

Maryam Taherzadeh

Santa Clara, CA 95054

☎ (+1) 650-417-1976

✉ mtaherzadeh696@alumni.sfbu.edu

🌐 [linkedin.com/in/maryam-taherzadeh](https://www.linkedin.com/in/maryam-taherzadeh)

GitHub: github.com/Maryam-Taherzadeh

Computational Scientist | Cheminformatics | AI in Drug Discovery

Summary

Computational Scientist focused on AI in drug discovery, with dual Master's degrees in Computer Science and Physical Chemistry and biotechnology industry experience. Develops Python and RDKit workflows for molecular data curation, feature engineering, QSAR/QSPR modeling, and bioactivity and solubility prediction. Applies computational methods to hit prioritization and molecular design relevant to hit-to-lead and lead optimization. Developed an MDM2-p53 inhibitor design workflow integrating QSAR potency prediction, reinforcement-learning-based molecular generation, predicted ADMET assessment, docking, and protein-ligand interaction analysis for iterative computational optimization. Experienced in protein-protein interaction (PPI) modeling, structural analysis, and building reproducible research pipelines and interactive scientific applications.

Technical Skills

Programming	Python, SQL, Bash/Linux, pandas, NumPy, Jupyter
Cheminformatics	RDKit, ChEMBL, molecular data curation, SMILES processing, molecular standardization, descriptors and fingerprints, chemical similarity analysis, physicochemical filtering
Molecular Machine Learning	QSAR/QSPR, IC50/pIC50 prediction, solubility and molecular property prediction, graph neural networks, scikit-learn, PyTorch, PyTorch Geometric, XGBoost, Random Forest, SVM/SVR
Data & Model Evaluation	Data preprocessing, feature engineering and selection, model validation, hyperparameter optimization, performance benchmarking
Structure & Interaction Analysis	Protein sequence and structure analysis, Biopython, protein-ligand and protein-protein interaction modeling, binding-site and interface analysis, AlphaFold, AlphaPulldown, MaSIF, SurfMAP, AutoDock Vina, PyMOL, PLIP
Generative Molecular Design	REINVENT4, DiffSBDD, reinforcement-learning-based molecular generation, diffusion-based design, reward design, compound prioritization, iterative computational optimization
Scientific Computing	Git/GitHub, Conda, reusable Python pipelines, reproducible workflows, Docker, Azure HPC, GCP, CUDA, SLURM, Singularity
Scientific Applications	Streamlit, Hugging Face Spaces, interactive molecular prediction tools, scientific visualization

Education

May 2019–May 2021	M.Sc. in Computer Science , <i>San Francisco Bay University</i> , California, USA, GPA: 3.67
Sep 2012–Feb 2015	M.Sc. in Physical Chemistry , <i>Shahrood University of Technology</i> , Semnan, Iran Thesis: Developed QSPR models to predict activity coefficients at infinite dilution for organic solutes in ionic liquids ([BMPYR][TCM]) using Support Vector Machines (SVM) and Artificial Neural Networks (ANN) .
Sep 2007–Nov 2011	B.A.Sc. in Applied Chemistry , <i>Azad University</i> , Khorasan Razavi, Iran

Work Experience

- May 2026–Present **AI in Drug Discovery Research Fellow**, *Algoverse AI Research Program*
- Developing an **MDM2–p53 inhibitor design workflow** integrating QSAR potency prediction, reinforcement-learning-based molecular generation, predicted ADMET assessment, docking, and structural evaluation.
 - Building reusable **Python and RDKit pipelines** for molecular data curation, descriptor generation, model training, performance evaluation, and compound ranking.
 - Designing and refining **REINVENT4 reward functions** to balance predicted potency, physicochemical properties, chemical similarity, and predicted ADMET endpoints during iterative molecular optimization.
 - Analyzing generated compounds using physicochemical filters, molecular docking, and **protein–ligand interaction analysis** to compare candidates and guide prioritization.
 - Collaborating with research mentors on computational study design and interpretation of results; prepared and submitted a manuscript to the **AI in Drug Discovery (AI4DD) workshop**.
- Nov 2024–Present **Independent Computational Drug Discovery Researcher**, *Computational Drug Discovery & Molecular Machine Learning*
- Developing reusable **Python and RDKit workflows** for molecular data curation, descriptor and fingerprint generation, feature selection, model validation, and compound prioritization.
 - Built and deployed an **MDM2–p53 compound prioritization application** integrating ChEMBL bioactivity data, XGBoost-based pIC₅₀ prediction, Lipinski filtering, and rule-based recommendations. [GitHub] [Demo]
 - Developed and benchmarked **QSAR/QSPR models** using Random Forest, XGBoost, SVR, molecular fingerprints, physicochemical descriptors, and ChemBERTa-based representations for molecular property and bioactivity prediction.
 - Implemented a **graph neural network for aqueous solubility prediction** using PyTorch Geometric, achieving $R^2 = 0.86$ and MAE = 0.35 on a **random test split**; deployed an interactive prediction application. [GitHub] [Live Demo]
- Aug 2023–Nov 2024 **Scientific Programmer / Computational Chemist**, *Revolution Medicines*, Redwood City, CA
- Developed and evaluated **small-molecule IC₅₀ prediction models** using RDKit descriptors, Random Forest, XGBoost, and SVR, comparing predictive performance against baseline models.
 - Supported **protein–protein interaction modeling** for oncology targets, including protein dataset preparation, structural prediction, interface evaluation, and interpretation of results.
 - Trained **MaSIF-based geometric deep-learning models** on internal protein surface datasets to learn surface fingerprints and identify potential binding sites and therapeutic hotspots.
 - Applied **AlphaPulldown** to bait–candidate screening and all-vs-all interaction prediction, evaluating structural confidence and interface quality using ipTM, DockQ, and PI-score.
 - Executed structural modeling workflows with AlphaFold, RoseTTAFold, AlphaPulldown, MaSIF, and SurfMAP on **Azure HPC**, using Docker, CUDA, SLURM, and Singularity.
 - Collaborated with computational and discovery scientists to interpret model predictions and structural analyses and support therapeutic hypothesis generation.
- Apr 2022–Jul 2023 **Bioinformatics Software Engineer**, *Abcam*, Fremont, CA
- Contributed to antibody discovery by developing bioinformatics and machine learning workflows for PTM detection and antibody sequence analysis.
- Analyzed antibody FASTA sequences to identify PTM-prone motifs associated with stability and therapeutic function
 - Extracted biological features from protein datasets using iFeatureOmega, BioPython, and UniProt annotations.
 - Developed ML models using SVM and Random Forest to predict Asn deamidation-prone sites.
 - Installed, debugged, and maintained reproducible bioinformatics workflows for PTM detection
 - Supported antibody engineering by identifying PTM-prone sites relevant to stability, specificity, and developability.
- Sep 2016–Jan 2019 **Private Physical Chemistry Teacher**, *Independent*, Iran
- Taught general, organic, and physical chemistry, strengthening students' foundational knowledge and problem-solving skills.
 - Prepared students for national chemistry olympiad exams and advanced chemistry coursework.

- Sep 2015–Sep 2016 **Data Analyst**, *Iran Lead & Zinc Industries & Mines Association*, Iran
- Developed statistical and predictive models to support operational efficiency and cost management across mining operations.
 - Analyzed historical sales data and market trends using time series forecasting to support demand prediction and inventory planning.
- Sep 2012–Jan 2015 **Computational Chemist & Thesis Advisor**, *Shahrood University of Technology*, Iran
- Calculated molecular descriptors using Dragon software and selected variables with stepwise regression (SR) and GA-PLS, narrowing down to 12 and 10 key descriptors, respectively.
 - Built predictive models using ANN and SVM models trained on curated descriptor sets to estimate activity coefficients across multiple temperatures.
 - Benchmarked model performance, with SR-ANN achieving the best results ($MSE = 0.0179$, $AAD = 2.43\%$), outperforming GA-ANN, SR-SVM, and GA-SVM.
 - Mentored two undergraduate students in computational chemistry workflows, data analysis, and QSPR/QSAR research projects.

Selected Projects

- MDM2–p53 Generative Molecular Design Integrated QSAR potency prediction, REINVENT4 molecular generation, predicted ADMET assessment, docking, and protein–ligand interaction analysis to prioritize compounds and guide iterative computational optimization. [GitHub]
- Structure-Based Molecular Generation with DiffSBDD Applied the pretrained DiffSBDD model to generate candidate molecules conditioned on the MDM2 binding pocket. Evaluated molecular validity, chemical diversity, and physicochemical properties to support compound prioritization. [GitHub]
- MDM2–p53 Compound Prioritization Built an interactive application combining ChEMBL bioactivity data, RDKit descriptors and fingerprints, XGBoost-based pIC_{50} prediction, physicochemical filtering, and rule-based recommendations. [GitHub] [Demo]
- GNN Solubility Prediction Developed a PyTorch Geometric model for aqueous solubility prediction and deployed an interactive application for molecular input, structure visualization, and prediction. [GitHub] [Live Demo]

Selected Certifications

Introduction to Physics and Machine Learning for ADMET Modeling
MITx, *Machine Learning with Python: From Linear Models to Deep Learning*
Advanced AI in Drug Discovery
Bioinformatics: Docking & Molecular Dynamics Simulation

Manuscripts & Conference Contributions

- 2026 **AI in Drug Discovery (AI4DD) workshop submission.** Manuscript describing an MDM2–p53 molecular design workflow integrating QSAR prediction, reinforcement-learning-based molecular generation, predicted ADMET assessment, and structural evaluation. *Status: Submitted.*
- 2015 H. Nemati, Z. Kalantar, N. Goudarzi, and M. Taherzadeh. “Prediction of activity coefficient at infinite dilution for organic solutes and water in ionic liquid [BMMOR][TCM] using Artificial Neural Network.” 18th Iranian Chemistry Congress, Semnan, Iran.
- 2014 M. Taherzadeh, Z. Kalantar, and N. Goudarzi. “Prediction of activity coefficient at infinite dilution for some organic solutes in ionic liquid of [BMPYR][TCM] by using GA-SVM and SR-SVM.” 17th Iranian Physical Chemistry Conference, Tehran, Iran.
- 2014 M. Taherzadeh, Z. Kalantar, and N. Goudarzi. “Prediction of activity coefficient at infinite dilution for organic solutes in ionic liquid of [BMPYR][TCM] by Artificial Neural Network.” 17th Iranian Physical Chemistry Conference, Tehran, Iran.