

Zhilin Yang

BASIC INFORMATION

Ph.D. Student
Language Technology Institute
School of Computer Science
Carnegie Mellon University, Pittsburgh, PA

Email: kimiyoung@yeah.net
Home page: kimiyoung.github.io

RESEARCH INTERESTS

Natural Language Processing, Data Mining, Machine Learning

EDUCATION

Tsinghua University, Beijing, China

B.E., Computer Science, 2011 - 2015

Advisor: Professor Jie Tang

- Major GPA **96.8/100**, Overall GPA **94.2/100**, ranked **1/124**
- Ranked **No.1** in the following courses:
 - **Programming**
Fundamentals of Programming, Foundation of Object-Oriented Programming, Programming and Training, Service Oriented Software Design and Development
 - **Artificial Intelligence**
Introduction to Machine Learning, Artificial Neural Network, Pattern Recognition, Principles of Signal Processing
 - **Mathematics/Computer Systems**
Probability and Statistics, Computer Organization, Principles and Techniques of Compilers, Introduction to High Performance Computing

RESEARCH PROJECTS

Learning Embeddings for Heterogeneous Data

Bayesian Embeddings for Learning Social Knowledge Graphs

- Connected 39 million researchers to 35 knowledge concepts
- Deployed online service on AMiner.org; attracted more than 200,000 visits
- Decreased the online error rate by 67%
- Jointly combined the advantages of topic models and word embeddings

Collaborative Embeddings for E-Commerce Buying Prediction

- A random walk based method to represent heterogeneous data in a shared semantic space
- Achieved top 3 in both stages of IJCAI 2015 machine learning competition

Markov Random Fields for Social Network Alignment

Energy Based MRFs for Heterogeneous Social Network Alignment

- Identified whether two accounts in different social networks refer to the same person in the physical world
- Proposed an MRF based model to leverage network alignment consistency
- Outperformed alternative methods by up to 10-30%

Active Learning for Social Network Prediction

Active Learning on Markov Random Fields

- Based the algorithm on non-progressive diffusion models
- Showed that non-progressive models are more suitable than progressive ones

Active Learning for Streaming Graphs

- Selected instances by sequentially minimizing the structural variability of an MRF

HONORS AND AWARDS	1st Place, Best Undergraduate Thesis Award, Dept. Computer Science, Tsinghua Univ.	2015
	3rd Place, Stage 2 of IJCAI 2015 Machine Learning Competition (out of 50 selected teams)	2015
	2nd Place, Stage 1 of IJCAI 2015 Machine Learning Competition (out of 800+ participants)	2015
	Finalist, Tsinghua Grand Scholarship (finalist 0.45%, highest honor of Tsinghua students)	2014
	China Computer Federation Outstanding Undergraduate Award ($\sim 1.6\%$)	2014
	Science and Technology Innovation Award (1.1%, awarded by Tsinghua University)	2014
	Professional Excellence Award ($\sim 4\%$, awarded by Tsinghua University)	2014
	National Scholarship ($\sim 2\%$, awarded by the Ministry of Education, China)	2013
SKILLS	C/C++, Java, Python, Ruby, Matlab, R, JavaScript, Pascal, VHDL, \LaTeX	
ACADEMIC ACTIVITIES	Reviewer of ACM Transactions on Knowledge Discovery from Data (TKDD)	
	External Reviewer of IJCAI'15, SDM'15, ICWSM'15, KDD'15, TKDE, etc.	
	Talks at WSDM'14, CIKM'14	
TEACHING	TA of a graduate course, <i>80240603-0 Advanced Machine Learning</i> , during undergrad at Tsinghua.	
PAPERS	Zhilin Yang , Jie Tang.	
	Multi-Source Bayesian Embeddings for Learning Social Knowledge Graphs.	
	<i>Preprint.</i> An extended version in Chinese won the best undergraduate thesis award. [PDF] [Slides]	
	Zhanpeng Fang*, Zhilin Yang* , Yutao Zhang. (* indicates equal contribution)	
	Collaborative Embedding Features and Diversified Ensemble for E-Commerce Repeat Buyer Prediction.	
	<i>To Appear In Proceedings of International Joint Conference on Artificial Intelligence (IJCAI) Workshop, 2015.</i> Invited talks, competition winners. [PDF] [Slides]	
	Yutao Zhang, Jie Tang, Zhilin Yang , Jian Pei, Philip Yu.	
	COSNET: Connecting Heterogeneous Social Networks with Local and Global Consistency.	
	<i>In Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2015.</i> Oral presentation, acceptance rate 19%. [PDF]	
	Zhilin Yang , Jie Tang, Yutao Zhang.	
	Active Learning for Streaming Networked Data.	
	<i>In Proceedings of ACM International Conference on Information and Knowledge Management (CIKM), 2014.</i> Full paper, acceptance rate 21%. [PDF]	
	Zhilin Yang , Jie Tang, Bin Xu, Chunxiao Xing.	
	Active Learning for Networked Data Based on Non-Progressive Diffusion Model.	
	<i>In Proceedings of ACM International Conference on Web Search and Data Mining (WSDM), 2014.</i> Plenary full-length presentation (5%), acceptance rate 18%. [PDF]	
	Yang Yang, Jianfei Wang, Yutao Zhang, Wei Chen, Jing Zhang, Honglei Zhuang, Zhilin Yang , Bo Ma, Zhanpeng Fang, Sen Wu, Xiaoxiao Li, Debing Liu, Jie Tang.	
	SAE: Social Analytic Engine for Large Social Networks.	
	<i>In Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2013.</i> Demo paper. [PDF]	