Project Guidelines

Prof. Oke

CEE 697M: Data Mining and Machine Learning for Engineers *04.14.2022*

Please use the following guidelines as you work on your project and prepare your presentations (date TBD).

Format

I will email the order of presentations in an Announcement a few days before. You will have **10 minutes** for your presentation [if you are presenting solo]; and **14 minutes** if you are in a group. Following each presentation, a few minutes will be allotted for questions. (Please keep responses brief and to the point.)

Structure

Your presentation should have the following sections. You may find it helpful to also have an Outline slide. For each item, the number of *recommended* slides has been placed in square brackets (note that these are recommendations, and not strict slide number requirements). You are free to reorganize the order of each section within reason, as well as increase/reduce the number of slides as you see fit. However, keep in mind your time constraints.

- **Introduction and/or Motivation** Background of your project (Why is the problem important? What has been done before, and what is yet unknown? Why do you care about it? Why should your audience care?) [1 Slide]
- **Objectives:** State the research questions/aims of your work. (What does your project accomplish?) [1 Slide]
- **Data:** Describe your data and sources. [1 Slide]
- **Methods:** Specify your model structure[s] (use equations where feasible, e.g. splines, and/or diagrams, e.g. trees/neural networks). [2 Slides]
- **Results:** Show estimates and results, including performance/errors. (What did you find? How good is/are your model[s]?) [2 Slides]
- **Conclusion:** Summarize what you did and what you discovered. Briefly state any limitations of your approach and potential future work. [1 Slide]
- **Author Contributions:** For those who worked in a group, include 1 slide stating how each member contributed to the project.

Use your slide real-estate wisely. *You may also wish to include slides in an appendix*, which you may refer to in response to post-presentation questions.

Submission Instructions

Following the class presentation, I will email you with comments. You should address any suggested corrections prior to the submission of your final presentation (submit as PDF), which will be due a few days afterward (TBD). Along with your PDF, you are also required to include the code you used in generating your results either as an R script/package, Python script/package or Jupyter Lab/Notebook. Your code should be portable. Thus, you should include your data folder in your submission. Your paths should all be relative. (I will not necessarily be running your code; so if your dataset is too large to upload, a small sample will suffice.) This means your code submission must be an archive (e.g. ZIP file). Submit this along with your final [revised] presentation.

Grading rubric

Component	Item to be evaluated	Points
Proposal	Clarity (all comments addressed)	10
Presentation	Organization (expected structure and clear objectives)	5
	Content (adequate model description and thoroughness of results)	5
	Methods (at least (a) 2 modeling approaches used (or 1 with prior approval) OR (b) 2 datasets used; correctness and appropriateness of framework)	10
	Ability to respond adequately to post-presentation questions	5
	Final version submitted with corrections addressed (if any)	5
Code	Organized and runnable; with reproducible results	10
Total		**50**