Advanced Data Analysis
Syllabus-Winter 2020

Instructor: Li-Pang Chen. Email: lchen723@uwo.ca

Office Hour: MWF 12:00-12:30 and 13:30-14:00.

Grading scheme:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Assignments</td>
<td>20%; Two assignments at 10% each.</td>
</tr>
<tr>
<td>Midterm Project</td>
<td>15%; A book review or research paper review article.</td>
</tr>
<tr>
<td>Final Project</td>
<td>65%; A presentation and written report.</td>
</tr>
</tbody>
</table>

References: References in this course contain textbooks and several research papers. Main textbooks are listed below:


Tentative Course Outline: In this course, I aim to provide some advanced methods for data analysis. The topics include

- Ch1: Brief introduction.
- Ch2: Linear models.
• Ch3: Nonparametric regression.
• Ch4: Machine learning: classification.
• Ch5: High-dimensional statistical analysis.
• Ch6: Analysis of graphical models.

Some details of grading schemes: Here are some detailed descriptions of two assignments, midterm project, and final project.

Assignments: Assignments aim to help students review basic materials in this course.

Midterm project: To make students have ability to explore new materials and summarize their comments to what they study, the purpose of midterm project focuses on writing a review article.

In midterm project, each student is free to find a book (or a research paper) EXCEPT for textbooks listed in References. It would be good to find “recent” book or research paper (e.g., published during 2017-2020). Book or research paper should be relevant to data analysis or data science. Materials can include, but not limit to our course outline listed above. If you are not sure whether a book or research paper you find is suitable, it is also welcome to discuss with the instructor.

After reading a book or a research paper, students should make detailed summaries (e.g., key ideas or important concepts in each chapter or each section) and make comments (e.g., what’s advantage/disadvantage of method? your opinions?)

Final project: Final project is about real data analysis. Students are encouraged to find any real dataset EXCEPT for datasets mentioned in this course. After finding the interesting dataset, students should apply any materials you learned in this course to analyze the dataset. Finally, you should present your finding and comment after obtaining numerical results. Students should prepare a “formal” report which may include motivation, data description, data analysis procedure, numerical results of analysis.

Meanwhile, students will give a 30 minutes presentation to present their data analysis results. Presentation of final project will be held in the last week of this term (Mar 30 - April 3). Of course, the detailed date depends on the number of students enrolled in this course.