Hari **Sharma**

Nashville, TN, 37122, USA

860-975-7070 | sharma.phys@gmail.com | github | linkedin | Google Scholar | U.S. Permanent Resident

Technical Skills

Programming	Python | SQL | BASH, C++
Data Science & Machine Learning	Deep Learning/Neural Networks (TensorFlow | PyTorch | Keras | Image Processing with CNNs) | Unsupervised Learning (K-Means | Clustering | PCA) | Linear & Logistic Regression (Scikit-Learn) | Data Processing & Visualization (Pandas | Matplotlib | Altair | Seaborn) | REST API Development | Big Data with PySpark | Cloud computing with AWS
Modeling Softwares	GROMACS | AMBER | MOE | ROSETTA | RDKit
Visualizing Softwares	VMD | Pymol | MOE
Industry Knowledge	Computational Drug Discovery | Computational Molecular Modeling | Cheminformatics | Structural Biology
Research Methods	Molecular Dynamics simulation (all atom and coarse-grained) | Docking | Comparative Modeling | Virtual High-throughput Screening | Pharmacophore Modeling
Enthusiastic about	Quantum Computing | Large Language Models | Diffusion Models | Geometric Deep Learning

Experience

THE DATA INCUBATOR *Phoenix, AZ*

Data Scientist Certification-Fellowship Program *May 2023*

- Carried out aspect-based sentiment analysis to extract keywords of a restaurant using the reviews provided for the restaurants in Nashville. Built a web application using Streamlit, GitHub and hosted on render.com.
- Web scraped and built a social graph of connections and determined influential people within the group using python NetworkX.
- Developed custom estimators and transformers using python and Scikit Learn to predict venue popularity via pipeline and feature unions. Finally, built an ensemble model combining smaller models to achieve better performance.
- Parsed, cleaned, and processed a 10 GB set of XML files of user actions on a Q&A website. Worked with RDD and DataFrames and implemented a ML pipeline with Spark ML.
- Built and trained series of neural networks using TensorFlow to perform image classification, including a multilayer perceptron, convolutional neural network and a network using transfer learning from a pre-trained model.

StemSynergy Therapeutics INC. *Franklin, TN, USA*

Senior Scientist *Oct. 2021 - Present*

- Performed various methodologies in computational drug discovery for instance ligand docking, virtual high-throughput screening (vHTS), pharmacophore search, binding site prediction etc. using MOE simulation software and subsequently validated the binding pose using molecular dynamics (MD) simulation with GROMACS leading to initial hits for 2 oncology drug development projects.
- Collaborated with experimental chemist and biologist and contributed towards the structure activity relationship (SAR) guided ligand design and suggested mutagenesis studies to be carried out.
- Analyzed the data obtained from MD simulations using self written python scripts.
- Wrote BASH shell scripts for performing fully automated molecular dynamics simulation using GROMACS software suite.
- Administered linux computer cluster nodes with GPU's and CPU's (4 nodes), installaed open source molecular modeling simulation softwares relevant for drug discovery and lead computational part of the research.
- Conducted unsupervised machine learning algorithms using k-means clustering on a custom build computer cluster. Implemented python script so that 10M+ small molecule compounds could efficiently be clustered into 14k+ clusters in less than 30 minutes.

Vanderbilt University *Nashville, TN, USA*

Postdoc Scholar *Sept. 2019 - Sept. 2021*

- Contributed towards the ligand docking algorithm in ROSETTA macromolecular modeling suite to carry out docking of ensembles of congeneric series of ligand simultaneously into protein and their mutant variants.
- Worked on multiple projects involving discovery of allosteric modulators of G-Protein coupled receptor for Muscarinic M5 receptor and the investigation of effect of mutation on the transmembrane domain dimerization of epidermal growth factor receptor kinases.
- Mentored rotating graduate student.

University of Connecticut *Storrs, CT, USA*

Graduate Research Assistant *Aug 2012 - Aug 2019*

- Carried out both all-atom and coarse-grained MD simulations for widely ranging applications from soft matter to biophysics.
- Implemented an algorithm using Graph theory and matrix algebra that is able to compute the end points of a cloud of data points oriented in a flexible contour where data points are dynamic.
- Assembled and administered linux computer cluster nodes with installation of open source molecular modeling software suites.
- Wrote python scripts from scratch to analyze the gigabytes of data obtained from MD simulations.

Education

The Data Incubator

Data Scientist Certification - Fellowship Program *Apr 2023 - May 2023*

University of Connecticut *Storrs, CT, USA*

PhD in Physics *Jan 2015 - Aug. 2019*

- Dissertation Title: Computer Simulations of Amphiphile Based Hybrid Nanostructures.

University of Connecticut *Storrs, CT, USA*

M. S in Physics *Aug 2012 - Dec. 2014*