

# Kuan-Hao Huang

✉ khhuang@cs.ucla.edu 🏠 ENG VI 368, UCLA, CA 🏠 <https://khhuang.me>

## RESEARCH INTERESTS

### Robust and Reliable Natural Language Processing

- Text representations and text understanding
- Cross-lingual transfer learning and multilinguality
- Information extraction

## EDUCATION

**University of California Los Angeles, CA**

Sep. 2018 - Jun. 2023

*Ph.D. in Computer Science*

Advisor: Kai-Wei Chang

**National Taiwan University, Taipei, Taiwan**

Sep. 2014 - Jun. 2016

*M.S. in Computer Science and Information Engineering*

Advisor: Hsuan-Tien Lin

**National Taiwan University, Taipei, Taiwan**

Sep. 2010 - Jun. 2014

*B.S. in Computer Science and Information Engineering*

## RESEARCH EXPERIENCE

**University of California Los Angeles, CA**

Sep. 2018 - Jun. 2023

*Graduate Student Researcher, Advisor: Kai-Wei Chang*

**Meta AI, Seattle, WA**

Jun. 2022 - Sep. 2022

*Research Intern, Manager: Ruty Rinott*

**Amazon Alexa AI, Manhattan Beach, CA**

Mar. 2022 - Jun. 2022

*Applied Scientist Intern, Mentor: Anoop Kumar and Aram Galstyan*

**Amazon AWS AI, Santa Clara, CA**

Jun. 2021 - Sep. 2021

*Applied Scientist Intern, Mentor: Kasturi Bhattacharjee and Rashmi Gangadharaiah*

**Tencent AI Lab, Bellevue, WA**

Jun. 2019 - Sep. 2019

*Research Intern, Mentor: Chen Li*

## CONFERENCE PUBLICATIONS

- [C17] **Kuan-Hao Huang**, Varun Iyer, I-Hung Hsu, Anoop Kumar, Kai-Wei Chang, and Aram Galstyan. ParaAMR: A large-scale syntactically diverse paraphrase dataset by amr back-translation. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023.
- [C16] I-Hung Hsu\*, **Kuan-Hao Huang\***, Shuning Zhang, Wenxin Cheng, Premkumar Natarajan, Kai-Wei Chang, and Nanyun Peng. TAGPRIME: A unified framework for relational structure extraction. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023. (\*equal contribution).
- [C15] I-Hung Hsu\*, Zhiyu Xie\*, **Kuan-Hao Huang**, Premkumar Natarajan, and Nanyun Peng. AMPERE: Amr-aware prefix for generation-based event argument extraction model. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023. (\*equal contribution).

- [C14] Tanmay Parekh, I-Hung Hsu, **Kuan-Hao Huang**, Kai-Wei Chang, and Nanyun Peng. GENEVA: Benchmarking generalizability for event argument extraction with hundreds of event types and argument roles. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023.
- [C13] Yixin Wan, **Kuan-Hao Huang**, and Kai-Wei Chang. PIP: Parse-instructed prefix for syntactically controlled paraphrase generation. In *Findings of the Association for Computational Linguistics: ACL 2023 (ACL-Findings)*, 2023.
- [C12] **Kuan-Hao Huang\***, Varun Iyer\*, Anoop Kumar, Sriram Venkatapathy, Kai-Wei Chang, and Aram Galstyan. Unsupervised syntactically controlled paraphrase generation with abstract meaning representations. In *Findings of the Association for Computational Linguistics: EMNLP 2022 (EMNLP-Findings)*, 2022. (\*equal contribution).
- [C11] I-Hung Hsu\*, **Kuan-Hao Huang\***, Elizabeth Boschee, Scott Miller, Premkumar Natarajan, Kai-Wei Chang, and Nanyun Peng. DEGREE: A data-efficient generation-based event extraction model. In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2022. (\*equal contribution).
- [C10] **Kuan-Hao Huang\***, I-Hung Hsu\*, Premkumar Natarajan, Kai-Wei Chang, and Nanyun Peng. Multilingual generative language models for zero-shot cross-lingual event argument extraction. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL)*, 2022. (\*equal contribution).
- [C9] **Kuan-Hao Huang**, Wasi Uddin Ahmad, Nanyun Peng, and Kai-Wei Chang. Improving zero-shot cross-lingual transfer learning via robust training. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2021.
- [C8] James Y. Huang, **Kuan-Hao Huang**, and Kai-Wei Chang. Disentangling semantics and syntax in sentence embeddings with pre-trained language models. In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2021.
- [C7] **Kuan-Hao Huang** and Kai-Wei Chang. Generating syntactically controlled paraphrases without using annotated parallel pairs. In *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2021.
- [C6] **Kuan-Hao Huang**, Chen Li, and Kai-Wei Chang. Generating sports news from live commentary: A chinese dataset for sports game summarization. In *Proceedings of the Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AACL)*, 2020.
- [C5] Sean T. Yang, **Kuan-Hao Huang**, and Bill Howe. JECL: Joint embedding and cluster learning for image-text pairs. In *Proceedings of International Conference on Pattern Recognition (ICPR)*, 2020.
- [C4] Pei Zhou, Weijia Shi, Jiayu Zhao, **Kuan-Hao Huang**, Muhao Chen, Ryan Cotterell, and Kai-Wei Chang. Examining gender bias in languages with grammatical gender. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2019.
- [C3] Yao-Yuan Yang, **Kuan-Hao Huang**, Chih-Wei Chang, and Hsuan-Tien Lin. Cost-sensitive reference pair encoding for multi-label learning. In *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, 2018.
- [C2] **Kuan-Hao Huang** and Hsuan-Tien Lin. A novel uncertainty sampling algorithm for cost-sensitive multiclass active learning. In *Proceedings of the IEEE International Conference on Data Mining (ICDM)*, 2016.
- [C1] **Kuan-Hao Huang** and Hsuan-Tien Lin. Linear upper confidence bound algorithm for contextual bandit problem with piled rewards. In *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, 2016.

## JOURNAL PUBLICATIONS

- [J4] Hong-Min Chu, **Kuan-Hao Huang**, and Hsuan-Tien Lin. Dynamic principal projection for cost-sensitive online multi-label classification. *Machine Learning (ECML PKDD Journal Track)*, 2019.
- [J3] **Kuan-Hao Huang** and Hsuan-Tien Lin. Cost-sensitive label embedding for multi-label classification. *Machine Learning (ECML PKDD Journal Track)*, 2017.
- [J2] Chun-Liang Li, Yu-Chuan Su, Ting-Wei Lin, Cheng-Hao Tsai, Wei-Cheng Chang, **Kuan-Hao Huang**, Tzu-Ming Kuo, Shan-Wei Lin, Young-San Lin, Yu-Chen Lu, Chun-Pai Yang, Cheng-Xia Chang, Wei-Sheng Chin, Yu-Chin Juan, Hsiao-Yu Tung, Jui-Pin Wang, Cheng-Kuang Wei, Felix Wu, Tu-Chun Yin, Tong Yu, Yong Zhuang, Shou-De Lin, Hsuan-Tien Lin, and Chih-Jen Lin. Combination of feature engineering and ranking models for paper-author identification in KDD Cup 2013. *Journal of Machine Learning Research*, 2015. (Extended first-place winner report of KDD Cup 2013 track 1).
- [J1] Wei-Sheng Chin, Yong Zhuang, Yu-Chin Juan, Felix Wu, Hsiao-Yu Tung, Tong Yu, Jui-Pin Wang, Cheng-Xia Chang, Chun-Pai Yang, Wei-Cheng Chang, **Kuan-Hao Huang**, Tzu-Ming Kuo, Shan-Wei Lin, Young-San Lin, Yu-Chen Lu, Yu-Chuan Su, Cheng-Kuang Wei, Tu-Chun Yin, Chun-Liang Li, Ting-Wei Lin, Cheng-Hao Tsai, Shou-De Lin, Hsuan-Tien Lin, and Chih-Jen Lin. Effective string processing and matching for author disambiguation. *Journal of Machine Learning Research*, 2014. (Extended first-place winner report of KDD Cup 2013 track 2).

## OTHER PUBLICATIONS

- [O2] Fei Wang, **Kuan-Hao Huang**, Anoop Kumar, Aram Galstyan, Greg Ver Steeg, and Kai-Wei Chang. Zero-shot cross-lingual sequence tagging as seq2seq generation for joint intent classification and slot filling. In *Workshop on Massively Multilingual Natural Language Understanding, EMNLP (MMNLU@EMNLP)*, 2022.
- [O1] Xueying Zhan, Qingzhong Wang, **Kuan-Hao Huang**, Haoyi Xiong, Dejing Dou, and Antoni B. Chan. A comparative survey of deep active learning. In *Workshop on Human in the Loop Learning, NeurIPS (Hill@NeurIPS)*, 2022.

## PREPRINTS

- [P2] **Kuan-Hao Huang**, Liang Tan, Rui Hou, Sinong Wang, Amjad Almahairi, and Ruty Rinott. Learning easily updated general purpose text representations with adaptable task-specific prefixes. *arXiv preprint arXiv:2305.13499*, 2023.
- [P1] Oscar Chew, **Kuan-Hao Huang**, Kai-Wei Chang, and Hsuan-Tien Lin. Understanding and mitigating spurious correlations in text classification. *arXiv preprint arXiv:2305.13654*, 2023.

## SELECTED AWARDS

|  |      |
|--|------|
| <b>Thesis Honorable Mention Award</b> , Taiwanese Association for AI | 2016 |
| <b>Fourth Place</b> , KDD Cup 2015                                   | 2015 |
| <b>Second Place</b> , ICASSP Signal Processing Cup                   | 2014 |
| <b>First Place</b> , Track 1 of KDD Cup 2013                         | 2013 |
| <b>First Place</b> , Track 2 of KDD Cup 2013                         | 2013 |

## PROFESSIONAL SERVICES

### Program Committee/Reviewer

- Natural Language Processing: ACL Rolling Review (2021-2023), ACL (2021-2023), EMNLP (2021-2022), NAACL (2022), EACL (2023)

- Machine Learning: ICML (2020-2023), NeurIPS (2021-2023), ICLR (2021-2023)
- Artificial Intelligence: AAI (2022-2023)

### **Student Volunteer**

- ACL 2020
- EMNLP 2021

### **Handbook Assistant**

- EMNLP 2018

## **TEACHING EXPERIENCE**

### **Guest Lecturer**, University of California Los Angeles, CA

- CS 146: Introduction to Machine Learning Fall 2022

### **Teaching Assistant**, University of California Los Angeles, CA

- CS 146: Introduction to Machine Learning Fall 2020
- CS 144: Web Applications Spring 2020
- CS 269: Fairness, Accountability, and Transparency in NLP Winter 2020

### **Teaching Assistant**, National Taiwan University, Taipei, Taiwan

- CSIE 5043: Machine Learning Fall 2013, Fall 2014, Fall 2015
- CSIE 1212: Data Structure and Algorithm Spring 2013, Spring 2015

## **MENTORING**

|   |                           |
|---|---------------------------|
| Di Wu, PhD student at UCLA  | Fall 2022 - Present       |
| Yixin Wan, MS student at UCLA                                     | Fall 2022 - Present       |
| George Pu, MS student at UCLA                                     | Spring 2022 - Present     |
| Oscar Chew, MS student at NTU                                     | Winter 2022 - Present     |
| Tanmay Parekh, PhD student at UCLA                                | Fall 2021 - Present       |
| Ashima Suvarna, MS student at UCLA                                | Fall 2021 - Spring 2023   |
| Fei Wang, MS student at USC (now a PhD student at USC)            | Fall 2021 - Summer 2022   |
| James Yipeng Huang, MS student at UCLA (now a PhD student at USC) | Summer 2020 - Spring 2021 |

## **REFERENCES**

### **Kai-Wei Chang**

Associate Professor, Computer Science Department, University of California Los Angeles  
[kwchang@cs.ucla.edu](mailto:kwchang@cs.ucla.edu)

### **Nanyun Peng**

Assistant Professor, Computer Science Department, University of California Los Angeles  
[violetpeng@cs.ucla.edu](mailto:violetpeng@cs.ucla.edu)

### **Aram Galstyan**

Amazon Scholar, Amazon / Research Professor, Information Sciences Institute, University Southern California  
[argalsty@amazon.com](mailto:argalsty@amazon.com)

### **Hsuan-Tien Lin**

Professor, Department of Computer Science and Information Engineering, National Taiwan University  
[htlin@csie.ntu.edu.tw](mailto:htlin@csie.ntu.edu.tw)