

CSE6240_Spring2020 Project Midterm Report Rubrics

Midterm Report Rubric: CSE 6240

Last updated on March 9, 2020

Template Settings:

- Number of pages = maximum 4, including references at the end
- Format = ACM
- Format is available at <https://cs.stanford.edu/~srijan/teaching/cse6240/spring2020/#project>

Total = 80 points

- **Dataset description [35 Points]**

- **Data preparation: [10 Points]**

- Source [2 points]
 - Describe how you got the dataset (e.g., crawling, API, from a website, etc.), give proper references wherever applicable (-2 for missing references of papers, websites, or APIs)
 - Data preprocessing steps and explanations [4 + 4 Points]
 - Explain how the data preprocessing, cleaning, imputation, and other processing was done.
 - Explain why this dataset is necessary and sufficient to achieve the goals of the project.
 - Please note: the dataset should be fixed and finalized [-10 if not]
 - Exception: if you are crawling dataset yourself, then you can continue to crawl the data beyond the midterm, as long as a big fraction of the dataset has been collected already.

- **Raw Data Statistics: Explain the dataset [10 points]**

- Describe the important properties of the dataset. How many data points are present? Describe the important features of the data and their basic statistics (range, mean, median, max). Details of ground-truth labels or dataset should be given, if applicable. Some examples are below:
 - Text data: vocabulary size, number of sentences per text entry, average number of tokens per sentences, and more.
 - Image data: number of images, categories of images, ground-truth labels

- Social media data: how many users are present in the dataset, how much information in total and per user is available, time range of data
 - Network data: number of nodes, edges, properties of nodes and edges, number of labels/classes
- Add at least 5 distinct features [$5 \times 2 = 10$ Points]
 - -2 points for each missing important detail
 - Should contain at least 1 relevant table [-5 if not present]
- **Data Analysis: Explore your data and talk about findings [15 points]**
 - Discuss relevant features, correlations, cluster visualizations, sentiment statistics, network statistics, centrality distributions from the data.
 - Add at least 5 insights or interpretations [$5 \times 2 = 10$ Points]
 - Add at least one visualization (figure or plot) [5 points]
 - All findings and figures should be quantifiable (-3 per non-quantifiable instance). Example: a word cloud is non-quantifiable.
 - -3 points for every non-relevant analysis below 5.
 - Please make sure to add only relevant visualizations and insights. Inserting vague plots and figures, for example, unrelated word cloud or generic network visualization will incur penalty (-5 points penalty).
- **Describe the Experimental Settings [10 Points]**
 - Explain the parameters of the experiment: what is the split used, cross-validation setting, evaluation metrics (for example, accuracy, AUC, or precision etc.), system settings (RAM, GPU, or CPU statistics).
 - -2 per missing relevant detail.
- **Baseline results and discussions [30 points]**
 - The report should have results from at least 2 baselines. At least one baseline should be from a published paper or preprint. Creating one reasonable baseline yourself (e.g., using feature engineering and standard ML classifier) is allowed. No additional points will be awarded for having more than 2 baselines.
 - **Baseline description [$7.5 \times 2 = 15$]**
 - Describe the baseline. Give a short technical description of the baseline, along with its reference, provide details of the kernel or hyper-parameters used, provide links to code repository used [5 points]
 - -2 per missing detail, reference, link
 - Why is this baseline suitable for the problem? [2.5 Points]
 - **Baseline result [10 points]**

- Results of the baseline on your dataset, presented in a table or figure (e.g., a bar chart) [5*2 = 10]
 - The baselines should be compared on the same metric [-5 if not]
- Result discussion [5 points]**
 - Compare the results of both the baselines. Why does one perform better than the other? If applicable, compare the result to the state-of-the-art reported in literature.
- Next steps [5 Points]**
 - Explain in detail your own proposed approach and what novelty or improvement you are adding over the baselines. [2.5 + 2.5 = 5 Points]
 - 2 for unclear explanation
 - For development projects, clearly describe what will be done by the final report and how exactly this will be achieved. Example, if you are creating an app, where do you plan to host it.

FAQ, common mistakes and penalty:

- Wrong format [-10]
- Use of irrelevant images, tables, sections [-10]
- Use of irrelevant terms, jargons, references, or statements [-2 for each violation]
- Missing project title or use of generic title (e.g., CSE 6240 Project) [-10]
- Length is over 4 pages [-10]
- Length is less than **2 pages** [-10]
- Incorrect references in text [-2 for each violation]. More details on how to use references are below:
 - Correct use:
 - Please see <https://www.acm.org/publications/authors/reference-formatting>
 - REV2 [1] finds fraudsters.
 - REV2 finds fraudsters [1].
 - Kumar et al. created REV2 [1].
 - Kumar et al. [1] created REV2.
 - Incorrect:
 - REV2 finds fraudsters. [1]
 - REV2 finds fraudsters [Kumar et al.].
 - REV2 finds fraudsters (Kumar et al.).
 - REV2 (Kumar et al.) finds fraudsters.
 - [1] created REV2.
 - Kumar created REV2. (missing numeric reference)

- Kumar et al. created REV2. (missing numeric reference)
 - Any other creative version you can think of!
- Numeric and last-name references should be provided only the first time it is introduced in each section.
- Use of abbreviation (sota, bn, mn, etc, eg) [-1 for each violation]
- Non-uniform references in the reference section [-2 for each violation]
- Which baselines can we use?
 - One baseline should be from a published or arxiv paper or report. The baseline should address the same or similar application task, or should be a general method applicable on your task.
 - The second baseline can be a reasonable baseline created by you to fit your application task.
- How should we evaluate the baseline?
 - All models should be run on YOUR dataset. These include the baseline models and your own model. Do not just copy the number from the baseline's original publication into your report, unless it is on the exact same dataset.
- Are there any size requirements for the dataset?
 - There is no minimum requirement on the size or number of datasets. We will not deduct points based on that.
- How do we use the baselines?
 - You should set up a task that answers the main question/aim of your project. Then the baseline should be used to solve this task. Finally, (after midterm), your model should be used to solve the same task.

Published by [Google Drive](#)—[Report Abuse](#)—Updated automatically every 5 minutes