

# Kuan-Hao Huang

---

Assistant Professor  
Department of Computer Science and Engineering, Texas A&M University  
Room 219, Peterson Building, 435 Nagle St, College Station, TX 77843

Tel: (979) 458-2205  
Email: [khhuang@tamu.edu](mailto:khhuang@tamu.edu)  
Website: <https://khhuang.me>

## RESEARCH INTERESTS

My research focuses on natural language processing and machine learning, with the goal of developing trustworthy and generalizable AI systems. I work on fundamental challenges in AI robustness, reliability, safety, fairness, and adaptability, aiming to ensure that large language models and vision-language models are reliable and socially responsible.

## ACADEMIC APPOINTMENTS

**Texas A&M University**, College Station, TX, USA Aug. 2024 – Present  
*Assistant Professor, Department of Computer Science and Engineering*

**University of Illinois Urbana-Champaign**, IL, USA Aug. 2023 – June. 2024  
*Postdoctoral Research Associate, Department of Computer Science*

## EDUCATION

**University of California Los Angeles**, CA, USA 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*

## SELECTED AWARDS

**Google Research Scholar Award** 2025

**Area Chair Paper Award** (*Outstanding Paper in the Semantics Track*), ACL 2023 2023

**Thesis Honorable Mention Award**, Taiwanese Association for Artificial Intelligence 2016

**Fourth Place**, KDD Cup 2015 2015

**Second Place**, ICASSP Signal Processing Cup 2014 2014

**First Place**, Track 1 of KDD Cup 2013 2013

**First Place**, Track 2 of KDD Cup 2013 2013

## CONFERENCE PUBLICATIONS

- C29.** Muhan Gao, Zih-Ching Chen, and Kuan-Hao Huang. The First Drop of Ink: Nonlinear Impact of Misleading Information in Long-Context Reasoning. In *Proceedings of the Forty-Third International Conference on Machine Learning (ICML)*, 2026.
- C28.** Taeyoon Kwon\*, Dongwook Choi\*, Sunghwan Kim, Hyojun Kim, Seungjun Moon, Beong woo Kwak, Kuan-Hao Huang, and Jinyoung Yeo. Embodied Agents Meet Personalization: Investigating Challenges and Solutions Through the Lens of Memory Utilization. In *Proceedings of the Fourteenth International Conference on Learning Representations (ICLR)*, 2026. (\*Equal contribution).
- C27.** Zhaotian Weng, Haoxuan Li, Xin Eric Wang, Kuan-Hao Huang, and Jieyu Zhao. What’s Missing in Vision-Language Models? Probing Their Struggles with Causal Order Reasoning. In *Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2026.

- C26.** Ren-Wei Liang\*, Chin-Ting Hsu\*, Chan-Hung Yu, Saransh Agrawal, Shih-Cheng Huang, Chieh-Yen Lin, Shang-Tse Chen, Kuan-Hao Huang, and Shao-Hua Sun. Adaptive Helpfulness-Harmlessness Alignment with Preference Vectors. In *Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2026. (\*Equal contribution).
- C25.** Yu Zhou\*, Bingxuan Li\*, Mohan Tang\*, Xiaomeng Jin, Te-Lin Wu, Kuan-Hao Huang, Heng Ji, Kai-Wei Chang, and Nanyun Peng. Contrastive Visual Data Augmentation. In *Proceedings of the Forty-Second International Conference on Machine Learning (ICML)*, 2025. (\*Equal contribution).
- C24.** Ziqi Wang, Hanlin Zhang, Xiner Li, Kuan-Hao Huang, Chi Han, Shuiwang Ji, Sham M. Kakade, Hao Peng, and Heng Ji. Eliminating Position Bias of Language Models: A Mechanistic Approach. In *Proceedings of the Thirteenth International Conference on Learning Representations (ICLR)*, 2025.
- C23.** Kuan-Hao Huang, I-Hung Hsu, Tanmay Parekh, Zhiyu Xie, Zixuan Zhang, Premkumar Natarajan, Kai-Wei Chang, Nanyun Peng, and Heng Ji. TextEE: Benchmark, Reevaluation, Reflections, and Future Challenges in Event Extraction. In *Findings of the Association for Computational Linguistics: ACL (ACL-Findings)*, 2024.
- C22.** Tanmay Parekh, I-Hung Hsu, Kuan-Hao Huang, Kai-Wei Chang, and Nanyun Peng. Contextual Label Projection for Cross-Lingual Structure Prediction. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2024.
- C21.** Tanmay Parekh, Anh Mac, Jiarui Yu, Yuxuan Dong, Syed Shahriar, Bonnie Liu, Eric Yang, Kuan-Hao Huang, Wei Wang, Nanyun Peng, and Kai-Wei Chang. Event Detection from Social Media for Epidemic Prediction. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2024.
- C20.** Oscar Chew, Hsuan-Tien Lin, Kai-Wei Chang, and Kuan-Hao Huang. Understanding and Mitigating Spurious Correlations in Text Classification. In *Findings of the Association for Computational Linguistics: EACL (EACL-Findings)*, 2024.
- C19.** 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. In *Findings of the Association for Computational Linguistics: EMNLP (EMNLP-Findings)*, 2023.
- C18.** Fei Wang, Kuan-Hao Huang, Kai-Wei Chang, and Muhao Chen. Self-Augmentation Improves Zero-Shot Cross-Lingual Transfer. In *Proceedings of the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (AACL)*, 2023.
- 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 61st Annual Meeting of the Association for Computational Linguistics (ACL)*, 2023. (Area Chair Award).
- 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 61st 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 61st 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 61st 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 (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 (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 2022 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 2022. (\*Equal contribution).

- C10. Kuan-Hao Huang<sup>\*</sup>, I-Hung Hsu<sup>\*</sup>, Premkumar Natarajan, Kai-Wei Chang, and Nanyun Peng. Multilingual Generative Language Models for Zero-Shot Cross-Lingual Event Argument Extraction. In *Proceedings of the 60th 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 2021 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 2021 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 16th 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 1st 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 the 25th International Conference on Pattern Recognition (ICPR)*, 2020.
- C4. Pei Zhou, Weijia Shi, Jieyu Zhao, Kuan-Hao Huang, Muhao Chen, Ryan Cotterell, and Kai-Wei Chang. Examining Gender Bias in Languages with Grammatical Gender. In *Proceedings of the 2019 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 *Advances in Knowledge Discovery and Data Mining - 22nd Pacific-Asia Conference (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 16th 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 *Advances in Knowledge Discovery and Data Mining - 20th Pacific-Asia Conference (PAKDD)*, 2016.

## JOURNAL PUBLICATIONS

- J5. Zhenhailong Wang, Joy Hsu, Xingyao Wang, Kuan-Hao Huang, Manling Li, Jiajun Wu, and Heng Ji. Visually Descriptive Language Model for Vector Graphics Reasoning. *Transactions on Machine Learning Research*, 2025.
- 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

- O10. Oanh N. Tran, Thanh Quoc Hung Le, Oscar Chew, Kuan-Hao Huang<sup>#</sup>, and Khoa D. Doan<sup>#</sup>. Look Before You Zoom: Adaptive Routing for the Resolution-Context Trade-off in Visual RAG. In *ICML Workshop on Efficient Multimodal Question Answering (EMM-QA@ICML)*, 2026. (<sup>#</sup>Jointly supervising).

- O9.** Oscar Chew\*, Serhii Honcharenko\*, Qian-Hui Chen, Patricia Lu, Dishant Zaveri, Khoa D. Doan, and Kuan-Hao Huang. Pop-Up Distractions Reveal Bag-of-Events Behavior in Video Large Language Models. In *ICML Workshop on Efficient Multimodal Question Answering (EMM-QA@ICML)*, 2026. (\*Equal contribution).
- O8.** Jaehoon Lee, Taewoo Ko, Kuan-Hao Huang, Jeewoong Park, Jung-ho Jeon, Brian A. Anderson, and Namgyun Kim. Accident Causality Extraction from Construction Accident Reports Using a Large Language Model. In *Proceedings of ASCE International Conference on Computing in Civil Engineering (i3CE)*, 2026.
- O7.** Jaehoon Lee, Taewoo Ko, Kuan-Hao Huang, JeeWoong Park, and Namgyun Kim. The Impact of Low-Quality Accident Reports on the Performance of Construction Safety AI Models. In *Proceedings of ASCE International Conference on Computing in Civil Engineering (i3CE)*, 2025.
- O6.** Saransh Agrawal and Kuan-Hao Huang. SHA256 at SemEval-2025 Task 4: Selective Amnesia – Constrained Unlearning for Large Language Models via Knowledge Isolation. In *Proceedings of the 19th International Workshop on Semantic Evaluation (SemEval)*, 2025.
- O5.** Thao Nguyen, Kuan-Hao Huang, Ge Liu, Martin D. Burke, Ying Diao, and Heng Ji. FARM: Functional Module-Aware Representations for Small Molecules Bridging SMILES with Natural Language and Molecular Graphs. In *NAACL Workshop on AI and Scientific Discovery: Directions and Opportunities (AISD@NAACL)*, 2025.
- O4.** Xiaomeng Jin, Jeonghwan Kim, Yu Zhou, Kuan-Hao Huang, Te-Lin Wu, Nanyun Peng, and Heng Ji. ARMADA: Attribute-Based Multimodal Data Augmentation. In *EMNLP Workshop on Advancing Natural Language Process for Wikipedia (WikiNLP@EMNLP)*, 2024.
- O3.** Ansel Blume, Khanh Duy Nguyen, Zhenhailong Wang, Yangyi Chen, Michal Shlapentokh-Rothman, Xiaomeng Jin, Jeonghwan Kim, Zhen Zhu, Jiateng Liu, Kuan-Hao Huang, Mankeerat Sidhu, Xuanming Zhang, Vivian Liu, Raunak Sinha, Te-Lin Wu, Abhay Zala, Elias Stengel-Eskin, Da Yin, Yao Xiao, Utkarsh Mall, Zhou Yu, Kai-Wei Chang, Camille Cobb, Karrie Karahalios, Lydia Chilton, Mohit Bansal, Nanyun Peng, Carl Vondrick, Derek Hoiem, and Heng Ji. MIRACLE: An Online, Explainable Multimodal Interactive Concept Learning System. In *Proceedings of the 32nd ACM International Conference on Multimedia: Technical Demos and Videos Program (ACMMM-Demo)*, 2024.
- 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 *EMNLP Workshop on Massively Multilingual Natural Language Understanding (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 *NeurIPS Workshop on Human in the Loop Learning (HiLL@NeurIPS)*, 2022.

## PREPRINTS

- P8.** Jiaqian Li, Yanshu Li, Boxuan Zhang, Ruixiang Tang, and Kuan-Hao Huang. TRACES: Proactive Safety Auditing for Multi-Turn LLM Agents via Trajectory-State Modeling. *arXiv preprint arXiv:2605.27690*, 2026.
- P7.** Chenghao Qiu, Chunli Peng, Yufeng Yang, Kuan-Hao Huang, and Yi Zhou. When Correct Demonstrations Hurt: Rethinking the Role of Exemplars in In-Context Learning. *arXiv preprint arXiv:2605.26350*, 2026.
- P6.** Rohan Surana, Gagan Mundada, Xunyi Jiang, Chuhan Wang, Zhenwei Tang, Difan Jiao, Zihan Huang, Yuxin Xiong, Junda Wu, Sheldon Yu, Xintong Li, Raghav Jain, Nikki Lijing Kuang, Sizhe Zhou, Bowen Jin, Zhendong Chu, Tong Yu, Ryan Rossi, Kuan-Hao Huang, Jingbo Shang, Jiawei Han, and Julian McAuley. Generate, Filter, Control, Replay: A Comprehensive Survey of Rollout Strategies for LLM Reinforcement Learning. *arXiv preprint arXiv:2605.02913*, 2026.
- P5.** Oscar Chew, Hsiao-Ying Huang, Kunal Jain, Tai-I Chen, Khoa D. Doan, and Kuan-Hao Huang. Is CLIP Cross-Eyed? Revealing and Mitigating Center Bias in the CLIP Family. *arXiv preprint arXiv:2604.05971*, 2026.
- P4.** Jiaqian Li, Yanshu Li, and Kuan-Hao Huang. Steering Vector Fields for Context-Aware Inference-Time Control in Large Language Models. *arXiv preprint arXiv:2602.01654*, 2026.
- P3.** Neeraja Kirtane and Kuan-Hao Huang. Language Steering for Multilingual In-Context Learning. *arXiv preprint arXiv:2602.02326*, 2026.
- P2.** Oscar Chew\*, Po-Yi Lu\*, Jayden Lin, Kuan-Hao Huang, and Hsuan-Tien Lin. PEPPER: Perception-Guided Perturbation for Robust Backdoor Defense in Text-to-Image Diffusion Models. *arXiv preprint arXiv:2511.16830*, 2025. (\*Equal contribution).
- P1.** Muhan Gao, Jash Shah, Weiqi Wang, Kuan-Hao Huang, and Daniel Khashabi. Science Hierarchography: Hierarchical Organization of Science Literature. *arXiv preprint arXiv:2504.13834*, 2025.

## ADVISING

### Ph.D. Students

Taeyoon Kwon, Computer Science	Incoming Fall 2026
Neeraja Kirtane, Computer Science	Incoming Fall 2026
Muhan Gao, Computer Science	2025 – Present
Oscar Chew, Computer Science	2025 – Present
Chan-Wei Hu, Computer Science ( <i>Rotating Student</i> )	2024 – 2024

### M.S. Students

Hung Nguyen, Computer Science	2026 – Present
Junggeun Do, Computer Science	2025 – Present
Bhaskar Ruthvik Bikkina, Computer Science	2025 – Present
Sicong Liang, Computer Science ( <i>Next: Ph.D. Student at UCF</i> )	2025 – 2026
Rusali Saha, Computer Science	2025 – 2026
Dishant Zaveri, Computer Science ( <i>Next: Technical Associate at Jefferies Group</i> )	2024 – 2026
Dheeraj Mudireddy, Data Science ( <i>Next: Data Scientist at Amazon</i> )	2025 – 2025
Rahul Baid, Computer Science ( <i>Next: Senior Software Engineer at Mubama Group</i> )	2024 – 2025
Saransh Agrawal, Data Science ( <i>Next: AI Engineer at Saptarishis Software</i> )	2024 – 2025

### Undergraduate Students

Serhii Honcharenko, Computer Science	2026 – Present
Kunal Jain, Computer Science ( <i>Next: Software Engineer at Google</i> )	2025 – 2026

### Visiting Students

Cheng-Yu Wang, M.S. Student at USC	2026 – Present
Xingrun Chen, Undergrad at UCLA	2026 – Present
Qian-Hui Chen, M.S. Student at NTU	2026 – Present
Jiongran Wang, Ph.D. Student at TAMU	2026 – Present
Hsin-Ling Hsu, Undergrad at NCCU	2025 – Present
Po-Yi Lu, Ph.D. Student at NTU	2025 – Present
Jiaqian Li, M.S. Student at Brown	2025 – Present
Anyi Wang, M.S. Student at LMU Munich ( <i>Next: Ph.D. Student at TAMU</i> )	2025 – 2026
Neeraja Kirtane, M.S. Student at UIUC ( <i>Next: Ph.D. Student at TAMU</i> )	2025 – 2026
Zihan Xue, Undergrad at UCLA ( <i>Next: Ph.D. Student at Northeastern</i> )	2025 – 2026
Zhaotian Weng, M.S. Student at USC ( <i>Next: Ph.D. Student at UC Santa Barbara</i> )	2024 – 2025

### Ph.D. Students (On Committee)

Mu-Ruei Tseng, Computer Science ( <i>Advisor: Ricardo Gutierrez-Osuna</i> )	Present
Lianhao Zhou, Computer Science ( <i>Advisor: Shuiwang Ji</i> )	Present
Yuyang Bai, Computer Science ( <i>Advisor: Yu Zhang</i> )	Present
Siyuan Yang, Computer Science ( <i>Advisor: Zhengzhong Tu</i> )	Present
Jaehoon Lee, Construction Science ( <i>Advisor: Namgyun Kim</i> )	Present
Hangxiao Zhu, Computer Science ( <i>Advisor: Yu Zhang</i> )	Present
Xiangbo Gao, Computer Science ( <i>Advisor: Zhengzhong Tu</i> )	Present
Xingyu Su, Computer Science ( <i>Advisor: Shuiwang Ji</i> )	Present
Chan-Wei Hu, Computer Science ( <i>Advisor: Zhengzhong Tu</i> )	Present

### M.S. Students (On Committee)

Arunima Chowdhury, Computer Science ( <i>Advisor: Ruihong Huang</i> )	Present
Dhawal Jiten Chaudhari, Computer Science ( <i>Advisor: Xin Li</i> )	Present
Pranati Tyagi, Computer Science ( <i>Advisor: James Caverlee</i> )	Present
Mihir Santosh Sanjay, Computer Science ( <i>Advisor: Shreyas Kumar</i> )	Present
Duc Nguyen, Computer Science ( <i>Advisor: Yu Zhang</i> )	Present
Yi Wen, Computer Science ( <i>Advisor: Meng Xia</i> )	2026
Mingyang Wu, Computer Science ( <i>Advisor: Zhengzhong Tu</i> )	2026
Michael Norman, Computer Science ( <i>Advisor: James Caverlee</i> )	2025
Divij Bajaj, Computer Science ( <i>Advisor: Ruihong Huang</i> )	2025

## TEACHING

**Instructor**, Texas A&M University, College Station, TX

CSCE 689: Special Topics in Advanced Natural Language Processing	Fall 2026
CSCE 638: Natural Language Processing: Foundation and Techniques	Spring 2026
CSCE 689: Special Topics in Trustworthy Natural Language Processing	Fall 2025
CSCE 638: Natural Language Processing: Foundation and Techniques	Spring 2025
CSCE 689: Special Topics in Trustworthy Natural Language Processing	Fall 2024

## SELECTED TUTORIALS AND TALKS

### **Challenges of Multimodal AI Foundation Models in Scientific Discovery**

Invited Talk, Course DATA 440/691, The College of William & Mary Apr. 2026

### **Controlling Large Language Models via Inference-Time Steering**

Invited Talk, Rice Workshop on Large Language Models, Rice University Mar. 2026

### **Tutorial on Large Language Models**

Tutorial, AI Bridge Summit, Texas A&M University Jan. 2026

### **Towards Robust and Trustworthy Large Language Models: Issues and Mitigation Strategies**

Tutorial, The 17th Asian Conference on Machine Learning (ACML) Dec. 2025

### **Toward Robust and Reliable Large Language Models**

Invited Talk, Graduate Seminar, Texas A&M University	Feb. 2026
Invited Talk, Course CS 6501, University of Virginia	Oct. 2025
Invited Talk, Data Science Colloquium, The College of William & Mary	Oct. 2025
Invited Talk, AI Seminar Series, Rice University	Oct. 2025
Invited Talk, Institute of Information Science, Academia Sinica	Jun. 2025
Invited Talk, National Tsing Hua University	Jun. 2025
Invited Talk, National Taiwan University	Jun. 2025

## PROFESSIONAL SERVICES

### **Organizing Committee**

Texas NLP Symposium 2026

### **Editorial Roles**

Action Editor, ACL Rolling Review (ARR) 2024 – Present

### **Conference Senior Area Chair**

Conference on Empirical Methods in Natural Language Processing (EMNLP) 2025 – 2026

### **Conference Area Chair**

Meeting of the Association for Computational Linguistics (ACL)	2024 – 2026
Conference on Empirical Methods in Natural Language Processing (EMNLP)	2024
Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL)	2025
Conference of the European Chapter of the Association for Computational Linguistics (EACL)	2026
International Conference on Machine Learning (ICML)	2026
Conference on Neural Information Processing Systems (NeurIPS)	2025 – 2026
International Conference on Learning Representations (ICLR)	2026

### **Conference Reviewer**

Meeting of the Association for Computational Linguistics (ACL)	2021 – 2023
Conference on Empirical Methods in Natural Language Processing (EMNLP)	2021 – 2023

Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL)	2022 – 2024
Conference of the European Chapter of the Association for Computational Linguistics (EACL)	2023 – 2024
Conference on Language Modeling (COLM)	2024 – 2026
International Conference on Machine Learning (ICML)	2020 – 2025
Conference on Neural Information Processing Systems (NeurIPS)	2021 – 2024
International Conference on Learning Representations (ICLR)	2021 – 2025
AAAI Conference on Artificial Intelligence (AAAI)	2022 – 2025
<b>Journal Reviewer</b>	
Computational Linguistics (CL)	2025 – Present
Journal of Machine Learning Research (JMLR)	2026 – Present
Transactions on Machine Learning Research (TMLR)	2024 – Present
<b>Handbook Assistant</b>	
Conference on Empirical Methods in Natural Language Processing (EMNLP)	2018
<b>Grant-Proposal Review Panelists</b>	
NSF CISE Directorate	2025 – 2026