CI Studio: Replicating Experimental Research Method (100 Points)

DUE: 11/23 SUNDAY 11:55PM

Goal: Demonstrate your understanding of the “experimental” research method in CI.

Team: Work in a team with 2 or 3 members.

Task: Repeat an experimental validation in a “computational” research article of your choosing, measure the performance metric yourself, compare your results against what was published in the original article, and report the findings.

Details: How to find a paper to replicate?

- Try relevant venues to find interesting papers with data set available: eg., JASIST, CACM, KDD, ICDM, SDM, ICWSM, WWW, WSDM. If data set is not available, contact authors for access--some may share all or subset of their data for class use.

- Try internet dataset repositories such as:
  - [http://archive.ics.uci.edu/ml/datasets.html](http://archive.ics.uci.edu/ml/datasets.html): Here, about 260 publicly available datasets are listed, along with papers where the datasets were used. You can browse through datasets of your interest, and try to replicate experiments of the corresponding papers, for instance.
  - [http://snap.stanford.edu/data/index.html](http://snap.stanford.edu/data/index.html): Stanford data mining group's public data sets (and associated papers are found in the group's publication page)
  - ICWSM's data sets:
    - [http://www.icwsm.org/data/](http://www.icwsm.org/data/)
  - Pick a computational paper that includes something you can “replicate” within the time frame. If an article contains algorithms too complicate to implement, or experimental environment too sophisticated, you may want to avoid such.

- You can choose any paper as long as: (1) it has computational nature, and (2) it uses experimental method as part of validation.
Pay attention to the specific settings of the experiment. For instance, assumptions, experimental questions to prove/disprove, data sets, evaluation measures, parameters, etc.

**Deliverable:**

- Using the ACM SIG Proceeding style in Word or LaTeX (available at [http://www.acm.org/sigs/publications/proceedings-templates](http://www.acm.org/sigs/publications/proceedings-templates)), prepare your findings in a report up to 5 pages

- Submit the PDF version of your report to the ANGEL dropbox “Studio”