Yichao Zhou

Last update on August 20, 2021

yichao.joey.zhou@gmail.com • https://yz-joey.github.io/

Research Interests

- Low resource information extraction and text mining
  - Information extraction with auxiliary supervision and word knowledge
  - Clinical case report annotation, indexing, and curation
  - Few-shot semi-structured information extraction from web documents
- Representation learning in NLP
  - Dynamic graph representation learning for pandemic forecast and risk factor discovery
  - Gender neutral embedding learning
  - Pronunciation-enhanced language modeling for humor detection
- Robustness in NLP
  - Adversarial attack detection and recovery
- Knowledge acquirement for Ads targeting and creative design
  - Conversion prediction and user funnel stage recognition
  - Cross-modal text generation and ranking approaches for Ad creative refinement

Education

University of California, Los Angeles — Los Angeles, CA, US
Ph.D degree in Computer Science (Data Mining and Natural Language Processing) *2017.9 – 2021.7*

University of California, Los Angeles — Los Angeles, CA, US
MS degree in Electrical Engineering (Signals & Systems) GPA:3.8/4.0 *2015.9 – 2017.6*

Southeast University — Nanjing, Jiangsu, P.R.China
BS degree in Electronics Engineering GPA:3.8/4.0 *2011.8 – 2015.6*

Experiences

Google AI — Mountain View, CA, US
Software Engineer *2021.8 – Now*

Google AI — Mountain View, CA, US
Research Intern *2020 Summer*

Verizonmedia, YAHOO! Research — Sunnyvale, CA, US
Intern Scientist *2018, 2019 Summer*

HELLA Co. — Lippstadt, NRW, Germany
Optical Design Intern for Undergraduate Thesis *2015 Winter and Spring*

Professional Activities

Reviewed 11 venues and 31 papers in total.

- Conference Chair
  1. NAACL 2021 Publication Chair *(Annual Conference of the North American Chapter of the Association for Computational Linguistics)*
  2. CIKM 2021 Session Chair *(ACM International Conference on Information and Knowledge Management)*
- Program Committee Member
  1. AAAI 2022 2021 *(AAAI Conference on Artificial Intelligence)*

2. CIKM 2021 *(ACM International Conference on Information and Knowledge Management)*
3. KDD 2021 2020 *(SIGKDD Conference on Knowledge Discovery and Data Mining)*
4. ACL 2021 *(Annual Conference of the Association for Computational Linguistics)*
5. AACL 2020 *(Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics)*

- Journal Reviewer

1. TKDE 2019 *(Transactions on Knowledge and Data Engineering)*

- Conference Sub-Reviewer:

1. ICDM 2019 2017 *(IEEE International Conference on Data Mining)*
2. KDD 2017 *(SIGKDD Conference on Knowledge Discovery and Data Mining)*
3. HealthNLP 2018 *(International Workshop on Health Natural Language Processing)*

Teaching Assistant Experiences

Taught 4 courses in 7 quarters and promoted as a teaching fellow at UCLA.

1. UCLA 2021 Spring & 2018 Winter, Prof.Wei Wang, **Data Mining**
   - This course introduces concepts, algorithms, and techniques of data mining on different types of datasets, which covers basic data mining algorithms, as well as advanced topics on text mining, recommender systems, and graph/network mining.
   - This course guides students to understand, develop, and use data mining systems to analyze large amounts of data.
2. UCLA 2020 Winter & Spring, Prof.David Smallberg, **Data structures and Algorithms** (C++)
   - This course teaches Object-oriented software development and abstract data type definition and use.
   - This course introduces Object-oriented view of data structures: stacks, queues, lists and algorithm analysis (e.g. Trees, graphs, and associated algorithms. Searching and sorting). Case studies and exercises from computer science applications.
3. UCLA 2019 & 2020 Fall, Prof.Todd Millstein, **Principles & Practices of Computing** (Python)
   - This course teaches students how to use computers as a tool for problem solving, creativity, and exploration via the design and implementation of computer programs.
   - This course covers key concepts including functional decomposition, usage of common data types, and usage of common control structures.
4. UCLA 2018 Fall, Prof.Kai-Wei Chang, **Introduction to Machine Learning**
   - This class introduces the fundamental concepts and algorithms in machine learning (supervised/unsupervised learning) and best practices in applying machine learning to practical problems.
   - This class covers a number of problem domains including computer vision, natural language processing, computational biology and robotics.

Honors, Awards, and Scholarships

- Student Volunteer/Travel Award

1. EMNLP 2019 *(Conference on Empirical Methods in Natural Language Processing)*
2. SIGKDD 2019 *(SIGKDD Conference on Knowledge Discovery and Data Mining)*

- UCLA Teaching Fellow, 2021
- UCLA Graduate Division Fellowships, 2018 2019 2020
- Excellent Undergraduate Thesis (2.7%), Southeast University, 2015

Publications and Patents

Publications

260+ Google Scholar citations in total and published at ACL, KDD, AAAI, WWW, EMNLP, CIKM, ICDE, etc.

1. **Yichao Zhou**, Learning Robust Representations for Low-resource Information Extraction, ***Doctoral Dissertation 2021***.

2. **Yichao Zhou**, Jyun-yu Jiang, Xiusi Chen, Wei Wang, #StayHome or #Marathon? Social Media Enhanced Pandemic Surveillance on Spatial-temporal Graphs, *The 30th ACM International Conference on Information & Knowledge Management*, ***CIKM2021***.

3. **Yichao Zhou**, Ying Sheng, Nguyen Vo, Nick Edmonds, Sandeep Tata, Simplified DOM Trees for Transferable Attribute Extraction from the Web, (Pre-print), ***arXiv2021***.

4. **Yichao Zhou**, Yu Yan, J. Harry Caufield, Kai-Wei Chang, Yizhou Sun, PeiPei Ping, Wei Wang, Clinical Temporal Relation Extraction with Probabilistic Soft Logic Regularization and Global Inference, *The 35th AAAI Conference on Artificial Intelligence*, ***AAAI2021***.

5. **Yichao Zhou**, Wei-Ting Chen, Bowen Zhang, David Lee, J. Harry Caufield, Kai-Wei Chang, Yizhou Sun, PeiPei Ping, Wei Wang, CREATe: Clinical Report Extraction and Annotation Technology, *IEEE 37th International Conference on Data Engineering*, (Demo), ***ICDE2021***.

6. Rujun Han, **Yichao Zhou**, Nanyun Peng, Domain Knowledge Empowered Structured Neural Net for End-to-End Event Temporal Relation Extraction, *The 2020 Conference on Empirical Methods in Natural Language Processing*, ***EMNLP2020***.

7. **Yichao Zhou**, Jyun-yu Jiang, Jieyu Zhao, Kai-Wei Chang, Wei Wang, "The Boating Store Had Its Best Sail Ever': Pronunciation-attentive Contextualized Pun Recognition", *2020 Annual Conference of the Association for Computational Linguistics*, ***ACL2020***.

8. **Yichao Zhou**, Shaunak Mishra, Manisha Verma, Narayan Bhamidipati, Wei Wang, "Recommending Themes for Ad Creative Design via Visual-Linguistic Representations", *The Web Conference*, (Short), ***WWW2020***.

9. Shaunak Mishra, Manisha Verma, **Yichao Zhou**, Kapil Thadani, Wei Wang, Learning to Create Better Ads: Generation and Ranking Approaches for Ad Creative Refinement, *The 29th ACM International Conference on Information & Knowledge Management*, ***CIKM2020***.

10. Cheng Zheng, Jyun-Yu Jiang, **Yichao Zhou**, Sean D. Young, Wei Wang, "Social Media User Geolocation via Hybrid Attention", *The 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, ***SIGIR2020***.

11. J. Harry Caufield, **Yichao Zhou**, Yunsheng Bai, David Liem, Garlid, A., Kai-Wei Chang, Yizhou Sun, Wei Wang, Peipei Ping, "Resources for comprehensive information extraction and data integration in biomedicine", *28th Conference on Intelligent Systems for Molecular Biology*, (Poster), ***ISMB2020***.

12. **Yichao Zhou**, Shaunak Mishra, Jelena Gligorijevic, T. Bhatia, Narayan Bhamidipati, "Understanding Consumer Journey using Attention based Recurrent Neural Networks", *SIGKDD Conference on Knowledge Discovery and Data Mining*, ***KDD2019***.

13. **Yichao Zhou***, Jyun-yu Jiang*, Kai-Wei Chang, Wei Wang, "Learning to Discriminate Perturbations for Blocking Adversarial Attacks in Text Classification", *2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing*, ***EMNLP-IJCNLP2019***.
(* equal contribution)

14. **Yichao Zhou**, J. Harry Caufield, Yizhou Sun, Peipei Ping, David Liem, D. Sigdel, Kai-Wei Chang, Wei Wang, "Learning Structured Knowledge from Clinical Case Reports", *27th Conference on Intelligent Systems for Molecular Biology and 18th European Conference on Computational Biology*, (Poster), ***ISMB/ECCB2019***.

15. J. Harry Caufield, **Yichao Zhou**, Yunsheng Bai, David Liem, Anders O. Garlid, Kai-Wei Chang, Yizhou Sun, Peipei Ping, Wei Wang, A Comprehensive Typing System for Information Extraction from Clinical Narratives, (Pre-print), ***medRxiv2019***.

16. **Yichao Zhou**, Chelsea Ju, J. Harry Caufield, Kevin Shih, Calvin Chen, Yizhou Sun, Kai-Wei Chang, Peipei Ping, Wei Wang. Clinical Named Entity Recognition using Contextualized Token Representations, (Pre-print), ***arXiv2019***.

17. Jieyu Zhao, **Yichao Zhou**, Zeyu Li, Wei Wang, Kai-Wei Chang, "Learning Gender-Neutral Word Embedding", *Conference on Empirical Methods in Natural Language Processing*, (Short), ***EMNLP2018***.

18. Tan Patrick, **Yichao Zhou**, Xinxin Huang, Wei Wang, "AZTEC: A Cloud-based Computational Platform to Integrate Biomedical Resources", *IEEE 33rd International Conference on Data Engineering*, (Demo), ***ICDE17***.

19. **Yichao Zhou**, Zai-Fa Zhou, Ninhuan Wang, "A novel residual stress test structure for MEMs thin films", *12th IEEE International Conference on Solid-State and Integrated Circuit Technology*, ***ICSICT2014***.

Patents

1. S. Mishra, R.H Tang, **Yichao Zhou**, V.R.Raithatha, L.N. Buamidipati, M. Abasi, S. Ravindran, "Stage-based content item selection and transmission", 2020