Hong Kong University of Science and Technology
COMP 5212: Machine Learning
Fall 2018

Programming Project

1 Objective

The objective of this project is twofold:

1. To propose and investigate a novel application of at least one machine learning method (either one covered or a variant of one covered in this course).

2. To acquire and practise the hands-on skills needed for solving real-world tasks using machine learning methods.

Please note that this is not intended to be a research project. Nevertheless, some studies that explore novel applications may inspire follow-up research later to develop new machine learning models or algorithms.

2 Assessment Components

The project has three components which will be assessed:

1. Project proposal
2. Project report and source code
3. Video presentation

This is a group project with each project group consisting of 2–3 members. One member of each group will submit the assessment components on behalf of the group, but the names of all members should be listed clearly in all the assessment components.

Note that this project cannot be used for earning credits in a different course. However, it can be part of your thesis research as long as it is not for earning course credits.

3 Project Proposal

3.1 Requirements

The first step is to identify a topic that you are interested in exploring more. Feel free to choose a topic that will only be covered later in the course as long as it interests you. It is also common that a project involves multiple topics covered in the course.

As novelty of the target application is a major assessment criterion, you should put more effort into this aspect by proposing a novel application with practical significance that ideally requires

nontrivial application of existing machine learning methods.

For inspiration, you may take a look at the Kaggle website (`https://www.kaggle.com/`) and many other resources available on the Web. Sometimes a data set originally used for one task may be used in a very different way for another task that has not been studied by others before.

Note that no extra computing resources beyond the standard facilities available in the university will be provided. So you should either keep the computational demand of your project reasonably low or make your own arrangement for additional computing support such as cloud computing services.[1]

In case you plan to use some more advanced machine learning methods not covered in the course for your project, please make sure that you also include the related methods covered in the course as baselines for comparison. Among other things, including the baselines will help to justify using more advanced methods.

Your proposal should cover at least the following aspects of the project:

- Project title
- Group members with full names, student IDs, and HKUST email addresses
- Description of the application and justification for its practical significance
- Formulation of the machine learning problems involved in the application
- Data set (and preprocessing, if needed)
- Machine learning methods
- Design of experiments and performance evaluation

The proposal should be no longer than four A4 pages.

3.2 Submission

The proposal will be due on **16 October 2018, Tuesday, 11:59pm**.

Like the programming assignments and the problem set, submission of the proposal should be done electronically using the Course Assignment Submission System (CASS):

`https://cssystem.cse.ust.hk/UGuides/cass/student.html`

The file should be in PDF format with filename `proposal.pdf`. When multiple versions with the same filename are submitted, only the latest version according to the timestamp will be used for grading.

[1] Since our department has joined the AWS Educate program, you can register and get free credits annually.

4 Project Report and Source Code

4.1 Requirements

The project report should be self-contained without requiring its reader to refer to the proposal in order to understand the project. It can be written like a conference paper. You are highly recommended to follow the NIPS 2018 paper style to format your report:

`https://nips.cc/Conferences/2018/PaperInformation/AuthorGuidelines`

You should have a separate section at the end of the report stating clearly the division of labor among all the group members. List the main duties and contributions of each member. You should try your best to ensure that the workload is divided fairly.

The report can have up to eight content pages including all figures and tables. Additional pages containing only the division of labor section and references are allowed.

Remember to include in the report a hyperlink to the YouTube video of your project (see the next section).

All the source code that you have written for this project should be submitted for grading. In case your code is modified from another source, you are expected to acknowledge it clearly in your report. Failure to do so is considered plagiarism. Note that if the project requires only minimum coding effort by your team, you may not score high on the technical quality of the project even if the application is novel and has practical significance. More information about the assessment criteria can be found below.

Data files should not be submitted to keep the submission file size small.

4.2 Submission

The project report and source code will be due on **30 November 2018, Friday, 11:59pm**.

Electronic submission using the CASS is required. Your submission should contain two files: report (`report.pdf`) and compressed source code (`code.zip` or `code.rar`).

5 Video Presentation

5.1 Requirements

You are required to prepare an oral presentation of your project in the form of a video. The video should be no longer than 15 minutes.

Note that the video is not a movie for entertainment or an advertisement. It is for a technical presentation like that in an academic conference. You should pay attention to both the technical content and the quality of your video.

5.2 Submission

When your video is ready, upload it to YouTube as an 'unlisted' (not 'private' or 'public') video and include its hyperlink in your report. The video should be ready by the time you submit the report.

6 Grading Scheme

This project will be counted towards 15% of your final course grade. The maximum scores for the three major assessment components are as follows:

1. Project proposal [**10 points**]
2. Project report and source code [**60 points**]
3. Video presentation [**30 points**]

Assessment will be based on multiple criteria including novelty and practical significance of the application, technical quality of the project, and relevance of the project to the topics covered in the course.

Please note again that this project cannot be used for another course. In case it is part of your own research, you should state it clearly in both the proposal and the report.

Late submission of each of the components will be accepted but with penalty. The late penalty is deduction of one point (out of a maximum of 100 points) for every minute late.

7 Academic Integrity

Please read carefully the relevant web pages linked from the course website.