Li Zhe

📞 139-8020-6966   ✉ zheli20@fudan.edu.cn
📍 Fudan University, 220 Handan Road, Yangpu District, Shanghai

EDUCATION BACKGROUND

Fudan University   Sep 2020 - Jun 2023
Statistics Master School of Data Science   Shanghai

- Research Areas: Distributed Algorithms, Spatial Econometrics, Complex Network Analysis
- Honors/Awards: Postgraduate Outstanding Student Scholarship Second Prize (2020-2021; 2021-2022); Sixth National Statistics Doctoral Academic Forum Thesis First Prize (Top5); Fudan University Big Data School Second Doctoral Student Forum Third Prize

Fudan University   Sep 2016 - Jun 2020
Physics (Data Science and Big Data Technology) Bachelor School of Data Science   Shanghai

- Honors/Awards: Undergraduate Excellence Scholarship Third Prize (2016-2019); Undergraduate Professional Scholarship (2017-2019); Graduate Scholarship (2019-2020)
- Core courses: Advanced Mathematics, Linear Algebra, Probability Theory and Mathematical Statistics, Statistical Computing, Machine Learning, Deep Learning and Neural Networks

RESEARCH EXPERIENCE

- Wu, S., **Li, Z.**, & Zhu, X. (2020). Distributed Community Detection for Large Scale Networks Using Stochastic Block Model. "Computational Statistic & Data Analysis", accepted, [joint first author]
- Ren, Y., **Li, Z.**, Zhu, X., Gao, Y., & Wang, H. (2022). Distributed Estimation and Inference for Spatial Autoregression Model with Large Scale Networks. "Journal of Econometrics", under revision, [joint first author]
- **Li, Z.**, Zhu, X., & Zou, C. (2022). Consistent Selection of the Number of Groups in Panel Models via Sample-Splitting. submitted to "Journal of Econometrics", [first author]
- Han, S., Zhou, J., Zhu, X., **Li, Z.**, Liu, J., Wang, H., & Gong, Y. (2022). High-Resolution Image Classification with Rich Text Information Based on Graph Convolution Neural Network. submitted to "IJCNN 2023"
- **Li, Z.**, Zhu, X., & Zou, C. Consistent Selection of the Number of Factors in Factor Model via Sample-Splitting (working paper).
- **Li, Z.**, Zhu, X., & Zou, C. Statistical Inference for the Number of Groups in Panel Models (working paper).

PROFESSIONAL EXPERIENCE

Byte Dance   Jul 2022 - Sep 2022
Algorithm engineer Data-Comments   Shenzhen

- According to the product manager's requirements, for the business line of products (trembles, watermelon video, etc.) sort of intervention, the use of power, power, filtering keywords and other strategy-based methods to improve the product review area quality rate, increase the review area users stay time.
- Using the Kafka framework, the shaking audio and video comment data is captured in real time, and descriptive analysis is used to find useful features for comment intervention.
- Using wind control data, identify the brush like, brush reply behavior, and further analysis of such comments whether there is a pornographic diversion, by analyzing the characteristics of the comments, the development of the corresponding suppression strategy and model, the Top 1000 comments of the somatosensory exposure index from 7.8 % to 3.7 %.

Microsoft Research Asia   May 2022 - Jul 2022
Algorithm engineer Software Analytics Group   Beijing

- Using Kusto to grab the company's server operation log, the Microsoft cloud server machine operation data descriptive statistics and analysis, focusing on the analysis of machine operation law and characteristics.
- uses Kusto to capture the company's server running logs, and by analyzing the server CPU, memory and other data, it is found that the data has obvious periodicity. Use ARIMA model to model and predict the subsequent utilization rate, and realize the "peak elimination and valley filling" of resources.
- Use Kusto to capture the company's server running logs, through the running logs, analysis of different nodes of the operating mode, and the use of nodes in a certain period of time the mean and standard deviation of the resources of the node classification, by identifying low utilization nodes to achieve resource migration, improve resource utilization.

PROFESSIONAL SKILLS

- Programming languages and software used: Python, R, Matlab, Shell, SQL, Java, Scala, LaTeX
- Deep learning frameworks: Pytorch, TensorFlow, DGL
- Distributed frameworks: Hadoop, Spark
- Foreign language ability: CET6, IELTS 6.0