

TokyoTech-POSTECH EconoSocioPhysics Workshop

19th Dec. Tuesday, 2023

Room 811, Building G5, Suzukake-dai Campus, Tokyo Tech

- 13:00- Woo-Sung Jung and Misako Takayasu
Opening remarks of this workshop
- 13:10- Jaehyuk Park (Assistant Professor, KDI School)
Labor space: A unifying representation of the labor market via large language models
- 13:40- Arthur Matsuo Yamashita Rios de Sousa (Assistant Professor, Tokyo Tech)
Predictive model for Sony TV weekly sales in the United States
- 14:10- Damin Lee (PhD Student, POSTECH)
Socioeconomic diversity and connectedness in human mobility + Scientific Mobility
- 14:40- Zhihua Zhong (Ph.D. Student, TokyoTec)
Novel approaches to urban human mobility: a physical analogy of electric circuit network and resulting gravity relations based on GPS data
- 15:10-15:30 Break
- 15:30- Sanggu Kang (PhD Student, POSTECH)
Urban green space inequality across urban scales
- 16:00- Qianyun Wu (Maggie) (Ph.D. Student, TokyoTech)
Unraveling collective responses during COVID-19 pandemic in Chinese and Japanese social media
- 16:30- Seonbin Jo (PhD Student, POSTECH)
Exploration Behaviors in Non-Fungible Token Collections
- 17:00- Jun'ichi Ozaki (Assistant Professor, Tokyo Tech)
Modeling of COVID-19 dynamics using GPS data: from human mobility to immunity
- 17:30-18:00 Discussion

Title: Labor Space: A Unifying Representation of the Labor Market via Large Language Models

Speaker: Jaehyuk Park (Assistant Professor, KDI School of Public Policy and Management)

Data: NAICS, O*NET, ESCO, Crunchbase (USA)

Abstract: The labor market is a complex ecosystem comprising diverse, inter-connected entities, such as industries, occupations, skills, and firms. Due to the lack of a systematic method to map these heterogeneous entities together, each entity has been analyzed in isolation or only through pairwise relationships, inhibiting comprehensive understanding of the whole ecosystem. Here, we introduce Labor Space, a vector-space embedding of heterogeneous labor market entities, derived through applying a large language model with fine-tuning. Labor Space exposes the complex relational fabric of various labor market constituents, facilitating coherent integrative analysis of industries, occupations, skills, and firms, while retaining type-specific clustering. We demonstrate its unprecedented analytical capacities, including positioning heterogeneous entities on an economic axes, such as 'Manufacturing–Healthcare'. Furthermore, by allowing vector arithmetic of these entities, Labor Space enables the exploration of complex inter-unit relations, and subsequently the estimation of the ramifications of economic shocks on individual units and their ripple effect across the labor market. We posit that Labor Space provides policymakers and business leaders with a comprehensive unifying framework for labor market analysis and simulation, fostering more nuanced and effective strategic decision-making.

Title: Predictive model for Sony TV weekly sales in the United States

Speaker: Arthur Matsuo Yamashita Rios de Sousa (Specially Appointed Assistant Professor, TokyoTech)

Data: Sales data of TV in USA

Abstract: The development of models to forecast the sales of a product is fundamental for any business. Accurate sales predictions allow leaders to adjust the company's strategy, maximizing profits and reducing costs. In this ongoing work in collaboration with Sony Corporation, we propose a model to predict TV weekly sales of Sony in the United States. The modeling in week resolution presents some challenges since this way to divide time is not regular (e.g., some years with 52 and others with 53 weeks) and there are special events that impact sales, such as the Super Bowl, which do not occur in the same week every year. After creating the concept of analogous weeks that enable us to compare weeks of different years, we construct our model by focusing on the weekly share of sales over a given period. The weekly shares vary less than the actual sales values and can be modeled using the Dirichlet distribution. In preliminary results, we obtained an average error of 10%~15% for years 2021 and 2022. We also present the prediction for the second half of 2023, which will be compared with the actual sales. This work contributes to the elucidation of sales dynamics, giving hints on consumer psychology and sales strategy effectiveness, and aims at a practical tool to be used in the business environment.

Title: Socioeconomic diversity and connectedness in human mobility

Speaker: Damin Lee (Ph.D. Student, POSTECH)

Data: SafeGraph human mobility data, US Census Bureau demographic data

Abstract: Socioeconomic segregation and mixing patterns in urban mobility indicate economic inequality in cities. We analyze the socioeconomic inequality patterns considering both individuals' selective movements and geographical constraints. We capture the individuals' movement propensity but also highlights the connections across different income-level areas within cities. Our research has the potential to improve our understanding of the intricate relationship between mobility, segregation, and equity in urban environments, Also, the efforts of understanding socioeconomic diversity and connectedness can promote greater social integration.

Title: Novel approaches to urban human mobility: a physical analogy of electric circuit network and resulting gravity relations based on GPS data

Speaker: Zihua Zhong (Ph.D. Student, TokyoTech)

Data: GPS data of mobile phones in Japan

Abstract: Human mobility in an urban area is complicated; the origins, destinations, and transport methods of each person differ. The quantitative description of urban human mobility has recently attracted the attention of researchers, and it highly related to urban science problems. Herein, combined with physics inspiration, we introduce a revised electric circuit model (RECM) in which moving people are regarded as charged particles. Analogical concepts of electromagnetism enable us to capture the characteristics of urban human mobility. We discover a nontrivial scaling relation, which is regarded as a new type of gravity model, to characterize the heterogeneous spatial distribution of urban human mobility. Furthermore, we propose a route generation model (RGM) to simulate a human flow pattern that automatically determines suitable routes between a given origin and destination as a source and sink, respectively. These discoveries lead to new approaches to the solution of many urban science problems.

Title: Urban green space inequality across urban scales

Speaker: Sanggu Kang (Ph.D. Student, POSTECH)

Data: Sentinel-2 satellite data, Copernicus built-up area data, US Census Bureau demographic data

Abstract: The motivation of this study is to investigate the unequal distribution of green spaces based on income and understand the factors contributing to this disparity. The key finding reveals a positive correlation between income and green space availability, with a higher degree of income inequality associated with more unequal access to green spaces in MSAs. This research underscores the importance of considering distribution patterns, rather than solely increasing the quantity of green spaces, when formulating policies to address this inequality.

Title: Unraveling collective responses during COVID-19 pandemic in Chinese and Japanese social media

Speaker: Qianyun Wu (Maggie) (Ph.D. Student, TokyoTech)

Data: Social media data (Chinese Weibo and Japanese Twitter)

Abstract: Observing the collective response to extraordinary events based on social media data is an important topic in social science. It helps us understand how humans react to crisis and interpret the underlying causes of the collective movements. We focus on two aspects of the topic: detect extraordinary changes based on macroscopic time series at group level, and then zoom into mesoscopic level to analyze the influence of communities on individual users. To detect the extraordinary changes, the bursts, we study the Chinese Weibo related to COVID-19 during January to October 2020. We segregate the long-term emotion time series into homogenous sections and classify them to different types of bursts based on statistical test. To analyze the influence of communities, we study the Japanese tweets and retweets related to COVID-19 vaccine during January 2020 to May 2022. We jointly study the shifts of individual opinions and the dynamic community affiliation (based on retweet network) to understand the influence of communities on changing individual's opinions towards vaccine.

Title: Exploration Behaviors in Non-Fungible Token Collections

Speaker: Seonbin Jo (Ph.D. Student, POSTECH)

Data: NFT image data, NFT trade data

Abstract: The Exploration-Exploitation dilemma is a fundamental concept in the decision-making process. We adapt the Lévy flight random walk model to NFT purchase history data to discern exploration patterns of wallets. As a result, we found significant variance in the market, but wallets generally tend to purchase similar collections. We also suggest that our Lévy flight-based model performs exceptionally well in predicting the similarity of the next collection, demonstrating potential applications in recommendation models.

Title: Modeling of COVID-19 dynamics using GPS data: from human mobility to immunity

Speaker: Jun'ichi Ozaki (Assistant Professor, TokyoTech)

Data: GPS data, Epidemic data

Abstract: COVID-19 has posed a new challenge to the globalized world. Like weather forecasting, the precise and comprehensive modeling of the infection spread is indispensable for living with COVID-19. In this study, we propose a model to predict the effective reproduction number by large-scale human mobility GPS tracking data. First, we categorize the daily activities observed in the GPS trajectories into four patterns to take account of the infection rate difference in each everyday scene. Second, we demonstrate that the effective reproduction number is theoretically expressed by a linear combination of the number of social contacts with a few approximations and assumptions. After the human-flow effect on the effective reproduction number is estimated, the other factors are considered comprehensively: the new variants, the vaccination, and immunity.