PLA6113 Exploring Urban Data with Machine Learning

Session 1 - Introduction

Spring 2023

Mauricio Rada

About me

- Solutions Analysis at UrbanFootprint

- Data Consultant (Global Internship Program) at the International Finance Corporation (private sector arm of WB)

- M.S. in Urban Planning at Columbia, B.S. in IE at Universidad del Pacifico [Peru]

- Research focused on applications of analytics and ML in data scarcity and urban design.

- Worked in projects related to analytics with the IADB, MIT CTL, CU DSI

Introduce yourself

Logistics

- Monday 9 - 11 am

- In-person (UP Lab)

- Email: mer2245@columbia.edu

- OH: Wednesday 10 - 11 am (Schedule)

About the course

- 3 credit elective for masters students

- Urban Analytics course aimed to provide a toolbox and quantitative thinking framework (and critical view) based on machine learning

- Applications are centered around urban topics including housing, transportation, city operations, etc.

- Course is motivated by real-world problems through assignments, case study, and the project.

Grading

- Assignments (8)   30%
  - Short exercises every week or every two weeks
- Case Study (1)   10%
  - 15 min presentation about a relevant case of application of Urban ML
- Project (proposal)   10%
- Project (midterm)   20%
- Project (final)   30%
- Participation   0%
  - Being engaged and a positive influence for the class
  - Ask questions and provide comments!
  - I would appreciate if you send a quick message if you cannot attend

Class structure

- 15 min of case study presentations

- Around 45 min of theory and discussion

    - Usually hand-written lecture, slides from time to time. Though they
      will be available.

- 45 min of practice in Python

- 10 min to decompress and reflect

Why take this class

- Start learning Python

- Understand judiciously apply different ML algorithms

- Develop a sense for exploring quantitative questions

- Implement (part of) the ML process

- Be able to interpret the results

- Use real-world data

- Start developing a programming portfolio

Why not take this class

- Bad choice for minimum workload

- Bad choice if you are not comfortable with data uncertainty

- Bad choice if you look for a theory-intensive machine learning/math/statistics class

Course structure

- Probability Theory (1)
- Supervised Learning Fundamentals (2)
- Unsupervised Learning Fundamentals (2)
- Advanced SL algorithms (2)
- Advanced UL algorithms (2)
- Guest lectures
- *Operations Research Basic Models

Summary

- This class aims to be fun, useful and challenging.

- Feedback is always welcomed

    - Send me an email or attend my OH!

- Please

    - Complete the survey

    - Register for the case study presentations

    - Install required software