FOTIC: Fourier Transform on Continuous-time Convolutions Model of event sequences. First project status report

Maria Ivanova [1]

Abstract

Event sequences data, specifically transactional data, occur in various domains, including e-commerce, healthcare, and finance. This data is typically large in terms of amount of available data and the length of event sequences per client, and this data is often sparse and non-uniform, thus, the classic approaches for time series processing are inapplicable.

To allow continuous time, the COTIC (Zhuzhel et al., 2023) method uses specific parametric intensity functions defined at each moment on top of existing models. Due to the parametric nature, these intensities represent only a limited class of event sequences. The COTIC method is based on a continuous convolution neural network suitable for non-uniform occurrence of events in time. In COTIC, dilations and multi-layer architecture efficiently handle dependencies between events.

In this project, we go further and would like to speed up the convolution operations in the COTIC network via the Fourier transform.

1. Introduction

Processing of transactional data plays a key role in e-comm and banking industry because of the significant financial implications.

In practice existing machine learning methods are designed to make this process more effective. In most methods transactions are grouped by one type of participants (e.g. users, clients) and each of them is presented as a sequence of transactions. Next, sequence encoder (e.g. RNN (Babaev et al., 2022)) is applied to extract features for prediction.

Although such an approach enables learning similarity of transaction participants, it doesn't take into account the nature of event sequences as temporary point processes which closely related to time series with non-uniform measurements.

2. Related work

Constant advances in deep learning techniques make researchers look for efficient neural architectures to model event sequences. The work (Mei & Eisner, 2017) adapts standard LSTM network to work with non-uniform sequences as this architecture efficiently handles long-term dependencies and deals with the effects not taken into account by the Hawkes process. Despite the fact that LSTM network is superior to standard RNNs in capturing long-term time dependencies in data, it has some drawbacks as well, for example, gradient explosion or problem with the parallelization of the computational process due to the sequential nature of data supplied to the model. To overcome these issues, the Transformer Hawkes Process (THP) (Zuo et al., 2020) was proposed. It adopts the attention mechanism for event sequence modelling. While efficiently identifying long-term dependencies, it can also be optimized in a parallel manner since THP doesn't have a recurrent structure. However, training time of this model is still relatively long.

The COTIC(Zhuzhel et al., 2023) method is based on one-dimensional continuous convolutional neural network architecture and it constructs representations of a temporal point process and predicts its' intensity. The COTIC model takes into account the continuous-time structure of the point process while learning its parameters and predicting target variables from representations at any point in time. The COTIC model doesn't imply any parametric structure for the probability interpolation between events, thus it is possible to model large variety of possible dependecies between events.

The training time of the COTIC approach is comparable to that of existing baselines but has better performance than others. The only model that provides comparable results but much slower than COTIC method, is NeuralHawkes, because it is based on a continuous recurrent neural network. CKConv method (Romero et al., 2021) formulates kernels as continuous functions to process sequential data, which

[1]Skoltech, Moscow. Correspondence to: Maria Ivanova <maria.ivanova@skoltech.ru>.

Selected Topics in Data Science, 2023.

can construct arbitrarily large kernels. CKConv speed up the convolution operations via the convolution theorem with F the Fourier transform.

In our project we would like to apply the Fourier transform for the convolution operations of the COTIC method and compare the results in terms of efficiency and performance.

References

Babaev, D., Ovsov, N., Kireev, I., Ivanova, M., Gusev, G., Nazarov, I., and Tuzhilin, A. Coles: Contrastive learning for event sequences with self-supervision. In *Proceedings of the 2022 International Conference on Management of Data*, pp. 1190–1199, 2022.

Mei, H. and Eisner, J. M. The neural hawkes process: A neurally self-modulating multivariate point process. *Advances in neural information processing systems*, 30, 2017.

Romero, D. W., Kuzina, A., Bekkers, E. J., Tomczak, J. M., and Hoogendoorn, M. Ckconv: Continuous kernel convolution for sequential data. *arXiv preprint arXiv:2102.02611*, 2021.

Zhuzhel, V., Grabar, V., Boeva, G., Zabolotnyi, A., Stepikin, A., Zholobov, V., Ivanova, M., Orlov, M., Kireev, I., Burnaev, E., et al. Continuous-time convolutions model of event sequences. *arXiv preprint arXiv:2302.06247*, 2023.

Zuo, S., Jiang, H., Li, Z., Zhao, T., and Zha, H. Transformer hawkes process. In *International conference on machine learning*, pp. 11692–11702. PMLR, 2020.