COMPUTATIONAL
Social Science@CUHK

WHEN SOCIAL SCIENCE MEETS COMPUTER SCIENCE...

Webinar of Computational Social Science Laboratory (CSSL@CUHK)

life2vec: Life trajectories in high dimensional spaces

14 November 2024 (Thursday), 09:30 - 11:00 (UTC+8, HKT)

Abstract

Here we represent human lives in a way that shares structural similarity to language, and we exploit this similarity to adapt natural language processing techniques to examine the evolution and predictability of human lives based on detailed event sequences. We do this by drawing on a comprehensive registry dataset, which is available for Denmark across several years, and that includes information about life-events related to health, education, occupation, income, address and working hours, recorded with day-to-day resolution. We create embeddings of life-events in a single vector space, showing that this embedding space is robust and highly structured. Our models allow us to predict diverse outcomes ranging from early mortality to personality nuances, outperforming state-of-the-art models by a wide margin. Using methods for interpreting deep learning models, we probe the algorithm to understand the factors that enable our predictions. Our framework allows researchers to discover potential mechanisms that impact life outcomes as well as the associated possibilities for personalized interventions.

Dr. Germans Savcisens
Postdoctoral Research Associate
Khoury College of Computer Sciences
Northeastern University

Biography

Germans is a postdoctoral research associate at the Khoury College of Computer Sciences, working with Professor Tina Eliassi-Rad. His research focuses on enhancing the transparency and fairness of machine learning algorithms used in network science, with a particular interest in mitigating biases in algorithms that model human behavior and social interactions. Germans obtained his PhD from the Technical University of Denmark, where he focused on using foundation models to study individual socioeconomic and health trajectories. His work explored the ability of such models to compress multimodal time series and provide interpretable predictions.

Webinar Links

CSSL@CUHK