Ziwen (Aaron) Li

7400 Boelter Hall, Los Angeles, CA 90095
zil105@g.ucla.edu (+1)3106968091

Research Interest

My research interest lies in Human-centered NLP, focusing on designing and developing robust and trustworthy human-AI decision-making systems to support activities of users from different communities such as lay users, domain experts and AI specialists.

Education

University of California, Los Angeles 10/2023–present
*M.S. degree in **Electrical and Computer Engineering***

University of Electronic Science and Technology of China (UESTC) 09/2019–06/2023
*B.Eng.degree in **Electronic Information Engineering** from **UESTC***
*B.Eng.(Honours) degree in **Electronics and Electrical Engineering** from **University of Glasgow (UoG)** (Dual Bachelor Degree Program)*

Coursework: Large-Scale Data Mining: Models and Algorithms (**GPA: 4.0**), Neural Networks and Deep Learning (**GPA: 4.0**), Advanced Neural Network and Deep Learning (**GPA: 4.0**)
Skills: Pytorch, Neo4j, Langchain, Javascript, React.js, Node.js, Java, Docker, microPython, Latex, Git/Terminal, MATLAB

Publications and manuscripts in review

GraPPI: A Retrieve-Divide-Solve GraphRAG Framework for Large-scale Protein-protein Interaction Exploration
Ziwen Li, Xiang 'Anthony' Chen, Youngseung Jeon.
Under review of NAACL'25

HAPPIER: An AI-based Scientific Discovery Support Tool for Facilitating Target Identification in Drug Discovery
Youngseung Jeon, **Ziwen Li**, Christopher Hwang, Jesus J. Campagna, Cohn Whitaker, Varghese John, Eunice Jun, Xiang 'Anthony' Chen
Under review of the 2025 CHI Conference on Human Factors in Computing Systems (CHI' 25)

CoSight: Incentivizing Sighted Viewers to Contribute to Online Video Accessibility
Ruolin Wang, Xingyu Liu, Biao Wang, Wayne Zhang, Ziqian Liao, **Ziwen Li**, Amy Pavel, Xiang 'Anthony' Chen.
Under review of the 2025 CHI Conference on Human Factors in Computing Systems (CHI' 25)

Decoupling Identification of Nonlinear Two-input Two-output Deterministic Systems based on Adaptive Regression and Causal Decomposition
Yi Zhang; Jinli Chen ; Zhirong Liu; Mingjun Xie; Zewen Dong; Mian Wu; **Ziwen Li**; Hailou Jin; Chanyin Yi; Siyu Xie; Steven Su; Peng Xu; Dezhong Yao
Under review of the IEEE Transactions on Industrial Electronics

Matlab Open Source Code: Noise-assisted Multivariate Empirical Mode Decomposition based Causal Decomposition for Causality Inference of Bivariate Time Series
Yi Zhang*, Guan Wang*, **Ziwen Li***(equal contribution), Mingjun Xie, Branko Celler, Steven Su, Peng Xu, Dezhong Yao (Jun, 2022).
Frontiers in Neuroinformatics, doi: 10.3389/fninf.2022.851645

Projects

HAPPIER: An AI-based Scientific Discovery Support Tool for Facilitating Target Identification in Drug Discovery 05-09/2024

- HAPPI is a **knowledge graph-based retrieval augmented generation** (RAG) system that support Protein-Protein Interactions (PPIs) pathway exploration that allowed domain experts to obtain an explainable PPI graph based on the requirements in input natural language queries and the information in existing dataset, providing a transparent and flexible PPI path recommendation to support proteins candidates selection in wet-lab sessions.

- Implemented natural language processing (NLP) methods such as **K-nearest neighborand (KNN) search** and **large language models (LLMs) inference** for different levels of recommendation due to the tradeoff between level of customization for recommendation and computational efficiency and constructed our graph database from STRING dataset suggested by experts in drug discovery.
- Used **Langchain, neo4j graph database** to secure the expanding, editing, customized path of the recommended graph of protein.
- Our database includes 18,767 proteins (nodes) and 2,955,220 PPIs (edges) which is the largest PPI graph database so far and **our system outperforms vanilla models in matrices of BERTScore, ROUGE-1, and ROUGE-L**.
- Undergoing **a pilot release in the UCLA Drug Discovery Lab** to support Alzheimer's drug development.

Colla: A Social Reading Tool for Promoting Accessibility of Research Paper Figures 10/2023-05/2024

- Colla is a **social reading tool based on LLM agents** that breaks paper reading into sub-tasks such as questions-answering, descriptive summary generation, highlighting, and making annotations on paper content and figures to support reading performance of both sighted and blind and low-vision (BLV) researchers.
- Used **React.js, node.js** to develop a full-stack chrome extension to coordinate the by-product information stored in a vector database created by **Langchain** from their annotations for content generation supporting of LLMs.
- Designed **a hybrid crowdsourcing solution** that combines machine functionality with the expertise of researchers within the same community, who share similar research interests and experiences, to offer a tailored approach to generating accurate and descriptive alt text for scientific figures.

Customized Article Summarization Improvement based on Large Language Model (LLM) Knowledge Distillation and Small Language Model Implementation 04/2023-06/2023

- Developed a custom text summarization model for handling specific structures and styles and investigated the limitations of LLMs like text-davincioo3 and turbo in handling custom **text summarization tasks** and explore the use of smaller models like BART.
- Utilized LLMs for preliminary text processing and feed the simplified data into smaller models for specific structural and stylistic transformations.
- Evaluated the system on three different datasets featuring customized characteristics: Strudel, MacSum, and QMSum with designed prompts for GPT-3.5 to handle preliminary data processing differently for each dataset.

Experience

Graduate Student Researcher 10/2023-present
UCLA HCI Lab
Supervisor: ***Prof. Xiang 'Anthony' Chen*** *and* ***Prof. Eunice Jun***

Teaching Assistant 10/2024-present
Courses: ECE 102 Systems and Signals
Lecturer: ***Prof. Jonathan Kao***

Research Intern 04/2023-06/2023
Westlake University
Supervisor: ***Prof. Yue Zhang***

Internship

Face (Mask) Recognition Access Control System, 5G + 4K Video Surveillance + AI Intelligent Analysis in the 5G Application Project of Anhui-Conch Group. 06-08/2022
Company: Tongling Branch of China Telecom (Global Top 500 Enterprise)
Image Processing Intern

- Utilized Yolo5 network to replace the original VGG-16CNN demo network to improve the recognition performance in pedestrian face capture with masks.
- Improved to the system performance in high crowd intensity scenario with high-resolution surveillance cameras by using image cropping techniques.
- Responsible for preliminary investigation, model training, debugging, and delivering the final presentation.