Ji Yang (Carl)

(+1) (217) 417-5987   Assistant Professor, Emory University
j.carlyang AT emory.edu   201 Dowman Drive, Atlanta, GA 30322

EDUCATION BACKGROUNDS

University of Illinois, Urbana Champaign   *Aug. 2014-Dec. 2020*
Ph.D. in Computer Science, Research Assistant, Teaching Assistant
GPA: 3.95/4.0; Research interests: graph data mining, network data science, applied machine learning.

Chu Kochen Honors College, Zhejiang University   *Sep. 2010-Jun. 2014*
B.Eng. in Computer Science & Technology
GPA: Major: 3.97/4.0, Overall: 3.86/4.0; Ranking: Top 2% of 201 students

RESEARCH BACKGROUNDS

Data Mining Group, UIUC   *Aug. 2016 - Nov. 2020*
Committee: Jiawei Han (chair), Jian Peng, Chengxiang Zhai, Jure Leskovec   *Urbana, Illinois, US*

- Research interests: graph data mining, network data science, applied machine learning.
- Coordinated the *SocialCube* research project with DARPA under Agreement No. W911NF-17-C-0099.
- Collaborated on the Intelligent Social Media and Sensor Stream Summarization and Situation Analysis research project with US Army Research Lab under Cooperative Agreement No. W911NF-09-2-0053.
- Revised Prof Han's popular textbook *Data Mining: Concepts and Techniques* for the 4th Edition.
- Won course project champion (CS412 Data Mining 2014 Fall UIUC) for fraud detection on Kaggle.
- Received Yunni & Maxine Pao Memorial Fellowship for research accomplishments.
- Received the UIUC 2020 Doctoral Dissertation Completion Fellowship (awarded to 20 candidates from 20 departments out from 74 nominations from 48 departments across the whole university).

State Key Lab of CAD&CG, Zhejiang University   *Feb. 2012 - Jun. 2014*
Advisor: Prof. Xiaofei He   *Hangzhou, Zhejiang, China*

- Developed novel manifold learning algorithms for dimension reduction and image retrieval.
- Won Chinese National Scholarship for Outstanding Merits, Zhejiang University (Top 1%).
- Won Chinese National Fellowship for Excellent Intellects in Research, Zhejiang University (Top 1%).

SELECTED PUBLICATIONS

- **Carl Yang**, Yuxin Xiao, Yu Zhang, Yizhou Sun, Jiawei Han: *Heterogeneous Network Representation Learning: A Unified Framework with Survey and Benchmark.* **TKDE**, 2021.
- **Carl Yang**, Jieyu Zhang, Jiawei Han: *Co-Embedding Network Nodes and Hierarchical Labels with Taxonomy Based Generative Adversarial Nets*, **ICDM**, 2020.
- **Carl Yang**, Aditya Pal, Andrew Zhai, Nickil Pancha, Jure Leskovec, Jiawei Han: *MultiSage: Empowering GCN with Contextualized Multi-Embeddings on Web-Scale Multipartite Networks.* **KDD**, 2020.
- **Carl Yang**, Jieyu Zhang, Haonan Wang, Bangzheng Li, Jiawei Han: *Neural Concept Map Generation for Effective Document Classification with Interpretable Structured Summarization.* **SIGIR**, 2020.
- **Carl Yang**, Jieyu Zhang, Haonan Wang, Sha Li, Myunghwan Kim, Matthew Walker, Yiou Xiao, Jiawei Han: *Relation Learning on Social Networks with Multi-Modal Graph Edge Variational Autoencoders.* **WSDM**, 2020.
- **Carl Yang**, Peiye Zhuang, Wenhan Shi, Alan Luu, Pan Li: *Conditional Structure Generation through Graph Variational Generative Adversarial Nets.* **NeurIPS**, 2019.

- **Carl Yang**, Mengxiong Liu, Frank He, Jian Peng, Jiawei Han: *cube2net: Efficient Quality Network Construction with Data Cube Organization.* **ICDM** (*PhD forum*), 2019.

- **Carl Yang**, Jieyu Zhang, Jiawei Han: *Neural Embedding Propagation on Heterogeneous Networks.* **ICDM**, 2019.

- **Carl Yang**, Lingrui Gan, Zongyi Wang, Jiaming Shen, Jinfeng Xiao, Jiawei Han: *Query-Specific Knowledge Summarization with Entity Evolutionary Networks.* **CIKM**, 2019.

- **Carl Yang**, Dai Teng, Siyang Liu, Sayantani Basu, Jieyu Zhang, Jiaming Shen, Chao Zhang, Jingbo Shang, Lance Kaplan, Timothy Harratty, Jiawei Han: *CubeNet: Multi-Facet Hierarchical Heterogeneous Network Construction, Analysis, and Mining.* **KDD** (*demo paper*), 2019.

- **Carl Yang**, Do Huy Hoang, Tomas Mikolov, Jiawei Han: *Place Deduplication with Embeddings.* **WWW**, 2019.

- **Carl Yang**, Kevin Chang: *Relationship Profiling over Social Networks: Reverse Smoothness from Similarity to Closeness.* **SDM**, 2019.

- **Carl Yang**, Yichen Feng, Pan Li, Yu Shi, Jiawei Han: *Meta-Graph Based HIN Spectral Embedding: Methods, Analyses and Insights.* **ICDM**, 2018.

- **Carl Yang**, Mengxiong Liu, Frank He, Xikun Zhang, Jian Peng, Jiawei Han: *Similarity Modeling on Heterogeneous Networks via Automatic Path Discovery.* **ECML-PKDD**, 2018.

- **Carl Yang**, Mengxiong Liu, Vincent Zheng, Jiawei Han: *Node, Motif and Subgraph: Leveraging Network Functional Blocks Through Structural Convolution.* **ASONAM**, 2018.

- **Carl Yang**, Xiaolin Shi, Jie Luo, Jiawei Han: *I Know You'll Be Back: Interpretable New User Clustering and Churn Prediction on a Mobile Social Application.* **KDD**, 2018.

- **Carl Yang**, Chao Zhang, Xuewen Chen, Jieping Ye, Jiawei Han: *Did You Enjoy the Ride: Understanding Passenger Experience via Heterogeneous Network Embedding.* **ICDE**, 2018.

- **Carl Yang**, Liyuan Liu, Mengxiong Liu, Zongyi Wang, Chao Zhang, Jiawei Han: *Graph Clustering with Embedding Propagation.* **BigData**, 2020 (Arxiv, 2017).

- **Carl Yang**, Lanxiao Bai, Chao Zhang, Quan Yuan, Jiawei Han: *Bridging Collaborative Filtering and Semi-Supervised Learning: A Neural Approach for POI Recommendation.* **KDD**, 2017.

- **Carl Yang**, Zhong Lin, Li-Jia Li, Jie Luo: *Bi-directional Joint Inference for User Links and Attributes on Large Social Graphs.* **WWW**, 2017.

- **Ji Yang**, Bin Xu, Binbin Lin, Xiaofei He: *Multi-Query Parallel Field Ranking for Image Retrieval.* Neurocomputing, 2014.

- Wenjie Hu, Yang Yang, Ziqiang Cheng, **Carl Yang**, Xiang Ren: *Time-Series Event Prediction with Evolutionary State Graph.* **WSDM**, 2021.

- Yuanzhen Xie, Zijing Ou, Liang Chen, Yang Liu, Kun Xu, **Carl Yang**, Zibin Zheng: *Learning and Updating Node Embedding on Dynamic Heterogeneous Information Network.* **WSDM**, 2021.

- Yanchao Tan, **Carl Yang**, Xiangyu Wei, Yun Ma, Xiaolin Zheng: *Multi-Facet Recommender Networks with Spherical Optimization.* **ICDE**, 2021.

- Yingkui Wang, Di Jin, **Carl Yang**, Jianwu Dang: *Integrating Group Homophily and Individual Personality of Topics Can Better Model Network Communities.* **ICDM**, 2020.

- Chanyoung Park, **Carl Yang**, Qi Zhu, Donghyun Kim, Hwanjo Yu, Jiawei Han: *Unsupervised Differentiable Multi-aspect Node Representation Learning.* **KDD**, 2020.

- Yang Liu, Xianzhuo Xia, Liang Chen, Xiangnan He, **Carl Yang**, Zibin Zheng: *Certifiable Robustness to Discrete Adversarial Perturbations for Factorization Machines.* **SIGIR**, 2020.

- Yiqing Xie, Sha Li, **Carl Yang**, Raymond Chi-Wing Wong, Jiawei Han: *When Do GNNs Work: Understanding and Improving Neighborhood Aggregation.* **IJCAI**, 2020.

- Jiayi Liu, Quan Yuan, **Carl Yang**, He Huang, Chao Zhang, Philip Yu: *DAPred: Dynamic Attention Location Prediction with Long-Short Term Movement Regularity.* **IAAI**, 2020.

- Yuxin Xiao, Zecheng Zhang, **Carl Yang**, ChengXiang Zhai: *Non-local Attention Learning on Large Heterogeneous Information Networks.* **BigData**, 2019.

- Aravind Sankar, Adit Krishnan, Zongjian He, **Carl Yang**: *RASE: Relationship Aware Social Embedding.* **IJCNN**, 2019.
- Yu Shi, Fangqiu Han, Xinwei He, Xinran He, **Carl Yang**, Jie Luo, Jiawei Han: *Preservation and Collaboration in Multi-View Network Embedding.* **TKDE**, 2019.
- Yu Shi, Xinwei He, Naijing Zhang, **Carl Yang**, Jiawei Han: *User-Guided Clustering in Heterogeneous Information Networks via Motif-Based Comprehensive Transcription.* **ECML-PKDD**, 2019.
- Jiaming Shen, Jinfeng Xiao, Yu Zhang, **Carl Yang**, Jingbo Shang, Jinda Han, Saurabh Sinha, Peipei Ping, Richard Weinshilboum, Zhiyong Lu, Jiawei Han: *SetSearch+: Entity-Set-Aware Search and Mining for Scientific Literature.* **KDD** (demo paper), 2018.
- Chao Zhang, Mengxiong Liu, Zhengchao, Liu, **Carl Yang**, Luming Zhang, Jiawei Han: Spatiotemporal Activity Modeling Under Data Scarcity: A Graph-Regularized Cross-Modal Embedding Approach. **AAAI**, 2018.
- Hanqing Lu, **Carl Yang**, Kevin Chang: *CONE: Community Oriented Network Embedding.* **IJCNN**, 2018, accepted.
- Binbin Lin, **Ji Yang**, Xiaofei He, Jieping Ye: *Geodesic Distance Function Learning via Heat Flows on Vector Fields.* **ICML**, 2014.
- Haifeng Liu, Zheng Yang, **Ji Yang**, Zhaohui Wu, Xuelong Li: *Local Coordinate Concept Factorization for Image Representation.* **TNNLS**, 2014.

TEACHING EXPERIENCES

- **Instructor**, CS584 Topics: Graph Data Mining, Emory University, 2020 Fall.
- **Teaching Assistant**, CS412: Introduction to Data Mining (online), Coursera, 2020 Spring.
- **Lead Teaching Assistant**, CS512: Data Mining: Principles and Algorithms, UIUC, 2019 Spring.
- **Lead Teaching Assistant**, CS512: Data Mining: Principles and Algorithms, UIUC, 2018 Spring.
- **Lead Teaching Assistant**, CS412: Introduction to Data Mining, UIUC, 2017 Spring.
- **Teaching Assistant**, CS511: Advanced Data Management, UIUC, 2016 Spring.
- **Teaching Assistant**, CS412: Introduction to Data Mining, UIUC, 2015 Fall.

INDUSTRY EXPERIENCES

Research Lab, Pinterest Inc _May. 2019 - Aug. 2019_
Research Intern, Supervisors: Dr. Aditya Pal, Prof. Jure Leskovec _San Francisco, CA, US_

- Empowered GraphSage for web-scale Pinterest graph embedding through efficient graph down-sampling and context-aware conditional aggregation on heterogeneous pin-board networks.

Places Data & AI Research, Facebook Inc. _May. 2018 - Aug. 2018_
Research Intern, Supervisors: Dr. Do Huy Hoang, Dr. Tomas Mikolov _New York, NY, US_

- Developed a two-step data-driven pipeline of feature generation and metric learning for place embedding to leverage ad-hoc place attributes and noisy training data towards efficient place deduplication.

Economic Graph Research, LinkedIn Co. _Aug. 2018 - Jul. 2019_
Remote Research Contractor, Supervisors: Dr. Myungwan Kim, Dr. Shipeng Yu _Sunnyvale, CA, US_

- Developed a relation profiling algorithm based on multiple signals including user attributes, link structures and diffusive messages in the social network with novel multi-modal graph autoencoders.

Big Data Research, Didichuxing Inc. _Jun. 2017 - Aug. 2017_
Research Intern, Supervisors: Prof. Xuewen Chen, Prof. Jieping Ye _Beijing, China_

- Constructed the transportation HIN (heterogeneous information network) based on DiDi's unique travel data and developed a pattern-aware HIN embedding algorithm for passenger experience prediction.

Research Lab, Snap Inc.
Research Intern, Supervisors: Dr. Jie Luo, Dr. Xiaolin Shi, Dr. Li-Jia Li

May. 2016 - Aug. 2016

Los Angeles, CA, US

- Developed a joint learning framework of user links and attributes for friend recommendation and interest targeting. Implemented a Spark pipeline and scaled up to millions of nodes and billions of edges.

Demographics ads serving, Google LLC
Software Engineer Intern, Supervisor: Dr. Tianyi Wu

May. 2015 - Aug. 2015

Seattle, WA, US

- Implemented data extraction and inventory analysis pipeline using flume C++. Implemented an online simulation of ads serving and an offline optimal algorithm based on max flow to analyze the inventory.

SERVICES AND ACTIVITIES

Paper Reviews
- The International World Wide Web Conference (**WWW**), 2021.
- The AAAI Conference on Artificial Intelligence (**AAAI**), 2021.
- The ACM International Conference on Web Search and Data Mining (**WSDM**), 2021.
- The Conference on Neural Information Processing Systems (**NeurIPS**), 2020.
- The IEEE International Conference on Data Mining (**ICDM**), 2020.
- The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (**ECML-PKDD**), 2020.
- The International Conference on Web Information Systems Engineering (**WISE**), 2020.
- The ACM Conference on Knowledge Discovery and Data Mining (**KDD**), 2020.
- The International World Wide Web Conference (**WWW**), 2020.
- The AAAI Conference on Artificial Intelligence (**AAAI**), 2020.
- The Conference on Neural Information Processing Systems (**NeurIPS**), 2019.
- The IEEE International Conference on Data Mining (**ICDM**), 2019.
- The International Conference on Web Information Systems Engineering (**WISE**), 2019.
- The ACM Conference on Knowledge Discovery and Data Mining (**KDD**), 2019.
- The International Conference on Machine Learning (**ICML**), 2019.
- The International World Wide Web Conference (**WWW**), 2019.
- The ACM International Conference on Web Search and Data Mining (**WSDM**), 2019.
- The ACM International Conference on Information and Knowledge Management (**CIKM**), 2018.
- The ACM Conference on Knowledge Discovery and Data Mining (**KDD**), 2018.
- IEEE Access, The Multidisciplinary Open Access Journal (**Access**), since 2018.
- IEEE Transactions on Neural Networks and Learning Systems (**TNNLS**), since 2017.
- IEEE Transactions on Knowledge and Data Engineering (**TKDE**), since 2017.

Research Seminars
- University of Sydney (virtual), Australia, Mar 2020.
- University of North Texas and University of Florida, US, Mar 2020.
- Emory University, US, Feb 2020.
- Zhejiang University, China, Jan 2020.
- University of British Columbia and Simon Fraser University, Canada, Dec 2019.
- Northwestern University, US, Nov 2019.
- Great Lakes Workshop on Data Science in the University of Notre Dame, US, Sep 2019.
- Fudan University and Shanghai Jiao Tong University, China, Nov 2018.
- Snap Inc., US, Mar 2018.
- Tsinghua University and University of Science and Technology, China, Aug 2017.