MINSI LU

Tel: (+86) 188-0013-9136 | Mail: minsilu0330@gmail.com | Website | Github | Google Scholar

EDUCATION BACKGROUND

Department of Pharmacy science, **Tsinghua University** **Beijing, China**

Bachelor of Science in Pharmaceutical Sciences *Aug. 2019 - Jul. 2024(expected)*

Bachelor of Science with minor in Software Engineering *Aug. 2021 - Jul. 2024(expected)*

PUBLICATIONS

Bowen Gao, Bo Qiang, Haichuan Tan, Minsi Ren, Yinjun Jia, **Minsi Lu**, Jingjing Liu, Weiying Ma, Yanyan Lan, **'DrugCLIP: Contrastive Protein-Molecule Representation Learning for Virtual Screening', NeurIPS 2023.**

Tangqi Fang, Yifeng Liu, Addie Woicik, **Minsi Lu**, Anupama Jha, Xiao Wang, Gang Li, Borislav Hristov, Zixuan Liu, Hanwen Xu, William S. Noble, Sheng Wang, **'Enhancing Hi-C Contact Matrices for Loop Detection with Capricorn', bioRxiv 2023.**

RESEARCH EXPERIENCE

Research Assistant, School of Computer Science & Engineering, University of Washington **Seattle, US**

Project: Resolution Enhancement and Time Dimension Modeling for Hi-C Data *Jun. 2023 - Oct. 2023*

Advisor: Prof. William Stafford Noble, Prof. Sheng Wang

- Aimed to integrate the time and other feature dimensions of Hi-C data for resolution enhancement.
- Independently integrated the temporal dimension of Hi-C data for better performance.
- Developed Capricorn, a tool for Hi-C resolution enhancement that incorporates high-order chromatin features as additional views of the input Hi-C contact matrix and leverages a diffusion probability model backbone to generate a high-resolution matrix. Capricorn outperforms the state-of-the-art in a cross-cell-line setting.
- Formed a related paper to introduce the novel resolution enhancement methodology

Research Assistant, Institute for AI Industry Research, Tsinghua University **Beijing, China**

Project: Contrastive Learning for Drug Virtual Screening *Feb. 2023 - Jun. 2023*

Advisor: Prof. YanYan Lan

- Aimed to develop a new method for virtual screening, which is a critical step in AI-assisted drug discovery.
- Proposed a novel contrastive learning framework, DrugCLIP, by reformulating virtual screening as a dense retrieval task and employing contrastive learning to align representations of binding protein pockets and molecules from a large quantity of pairwise data without explicit binding-affinity scores.
- Introduced a biological-knowledge-inspired data augmentation strategy using homologous counterparts to learn better protein-molecule representations.
- Formed a related paper to introduce the novel feature representation methodology, was submitted to NeurIPS.

Research Assistant, Institute for AI Industry Research, Tsinghua University **Beijing, China**

Project: Generative Model for Structure-Based Drug Design *Oct. 2023 - Present*

Advisor: Prof. Hao Zhou

- Aimed to propose a novel approach to generate a new drug molecule atom by atom based on the structure features of its binding pocket.
- Independently introduced the DrugCLIP framework for protein-molecule representation and designed the methodology of molecule generation.
- Compared this new method to some state-of-the-art methods and discussed the further improvement.

Research Assistant, Integrated Science Research Center, Peking University **Beijing, China**

Project: High-throughput Genome-wide Functional Screening Based on CRISPR *Jul. 2021 - Sep. 2021*

Advisor: Prof. Wensheng Wei

- Aimed to use CRISPR-Cas9 functional screening tools to discover the genes related to the IFN-α-induced apoptosis pathway.

- Independently introduced and designed sgRNA sequences, and construct sgRNA plasmids for gene screening.
- Constructed GoldenGate point mutation on plasmids.
- Carried out cell culture, virus extraction, and titer determination.

Research Assistant, School of Pharmaceutical Science, Tsinghua University **Beijing, China**

Project: Research on Her2 and CXCR4 Bispecific Antibodies *Oct. 2020 - Feb. 2021*

Advisor: Prof. Juanjuan Du

- Aimed to design a novel bispecific antibody to treat breast cancer.
- Independently screened the strain, did the transformation, and plasmid transfection, determined the antibody Kd value.

PROJECTS

Tsinghua University **Beijing, China**

Project: Multiomics Integration Via Graph Learning *Sep. 2022 - Jan. 2023*

Advisor: Prof. Jianyang Zeng

- Aimed to integrate scRNA-seq data and scATAC-seq data through graph learning based on GLUE.
- Independently decayed the weight of edges according to the distance of the ATAC peak from the TSS and used a set of independent aggregators to combine the messages in each GCN layer.
- Our method achieved comparative performance with GLUE on metrics, and had the ability to identify more genes with regulatory relationships than GLUE.
- Repo: https://github.com/minsilu/Multiomics-Integration-via-Graph-Learning

Project: PPI Network Guided Driver Target Discovery *Sep. 2022 - Jan. 2023*

Advisor: Prof. Jianyang Zeng

- Aimed to design a new method for driver target discovery using single-cell RNA-seq data by introducing a protein-protein interaction (PPI) network.
- Independently introduced PPI to re-calulate regularization loss in the sc-ETM model and design a 2-layer MLP for target prediction.
- Achieved better cell clustering performance than sc-ETM and better driver target discovery performance thanked to the introduction of a PPI network for gene embedding. Biological case studies on our discovered targets also showed the effectiveness of our proposed framework.
- Repo: https://github.com/minsilu/PPI-Guided-Driver-Target-Discovery

Project: Text Sentimental Classification Based on the Twitter Dataset *Feb. 2023 - Jun. 2023*

Advisor: Prof. Mingsheng Long

- Spearheaded a comprehensive text classification project utilizing a Twitter sentiment analysis dataset, focusing on advanced preprocessing and diverse model comparison.
- Independently used TF-IDF methods for feature extraction; implemented and optimized a variety of machine learning models, including Decision Trees, Random Forests, MLPs, ResNet, and BERT, for nuanced sentiment analysis.
- Achieved breakthrough accuracy, particularly with the BERT model, outperforming other models in classifying sentiments into four categories.
- Repo: https://github.com/minsilu/Text-Sentimental-Classification-based-on-Twitter-Dataset

Project: Gomoku Genius: AI Search Strategies for Classic Board Game Mastery *Feb. 2023 - Jun. 2023*

Advisor: Prof. Mingsheng Long

- Aimed to investigate and implement five distinct AI agents for board games: Minimax Search, Alpha-Beta Pruning, Truncated Search, Monte Carlo Tree Search (MCTS), and Evaluation Function-based AlphaZero, addressing the challenges of search complexity and function design.
- Independently: Optimized Minimax for Tic-Tac-Toe. Applied Alpha-Beta Pruning to reduce search space. Implemented Truncated Search in Go-Moku, balancing time constraints with strategic depth using pattern evaluation. Conducted gameplay analysis comparing naive MCTS and Alpha-Beta. Enhanced MCTS with evaluation function, significantly improving board state assessments and decision-making.
- Repo: https://github.com/minsilu/Gomoku-Genius--AI-Search-Strategies-for-Classic-Board-Game-Mastery

SKILLS

Mathematics: Calculus, Linear Algebra, Probability and Statistics, Discrete Mathematics, Data Structure and Algorithms.

Programming Skills: Skilled at C/C++, Java, Python, R, SQL. Familiar with Machine Learningand Deep Learning, including several Deep Learning Frameworks. Familiar with Linux/Unix.

Experimental skills: pharmacy, chemistry and biology

AWARDS & HONORS

Scholarship for Excellent in Academic, **University Scholarship** (10%) *2023*

First Prize in 16th "Novozymes Cup" Life Science Culture Festival, **Municipal Award** (10%) *2021*

Third Prize in 39th "Challengs Cup" Student Academic Competition, **University Award** *2021*

Scholarship for Progress in Academic, **University Scholarship** *2021*

Third Prize in Winter Vacation Social Practice for Students, **University Award** *2021*

Gold Medal in 27th China Biology Olympiad(CBO), **National Award** *2018*

SOCIETY ROLES

Leader at Student Winter Vacation Social Practice Team. *Dec. 2020 - Jan. 2021*