PHYO PHYO KYAW ZIN

Cheminformatician, Data Scientist & Software Developer

phyophyokyawzin1@gmail.com| 919.771.5803 | Raleigh NC, 27606 | Blog | phyophyokyawzin | zinph | zinph

EXPERIENCE

Terray Therapeutics Inc. **Raleigh, NC**

***Cheminformatics Scientist* (August 2022 – Present)**

- Design diverse and focused screening libraries which contain up to hundreds of millions of molecules to be synthesized on chips.
- Develop rapid, reaction-based enumeration methods for enumerating virtual libraries.
- Analyze chemical data to extract valuable SAR insights for interactive drug design and optimization.
- Develop a fully automated QSAR analysis pipeline for tArray (and IC50) data to enhance lead optimization and streamline resynthesis selection.
- Collaborate with building block vendors to catalog and facilitate purchases, further expanding the Terray chemical space.

Simulations Plus Inc. **Raleigh, NC**

Scientist 1 – Cheminformatics (May 2021 – August 2022)
Postdoc – Cheminformatics (July 2020 – April 2021)

- Extract biological data including ADMET properties of small molecules from literature, public and commercial databases.
- Develop automated pipelines to efficiently process and curate chemical data.
- Build machine learning models to predict biological endpoints of molecules (e.g., sites of metabolism for CYP450 isoenzymes, mutagenicity across AMES strains).

Collaborations Pharmaceuticals Inc. **Raleigh, NC**

Cheminformatics Intern (June 2019 – August 2019)

- Build MacrolactoneDB, a web application hosting macrolactones with bioactivity information mined & curated from public repositories.
- Build QSAR regressors for three protein targets with high frequency across MacrolactoneDB using 7 ML methods and 7 descriptor sets to predict biological endpoints of macrolides.

SKILLS

Python, R, Keras, Tensorflow, Linux, RDKit	8/10
Machine Learning, Deep Learning	7/10
Multi-GPU computing, Big Data Analysis	8/10
Chemical Data Mining, Curation & Data Visualization	9/10
Automation scripts, web scraping	8/10
Virtual Chemical Manipulation & Library Design	9/10
Web Development, PHP, HTML, CSS	6/10
Technical Writing & Presentations	8/10

FELLOWSHIPS

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

EDUCATION

PhD, Chemistry (Specialty: Cheminformatics), 2016 - 2020
North Carolina State University, Raleigh, NC
Adviser: Denis Fourches

B.A., Computer Science & Chemistry (Dean's List), 2012 - 2016
Berea College, Berea, KY

Publications can be found on Google Scholar.

ACHIEVEMENTS & AWARDS

- Best Use of Google Cloud Platform at PackHacks2019 (Apr 2019)
- Second-Best Overall Project at DiamondHacks2019 (Mar 2019)
- Best Use of AI/ML at PearlHacks2019 (Feb 2019)
- CINF Scholarship for Scientific Excellence at ACS (Aug 2017)
- Storytelling prize at ETHDenver (Feb 2018)
- Class of 1958 Chemistry Award at Berea (Apr 2016)
- 2nd place in poster presentation at TRIWiC, Ohio (Feb 2016)
- Pugsley Award for Excellence in Mathematics at Berea (Dec 2012)

PROJECTS

SoM (ADMET Predictor): Mined metabolism data from literature, public and commercial databases, extensively curated data, and built machine learning models to predict atomic sites of metabolism in the molecules along with the resulting metabolites.

Ames (ADMET Predictor): Updated 11 existing Ames models from ADMET Predictor with new NIHS data, which were preprocessed and curated in-house. Built classification models to predict mutagenicity of new compounds across major strains.

PKS Enumerator **(PhD Thesis):** Developed Structural Motif (SM) based enumeration algorithm to generate *in-silico* chemical libraries of macrolide scaffolds with multiple user constraints

SIME **(PhD Thesis):** Developed scaffold-based enumeration method that integrates SMs and sugars of interest at specified places to create virtual libraries of fully assembled macrolides

MacrolactoneDB **(PhD Thesis):** See Cheminformatics Intern at Collaborations Pharmaceuticals Inc. in **Experience**

CryptoChem **(PhD Thesis):** Developed a novel cheminformatics-based encryption system to securely store and transfer information with chemicals. (Funded by DARPA)

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 generated chemical network visualizations of PFAS derivatives from EPA database

River Blindness**:** Identified new potent Closantel analogues via drug repositioning, cheminformatics analysis, QSAR modeling and structure activity-based docking

Sustainability Dashboard Management System**:** Developed an integrated platform for displaying sustainability data in meaningful graphs, along with an administrative backend to effectively store, organize and manipulate large data

GhostWriter: Devised an NLP-based ML system that adaptively co-authors movie scripts with images and creates audiobooks

RecipeGenie**:** Designed an AI-based software to suggest dishes along with recipes by taking a snap of food in the pantry/fridge

ChemX**:** Developed a program that takes a molecule and generates its analogues by replacing its fragments with the building blocks sharing similar chemical properties

MyPueblo**:** Built an AI system that predicts real-estate property values, classifies bank transaction data into relevant finance categories & visualizes total cost breakdown for each category