firms to run online experiments (aka A/B test). See

`https://en.wikipedia.org/wiki/Multi-armed_bandit` for a general description of multi-armed bandit and
`https://support.google.com/analytics/answer/2844870?hl=en&ref_topic=1745207`
for a description of Google's online experiment platform using the multi-armed bandit approach. We will also devote one session of our class to introduce the basics of multi-armed bandit. Your task is to build a multi-armed bandit protocol for online experiments. Use simulation to demonstrate that multi-armed bandit could identify the causal effect of multiple treatments more efficiently (faster and with fewer subjects) than the standard randomized experiments (A/B test).

Elevator Schedule and Routing Design

The elevators at skyscrapers are generally very crowded during rush hours. Your task is to design the schedule and routing of elevators, with the objective that all employees working in the skyscraper could get to their workspace as fast as possible. Alternatively, you may optimize the waiting time of the employees. It is suggested that you to consider the uncertainties in both the origin and destination of each employee, as well as the randomness in arrival time. You may compare different strategies such as some elevators only stop at certain levels or some other possible stopping roles.

Money Ball

Data analytics have been extensively used by professional sports teams to train and evaluate sports players and build the team roster. MIT Sloan even holds an annual conference to promote the communications of sports analytics: `http://www.sloansportsconference.com/` You may refer to Chapter 4 of Analytics Edge for a mild introduction of the monetary ball theory. Your task is to build your own "money ball theory". Specifically, you choose a professional sports of your own interest. Then, you need to build model(s) to understand the players' past performances as well as their contribution to the team. Eventually, the final output of the project is, given a budget, constructing a team with the highest probability to win the next season's title of the sports league you consider (assuming the salary of each player is the same as the current season). You may obtain data from the sports reference website: `https://www.sports-reference.com/`

Reference

Athey, S., G. Imbens. 2015. Recursive partitioning for heterogeneous causal effects. *PNAS*. **117**(27) 7353-7360.