PHYO PHYO KYAW ZIN

Cheminformatics Scientist, Data Scientist & Software Developer

phyophyokyawzin1@gmail.com| 919.771.5803 | North Carolina | Website | LinkedIn | GitHub| DevPost | Google Scholar

PROFESSIONAL EXPERIENCE

Terray Therapeutics Inc. (Remote)
Cheminformatics Scientist (August 2022 – Present)

- Designed and optimized large virtual chemical libraries containing millions of molecules for chip-based synthesis, advancing compound diversity, SAR understanding, and hit optimization.
- Developed and automated rapid enumeration methods and workflows for virtual libraries, improving library enumeration and analysis speed by 73%.
- Curated and analyzed large-scale chemical data, extracting actionable structure–activity relationship (SAR) insights to improve drug design and optimization.
- Built and managed databases and metabase views for tracking compound registration, protein targets, and results, enabling real-time data access and decision-making.
- Collaborated with building block vendors to catalog and procure materials, expanding and diversifying Terray's chemical space.

Simulations Plus Inc. (Remote)
Scientist 1 – Cheminformatics (May 2021 – August 2022)
Postdoc – Cheminformatics (July 2020 – April 2021)

- Led efforts in mining, curating, and analyzing ADMET properties from biological data and literature sources to predict metabolism sites, mutagenicity, and other key endpoints.
- Built machine learning models to predict CYP metabolism sites, and metabolites.
- Developed machine learning models for mutagenicity prediction using major bacterial strains, improving the accuracy and application of Ames test predictions.
- Contributed to the development and optimization of ADMET Predictor for better chemical safety evaluations, including predicting key pharmacokinetic and toxicological endpoints.

Collaborations Pharmaceuticals Inc. (Hybrid)
Cheminformatics Intern (June 2019 – August 2019)

- Developed the MacrolactoneDB web application, hosting macrolactones with curated bioactivity data, enhancing knowledge sharing in drug design.
- Built QSAR models for protein targets using machine learning algorithms, advancing predictions of biological endpoints for macrolides.
- Developed reliability & applicability domain assessment methods to quantify model performance & improve accuracy.

SKILLS & EXPERTISE

- **Computational Toxicology & ADMET**: Expertise in modeling CYP metabolism, mutagenicity, solubility, & other ADMET properties.
- **Cheminformatics & Data Science**: Data mining and building predictive models for chemical safety, including ADMET properties and metabolism.
- **Machine Learning & QSAR Modeling**: Applying clustering, classification, regression, and ensemble methods for biological and toxicological endpoint prediction.
- **High-Throughput Data Analysis**: Integrating and analyzing screening and assay data for chemical toxicity and safety.
- **Programming & Automation**: Advanced Python (pandas, numpy, scikit-learn, RDKit), SQL, R; experience with workflow automation and cloud computing.
- **Data Visualization & Reporting**: Proficient in matplotlib, seaborn, ggplot2, Spotfire, Tableau, and Metabase for effective data visualization and reporting.
- **Scientific Communication**: Strong skills in technical documentation, presentations, and creating dashboards for cross-functional collaboration.

SELECTED PROJECTS

- Ames (ADMET Predictor): Updated Ames models with new data, built classification models for predicting mutagenicity across major strains.
- CYP SoM / Substrate / Inhibitor (ADMET Predictor): Mined and curated extensive metabolism data to develop ML models for predicting CYP substrates, inhibitors, and specific sites of metabolism, along with resulting metabolites.
- **ROSE (Resynthesis Selection and Optimization Engine):** Developed a Streamlit app for suggesting compounds for resynthesis from large chemical libraries, leveraging multiple data-driven, SAR-derived algorithms and user-defined parameters for optimal results.
- **BB Mapper (Building Block Mapper):** Created a framework for mapping off-platform resynthesis compounds to on-platform compounds with various chemical tethers and locations, ensuring accurate mapping of building blocks and quality control.
- **Library Design Workflow:** Led the development and implementation of a comprehensive library design workflow, building core components and overseeing the process from design to registration and hand-off.
- **Library Enumeration Project:** Developed a reaction-based enumeration system and implemented an automated workflow to track, enumerate, and notify the team of library enumeration statuses via Slack, including detailed logs on successes and errors.
- SIME (PhD Thesis): Created a scaffold-based enumeration method integrating SMs and sugars to build virtual libraries of macrolides.
- MacrolactoneDB (PhD Thesis): Developed a web application for macrolactones, including bioactivity information curated and mined from public repositories and literature.
- Imatinib DLNN (PhD Thesis): Implemented rigorously validated and highly predictive QSAR regression models using 2D/3D/4D molecular dynamics derived & time-series based descriptors for Imatinib analogues.
- **Virtual Bio-profiling of PFAS Derivatives:** Conducted cheminformatics analysis and visualization of PFAS derivatives, enhancing understanding of environmental contaminants.

EDUCATION

- **PhD, Chemistry (Specialty: Cheminformatics)**
  North Carolina State University, Raleigh, NC
  Adviser: Denis Fourches (2016 – 2020)
- **B.A., Computer Science & Chemistry (Dean's List)**
  Berea College, Berea, KY
  (2012 – 2016)

SELECTED AWARDS & ACHIEVEMENTS

- **Best Use of Google Cloud Platform**, PackHacks 2019
- **Best Use of AI/ML**, PearlHacks 2019
- **CINF Scholarship for Scientific Excellence**, ACS 2017
- **2nd Place, Poster Presentation**, TRIWiC, Ohio 2016
- **Pugsley Award for Excellence in Mathematics**, Berea College 2012
- **Class of 1958 Chemistry Award**, Berea College 2016

FELLOWSHIPS & SCHOLARSHIPS

- **AAUW International Doctoral Fellowship** ($20k), 2019-2020
- **Olive Ruth Russell Fellowship** ($6k) (2016, 2017, 2019)
- **4-year Tuition Scholarship** ($130k), Berea College, 2012-2016
- **Scholarship ($3.5k),** Pre-Collegiate Program of Yangon, 2011