<div align="center">

BT4222 Mining Web Data for Business Insights

Academic Year 2022/2023 Semester 2

Lecturer: Dr. Dandan Qiao

Project Guidelines and Grading Criteria
Updated as of 30th March 2023

</div>

Project Description

The course project is aimed at allowing students to gain some hands-on experience in solving text mining / machine learning problems. In this project, you are tasked to apply relevant text mining / machine learning tools and techniques that you have acquired through the course (of course, you are welcome to go beyond the scope of the course) to mine useful insights and arrive at meaningful conclusions for a real-world application. You are required to complete this project in groups of 4-5, delivering three (3) components: Group Project Presentation Slides, Group Project Final Report (pdf), and Group Project Programming Codes. The **entire** Group Project will contribute **50%** to your overall grade. Details of each component, its requirements, and the respective grading criteria will be specified below.

Disclaimer:

Individuals in the same group will generally receive the same scores for all components of the Group Project unless feedback is received that a particular member is only superficially participating and not doing actual work. Please voice out your concerns to the lecturer if you think it is necessary.

AI Policy:

You are allowed to use AI (tools like ChatGPT, at a minimum), in this class. However, please acknowledge using. Please include a paragraph at the end of the report explaining what you used the tool for, what prompts you used to get the results, and also provide a version prior to the use of this tool. Failure to do so is in violation of academic honesty policies, and will be penalized in grading.

Important things to note:

Please name your files in the following format -*ReportXX.pdf*[1] for Group Project Final Report, *PresentationXX.pptx* for Group Project Presentation and *CodesXX*.zip for Group Project Completed Codes, where **XX** is your group number.

Project Component	Due Date and Time	Submission Items
Group Project Presentation	[Week 13] 14 Apr 2023 @ 23:59	Presentation Slides
Group Project Final Report	[Week 13] 14 Apr 2023 @ 23:59	PDF Report

---

[1] If you have used ChatGPT or any other AI tools in your project, please acknowledge and summarize what prompts have been used at the end of "Report_ChatGPT.pdf". This summary will not be counted towards the page limit (8~10 pages).

Group Project Codes	[Week 13] 14 Apr 2023 @ 23:59	Completed Codes (==Either Jupyter notebooks or py files are okay. If there are multiple files, please compress to a zip file.==) ==If there is data, please compress the code files and dataset together into the zip file.== The zip file should be named in the way as required.

Group Project Final Report

You are to submit an **8-10 pages of PDF report** for your Group Project. A suggested outline of the Group Final Report includes:

1. A title page, which should include the title, the author information (name and student metric number), and the URL that directs to your completed codes.
   <u>This page will not be counted in the page limit.</u>

2. A description of the problem
   - Background & Motivation
   - Clarify the problem you are going to address.
   - Highlights of your findings and implications

3. Some basic facts of the dataset(s)
   - Source of the dataset(s)
   - Data collection methods:
     - If you scraped the data, write about how you made use of scraping tools to get the desired data from for example Facebook/Twitter.
     - If you downloaded the data, gather the basic information on how the dataset was assembled by the creator.
   - Describe the dataset(s) - E.g., number of observations, number of variables, type of variables (string, integer, etc.) [You may present the variables in the form of a table.]

4. A detailed discussion of the following:
   - Pre-processing steps to sanitize/manipulate/combine your dataset(s)
   - Machine learning models that you have used
   - The insights that you have gathered from applying machine learning methods on the dataset(s)

5. Conclusions

<u>Grading Criteria</u>
The Group Project Final Report weighs **25%** of your overall final grade. Your report will be evaluated based on:

(1) Clarity and completeness of the report
(2) Appropriateness of the models and methods applied for analysis
(3) Coherence and consistency of proposed methods with hypotheses
(4) Usefulness of your analysis in other similar real-world problems
(5) Concise summarization of your work
(6) Reasonability of the discussion of advantages and limitations of applied methods to the defined problem

Group Project Presentation

What to present: 11-minute worth of slides containing highlights of your analysis
When to submit: [Week 13] 14 April 2023, Friday at 23:59
What to submit: Presentation Slides
*Where to submit: **Canvas > Final Project Submission***

Guidelines for Presentation
You are to prepare 11-minute worth of slides for the presentation. The flow and content (please select the highlights of your analysis) of the presentation can follow that of the Group Project Final Report.

Grading Criteria
The Group Project Presentation weighs **20%** of your overall grade. The grading of the presentation will be dependent on:

(1) Clarity in highlighting the background and research questions
(2) Logical and reasonable explanations of selected methods of analysis
(3) Comprehensiveness and coherence of the presentation
(4) Readiness and confidence of presenters
(5) Self-explainability of slides' content
(6) Good time management during the presentation

Group Project Codes

You are to submit your codes (preferably in Jupyter notebook) together with the Group Project Final Report in Week 13 (14 April 2023) by 23:59. As long as your codes are complete and error-free, you will be awarded 5% (specifically, the codes shall be able to run smoothly and produce the required results).

All the best for your project!

- End -