Shunyang Wang, Ph.D.

lygwangshunyang@gmail.com| | (530)298-6828 | https://shunyang.xyz/

EDUCATION

Ph.D. in Chemistry and Chemical Biology **09/2018 – 12/2022**

University of California, Davis
Advisors: Dr. Oliver Fiehn, Dr. Dean J Tantillo, Dr. Tobias Kind
Dissertation: *Quantum chemistry explorations on compound identification in metabolomics*

RESEARCH & EXPERIENCE

Machine Learning Scientist, Computational Metabolomics, Brightseed **02/2023 – Current**

Metabolomics pipeline automation

- Collaborated with a multidisciplinary team spanning metabolomics, natural products, business development and computational experts to develop innovative solutions.
- Designed and implemented a fully automated pipeline integrating sample intake, data processing, bioactive compound prediction and commercial report generation.
- Developed and maintained the backend infrastructure powering the company's core products.

Bioactive Natural Product Discovery

- Developed a mass spectrometry-based algorithm for bioactive prediction.
- Validated the model through collaboration with experimental teams, doubling the success rate and reducing screening time from months to weeks.
- Contributed numerous leads to preclinical and clinical studies.

Metabolomics Spectral Database

- Developed a fully automated data processing pipeline, serving as the core of company products.
- Built and maintain a spectral database of over 10 million entries for plant profiling.
- Optimized database search time, reducing it from hours to microseconds.

Retention Time Prediction of Small Molecules Using Machine Learning

- Created a graph neural network model, achieving a 100-fold increase in prediction speed.
- Achieved a reduction in mean absolute error to within 10 seconds for a 10-minute method.
- Deployed a scalable machine learning model in a Docker environment, facilitating automated retention time prediction in experimental workflows.

Doctoral Research, UC Davis **12/2018 – 12/2022**

Quantum Chemistry Simulation of Electron Ionization Mass Spectra

- Developed the world's first in-silico mass spectral database using molecular dynamics.
- Created a cutting-edge excited-state molecular dynamics algorithm.
- Developed and implemented an automated MD pipeline to simulate thousands of molecules on HPC resources, significantly accelerating data generation for mass spectral predictions.
- Enhanced mass spectral prediction accuracy, surpassing all machine learning methods.

Reaction Dynamics and Selectivity in Photochemical Reactions

- Investigated the dynamical origins of selectivity in a classic photochemical reaction using machine-learning-assisted non-adiabatic molecular dynamics.
- Demonstrated the dominant role of momentum in controlling product selectivity, published in *Nature Chemistry 16, 615–623 (2024)*.

SKILLS

Computation Skills

- **Programming & Scripting**: Python, Fortran, SQL
- **Machine Learning & Cheminformatics**: PyTorch, Tensorflow, RDKit, Chemaxon
- **High-Performance Computing**: AWS, Docker, Slurm
- **Computational Chemistry**: Gaussian, ORCA, molecular docking, OpenMM

Laboratory Skills

- **Analytical Techniques**: GC-MS and LC-MS operation, metabolomics sample preparation.