

SEEM2460 Introduction to Data Science (Spring 2020)

Guidelines to Course Project

Theme: Understanding and Fighting COVID-19

We are facing the pandemic which is spreading across the world. The number of cases of COVID-19 has been growing exponentially, reaching 1.4 million with over 81,000 deaths [source: worldometers.info]. We wish to leverage the creativity and innovative thinking of our students in the use of data science to help understand and fight COVID-19. Hence we have this as the theme of this year's course project of SEEM2460.

Each course project team can include up to 3 members. Please try to come up with a course project topic to be studied by your team. Some examples are suggested below, based on the topics covered throughout the course, to help you formulated your ideas. You may also come up with a totally new topic. The suggestions include with some references are included below (with table and references).

The final project includes the selection of a case study where data science is applied. The student team is required to write a report that describes the problem addressed, the application domain, the data used, the data processes involved, how various techniques and technologies are used, the outcome or impact, and the student's own assessment of the application of data science. The student is required to write a final project report (not exceeding 10 A4 pages with font size 12, single-line spacing, 1-inch margin all-around). Please submit your paper with the signed acknowledgement statement obtained from CUHK VeriGuide system. You could attach the scanned copy of the statement with your signature to the paper submission.

To facilitate group formation, we will be putting up a Google Sheet (see BlackBoard). Please enter your Name, Year of study, Major of Study, Topic(s) of Interest (based on the 'Abbreviations' in Table 1), Contact Information (e.g. email) and any Remarks. You may refer to this Google Sheet to search for team-mates.

The course project proposal will be due on April 27, 2020. Please submit a 1-2 page proposal with the following information:

- Problem Statement (5 points)
- Motivation (5 points)
- Data Sources (5 points)
- Data Processing Approach(es) (5 points)
- Planned Analysis/Implementation / Experimentation (5 points)
- Potential Conclusions (5 points)

The course project report will be due on 20 May 2020. Evaluation Criteria include

- Organization (15) -- This includes:
 - structure of the content (5),
 - logical flow of ideas (5),
 - exposition with appropriate breadth of coverage and depth of explanation (5)
- Clarity of Writing -- graded individually (5) -- Since this part is graded individually, the report should indicate clearly the portion of the writing of each group member
- Technical Correctness (10)

- Literature Survey (5)
- Overall Readability (5)
- Innovative Aspects (10)

Table 1. Example topics.

Abbreviation	Suggested Example
COLLECTION	Design an application that collects useful data for analysis and explain the analysis.
VISUALIZE	Select relevant pandemic data and design a visualization that uncover certain trends.
DECISION	Select relevant data and build a decision tree.
ANALYZE	Consider the measures taken by different countries to fight the pandemic and their effectiveness.
PREDICT	Formulate a prediction problem based on the trends discovered from data.
FINANCE	What were the trends in the financial markets in response to the pandemic?
LOGISTICS	Apply logistics modeling in resource distribution and service provision.
OTHERS	Please state the other topics that you are considering

References:

<http://www.coronavirus.gov.hk> (Hong Kong Government's information)
https://www.chp.gov.hk/files/pdf/statistics_of_the_cases_novel_coronavirus_infection_en.pdf
<https://www.google.com/covid19/> (especially section on data and trends)
<https://www.worldometers.info/coronavirus/>
<https://systems.jhu.edu/research/public-health/ncov/>
<https://nextstrain.org/ncov>
<https://www.tableau.com/covid-19-coronavirus-data-resources> (with a workbook for trial)
<https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/> (Imperial College Coronavirus Projection)
<https://github.com/CSSEGISandData/COVID-19>