Peer Rewiew №1: OMS on Cohortney

Nikolay Kozyrskiy, Gleb Balitskiy

April 2021

1. **Are project objectives clearly explained? If not write what was not clear. Provide your opinion on the problem.**

   The problem definition is stated clearly. A pithy introduction to the clustering of event sequences helps to get involved into the problem easily and quickly. This task is of great demand as clustering of event sequences is widely applicable in domains such as healthcare, marketing, and finance. The analysis of visitors to websites, hospitals, or bank transactions are among the use cases. It is clear from the introduction what is the aim of the project and what authors are going to do.

2. **Is proposed baseline solution relevant to the problem? Could you recommend something to the team?**

   The proposed solution coincides completely with the stated problem. The authors are planning to combine best six methods into one user-friendly framework with API. They clearly understand the difficulty of creating one common API as the general interfaces of all proposed methods are different. So our suggestion is to think carefully about the architecture of the framework before starting the implementation.

3. **Is it clear from the report how the team is going to test / evaluate the results?**

   The evaluation procedure is clearly stated. The authors will take the same data as was used in Cohortney paper, which is a good choice in our opinion as it will be easy to compare the results with the ones stated in the paper. The synthetic data can also provide an explicit evidence of the quality of implemented algorithms.

4. **Is it possible to guess next project development steps from the report and github repository? What a team is going to do next, provide with 3 next possible steps?**

   The first three possible steps can be seen in the report. We assume the following:

   - Thinking over a structure of the framework. In order to save time the resulting structure can be even drawn on the piece of paper by hands.

- Implementation of the first mentioned method and implementation of all the necessary environment for the rest ones.
- Implementation of other 5 methods and creating a common API.

5. **Would you recommend to improve the report and how? Or it is all good.**

   The report is very good in general. The problem statement, the aim of the project and the proposed solutions are clear and easy to understand. The only recommendation is to add some short descriptions (at least references to the original papers) for each algorithm in README.md file on GitHub in order to facilitate the use of the framework.

6. **Is project github repository easy to follow?**

   The structure of the GitHub repository is very clear and easy to follow by now.