

Dr. Ziyue LI (Bonald)

Ph.D. in Data Science and Machine Learning
Department of Industrial Engineering and Decision Analytics (IEDA)
The Hong Kong University of Science and Technology

Email: zlibn@connect.ust.hk

Website: <https://bonaldli.github.io>

Tel: +852 69964596

Education

Ph.D. in Industrial Engineering and Decision Analytics *09.2017-09.2021*
The Hong Kong University of Science and Technology Hong Kong

- **Concentration:** Data Science; Machine learning models & algorithms; Statistics.
- **Others:** Interpretable Machine Learning, NLP, Trustworthy AI.
- **Advisor:** Prof. Fugee Tsung, collaborating with Prof. Hao Yan (Arizona State University, U.S.A), and Prof. Chen Zhang (Tsinghua University, China)

Exchanging in Mechanical and Manufacturing Engineering *2015-2015*
University of New South Wales Sydney, Australia

- China Scholarship Council (CSC) Scholarship (Distinguished)

Bachelor of Engineering in Mechanical Engineering *2013-2017*

Bachelor of Economics in Finance *2015-2017*

Xi'an Jiaotong University Xi'an, China

- GPA: 3.76/4.3 Ranking: 6th/255, Top 5%

- Outstanding Graduates Award, National Scholarship

Research Interests

My research interests focus on big data analytical methodologies for real-world problems. The goal is to build up novel models that preserve the innate data structure, and **combines the machine learning models with domain/human knowledge**, for higher accuracy, efficiency, explainability, and interpretability. My current research includes transfer learning (multi-task learning), anomaly detection, spatiotemporal big data and tensor techniques (tensor decomposition, completion), knowledge graph and topic models (NLP), and so on. Those works have been published in top-tier scientific venues (AAAI, IEEE ICDE, CASE, KDD, TKDD), awarded with **four best paper awards**, and applied to smart mobility, NLP, and manufacturing. I will be joining University of Cologne (Germany) as an Assistant Professor.

Selected Awards

- **Best Theoretical Paper Award**, INFORMS Data Mining and Decision Analytics Workshop (2021): Runner-up.
- **Best Applied Paper Award**, INFORMS Data Mining and Decision Analytics Workshop (2021): Finalist.
- **2nd Runners-up**, HKUST Three Minute Thesis Competition (2021).
- **Best Student Paper Award**, INFORMS Data Mining Section (2020): Finalist, selected out of 40 submissions.
- **Best Student Paper Award**, INFORMS Quality, Statistics, and Reliability (QSR) Section (2020): Finalist, selected out of 26 submissions.
- **Best Conference Paper Award**, IEEE International Conference on Automation Science and Engineering (CASE) (2020): Selected out of 500 submissions.
- **1st Runners-up, Audience Award**, Hackathon@UST - Uber Smart Transportation (2018).

- **Hong Kong Ph.D. Fellowship Scholarship Award** (2017-2020): Highly-selective and prestigious, 2 recipients in IEDA.
- **National First Prize**, “Challenge Cup” National Curricular Academic Science and Technology Contest (2015): A data-driven study for land transfer in rural China (Applied Economics).

Research Experience

Artificial Intelligence & Ethics 04.2021 - 07.2021

RWTH Aachen, Germany; TU Delft, Netherlands; Politecnico di Milano, Italy.

- Research in responsible trustworthy AI for social good (algorithm fairness, diversity, non-discrimination).

Graph-regularized / Graph embedding tensor topic model 2019 - present

Keywords: interpretable ML, NLP, generative model, big data, graph-structure

- Developed high-dimensional tensor topic models (LDA, DMM model)
- Incorporated external information as graph to improve model interpretability
- Sped up learning via online variational EM algorithm with stochastic gradient

Distributed Machine Learning System @ Bell Labs Germany 2019-2020

Keywords: big data system, cloud computing, serverless machine learning

- System latency analysis and component profiling
- Designed and executed machine learning model training and inference in a serverless system, with data/model parallelism.

Tensor decomposition & completion prediction on graph data 2018 - 2020

Keywords: smart transport, big data, time-series data, spatiotemporal data, graph

- Developed tensor decomposition based prediction with 2D-ARIMA model, and low-rank tensor completion based prediction for spatiotemporal data
- Incorporated graphs structure to improve the prediction accuracy
- Adopted block coordinate descent algorithm for efficient learning

Transfer learning for anomaly detection

2017 - 2019

Keywords: small data analysis, cold-start, transfer learning, multi-task learning

- Proposed hybrid transfer learning model with parameter transfer and feature representation transfer approach
- Proposed a decomposition model for data profile to detect anomaly accurately

Conference Publications

[C1] **Z. Li**, H. Yan, C. Zhang, and F. Tsung, “Choose A Table’: Tensor Dirichlet Process Multinomial Mixture Model with Graphs for Passenger Trajectory Clustering”. the 36th AAAI Conference on Artificial Intelligence, 2022, *under review*.

[C2] **Z. Li**, “Tensor Topic Models with Graphs and Applications on Individualized Travel Patterns”, 2021 IEEE 37th International Conference on Data Engineering (ICDE), Crete, Greece, 2021, accepted (*oral*) and published.

- **Best Applied Paper Award**, INFORMS Data Mining and Decision Analytics Workshop 2021, **Finalist**.
- ICDE: Top-tier conference in data mining (acceptance rate 19.1%).

[C3] **Z. Li**, H. Yan, C. Zhang and F. Tsung, “Long-Short Term Spatiotemporal Tensor Prediction for Passenger Flow Profile” in IEEE 16th International Conference on Automation Science and Engineering (CASE), 2020, published.

- **Winner** of IEEE CASE 2020 **Best Conference Paper Award**.

[C4] **Z. Li**, N. D. Sergin, H. Yan, C. Zhang, and F. Tsung, “Tensor Completion for Weakly-Dependent Data on Graph for Metro Passenger Flow Prediction” Proceedings of the AAAI Conference on Artificial Intelligence, 2020, accepted (*oral*) and published.

- **Best Student Paper Award**, Quality, Statistics, and Reliability (QSR) Section, INFORMS 2020, **Finalist Award**.
- AAAI: Top-tier conference in machine learning, AI (acceptance rate 20.6%).

[C5] D. Li, Z. Yu, **Z. Li**, R. Zhao, “RANG-AT&T: RoAd Network Graph Augmented Transformer for Trajectory Representation Learning”. the 36th AAAI Conference on Artificial Intelligence, 2022, *under review*.

[C6] Z. Yu, D. Li, **Z. Li**, R. Zhao, “Trajectory-guided Contrastive Representation Learning on Road Network Graph”. the 36th AAAI Conference on Artificial Intelligence, 2022, *under review*.

Journal Publications

[J1] **Z. Li**, H. Yan, C. Zhang, and F. Tsung, “Individualized Passenger Travel Pattern Multi-Clustering based on Graph Regularized Tensor Latent Dirichlet Allocation”. Data Mining and Knowledge Discovery, Springer, *accepted and incoming*.

- **Best Student Paper Award**, Data Mining, INFORMS 2020, **Finalist Award**.

[J2] **Z. Li**, K. Zhang, H. Yan and F. Tsung, “Profile Decomposition based Hybrid Transfer Learning for Cold-start Data Anomaly Detection”. The ACM Transactions on Knowledge Discovery from Data (TKDD), 2021, *accepted and incoming*.

[J3] **Z. Li**, H. Yan, C. Zhang and F. Tsung, “Long-Short Term Spatiotemporal Tensor Prediction for Passenger Flow Profile” IEEE Robotics and Automation Letters, 2020.

[J4] F. Tsung, **Z. Li**, “Discussion of ‘A novel approach to analysis of spatial and functional data over complex domains’” Quality Engineering, 2020, published.

Working Papers

[W1] **Z. Li**, H. Yan, C. Zhang, and F. Tsung, “Tensor Dirichlet Process Multinomial Model with Graph Embedding for Individual Passenger Clustering,” working paper.

[W2] N. stoimenova, **Z. Li**, A. Angelucci, F. Berardinucci, “Conceptual Exploration into the Fairness Challenges of Developing AI Systems: a Case Study on Corporate Wellness Programs”, working paper.

[W3] M. Li, **Z. Li**, and F. Tsung, “Spatiotemporal Attention-based Auto-Encoder for Dynamic Metro Station Clustering,” working paper.

Teaching Experience

Teaching Assistant, MATH5470-Statistical Machine Learning (PG) *Spring, 2021*

- Took charge of tutorials covering contents of supervised learning such as regression, classification, regularization, neural networks and support vector machine; and unsupervised learning such as random forests.

Teaching Assistant, IEDA2200-Engineering Management (UG) *Spring, 2020*

- Took charge of all tutorials covering contents of production line optimization, time-series forecasting, inventory management and manufacturing strategy.

Co-instructor, IEDA3270-Quality Engineering (UG) *Spring, 2019*

- Led all lab experiments including statistics, factorial design, control charts, and regression models.

	Teaching Assistant , EEMT5220-Six Sigma Quality Management (PG) <i>Fall, 2018</i>	
	<ul style="list-style-type: none"> • Took charge of all tutorials covering contents of statistical analysis, hypothesis testing, statistical process control. 	
Industry Experience	<i>Data Mining Researcher</i>	<i>09.2021 - 12.2021</i>
	SenseTime, Hong Kong Science and Technology Park	Hong Kong
	<ul style="list-style-type: none"> • Lead research in reinforcement learning and deep learning for combinatorial optimization. • Guided research in Machine learning research in smart cities and manufacturing. 	
	<i>Project Manager</i>	<i>09.2020 - present</i>
	Hong Kong Metro MTR Co.-HKUST R&D Project	Hong Kong
	<ul style="list-style-type: none"> • Research collaboration with MTR and Computer Science Dept, HKUST, developing an Intelligent Transport System (ITS) based on data-driven methods. 	
	<i>Cloud Computing Scientist (Intern)</i>	<i>09.2019 - 02.2020</i>
	Nokia Bell Labs	Stuttgart, Germany
	Keywords: <i>cloud computing, distributed system, FaaS, machine learning system</i>	
	<ul style="list-style-type: none"> • Research in serverless computing, machine learning systems based on Amazon Web Service (AWS), and Bell Labs KNIX MicroFunctions. • Conducted data transmission latency analysis and severless system component profiling and optimization; • Conducted serverless machine learning inference and training (regression, NLP, image recognition based on ResNet, MXNet, TensorFlow) in AWS and KNIX. 	
Competitions	Finalist Award in Belt & Road Justice LawTech Hackathon	<i>08.2018 - 10.2018</i>
	HK Law Society	Hong Kong
	<ul style="list-style-type: none"> • Legal service recommender system based on NLP and Semantic & Cognitive analysis for domestic workers. 	
	Top 10 in Global Big Data Competition	<i>07.2018 - 08.2018</i>
	JD.com	Beijing, China
	<ul style="list-style-type: none"> • Quantile boosting time-series prediction model with effective feature engineering and representation methods and improved forecast accuracy by 72%. 	
	1st-Runners-up, Audience Award in Hackathon@UST	<i>04.2018 - 05.2018</i>
	Uber Smart Transportation	Hong Kong
	<ul style="list-style-type: none"> • Developed facial recognition model and integrated into vehicles to achieve facial car door unlocking, underage driving detection, and drowsy driving detection. 	
Invited Talks	1. Session Chair , " <i>Knowledge-integrated Data-driven Smart Transportation</i> "	<i>10.2021</i>
Session Chair	INFORMS Data Mining section 2021	Anaheim, California, U.S.A
	2. Invited Presenter , " <i>Individualized Passenger Travel Pattern Multi-Clustering based on Tensor Latent Dirichlet Allocation with Graph Structure</i> "	<i>11.2020</i>
	INFORMS Annual Meeting 2020	National Harbor, Maryland, U.S.A
	3. Invited Presenter , " <i>Tensor Completion for Weakly-dependent Data on Graph for Metro Passenger Flow Prediction</i> "	<i>10.2019 and 01.2020</i>
	Data Science Symposium Waseda University	Tokyo, Japan
	INFORMS Annual Meeting 2019	Seattle, Washington, U.S.A
	4. Invited Presenter , " <i>Transfer-learning-based Anomaly Detection for Monitoring</i> "	

**Other
Awards**

- **Honored Graduate Award**, Xi'an Jiaotong University 2017
- **Regional First Prize**, 7th Mechanical Innovation & Design Competition 2016
- **National Scholarship**, Xi'an Jiaotong University 2016
- **1st-Runners-up Award**, XJTU Entrepreneurship Competition 2015
- **National Encourage Scholarships**, Xi'an Jiaotong University 2014, 2015

Skills

- Python • R • Matlab • AWS • C++ • Hadoop, Map-Reduce
- TensorFlow • PyTorch • Keras • Scikit-learn
- Psychology • Consulting • Public Speaking
- **English** (Fully Professional) • **German** (B2, Fluent) • Chinese (Native)

Reference

- **Prof. Fugee Tsung** (Ph.D. Supervisor)
Chair Professor, Acting Dean
Department of Industrial Engineering and Decision Analytics
The Hong Kong University of Science and Technology, Hong Kong
Email: season@ust.hk, Phone: +852 2358-7097
- **Prof. Hao Yan** (Research Collaborator, Co-Supervisor)
Assistant Professor
School of Computing, Informatics & Decision Systems Engineering
Arizona State University, U.S.A
Email: haoyan@asu.edu, Phone: +1 (480) 727-0556
- **Prof. Chen Zhang** (Research Collaborator, Co-Supervisor)
Associate Professor
Department of Industrial Engineering, Tsinghua University, China
Email: zhangchen01@tsinghua.edu.cn, Phone: +86-10-62796135