EXPLORING DIRECTED VS. UNDIRECTED GRAPH-BASED TOPOLOGICAL DATA ANALYSIS OF TRANSFORMER ATTENTION MAPS

Selected Topics in DS 2024 Course

Author
Vasily Ivanov
2024

1 Project status report

Here I would describe what is good and what is bad in the report.

- Main problem is described clearly.
  As far as I understand, the problem is that attention maps of trained transformer can be a good features for further models that are based on those features. These attention maps can be analyzed as directed or undirected graphs and Topological Data Analysis method is proposed for attention maps as directed graphs.
  However it is not clear what does it means to represent attention map as directed graph and this information is unfortunately missing.

- Main idea of the approach is described in form of literature review. But I kinda missed what will be different between barcodes on directed and undirected graphs.

- As I understand in further work comparison will be provided between baseline / directed graphs / undirected graphs methods and in a scope of the student project task it seems sufficient. However basic ideas of these methods and general idea why it should work better at all is not highlighted in the text.

In my eyes general structure is good but lacks consistency or thematic glue between points. Some literature things can be reduced or removed without harm. It would be nice to add some transitions to underline logic of the report. Also I missed some key information that would help me to understand what is going on - about undirected and directed graphs topological features. There in the literature review mentioned that Kushnareva's work used some new TDA features to achieve exceptional performance in NLP tasks. Probably it would be really nice to cover what kind of features those were.
Further experiments seems reasonable and interesting to me. It is at early stage and results are not yet concluded. Actually there are metrics for the baseline only.
Presented results and metrics are reasonable although it would be interesting to know why these were selected. Probably could be briefly outlined in the dataset's description.

2 Repository

Github is not very well structured and lacks requirements.txt, or code files or clear pipeline of experiments. Dataset is mentioned in the report but not in the README.md. It has single Jupyter notebook but it is saved without outputs and content of the file is a bit messy. It contains commented code and useless parts that were kept there for diagnostics. There are unique, meaningful comments that should be put elsewhere - in the report or README.md.
I wasn't able to run the code because environment info and requirements are not present.
In my opinion each experiments should be separated into single script that runs that experiment and there should be clear instruction how to reproduce with clear environment settings. This can be achieved via Docker.

3 Conclusions

To me it seems that repo is in the early stage and can become really good with time and some love. I would suggest to focus on reproducibility and to underline key things in the report part. Also I would suggest to drop Jupyter Notebooks entirely because it is really hard to make them clean and production-quality and because they don't like version control and it is hard to track and implement code fixes using them.