

Zexi Huang

MACHINE LEARNING SCIENTIST AT TIKTOK

✉ zexihuang.phd@gmail.com | 🏠 zexihuang.com | 📧 zexihuang | 🌐 zexihuang | 🗣️ Zexi Huang

Education

College of Engineering, University of California, Santa Barbara (UCSB)

Santa Barbara, USA

DOCTOR OF PHILOSOPHY AND MASTER OF SCIENCE IN COMPUTER SCIENCE

Sep. 2018 - Apr. 2023

- Course GPA: 4.0/4.0, UCSB Computer Science *Outstanding Scholar* Fellow (top 4 out of 63 PhD students).
- Advisor: Prof. Ambuj Singh. Dissertation: Learning Representations for Information-rich Graphs.

Yingcai Honors College, University of Electronic Science and Technology of China (UESTC)

Chengdu, China

BACHELOR OF ENGINEERING IN COMPUTER SCIENCE AND TECHNOLOGY (HONORS)

Sep. 2014 - Jun. 2018

- GPA: 3.96/4.0, Avg. Score: 92.79/100 (1/90 in freshman year, 1/87 in sophomore year, 1/93 in junior year).
- Recipient of the *Most Outstanding Students Awards* of UESTC (top 10 out of 5,000 seniors).
- Advisors: Prof. Junming Shao and Prof. Sinno Jialin Pan. Thesis: Transfer Learning for Community Detection.

Work Experience

Machine Learning Scientist, Live US, TikTok

San Jose, USA

REPRESENTATION LEARNING FOR TIKTOK LIVE RECOMMENDATION

May. 2023 - Present

- Optimizing user, content, and feature embeddings with deep representation learning models for the recommendation system of Live.

Applied Scientist Intern, Books Tech, Amazon

San Diego, USA

STOCHASTIC INVENTORY MANAGEMENT FOR PRINT-ON-DEMAND AND GRAPH-BASED TEXT CLASSIFICATION FOR CONTENT INTELLIGENCE

Jun. 2022 - Sep. 2022

- Developed a stochastic inventory management model based on dynamic programming for Amazon's Print-On-Demand business, optimizing the ordering strategy for 42.2M units of books and realizing an annual saving of \$10.4M.
- Designed and implemented a graph-based NLP model for long text embedding and classification using graph neural networks, leading to superior performance to state-of-the-art transformer-based models for Kindle book contents.

Applied Scientist Intern, Books Tech, Amazon

San Diego, USA

GRAPH-BASED FRAUD DETECTION IN KINDLE DIRECT PUBLISHING

Jun. 2021 - Sep. 2021

- Designed graph-based machine learning models for fraud detection based on multi-modal signals in Kindle Direct Publishing.
- Implemented the models in an end-to-end fashion and deployed them to graphs with millions of nodes and billions of edges in production.
- Validation results show that the models can surface fraud rings undetected by existing processes with an estimated annual value of \$2.4M.

Applied Scientist Intern, Books Tech, Amazon

San Diego, USA

GRAPH-BASED FRAUD DETECTION IN KINDLE DIRECT PUBLISHING

Jun. 2020 - Sep. 2020

- Proposed to augment existing fraud detection methods with graph-based machine learning models for Kindle Direct Publishing.
- Designed and implemented various heuristics and an embedding framework for attributed heterogeneous multiplex networks.
- The models are deployed into production and results show that they surface up to 15 times more fraud compared to the existing processes.

Research Intern, Computational Intelligence Lab, Nanyang Technological University

Singapore

TRANSFER LEARNING FOR COMMUNITY DETECTION IN MULTIPLEX NETWORKS

Sep. 2017 - Feb. 2018

- Proposed to refine community detection results in some layers with transferred knowledge from other layers in multiplex networks.
- Designed a representation-based community detection framework and implemented it with an extended symmetric NMF approach.
- Our algorithm outperforms other representation-based community detection algorithms, especially when the target layer is noisy.

Publications

[WSDM-2023] Zexi Huang*, Mert Kosan*, Sourav Medya, Sayan Ranu, Ambuj Singh. **Global Counterfactual Explainer for Graph Neural Networks**. *ACM International Conference on Web Search and Data Mining*, 2023. (*: equal contribution)

[WSDM-2022] Zexi Huang, Arlei Silva, Ambuj Singh. **POLE: Polarized Embedding for Signed Networks**. *ACM International Conference on Web Search and Data Mining*, 2022.

[KDD-2021] Zexi Huang, Arlei Silva, Ambuj Singh. **A Broader Picture of Random-walk Based Graph Embedding**. *ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 2021.

[Preprint] Zexi Huang, Mert Kosan, Arlei Silva, Ambuj Singh. *Link Prediction without Graph Neural Networks*. arXiv preprint arXiv:2305.13656, 2023.

[Preprint] Wei Ye, Zexi Huang, Yunqi Hong, Ambuj Singh. *Graph Neural Diffusion Networks for Semi-supervised Learning*. arXiv preprint arXiv:2201.09698, 2022.

Research Experience

Link Prediction without Graph Neural Networks

Santa Barbara, USA

ADVISOR: PROF. AMBUJ SINGH, COLLABORATORS: MERT KOSAN, ARLEI SILVA

May. 2021 - Feb. 2023

- Scrutinized the training and evaluation of supervised link prediction methods and identify their limitations in handling class imbalance.
- Proposed a novel topology-centric framework that combines graph learning, topological heuristics, and an N-pair loss for link prediction.
- Results showed that the proposed method is 145% more accurate and trains/infers 11/6,000 times faster than the state-of-the-art methods.

Global Counterfactual Explanation for Graph Neural Networks

Santa Barbara, USA

ADVISOR: PROF. AMBUJ SINGH, COLLABORATORS: MERT KOSAN, SOURAV MEDYA, SOURAV MEDYA, SAYAN RANU

Oct. 2021 - Aug. 2022

- Formulated the novel problem of global counterfactual reasoning/explanation of graph neural networks for graph classification.
- Proposed GCFExplainer, the first global explainer powered by vertex-reinforced random walks on an edit map with a greedy summary.
- Results showed that GCFExplainer not only provides crucial high-level insights but also outperforms existing methods in recourse quality.

Signed Embedding for Polarized Graphs

Santa Barbara, USA

ADVISOR: PROF. AMBUJ SINGH, COLLABORATOR: ARLEI SILVA

Oct. 2020 - Aug. 2021

- Designed a novel polarization measure for signed graphs and showed that existing methods fail in polarized signed link prediction.
- Proposed a polarized embedding algorithm that captures both topological and signed similarity jointly via signed autocovariance.
- Extensive experiments showed that the proposed model outperforms state-of-the-art methods by up to one order of magnitude.

Multiscale Graph Convolution via Neural Diffusions

Santa Barbara, USA

ADVISOR: PROF. AMBUJ SINGH, COLLABORATOR: WEI YE, YUNQI HONG

Apr. 2020 - Oct. 2021

- Interpreted the layer-wise propagation rule of GCN from the perspective of power iteration and analyzed its converging process.
- Designed a novel GCN architecture that learns to aggregate multiscale information based on graph diffusions with a neural network.
- Illustrated the effectiveness and efficiency of the proposed model by extensive comparative studies with state-of-the-art methods.

Graph Representation Learning Based on Random-walks

Santa Barbara, USA

ADVISOR: PROF. AMBUJ SINGH, COLLABORATOR: ARLEI SILVA

Sep. 2018 - Feb. 2021

- Presented a unified view of embedding, covering different random-walk processes, similarity metrics, and embedding algorithms.
- Showed both theoretical and empirical evidence of the superiority of the novel autocovariance embedding in link prediction.
- Illustrated ways to exploit the multiscale nature of random-walk similarity to further optimize embedding performance.

Teaching Experience

Lead Teaching Assistant, Department of Computer Science, UCSB

Sep. 2020 - Sep. 2021

Teaching Assistant, CS 130A, CS 8, CS 174A/174N, Department of Computer Science, UCSB

Sep. 2018 - Dec. 2019

Instructor, Machine Learning Workshops for LMU/UCSB Nanotech PhD Symposium

May. 2019 - Jun. 2019

Honors & Awards

Top 10 best papers of WSDM'23 and best paper in the MLoG workshop

Feb. 2023

WSDM'22 NSF Travel Award

Feb. 2022

UCSB Computer Science Lead Teaching Assistant Fellowship

Sep. 2020

UCSB Computer Science Outstanding Scholar/Academic Excellence Fellowships

Sep. 2018

Outstanding Bachelor Thesis Award of UESTC

Jun. 2018

Most Outstanding Students Award of UESTC/Honorary Graduate of UESTC/Sichuan Province

Dec. 2017

National/People's/Tang Lixin Scholarships

Dec. 2015 - Jun. 2018

First Place in Sichuan Province, the 7th Chinese Mathematics Competitions

Nov. 2015

Academic Services

Registration Chair KDD'23

Program Committee AAAI'23, KDD'22, SDM'22

Reviewer Neural Networks'23, NeurIPS'22, ICLR'22, TIST'22, TKDD'21-22, KDD'20-21, WebConf'20-21