Yue FENG

Cell phone: (+44) 077-8361-0301 (+86) 156-0011-5166 Email: yue.feng.20@ucl.ac.uk
Homepage

EDUCATION

University College London (UCL), United Kingdom Sep 2020 - now
Department of Computer Science.
PhD in Foundational Artificial Intelligence; Advisor: Prof. Emine Yilmaz.

University of Chinese Academy of Sciences (UCAS), China Sep 2016 - Jul 2019
Institute of Computing Technology.
Master in Computer Science; Major GPA: *3.8/4.0*; Advisor: Prof. Jun Xu and Prof. Jiafeng Guo.

Harbin Institute of Technology(HIT), China Sep 2012 - Jul 2016
School of Software Engineering.
Bachelor in Computer Science; Major GPA: *3.85/4.0*.

RESEARCH INTEREST

Natural language processing, information retrieval, data mining and machine learning, with an emphasis on text mining, deep learning, ranking and relevance, question answering, text matching, dialogue systems, and knowledge graph. I published papers in top conferences in IR (SIGIR, CIKM, WSDM) and NLP (ACL, AAAI).

WORK EXPERIENCE

ByteDance AI Lab, ByteDance Inc, China. Apr 2020 - Sep 2020
Research Intern: Dialogue Systems; Mentor: Dr. Hang Li.

- I worked on a project of dialog state tracking. We proposed a sequence-to-sequence framework for dialogue understanding to facilitate the simple integration of new services without requiring additional training data. A research paper is under review.

Baidu Research, Baidu Inc, China. Oct 2019 - Mar 2020
Research Intern: Question Answering; Mentor: Prof. Ping Li.

- I worked on a project of E-commerce question answering. We try to grasp logical information and construct reasoning among the reviews to generate an explainable natural answer for the complex question. A research paper is under review.

Baidu Research, Baidu Inc, China. Apr 2017 - Sep 2017
Research Intern: Semantic Mining, Text Matching, Knowledge Graph; Mentor: Prof. Ping Li.

- I worked on a project of semantic mining of reviews in E-commerce. We proposed a multi-task neural learning architecture for identifying helpful reviews to automatically reconstruct effective features and improve generalization ability. Further, to solve the lack of annotated reviews and distributional bias, a reinforced semi-supervised neural learning method is proposed to select high-related unlabeled reviews to help training. Papers are published in ***ASONAM'18*** and ***CIKM'20*** respectively.
- I worked on a project of paraphrase identification. We proposed a multi-granular matching layer to capture all different granular information of semantic matching. This paper is published in ***CIKM'18***.
- I worked on a project about open information extraction and designed an attention-based sequence-to-sequence model with copy mechanism and coverage mechanism to transform natural text into facts directly. Further, the Bayesian Network is proposed to learn interpretable relationships from open domain facts to enrich and refine concept graphs. Papers are published in ***WSDM'18*** and ***ACL'20*** respectively.

RESEARCH EXPERIENCE

University of Chinese Academy of Sciences, China. Sep 2016 - Jul 2019

Research Assistant: Information Retrieval, Question Answering; Advisor: Prof. Jun Xu and Prof. Jiafeng Guo.

- We proposed a Monte Carlo Tree Search enhanced Markov Decision Process to alleviate greedy selection problem in search result diversification. It adopts Policy-Value Networks to model environment and explore the possible rankings at the subsequent positions to achieve better decision-making. This paper is published in ***SIGIR'18***.
- We proposed an adapted Markov Decision Process to simulate human reading process and identify the main content of passages to improve answer span prediction in machine reading comprehension. This paper is published in ***AAAI'19***.

PUBLICATIONS

- ***Yue Feng***, Miao Fan, Mingming Sun and Ping Li. A Reinforced Semi-supervised Neural Network for Helpful Review Identification. Proceedings of the 29th ACM International Conference on Information and Knowledge Management (***CIKM'20***).
- Jingyuan Zhang, Mingming Sun, ***Yue Feng***, and Ping Li. Learning Interpretable Relationships between Entities, Relations and Concepts via Bayesian Structure Learning on Open Domain Facts. The 2020 Annual Conference of the Association for Computational Linguistics (***ACL'20***).
- Zhaohui Li, ***Yue Feng***, Jun Xu, Jiafeng Guo, Yanyan Lan, and Xueqi Cheng. Teaching Machines to Extract Main Content for Machine Reading Comprehension. Proceedings of the 33th AAAI Conference on Artificial Intelligence (***AAAI'19***).
- ***Yue Feng***, Jun Xu, Yanyan Lan, Jiafeng Guo, Wei Zeng, and Xueqi Cheng. From Greedy Selection to Exploratory Decision-Making: Diverse Ranking with Policy-Value Networks. Proceedings of the 41st annual international ACM SIGIR conference on Research and development in information retrieval (***SIGIR'18***).
- Miao Fan, Wutao Lin, ***Yue Feng***, Mingming Sun, and Ping Li. A Globalization-Semantic Matching Neural Network for Paraphrase Identification. Proceedings of the 27th ACM International Conference on Information and Knowledge Management (***CIKM'18***).
- Jingyuan Zhang, Xin Wang, ***Yue Feng***, Mingming Sun and Ping Li. FastInput: Improving Input Efficiency on Mobile Devices. Proceedings of the 27th ACM International Conference on Information and Knowledge Management (***CIKM'18***).
- Mingming Sun, Xu Li, Xin Wang, Miao Fan, ***Yue Feng***, and Ping Li. Logician: A Unified End-to-End Neural Approach for Open-Domain Information Extraction. The 11th ACM International Conference on Web Search and Data Mining (***WSDM'18***).
- Miao Fan, ***Yue Feng***, Mingming Sun, Ping Li, Haifeng Wang, Jianmin Wang. Multi-Task Neural Learning Architecture for End-to-End Identification of Helpful Reviews. The 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (***ASONAM'18***).

PROJECTS

- **Easy Machine Learning**: This platform presents a general-purpose dataflow-based system for easing the process of applying machine learning algorithms to real world tasks. https://github.com/ICT-BDA/EasyML.
- **Tech Leader Gemini @ Baidu Inc**: A project uses ensemble learning based on deep neural architecture for large-scale paraphrase identification in Baidu Knows. (+4.2% precision, +7.36% recall, and +5.62% AUROC)
- **Fast Input @ Baidu Inc**: A system which includes layout modeling, instant mistouch correction and user input text prediction to enhance input efficiency of Baidu Input System for mobile devices. (+7.8% precision)

HONORS AND AWARDS

- CIKM 2020 Student Travel Award
- ACL 2020 Student Travel Award
- AAAI 2019 Student Travel Award

- SIGIR 2018 Student Travel Award
- 2016 National Scholarship in China (top 0.2%)
- 2015 1st Prize in National Information Security Contest in China (top 0.1%)
- 2014 Excellent Student Cadre Model in Harbin Institute of Technology (top 0.1%)

SKILLS

- **Programming**: Python, Java, C++, SQL, MATLAB, Shell Scripting, Qt, LaTeX.
- **Theory**: Foundation of algorithms, math, NLP, IR, machine learning and data mining.
- **System**: Linux/Unix, Mac OS X, Windows, MySQL/SQL Server.
- **Tools**: Weka, LibSVM/SVMlight, Mallet, Numpy&SciPy, Pandas, Scikit-learn, Theano, TensorFlow, Keras, PyTorch, MXNet, NLP pipelines (Stanford NLP, OpenNLP), NLTK, LP_solver, CPLEX & many optimization toolkits.