CS732/DS732: Data Visualization -- Course Evaluation Guide

Jaya Sreevalsan Nair
International Institute of Information Technology Bangalore
Oct 11, 2023[1]

Grading scheme, as announced on August 1, 2023:
- 20% of the final grade for each of 3 assignments (A1, A2, A3)
    - Each of the assignments has a demo for 5%, which may be submitted as a video as per instructions
- 5% for reading-writing assignment (RWA) (or a take-home exam)
- 15% for mid-term
- 15% for end-term
- 5% for class attendance

Submission instructions:

- Programming assignments and RWA: 65% of the final grade
    - It is compulsory to attempt all 3 assignments and the RWA.
    - The assessment of the programming assignment is based on deliverables, i.e. code+report+folder of images+demo for each assignment.
        - The report is where one can elaborate on the data used, hypothesis, analytical methodology, inferences, and responses to any questions posed in the assignment.
        - The code submissions must be source code in text format, say Python code in .py text file format, and not as Jupyter notebooks.
        - The video, if included, must indicate how you have been able to demonstrate visualizations, interactions with the application, and other inferences that can be pointed out/highlighted/annotated.
        - For time-varying data visualization, gif images may be generated from frames generated for different time steps. It is important to provide metadata in the plot on the time instance, variable type, etc.
        - It is important to submit a folder of images.
        - A README can be included with instructions on how to run the code, the annotation for images in the folder, and other details related to the execution of your code.

---

[1] Previous versions on Sep 12, and Aug 24, 2023

- All submissions must be done on LMS.
  - If your entire submission is larger than the permissible size for LMS submissions, upload your submission on Outlook OneDrive; and submit a document containing the URL to the submission.
  - It is the onus of the student to ensure the correct access permissions are provided in the repository, if the submission files reside outside of LMS so that there is no difficulty in accessing the files for assessment. In such cases, these repositories have to be accessible until the course grades are announced.
- The scheduling of assignments provides **2-3** weeks to complete each of the programming assignments [A1, A2, A3], and 2-4 weeks for RWA.
  - Assignment announced by Friday midnight IST.
  - All assignment submissions must be done by Monday midnight IST, as per schedule.
  - A3 and A4 is a combined group project with groups of 3 members. The TAs will help with the logistics of team creation, dataset selection for the team, etc.
  - Assignment announcement date and submission deadline:
    - **A1: Aug 25, Sep 18;**
    - **A2: Oct 13, Nov 06;**
    - **A3: Nov 03, Nov 27.**
    - **RWA: Sep 12 (paper selection by Sep 18), Oct 16.**
- 2 written exams - 30% of the final grade
  - Proctored exams in person during mid-term and end-term weeks.
  - Open notes, only hand-written notes.
- 1 report-writing - 5% of the final grade
  - This can be based on a research paper or a theme. It will be allocated on a first-come-first-serve basis.
    - The topics will be published on October 06, 2023.
    - The choice of topic/paper by the student must be communicated by October 13, 2023. If not received by the deadline, a randomly picked topic/paper by the instructor by October 14, 2022.
    - The report is due on October 30, 2023.
  - If it is based on a research paper, the report must say why the method is important, and its impact on the research community (using the papers that have

cited the paper, a state-of-the-art paper that explains the value of the paper), etc.

- ○ The technical report is to be written in the IEEE conference paper format.
- ○ This also includes references to papers/articles/etc. by citing them appropriately in-place in the report. There will be negative points for not doing citation references within the article and bibliography properly.
- ○ A more detailed description of how to prepare the report will be posted at the time of the announcement of the papers.

A1: Visual Exploration Using Visual Analytics Tools

Date of announcement: August 25, 2023 (Friday)

Date of submission: 11:59 pm IST, September 18, 2023 (Monday)

Summary: A group project on visual exploration

Dataset: As chosen by the group from the list provided by the instructor and TAs

A1 tasks and requirements:

1. Read the selected dataset, process the same, and explore the data using visualization and optionally, other simple data analysis techniques, such as statistics.
2. Since it is visual exploration, multiple visualizations should collectively lead to knowledge discovery from the dataset.
   a. A data story must preferably evolve from multiple visualizations, instead of presenting several visualizations.
3. The assignment implementation must start with a question that can be answered by a list of visualization tasks implemented on the dataset. Visualization tasks include exploration and summary, along with descriptive verbs such as "overview", "trends", "search", etc. (If you are interested more in understanding tasks, please refer to Schulz et al. [1]). Your solution should solve n sets of tasks for an n-member team.
   1. The visualizations may be generated on independent tools or an integrated tool, e.g., Tableau, PowerBI, etc. We recommend the use of Tableau.
   2. You are advised to not split/fragment a task across different team members, as much as possible, and also ensure that all team members have tasks that will enable all in the team to make equal contributions.
   3. You may create tasks based on cohesive subsets of variables in the dataset.
4. The report shall contain a detailed description of the tasks you have tried, and eventually implemented, who did what in the team, etc. Specifically, mention the contributions of each member.

A1 assessment:

1. Assessment is based on the report, images folder, and a video demo of the data stories.
   a. The report should contain details on dataset description, tasks, visualizations, and inferences/conclusions. The report should have a section on Author Contributions where the contributions of the team members are explicitly mentioned.
   b. The video demo should be crisp and at most 5 minutes long, where each member explains their task, visualization solutions, and inferences. The first minute can be used by the team leader or the data processing contributor to mention the preprocessing of the data.
   c. The visualizations are expected to be simple plots that are understandable by laymen.

d. The folder of images must be organized properly, and all images in the folder must be featured in the report. The images in the folder may be named as Fig<number>.jpg/png/<image file extension> where the number is the index of the image in the report. You may add additional images that are not contained in the report, in which case, include a readme file to indicate the significance of such images.

References:

[1] Schulz, H. J., Nocke, T., Heitzler, M., & Schumann, H. (2013). A design space of visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2366-2375.

RWA: Reading-Writing Assignment

Date of announcement (of paper allocation): 11:59 pm IST, September 12, 2023 (Tuesday)
Deadline for paper selection: 11:59 pm IST, September 18, 2023 (Monday)
Date of submission: 11:59 pm IST, October 09, 2023 (Monday)
Summary: The assignment is reading a visualization research paper, understanding, and writing a report on the same.
Paper: As selected from a list published by the instructor/TA.

Guidelines:

1. Read the allocated paper thoroughly.
2. Write a technical report using LaTeX (use overleaf to help your case) and IEEE conference proceeding 2-column format, with appropriate bibliography, citing references in-place and as required. Follow appropriate BibTeX to format the bibliography.
   a. Use the citations given in BibTeX format in Google Scholar to have consistent referencing.
   b. Use the standard sectioning used in Computer Science papers: abstract, introduction, related work including the gaps addressed by the chosen paper, methodology in the paper, and the impact of the paper (use Google Scholar to find around 2-3 influential papers citing the chosen paper) including state-of-the-art, and conclusions.
   c. Refrain from using first-person narratives in the paper, refer to the authors of the paper. If there are up to two authors, refer to them by their last names, e.g., "Doe and Smith [21] have discussed ... ". In case there are 3+ authors, use the last name of the first author followed by the phrase "et al.", e.g., "Doe et al. [42] have proposed ... "
   d. In the case of paper review, the title of the paper must be A Critical Review of "<topic-title>". In the case of the exposition of a topic, the title of the paper must be The State-of-the-Art in "<topic-title>". You could come up with variants of these titles.
3. Your report may contain figures and tables borrowed from other papers or created on your own. If borrowed, mention the source as "Image courtesy: [32]" in the figure caption, and likewise for tables. For Tables, Equations, etc., regenerate the same in your article and avoid including screenshots of the same from the original papers.
4. The figure caption goes below the figure, and the table caption goes above the table. All figures and tables must be referred to in your text content. e.g. "Figure 2 shows how ...." or "Table 2 gives a comprehensive analysis of ..."

5. The technical report writing exercise is purely a reading-writing exercise, and no coding is expected. That said, if you would like to run small experiments to improve your paper, you are welcome.
6. Run a spelling and grammar-checking tool(s) on your document and edit your report appropriately before submission.
7. Do not have more than 8 contiguous words verbatim from any source you are citing, otherwise, it would be flagged as plagiarized material.

A2: Visualization Methods for Data Types

Date of announcement: October 13, 2023 (Friday)
Date of submission: 11:59 pm IST, November 06, 2023 (Monday)
Summary: A group project on scientific visualization (Scivis) and information visualization (InfoVis) method implementation on selected datasets
Dataset: As assigned by the instructor and TAs, and A1 datasets

A2 tasks and requirements:
1. Use Python to implement SciVis methods, namely, color mapping, contour mapping, and quiver plots (optionally, experiment with streamlines), and generate animation gifs using 5-10 time instances (i.e. at least 5 and at most 10 temporal samples) of sea surface data.
2. Use Gephi or Networkx to implement a node-link diagram of a sufficiently complex network and compare at least 3 graph layout algorithms.
3. Use Javascript to implement Infovis methods as browser visualizations, namely, the parallel coordinates plot with user interactions and treemap visualizations with at least 3 different experiments.

Guidelines and assessment:
1. Assessment is based on the report, images folder, and codebase, as applicable.
   a. The report shall have images with appropriate captions, and the inferences in the text shall be written about these images.
   b. The images in the folder may be titled using their figure indices in the report so that the high-resolution images are available for assessment. It is sometimes hard to see the relatively lower-resolution images included in the report.
   c. The screenshots from Gephi and other GUIs shall be complete and of high resolution.
   d. The codebase shall include a README for implementation.
   e. A video demo is not required for this assignment. Hence, it is important to include sufficient inferences in the report.
   f. The report shall be professionally prepared as done for RWA. Inclusion of appropriate references, as possible, is encouraged.
2. There are 3 tasks for each of the visualization domains, i.e., Scivis and Infovis. Each student shall implement at least one Scivis and Infovis visualization task. The combined outcome of the team can be proportional to the team's strength.
   a. The report shall clearly state the contributions of each team member.
   b. Apart from data cleaning and processing (as required), there cannot be joint ownership of the methods. If a team violates this rule, marks will be deducted accordingly.
3. The contour mapping shall be done using the Marching Squares algorithm and/or contour fill algorithm. Argue for your choice in your report.

4. The color mapping shall include experimentation with different colormaps, e.g. continuous, discrete, logarithmic, etc., and different color palettes. As appropriate, you could comment about the use of sequential and diverging colormaps.
5. The color mapping shall comment about parametric mapping using global vs local maxima and minima, across different timesteps.
6. The quiver plot can experiment with one of the following:
    a. Grid sampling experiments along with a comparison of magnitude-proportional and same-sized vectors.
    b. Compare and contrast between the streamline method and quiver plots, and consider experiments with different seeding strategies for the streamline implementation.
7. Node-link diagram experiments shall include different graph layouts, and optionally, any desired node filtering implementation.
8. Parallel coordinate plot implementation is complete only with user interactions, namely, axes reordering and brushing.
9. Treemap experimentations shall include different remodels of tabular data to the tree data structure. Optionally, different spatial partitioning strategies can be experimented with.