CS 489/689 – Project Guide

For undergraduate students:

- Three or four can make a team
- A project topic should be interesting
- Focus on solving interesting problems with appropriate machine learning techniques
- Performance comparison with other benchmark methods is NOT necessary

For graduate students:

- For master students, one or two can make a team
- For Ph.D. students, a project must be an individual work
- Performance comparison with other benchmark methods must be presented (e.g. k-fold cross-validation)
- Experimental settings and results must be well presented

The goal of this term project is to apply machine learning techniques appropriately for real-world problems. You can use any libraries or built-in functions.

The project will be evaluated by:

1. Whether the motivation of the project is well described.
2. Whether computational challenges are well described.
3. Whether the dataset is well described.
4. Whether choice of machine learning techniques is well justified.
5. Whether an appropriate approach is applied to realistic data sets that may be being used in practice.
6. How well the experimental settings are described (especially for graduate students)
   a. What machine learning methods did you consider for benchmark? How do you tune the optimal parameter?
   b. What are evaluation strategies (e.g., k-fold cross validation) for the performance comparison?
   c. How many times are the experiments repeated?
7. How well the experimental results are discussed.
a. How can the experimental results be interpreted?

8. Conclusion
   a. What obstacles have you overcome?
   b. Any issues that you want to share with others?
   c. Any future plans in related research?

Find a dataset that you may be interested. Kaggle.com and UCI Machine Learning Repository contain a lot of interesting datasets and well-defined problems. If you have any specific related dataset, you can suggest it too. If you generate a dataset by yourselves, it will be plus (especially for undergraduate students).

The followings must be included in the proposal:

1. Problem definition
2. List of the method(s) that you will implement (or compare)
3. Brief description of the data (e.g., # of samples and features)
4. Description of Data settings or preprocessing (if need)
5. Experiment Design

- **Proposal Deadline:**

You must submit a proposal for your project to WebCampus by **Sunday, March 15, 2020**