# Yujing (Kelly) Wang

**ADDRESS**
China University of Geoscience
Beijing, China 100083

**CONTACT**
Mobile: (+86)13717701736
Email: kelly1988@sina.com

## EDUCATION

- **Peking University**, Beijing, China — September 2010 - July 2013
  Master of Computer Science,
  Specialty: Machine Learning, Data Mining, Document Understanding

- **Peking University**, Beijing, China — September 2006 - July 2010
  Bachelor of Computer Science,
  Overall GPA 3.59/4.0, Rank top 10%

## WORK EXPERIENCE

**Research Software Developer,** Oct. 2016 - Present
**Ads Intelligence, Hulu**

- Invented and developed inventory reservation and online allocation algorithms for guaranteed ads contracts.
- Designed and implemented the simulation framework for ad server.
- Research in real-time-bidding scenario in Hulu.

  **Impacts**

  - *Reduced 10+% under-delivery by applying the new reservation and online allocation algorithms.*
  - *The simulation framework can efficiently compare the performance of different settings and algorithms without affecting online business via A/B test.*

**Senior Algorithm Development Lead,** Mar. 2015 - Sep. 2016
**iPinyou**

- Leading a team of 6+ engineers that implements and optimizes Real-Time Bidding (RTB) algorithms in iPinyou, including real-time CTR prediction, CPC control, behavioral targeting, lookalike audience modeling, bid price optimization, creative selection, frequency capping, traffic forecasting, and campaign allocation.

  **Impacts**

  - *Improved the overall CTR by 20%+ with an enhanced CTR prediction model.*
  - *Increased the overall ROI of re-targeting campaigns by 30%+, through precisely estimating purchase intents via a behavioral targeting model.*
  - *Designed and developed an accurate feedback controller that dynamically keeps the variance of CPC under 10% for 95%+ ads campaigns.*

**Associate Researcher,** July. 2013 - Feb. 2015
**Machine Learning Group, Microsoft Research Asia**

- Crowd-based Tagging of Queries and Documents
  - Constructed a web-scale corpus of document-tag annotations from the wisdom of crowds.
  - Designed and implemented an algorithm of training phrase-tag associations based on an word embedding model for text tagging.

  **Impacts**

  - *Shipped to Microsoft Contextual Ads Platform for query-ads matching, achieving 5% CTR lift and 50% RPM increase.*
  - *Shipped to Bing News Portal for news tagging, achieving 98%+ precision and 80%+ recall.*

- Large-Scale Distributed Multi-Label Classification of Entities
  - Designed a high-precision classification algorithm that classifies entities with text descriptions into 200+ pre-defined types (e.g., Actor, Film, Book) by appropriately leveraging relationships among types.

  **Impacts**

  - *Shipped the entire pipeline to Microsoft Satori Knowledge Graph, achieving 99%+ classification accuracy for more than 200 types. The pipeline is re-run once a day to assign possible types to new entities.*

- Heterogeneous Graph-based Intent Learning
  - Constructed a heterogeneous graph of queries, web pages and Wikipedia concepts.
  - Invented a learning algorithm that improves search relevance by inferring search intents via soft clustering on the heterogeneous graph. The search intent is less ambiguous and human-understandable.
  - Implemented a joint embedding algorithm on the heterogeneous graph.

**Research Intern,** Sep. 2010 - July. 2013
**Machine Learning Group, Microsoft Research Asia**

- Browsing to Search through Context-Aware Query Reformulation
  - Worked on Browsing to Search, a feature allowing a software user to highlight his/her interested words and to see related search results.
  - Designed and built a Browsing to Search system that predicts context-aware search intents by the words a user selects and shows to him/her the corresponding queries.
- Dependency Graph-based Document Representation Model
  - Proposed a novel document representation model that captures the semantic relationships between words within a document by a dependency graph.
  - Designed a similarity function on the graph that achieves higher precision/recall than the state-of-the-art models.
- Task Assistant
  - Classified the task intent using a supervised learning model.
  - Recognized the information slots in the natural language sentences using Conditional Random Field.

**Software Development Engineer Intern** Jun. 2009 - Jan. 2010
**Bing News, Microsoft Search Technology Center**

- Realtime Discovery and Crawling of News Entries
- Template-based News Content Extraction
  - Proposed a new similarity metric and performed clustering of news content based on it.
  - Generated a template for each cluster and applied template-based information extraction.
  - The outstanding thesis of Peking University in 2010.

## SKILLS

- C#, Java, Python, C++, Scala, Hadoop, Spark
- Microsoft SQL Server, MySql, Non-SQL databases.
- 7+ years experience in machine learning, text mining and document understanding.
- 4+ years experience on computational advertising.
- 1+ years experience of leading a team.
- Extensive knowledge on automatic trading system and trading strategy.

## PUBLICATIONS

- Xiang Ren, **Yujing Wang**, Xiao Yu, Jun Yan, Zheng Chen, Jiawei Han: Heterogeneous graph-based intent learning with queries, web pages and Wikipedia concepts. WSDM 2014: 23-32
- Zhongqi Lu, Yin Zhu, Sinno Jialin Pan, Evan Wei Xiang, **Yujing Wang**, Qiang Yang: Source Free Transfer Learning for Text Classification. AAAI 2014: 122-128
- **Yujing Wang**, Yunhai Tong, Ming Zeng: Ranking Scientific Articles by Exploiting Citations, Authors, Journals, and Time Information. AAAI 2013: 933-939
- Bingyue Peng, **Yujing Wang**, Jian-Tao Sun: Mining mobile users' activities based on search query text and context, PAKDD 2012.
- **Yujing Wang**, Xiaochuan Ni, Jian-Tao Sun, Yunhai Tong, Zheng Chen: Representing document as dependency graph for document clustering. CIKM 2011: 2177-2180
- Ce Zhang, Bin Cui, Gao Cong, **Yujing Wang**: A Revisit of Query Expansion with Different Semantic Levels. DASFAA 2009: 662-676
- Ce Zhang, **Yujing Wang**, Bin Cui, Gao Cong: Semantic similarity based on compact concept ontology. WWW 2008: 1125-1126

## PATENTS

- Ashish Sharma, Jianwen Zhang, Siarhei Alonichau, Woonyeon Yoo and **Yujing Wang**, High Precision Limited Supervision Relationship Extractor US20160098645. 2016.

## HONORS AND AWARDS

- **Data Mining Contest of China Postgraduate Students Second place, 2011**
- Outstanding Graduates of Peking University 2010
- The (International) Mathematical Contest in Modeling Meritorious Winners, 2009
- Wusi Scholarship of Peking University 2009
- Computer World Scholarship of China Undergraduates 2008
- GuangHua Scholarship of Peking University 2007
- **Contemporary Undergraduate Mathematical Contest in Modeling First prize, 2007**
- **Intel International Science and Engineering Fair 4th place in mathematics, 2006**
- China Yough Scientist of Tomorrow Second price, 2005