15-826, SCS CMU

Christos Faloutsos

Spring 2017

Partial Credit Grading Scheme - default project

Phase 1

Shortcomings: -1 if listed but not elaborated/vague.

[From Christos:]
2-person teams - for each paper (worth 15pts of the 100)
1. 2pts for problem definition
2. 10 points for main idea /summary
3. 1pts why/why-not it is useful for your project
4. 2pts shortcomings of the paper; your ideas on how to improve it

1-person teams: 33.3% for each paper, 0% for coordination - the points would be 5-20-3-5 for the 4 items above

Spelling /grammar:
- -1pt if there are a 1-5 spelling errors (spell-checking should be a habit!)
- -5pt if there are many spelling errors (2-3 per page)
- no penalty if there are 1-5 grammar errors
- -5pt if there are >5 grammar errors

Split the grade: if one person did a good job, but the team-mate did not, give two different grades, one to each team-mate.

Phase 2

The bold parts, are as announced   - the rest are detailed, here.

70% implementation

       60% if their code runs (-5%, if there is no 'makefile')

             No penalty for too long

      6% for handling (even incorrectly) the

             3 density measures * 2 dimension-selection policies

             (single-member teams need only do the 'geometric' density measure and

              the 'max cardinality' selection)

              [we'll accept any bucket-size, for the 'time' aspect]

       4% for unit tests (at least 3 settings - eg., a single 2-mode block; 3 blocks in 2-modes;

             3   blocks in 3-modes)

25% design of experiments

       20% for handing in the phase2 report (with corrections in phase1, spell-checked, etc)

       5% for listing optimizations that they may consider for phase 3

             (like, eg. clustering index on attribute 'source'; temporary table for <bla>; etc)

  5% plan of activities

- Small penalties (-1 to -3) for small omissions (typos, forgetting the graded phase1)
- -10 for no or insufficient instructions/Readme
- -30 if code is not fully correct and/or requires edits and changes by the grader to run

For regrades: Here is who graded what:
- A-D: NAHTA *mnahta @ andrew*
- F-K: CHANDRASEKARAN *sanjayc @ andrew*
- L-O: FERNAU *rfernau @ andrew*
- P-Z: VARMA *tvarma @ andrew*

Phase 3

- [-1pt] if they do ROC on a *sample*, because it is too slow (it shouldn't be so slow) -
  - no penalty, if they give a good explanation with documentation (eg., output of 'explain select', or 'uptime' etc) [*christos*]
- No penalty, if they do ROC and/or wall-clock times, on a *bucketized* version - any reasonable bucketization, is fine ('reasonable' = it results in >50 buckets) [*christos*]
- No penalty if they give the AUC, stopping earlier than x=1.0, and they email the plots and AUC afterwards

*copy pasted from the website rubric
- Writeup
  - [2%] Introduction - Motivation
  - [3%] Problem definition
  -
  - [5%] Survey
  - Implementation - Task 3: optimization
    - [20%] For T3.1 ('copy' method)
    - [20%] For T3.2 ('mark' method
    - [10%] For an optimization of your own (describe idea; give experimental results)
    -
  - Experiments - Task 4: Evaluation
    - [10%] For T4.1: report info about the top k=5 blocks for each of the two datasets
    - [0%] (optional) For T4.2: describe briefly whether you think they are suspicious or not
    - [10%] for T4.3: give the ROC curves and AUC for each curve
  - Experiments - Task 5: Anomaly detection
    - [5%] For T5.1: report info about the top k=5 blocks for each dataset
    - [0%] (optional) For T5.2: describe briefly whether/which ones are suspicious.
    -
  - [5%] Conclusions
- Software (testing, packaging and documentation) [10%]
  - -2 if port/user is not specified in the README and this causes an issue when running

For regrades: Here is who graded what:
- A-D: NAHTA *mnahta @ andrew*
- F-K: CHANDRASEKARAN *sanjayc @ andrew*
- L-O: FERNAU *rfernau @ andrew*
- P-Z: VARMA *tvarma @ andrew*