CIS 5930: Data Mining: Fall 2019

Department of Computer Science

Florida State University

Course Project Details

Project Proposal: Due Oct 04, 2019, 11:59 pm

Students enrolled in CIS 5930 are invited to submit proposals for their course projects. This is a group project; each group should have either two or three students. You may use the discussion board in Canvas to look for project partners.

The project must stay within the theme and spirit of the course and must be in the general area of data mining. Projects that do not adhere to this spirit might not be accepted. The goal of the project is to work on a challenging, real-world data mining problem. The implementation must be done in Matlab or Python.

The project proposal must contain the following components:

(i) Project title and the names of the team members

(ii) Detailed description of the research question

(iii) A thorough literature survey with appropriate references

(iv) Details of the algorithm(s) to be implemented

(v) Expected experiments and analysis to be performed, including a list of datasets to be used for the experiments

(vi) A timeline of significant steps in the project

You may consider reading papers in data mining (IEEE International Conference on Data Mining, SIAM Data Mining Conference, ACM Conference on Knowledge Discovery and Data Mining) or computer vision (IEEE Conference on Computer vision and Pattern Recognition, IEEE International Conference on Computer Vision) conferences to come up with project ideas.

If the material from this course project is expected to be used currently or was used in previous terms in other courses, or as part of a thesis or dissertation, the student must mention details about the same in the proposal. The instructor will discuss this with the students to judge the acceptance in these cases on a case-by-case basis. On most cases, such joint material will be allowed, but not informing the instructor of this could lead to a case of this being identified as self-plagiarism.

The instructor holds the right to reject any project proposal in situations where the material is not related to the course, or if the project is not worth 40% of the grade or is too ambitious to be completed within the project deadline. The instructor might schedule one-on-one interviews with

the students in these circumstances to further discuss the proposals. Students are strongly encouraged to discuss their project ideas with the instructor before starting to write the proposal.

The proposal should be no more than 3 pages (single space with font size 10, 11 or 12) excluding references. Submit your proposal as a PDF file through Canvas. **Only one submission is required from each project team.**

Final Project Report: Due Dec 10, 2019, 11:59pm

Report: A final report should be prepared for the course project. The report should contain the following components:

(i) An introduction to the problem in question including potential applications

(ii) A thorough literature survey with appropriate references

(iii) Description of the methodology used

(iv) Description of the implementation, including potential problems encountered during the implementation, assumptions made and reasons for such assumptions and such

(v) Experimental results containing figures and tables with appropriate discussions

(vi) Conclusion and future research directions

The report should be no more than 8 pages (single space with font size 10, 11 or 12) excluding references.

Code: You must submit all the Matlab / Python code files used and developed during and for this project. The codes are expected to run in sequential fashion, that is, one trigger must make the whole code run in sequence without any disruption, until results (visual and metric) are produced. Please include an estimated time for completion and display real-time progress on the code (such as "Training Completed" etc.). Please also provide adequate comments in the code files.

Include a README file containing detailed instructions on how to run the code. Also include any libraries that are used in your implementation. The instructor should be able to compile and run the code in his machine and generate the results produced in the report (with a single function call).

In any cases of borrowed code, be it from forums or external sources, appropriate citation must be provided. Borrowing code without appropriate citation will be considered a violation of academic integrity.

Submit the code files and the project report (in PDF format) as a single zip file through Canvas. **Only one submission is required from each project team.**

Project Presentation: Students are required to present their projects in class. Details will be posted later in Canvas.

Academic Integrity: You are expected to maintain the utmost level of academic integrity in the project. Any violation of the code of academic integrity will be reported to the dean for official actions. It is an academic violation to copy, to include text from other sources, including online sources for both material and code, without proper citation and licensing. Please see the complete explanation of the Academic Honor Code. Evidence of plagiarism or academic dishonesty will likely result in failing the course and at worse can lead to disqualification from your degree program. Please contact the instructor before borrowing material when unsure.