Weijiang Li

Email : wli27@nd.edu
Website : https://vickyli99.github.io

EDUCATION

- **University of Illinois Urbana-Champaign**   **Master of Computer Science**
  GPA: 4.00/4.00   *Aug 2022 - Dec 2023*

- **University of Illinois Urbana-Champaign**   **Bachelor of Science, Computer Science and Chemistry**
  GPA: 3.76/4.00   *Aug 2018 - May 2022*

RELATED COURSEWORK

- **Graduate-Level Courses**: Advanced NLP, Meta-Learning, Transfer Learning, Statistical Modeling
- **Undergraduate-Level Courses**: Deep Learning, Bioinformatics, Text Information Systems, Database Systems, Inorganic Chemistry, Organic Chemistry, Physical Chemistry

PUBLICATIONS

- **From Text to Translation: Using Language Models to Prioritize Variants for Classification**:
  Authors: Weijiang Li, Xiaomin Li, Ethan Lavallee, Alice Saparov, Marinka Zitnik, Christopher Cassa
  Under review at Genome Medicine [PDF]

- **Emails by LLMs: A Comparison of Language in AI-Generated and Human-Written Emails**:
  Authors: Weijiang Li, Yinmeng Lai, Sandeep Soni, Koustuv Saha
  Published at 17th ACM Conference on Web Science (WebSci 2025) [PDF]

- **Assessing the Two-Sided Impact of Jailbreak on Autonomous Agents in Scientific Domain**:
  Authors: Chenhui Zhang*, Weijiang Li*, Aswathy Ajith, Zinan Lin, Ian Foster, Dawn Song, Bo Li

- **Chemical-Reaction-Aware Molecule Representation Learning**:
  Authors: Hongwei Wang, Weijiang Li, Xiaomeng Jin, Kyunghyun Cho, Heng Ji, Jiawei Han, Martin D. Burke
  Published at The Tenth International Conference on Learning Representations (ICLR 2022) [PDF]

- **Fine-Grained Chemical Entity Typing with Multimodal Knowledge Representation**:
  Authors: Chenkai Sun, Weijiang Li, Jinfeng Xiao, Nikolaus Parulian, ChengXiang Zhai and Heng Ji
  Published at IEEE International Conference on Bioinformatics and Biomedicine (BIBM 2021) [PDF]

- **Towards out-of-distribution generalizable predictions of chemical kinetics properties**:
  Authors: Zihao Wang, Yongqiang Chen, Yang Duan, Weijiang Li, Bo Han, James Cheng, and Hanghang Tong
  Published at AI for Scientific Discovery: From Theory to Practice (NeurIPS 2023) [PDF]

RESEARCH EXPERIENCE

- **Division of Genetics at Brigham and Women's Hospital, Harvard Medical School**   Boston, MA
  Research Assistant   *Jun 2023 - Present*
  - **Reclassifying Genetic Variant Pathogenicity with Language Models**:
    - Extracted and processed metadata from the ClinVar database, built data cleaning and filtering methods and pipelines for ClinVar submission summary text data, and analyzed the corpus based on semantic features
    - Trained a sequence classifier for ClinVar submission summary data, validated the model using clinical data, and evaluated trained models via UMAP and attention visualization to validate interpretability on model predictions
  - **Extracting Forms of Evidence of Pathogenicity from Clinical Text Summaries**:
    - Processed ClinVar dataset and distilled a gold standard set with explicit ACMG evidence code mentions, and leveraged LLMs for annotation on the dataset, resulting in a labeled dataset for evidence type classification models training
    - Using clinical data, such as functional score from deep mutational scanning data (DMS) and population data, for validating our trained classification models on extracting and labeling evidence types that are not mentioned in the text

- Using clinical data, such as functional score from deep mutational scanning data (DMS) and population data, for validating our trained classification models on extracting and labeling evidence types that are not mentioned in the text

Blender Lab Directed by Professor Heng Ji — Champaign, IL
Research Assistant — *Jan 2021 - Aug 2022*
- **Chemical Representation Learning**:
  - Processed existing chemical reaction datasets using RDKit; adapted and implemented baseline evaluation pipeline using models such as MolBERT and Mol2Vec, and obtained results for both reaction prediction and property prediction tasks
- **Fine-Grained Chemical Entity Typing**:
  - Constructed an organic chemical compound classification and naming ontology, organized and conducted chemistry annotation, and developed into the first open-sourced chemical entity typing evaluation dataset

Teaching Experience

University of Illinois Urbana-Champaign — Champaign, IL
Graduate Teaching Assistant — *Aug 2022 - Dec 2023*
- **CS 361, Probability & Statistics for Computer Science**: Fall 2022, Spring 2023, Fall 2023
  - Provided academic support to a large undergraduate class of over 350 students, by leading weekly recitation sessions to facilitate course content review, hosting office hours each week, and designing homework and exam questions and rubrics
- **CS 421, Programming Languages & Compilers**: Summer 2023
  - Assisted 150 undergraduate and graduate students during summer session through hosting weekly office hours and organizing quiz and exam reviews

University of Illinois Urbana-Champaign — Champaign, IL
Undergraduate Course Assistant — *Jan 2019 - May 2019*
- **CS 125, Intro to Computer Science, led by Professor Geoffrey Challen**:
  - Help with a class of 800+ undergraduate students on a programming course based on Java by helping students during weekly lab sections, hosting office hours and tutoring sessions, and answering questions on the class forum

Professional Experience

COUNTRY Financial — Champaign, IL
Program Manager and Software Development Intern — *May 2022 - Aug 2022*
- **Infrastructure as Code**:
  - Headed the research and execution of an Infrastructure as Code initiative, leveraging Ansible and GitLab CI/CD pipelines, which resulted in the implementation of several automated infrastructure development features within the organization
  - Conducted weekly meetings with project mentors and organized frequent internal meetings to monitor project progress; utilized GitLab README to document notes and issues for future reference

Leadership Experience

Built by Girls at UIUC — Champaign, IL
Corporate Chair — *Aug 2020 - May 2022*
- Collaborated with industry partners to design, coordinate, and deliver workshops and recruitment events for women in technology; organized an annual Project Summit featuring industry mentors to guide and mentor participants

Reflections | Projections — Champaign, IL
Content & Social Committee Staff — *Feb 2020 - May 2022*
- Arranged and hosted keynote talks and workshops for RP with speakers from industry and academia, and set up social events with partnered companies for students and staff members in the CS department during the conference

Alpha Phi Omega, National Service Fraternity — Champaign, IL
Volunteer Service Project Coordinator, Leadership & Tech Committee Member — *Jan 2019 - Jan 2020*
- Coordinated 150+ students and staff members in the organization, managed 30+ weekly volunteering events throughout the semester for the ETC Coffee House Volunteer Program, and maintained service events and logistics

Honors and Awards

- Best Reviewer Award, AAAI 2025 AI4Research Workshop, 2025

- Life + Career Design Scholarship from the College of Liberal Arts & Sciences, Spring 2021

- Edmund J James Scholar 2019 - 2021

- Eunice S. Wu Memorial Scholarship from the Department of Chemistry 2018 - 2019, 2019 - 2020

Posters and Talks

- Poster presentation at ASHG Annual Meeting, Oct 2025

- GRIP Seminar at the Division of Genetics at Brigham and Women's Hospital, Harvard Medical School, Oct 2025

- GRIP Seminar at the Division of Genetics at Brigham and Women's Hospital, Harvard Medical School, May 2024

- Poster presentation at the AI Cures Conference 2024 hosted by MIT and MGB, April 2024

Academic Service

- Reviewer for ACL 2024 Language + Molecule Workshop, ACL Demo Track 2025, EMNLP Demo Track (2024, 2025), ICLR 2025, AAAI 2025 AI4Research Workshop, Neurips SEA Workshop 2025, Program Committee Member, EACL 2026