

The Tools of Topological Data Analysis

Guidelines for Final Projects

Online Course - Spring 2020

PROF. P. CLARE

TA: Kunlin Yang

The purpose of the final project is to demonstrate the knowledge you have acquired about Topological Data Analysis (TDA) methods, by examining and analyzing recent papers written by professional researchers.

The notes below are meant to guide your study.

1. **Title:** this should provide a sharp focus for your topic or theme
2. **Abstract:** this should be a summary of your argument (200 words)
3. **Introduction:** this leads the reader into your assignment, it should announce and outline the main questions discussed in your paper (200-300 words). *This is the most important part of the paper:* it should be non-technical and convince the reader that they should invest the time and energy to read the rest of the paper.
4. **Body of the paper:** make sure to give all details on the assumptions you are making and the theorems and methods you are using. Definitions for the objects you will use must be stated with precision. It is usually not necessary to prove theorems (unless they are new), but you should state them completely and accurately and give a reference where the reader can look them up.
5. **Conclusion:** summarize your work and comment on it. Indicate possible avenues of continuation and future work on the same topics.

Scope and Format of Paper

The number of words in the research paper about 2500 words. It will be including but not limited to the following components:

1. Title Page
2. Abstract Page
3. Introduction page
4. Main Text (sections and subsections)
5. Conclusion
6. List of References
7. Tables and Figures (if relevant)

Papers or books cited as references can be listed in the following format:

Name, "Title", Journal title/Book title, Volume No. (for papers), Issue No. (for papers) Publisher (for books), Date, Pages.

Example:

Kolmogorov, A., "Confidence limits for an unknown distribution function", Ann. Math. Statistics 12 (1941), 461-463 The questions below are meant to guide your study and help you structure your analysis.

Asking and answering the questions below should help you give structure to your analysis.

What is the general problem studied in the paper?

Give the general context of the article: indicate what concrete problem the authors are studying. Explain what they are trying to measure or to determine.

The mathematical framework

Describe in detail the variables of the problem and the origin of the data, explain the notation used by the authors.

Most situations where TDA is applied are related to statistics and the handling of large data sets. Explain the difficulties generated by the situation under discussion (the reason why new methods need to be introduced).

Topological background

Introduce the basic notions of topology the authors use to approach the problem: simplicial complexes, homology groups, persistent homology... (For this part, you should use what was discussed in class and the documents we studied together).

The TDA methods

Explain what the authors did to solve their problem by means of TDA: use of persistent homology, persistence landscapes, barcodes...

Results

What did the authors observe?

Discussion

Are the results obtained using TDA significantly better than purely statistical methods? Do you think they could be improved? Can you think of other situations where the same methods could apply?

Suggested Paper Topics

Various papers will be sent out separately. Students should feel free to explore recent literature on their own, for instance on [arXiv.org](https://arxiv.org).