Xingran Chen

Ann Arbor, MI ◇ Phone: +1 734-510-0713 ◇ chenxran@umich.edu ◇ www.chenxingran.com

EDUCATION

University of Michigan — Ann Arbor, MI, USA
Master of Science in Biostatistics — *Sept. 2022 - Present*

- GPA: 4.3/4.3 (all A^+)
- Outstanding First-Year Masters Student Award

Shanghai University of Finance and Economics — Shanghai, China
Bachelor of Science in Statistics (Data Science track) — *Sept. 2017 - Jun. 2021*

- Supervisor: Wanyun Cui
- Knowledge Graph, Rule Mining, Rule-based Reasoning, Natural Language Processing

RESEARCH EXPERIENCE

Research Assistant (Supervisor: Zhenke Wu) — *Jan. 2022 - present*
University of Michigan

- Identified important research questions in imputing missing covariates with machine learning models for statistical inference.
- Provided theoretical support and algorithms to generalize prediction-powered inference to arbitrary missing pattern scenarios.

Research Assistant (Supervisor: Stan Z. Li) — *Jul. 2022 - present*
Westlake University

- Led the design of a generalized Mixup approach for general-purpose data augmentation, along with an in-depth theoretical analysis of its generalization bound, and corresponding mixup evaluation framework.
- Participated in the construction of a protein design benchmark for providing a solid and trustworthy baseline and comparison for future research (ProteinBench). (NeurIPS 2023)
- Initiated the development of a SOTA deep learning framework for molecular editing, yielding cutting-edge performance in retrosynthesis prediction (MotifRetro).

Research Assistant (Supervisor: Wanyun Cui) — *Jul. 2020 - Jun.2022*
Shanghai University of Finances and Economics

- Introduced a novel instance-based reasoning method (IBLE) for improving knowledge graph completion. (NeurIPS 2022)
- Proposed a new rule induction system utilizing implicit knowledge stored in language models (Orion). (NeurIPS 2021)
- Designed an efficient approach for textual knowledge integration (OK-Transformer). (ACL 2022 Findings, ACL 2023)

PUBLICATIONS

Free Lunch for Efficient Textual Knowledge Integration in Language Models.
Wanyun Cui, **Xingran Chen**. *ACL 2023*

Exploring Automatically Perturbed Natural Language Explanations in Relation Extraction.
Wanyun Cui, **Xingran Chen**. *ACL 2023 Findings*

Instance-based Learning for Knowledge Base Completion.
Wanyun Cui, **Xingran Chen**. *NeurIPS 2022*

Open Rule Induction.
Wanyun Cui, **Xingran Chen**. *NeurIPS 2021*

Enhancing Natural Language Representation with Large-Scale Out-of-Domain Commonsense.
Wanyun Cui, **Xingran Chen**. *ACL 2022 Findings*

Enhance Causal Span Detection via Beam-Search-based Position Selector.
Xingran Chen, Ge Zhang, Adam Nik, Yuming Li, Jie Fu. *CASE Workshop in EMNLP 2022 (Oral)*

On the Effectiveness of Speech Self-supervised Learning for Music
Yinghao Ma, Ruibin Yuan, Yizhi Li, Ge Zhang, **Xingran Chen**, Hanzhi Yin, Chenghua Lin, Emmanouil Benetos, Anton Ragni, Norbert Gyenge, Ruibo Liu, Gus Xia, Roger Dannenberg, Yike Guo, Jie Fu. ISMIR 2023

MARBLE: Music Audio Representation Benchmark for Universal Evaluation.
Ruibin Yuan, Yinghao Ma, Yizhi Li, Ge Zhang, **Xingran Chen**, Hanzhi Yin, Le Zhuo, Yiqi Liu, Jiawen Huang, Zeyue Tian,

Binyue Deng, Ningzhi Wang, Wenhu Chen, Gus Xia, Wei Xue, Si Liu, Shi Wang, Ruibo Liu, Yike Guo, Jie Fu. *NeurIPS 2023 Dataset and Benchmark Track*

ProteinInvBench: Benchmarking Protein Inverse Folding on Diverse Tasks, Models, and Metrics
Zhangyang Gao, Cheng Tan, Yijie Zhang, **Xingran Chen**, Lirong Wu, Stan Z. Li. *NeurIPS 2023 Dataset and Benchmark track*

Leveraging Self-Training in Causality Classification of Socio-Political Event Data.
Adam Nik, Ge Zhang, **Xingran Chen**, Yuming Li, Jie Fu. *CASE Workshop in EMNLP 2022*

PREPRINT

TPDM: Selectively Removing Positional Information for Zero-shot Translation via Token-Level Position Disentangle Module.
Xingran Chen, Ge Zhang, Jie Fu. *arXiv preprint arXiv:2305.19857*

MotifRetro: Exploring the Combinability-Consistency Trade-offs in retrosynthesis via Dynamic Motif Editing.
Zhangyang Gao*, **Xingran Chen***, Cheng Tan*, Stan Z Li. *arXiv preprint arXiv:2305.15153*

MERT: Acoustic Music Understanding Model with Large-Scale Self-supervised Training.
Yizhi Li, Ruibin Yuan, Ge Zhang, Yinghao Ma, **Xingran Chen**, Hanzhi Yin, Chenghua Lin, Anton Ragni, Emmanouil Benetos, Norbert Gyenge, Roger Dannenberg, Ruibo Liu, Wenhu Chen, Gus Xia, Yemin Shi, Wenhao Huang, Yike Guo, Jie Fu. *arXiv preprint arXiv:2306.00107*

PROJECTS & WORKSHOPS

Multimodal Art Projection *Mar. 2023 - present*
contributor

- Participated in the advancement and discussion of training strategies for music generative foundational model.
- Contributed code to the construction of MIR evaluation benchmark MARBLE and training framework for music generative foundational model.

Enhance Causal Span Detection via Beam-Search-based Position Selector *Jul. 2022 - Sept. 2022*
The 5th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE@EMNLP 2022)

- Lead the development of a novel algorithm based on beam search for multi-spans detection in text. The performance of the proposed system obtained the **SOTA** in the workshop.

Enhance Language Models in Winograd Schema Challenge (WSC) *Sept. 2020 - Dec. 2020*
Shanghai University of Finances and Economics (Machine Learning course project)

- Explored the pattern of WSC and constructed the CausalWiki dataset based on prior work of WikiCREM.
- Used the dataset to fine-tune BERT and RoBERTa. The results obtained the SOTA on WSC273, PDP, and increased by on average 1.5% on DPR and WinoGender datasets.

AWARDS

- Outstanding First-Year Masters Student Award (UMICH), 2023
- Third Prize of Shanghai University Science and Technology Innovation Contest, 2021
- Provincial Third Prize of Contemporary Undergraduate Mathematical Contest in Modeling, 2019
- Third Prize of Statistical Contest in Modeling (SHUFE), 2018
- Shanjia Scholarship (SHUFE). 2018

SERVICES

Teaching
Teaching Assistant: BIOSTAT 602 - Biostatistical Inference (UMICH) Winter 2024
Teaching Assistant: BIOSTAT 650 - Theory and Application of Linear Regression (UMICH) Fall 2023
Academic Reviewer
ACL Rolling Review Oct. 2023
ICLR 2024
NeurIPS 2023
ACL 2023

EMNLP 2022, 2023

Department Service

Organizer, Graduate Student Academic Workshop, Department of Biostatistics, University of Michigan 2022, 2023

Social Activity & Volunteer

Volunteer, Primary School Dining Hall, Nagasaki, Japan 2022

Captain, SUFE Association of Baseball, Shanghai University of Finance and Economics, 2020-2021

SKILLS

Programming	`Python, R, SAS, SQL, C++`
Frameworks	`pytorch, huggingface, fairseq, lightning, pykeen, llm-foundry`
Toolkits	`Git, LaTex, WandB`
Languages	`Mandarin, English, Japanese (JLPT N2 150)`
Hobbies	`Chess, Hiking, Baseball`