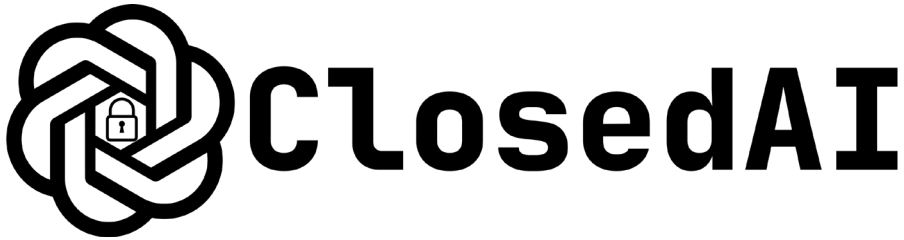
Machine Learning course 2023. Final Project 11.



Contrastive Learning for Event Sequences with Self-Supervision on multiple domains

Replication study

Denis Grankin Ekaterina Andreichuk Ivan Apanasevich

Irena Gureeva Mikhail Konenkov

Motivation

Learning best algorithms for lifestream data embeddings

Extracting embeddings from event sequences is a useful approach that can be used for learning users preferences, advising customers products based on their clickstream data and many other real-world applications. However, classical methods for analyzing event sequences that are attributed to a person and capture their regular and routine behavior do not capture relationships between single object and its immediate neighborhood very well. Also, it usually requires a labeled data to operate this.
Therefore, new methods for data augmentation and event sequences extraction is needed.

Related work

CoLES: Contrastive Learning for Event Sequences with Self-Supervision

Dmitrii Babaev
AIRI
Sber AI Lab
Moscow, Russia

Nikita Ovsov
Ivan Kireev
Maria Ivanova
Sber AI Lab
Moscow, Russia

Gleb Gusev
Sber AI Lab
MIPT
Moscow, Russia

Ivan Nazarov
AIRI
Moscow, Russia

Alexander Tuzhilin
New York University
New York, USA

The paper on which our replication study is based on.

Event sequence metric learning

Dmitrii Babaev*
Sberbank AI Lab

Ivan Kireev
Sberbank AI Lab

Nikita Ovsov
Sberbank AI Lab

Mariya Ivanova
Sberbank AI Lab

Gleb Gusev
Sberbank AI Lab

Alexander Tuzhilin
New York University

Metric learning method for producing embeddings of complex event sequences. The main advantage is that almost no pre-processing is needed for complex even streams to get their compact embedding.

Attributed Sequence Embedding

Zhongfang Zhuang, Xiangnan Kong, Elke, Rundensteiner
Worcester Polytechnic Institute
{zzhuang, xkong, rundenst}@wpi.edu

Jihane Zouaoui, Aditya Arora
Amadeus IT Group
{jihane.zouaoui, aditya.arora}@amadeus.com

Study about extracting data embeddings from attributed sequences.

A Framework For Contrastive Self-Supervised Learning And Designing A New Approach

Technical Report

William Falcon
New York University, NY
Lightning Labs, NY
waf251@nyu.edu

Kyunghyun Cho
New York University, NY
kc119@nyu.edu

Study with analysis of contrastive learning approaches and conceptual framework for Contrastive Learning.

CoLES framework

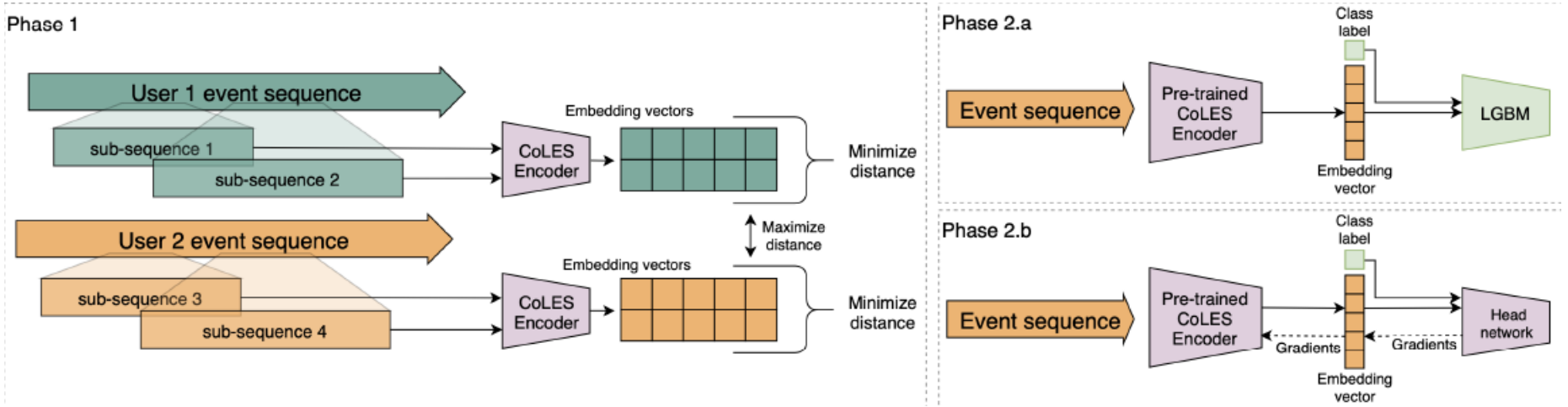


Figure 1: General framework. Phase 1: Self-supervised training. Phase 2.a Self-supervised embeddings as features for supervised model. Phase 2.b: Pre-trained encoder fine-tuning.

DataFusion 2022 datasets

ClosedAI

Hyperparameters

Conducted Experiments

Analyzed methods

ClosedAI

Skoltech

General pipeline

Transactions

Clickstreams. PySpark

ClosedAI

Results

Results

Contribution

Conclusion