First project peer review at STDS course

Project name: FOTIC: Fourier Transform on Continuous-time Convolutions Model of event sequences

Project author: Maria Ivanova

Reviewer: Petr Mokrov

The project under consideration deals with the problem of transactional data processing. This problem appears in various industrial applications like e-commerce, finances and so on. That is why developing of methods dealing with such data efficiently is of high importance, and the general problem statement proposed in the project is worth to be presented in Selected Topics in Data Science course.

In order to solve the problem posed, the authors propose to improve COTIC method [3]. This method takes the advantage of Continuous Convolution Neural Networks (CCCNs) [2], which are able for non-uniform occurrence of events in time. And the main idea behind the project under consideration is to speed up CCCNs with the Fourier transform [1]. Thanks to such an improvement, called FOTIC, temporary data in the corresponding applications will be processed faster and people from business and commerce will obviously benefit from the proposed improvement.

Both the problem statement and the main idea, briefly described above, are accurately presented in the project proposal manuscript under consideration in the Abstract and Introduction sections. The motivation (speeding up the computation time) and the methodology by which the desired goals are planned to be achieved (empowering CCCNs with Fourier transform) are clear and stated explicitly. In general, the style and structure of the manuscript follows scientific best-practices. The only tiny concern I found (not concern, but, better-to-say, suggestion) is to divide the Introduction section into two subsections: the first one describes the motivation and general problem description while the second one is purely about COTIC.

Concerning the Related works section, it seems to be actual and gives the full picture of the research area. The authors of the project provide not only the alternative methods itself, but point out the differences of these methods from the proposed approach and limitations. In other words, the structure of the Related Works section is in concordance with those of good and internationally-recognized scientific researches.

As for the recommendations, apart from those stated above, I have no specific concerns. I know, that the team project is quite experienced in writing scientific papers and I look forward to future revisions of their manuscript with detailed description of methodology, experiment protocol and results. Of the special interest is the time comparison investigation, which demonstrates, that the proposed method is indeed capable to overcome the baselines and alternative approaches in terms of time consumption by statistically significant margin.

At the current early stage of the project, the code is not provided. Yet, I guess that the research under consideration is planned to be submitted to a Computer Science conference/paper, that is why I am not sure if I and STDS team have a right to claim for the code.

In conclusion, I want to emphasize one more time, that the project is interesting from practical point of view and potentially applicable in various industrial applications.

Bibliography

[1] David W. Romero et al. "CKConv: Continuous Kernel Convolution For Sequential Data". In: *International Conference on Learning Representations*. 2022. URL: https://openreview.net/forum?id=8FhxBtXSl0.

[2] Hui Shi et al. "Continuous cnn for nonuniform time series". In: *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2021, pp. 3550–3554.

[3] Vladislav Zhuzhel et al. "Continuous-time convolutions model of event sequences". In: *arXiv preprint arXiv:2302.06247* (2023).