Bomin (David) Wei

(609) 356 3006 davidwei23@ucla.edu GitHub Linkedin

EDUCATION

University of California, Los Angeles, Los Angeles, CA Sep. 2023 – Jun. 2027
Freshmen, Linguistic and Computer Science
Core Courses Taken:
AP Calculus BC (5), AP Computer Science A (5), Linear Algebra, Multivariable Calculus, Math 32A, CS 31

RESEARCH & INTERSHIP EXPERIENCE

Deep Learning-based Model for Drug Repurposing Mar. 2021 - Oct. 2022
University of Utah
- Designed an embedding method based on **Word2Vec** to obtain complete drugs and proteins representation
- Designed **CNN** + **LSTM** Deep Neural Network to extract non-local features for binding interactions
- Developed **novel testing methods** with special data splitting to evaluate models' performance in real-world applications; and automated data cleaning and pre-processing of 2 million text string data; optimized the **classification criteria** for better model selection by **statistical analysis** on data distributions.
- Presented in professional conferences **Intelligent Systems for Molecular Biology** (ISMB, July. 2022) and the **IDWeek** (Oct. 2022); and **published (first author)** on Scientific Reports.
- https://github.com/David-BominWei/DeepLPI

SARS-COV-2 Genetic Mutation Modeling Prediction Oct. 2020 - Mar. 2021
Hong Kong University
- Modeled and visualized the SARS-CoV-2 mutation based on global geographical regions and time
- Used ARIMA model to analyze SARS-CoV-2 RNA sequence mutations as time series and predicted most likely mutation sites for different regions.

Machine Learning Algorithm (NLP Group) Intern Jul. 2023 – Sep. 2023
4Paradigm, Beijing
- Compared and analyzed existing passage retrieval methods, such as BM25, using an efficient self-designed dataset; compiled query similarities datasets with a focus on banking question answering and integrated difficult negative samples; designed and finetuned a new query similarity model now used by the group.
- Finetuned and evaluated the basic model and found a problem that cannot distinguish proper nouns and extract important words. Finetuned the embedding module to reached better model result.

INTERESTS PROJECTS

Volunteer Management & Information Platform Sep. 2020 – Dec. 2022
- Conducted assessment for Little Oaks Charity Center, leading to a customized volunteer management system.
- Achieved 1,614 users and 320 project registrations within a year, significantly improving operational efficiency.

Q2Q Dataset (https://github.com/David-BominWei/Q2QDataset) Jul. 2023 – Aug. 2023
- Created an open-source Chinese query similarity dataset to train question-answering models.
- Utilized the BM25 model to construct negative samples for training.

Personal Blog Designer and Developer Sep. 2020 – Sep. 2023
- Developed a blog platform using Hexo and deployed it on GitHub Pages.

Machine Learning for Movie Recommendation Mar. 2023 – Jun. 2023
- Built a movie recommendation engine using inception networks, compared with a bag-of-words model.

Computational Graphics Mar. 2023 – Jun. 2023
- Developed a 3D GUI for matrix projection using Pygame, applying object-oriented programming.

SKILLS

- **Programming Languages:** Python (PyTorch, Scikit-learn, Pandas), C++, Java, R
- **Frameworks:** LSTM, ResNet, Transformer, CNN, MLP; Hexo, WordPress
- **Software:** Fusion360, OnShape, AutoCAD, Premiere Pro

HONORS AND AWARDS

- **First Place** in the Computer Science category and Air Force Research Laboratory Award, Mercer Science and Engineering Fair (ISEF affiliated) 2023
- **Silver Medal**, S.-T. Yau High School Science Award, USA Regional; **ranked 2nd** in CS category 2022
- **1st Place & Best Poster** (in Biology and Medicine) at IEEE-ISEC 2021 Conference 2021
- **Gold Division**, United States of America Computing Olympiad (USACO) 2022
- **1st Place** Kaggle Science Olympiad National Invitational 2022 Machine Learning Event 2022
- **6th place in total & 3rd place in AI** round in the hackathon CMU Info & Math Competition (CMIMC) 2022

PUBLICATIONS

1. Wei, B., et al. *"DeepLPI: A Deep Learning Model for Drug Repurposing."* Sci Rep 12, 18200 (2022) https://www.nature.com/articles/s41598-022-23014-1
2. Wei, B., et al. *"Modeling SARS-CoV-2 Mutations Based on Geography and Time."* bioRxiv (2021) https://www.biorxiv.org/content/10.1101/2021.08.11.455941