PLA6113 Exploring Urban Data with Machine Learning

Final guideline

- Final presentation: **04/27/2022** in class
- Final packet (paper, presentation slides, and technical documentation): **05/04/2022 11:59pm**

For the final, you will be asked to do the following:

- **Main analysis** - applications of machine learning
  The analysis must implement at least one machine learning algorithm discussed in our class. You may also use techniques we don't explicitly cover, but do not stray too far from the course topics.

- **Your full paper OR poster AND the presentation should include the following** (You can use some/all paragraphs of your midterm paper to the final):
  - Introduction: introduction to the urban topic/problem at-hand you are aiming to solve/answer
  - Literature review and theoretical framework
  - Methodologies: clear description of the specific question(s) being asked/hypotheses being tested, data sources, variables, and methodology
  - Results and implications: results (visualized in some manner) and insights/explanations from the analyses. You should also address the implications of your findings for urban operations, policy, and/or planning.
  - Conclusions and Next Steps: conclusion synthesizing your analysis/insights, laying out the limitations and commenting on how you could improve the analysis, etc. with additional data and opportunities for future analysis
  - References: bibliography of works cited in APA format

  A final paper is expected to be **a total of 15 pages**, including figures, tables, maps, and reference. Also, you will be asked to give a **8 minutes presentation** - if you don't keep time, points will be deducted.

- **Technical documentation** (**data and code**)
  As we have discussed, reproducibility is an important concept of data analytics. As a proper practice, you should try to organize your data analytic environment. Please assume that you will collaborate with someone and share a structured technical documentation including data and Jupyter notebooks with appropriate markdown or notes. Please submit your folder (compressed) used for the final project analysis. This folder should include at least a data folder and script folder, and Jupyther notebooks should be run without error.

Rubric

The final will be graded on the following criteria:

- Methodological framework and application of machine learning techniques (15)
  - Clear description of the specific question(s) being asked/hypotheses being tested (2)
  - Data sources and variables (3)
  - Clear description and implementation of ML algorithms (5)
  - Model evaluation (3)
  - Findings and implications (2)
- Is paper clearly written and intelligently formulated? (5)
- Communication through the presentation (5)
- Is the technical documentation suitable for reproducibility? (5)