

Hong Kong Baptist University  
Department of Computer Science  
COMP4057 Social Computing and Web Intelligence /  
COMP7630 Web Intelligence and Its Applications  
Semester 2, 2017-18

Group Project / Group Term Paper

**1. What is the aim of the group project / group term paper?**

The group project's aim is to implement and apply social computing and web intelligence concepts and algorithms learned in this course to some real-world data. (Essentially COMP4057)

The group term paper's aim is to carry a literature review on the technological development in terms of algorithms and applications of a particular social computing and web intelligence related topic. (Essentially COMP7630)

**2. What are to be covered and submitted for the group project? (COMP4057)**

Groups (with 3 to 4 students each) have been formed. Each group is expected to work on a topic related to social computing and web intelligence, or some related topics (subject to the approval of the course instructor).

For the group project, a number of steps are expected to be accomplished, including problem identification, data collection (e.g., via crawling) from online social media, data pre-processing, feature engineering, data mining techniques being used, and presentation of the analysis results. The following lists some projects of last year for your reference:

- P-1 Collaborative & Content Based Movie Recommendation System for DouBan
- P-2 Building Text Sentiment Classifier for Comments in Social Media Network (9GAG)
- P-3 Public Impression on Companies based on Comments on Facebook
- P-4 Clustering stocks with similar performance based on data via Yahoo finance API

Using P-1 as an example,

- Identified problem is movie recommendation

- Data source is DouBan (web scraping)
- Data pre-processing includes removing the unused tags and so on
- Features being extracted include rating scores, reviewing comments, casting, type of movies, etc.
- Data mining techniques include content-based and collaborative filtering algorithms, and
- Result presentation (performance comparison in terms of different performance metric, visualization of 2-D projection of users and items)

Project deliverables include:

#### i) Group Project Report

(submit softcopy using the filename convention “GroupX\_ProjectReport.pdf” one per group together with the presentation slides and the prototype materials by April 25 2018 23:59)

The report should cover the details of all the steps you have carried out as well as the discussion on the results obtained AND your responses to the comments from your peers (see the Oral Presentation section). There is a page limit of 12 pages (1.5 spacing; 12pt font size; 1 inch for all margins; title, table of contents, and references are all excluded from the 12 pages).

#### ii) Individual Report

(submit softcopy using the filename convention “IDXXXXXXXXX\_Report.pdf” one per student by April 25 2018 23:59)

It should be a one-page report (1.5 spacing; 12pt font size; 1 inch for all margins) including your own *insights gained, lessons learned* as well as *new possibilities* that you can think of after carrying out this group project.

### **3. What are to be covered and submitted for the group term paper? (COMP7630)**

Groups (with 3 to 4 students each) have been formed. Each group is expected to prepare a group term paper on a topic related to social computing and web intelligence. The term paper should be a literature review paper presenting the technological development and applications of a particular area in social

computing and web intelligence. For example, it can be web search ranking, community detection, link prediction, recommender systems, targeted advertisement, etc. In order to get a high mark, each group is suggested to pick at least 3 x [group size] papers which are good and recent enough, read them in detail, describe the key ideas, compare and contrast them, and discuss possible future work and new opportunities of application. The sources of the papers can be a combination of web resources, news/magazines and journals/conferences articles (e.g. using scholar.google.com). The selected papers that you read and mentioned should be included in the reference section of your group term paper and in-text citations should be included throughout your report.

To gain insights on more trendy topics, you can scan through the papers published in the following conferences (but not limited to) to find the papers you are interested in and yet readable to you:

- <http://www.wsdm-conference.org/2017/proceedings/>
- <https://www.ijcai.org/proceedings/2017/>

Project deliverables include:

#### i) Group Term Paper

(submit softcopy using the filename convention “GroupX\_TermPaper.pdf” one per group together with the presentation slides by April 25 2018 23:59)

The group term paper should cover what being mentioned above AND your responses to the comments from your peers (see the Oral Presentation section). There is a page limit of 15 pages (1.5 spacing; 12pt font size; 1 inch for all margins; title, table of contents, and references are all excluded from the 15 pages).

#### ii) Individual Report

(submit softcopy using the filename convention “IDXXXXXXXX\_Report.pdf” one per student by April 25 2018 23:59)

It should cover a new business related idea related to the technology covered in the term paper. Novelty is what being emphasized here (even though it is of your own and not mentioned by others before).

### **4. Oral Presentation (COMP4057 and COMP7630)**

All group members of both courses must participate in the oral presentation to be held on April 16 and April 23 2018. Presentation of COMP4057 and COMP7630 will be mixed together on both presentation dates. For each presentation, there will be 6 minutes for oral presentation and 6 minutes for Q&A session. In addition, each group should identify a representative to collect comments/questions from its group members on the presentations of other groups within the same day of the presentation session (detailed logistics to be released). The quality of those comments/questions will be marked.

## **5. Evaluation (COMP4057 and COMP7630)**

Group Project Report / Term Paper	70%
Group Presentation	15%
Group-based comments on other groups	5%
Individual Report	10%