Amitesh Badkul

@ Email | LinkedIn | GitHub | Website

Education

Birla Institute of Technology and Science, Pilani — Hyderabad, India
Bachelor of Engineering in Electrical and Electronics Engineering — *August 2018 - June 2023*
Master of Science in Chemistry — *Dual Degree Program*

- **Thesis:** In-silico prediction of protein-ligand binding affinity using deep learning. [Link]

Experience

Graduate Center, CUNY — New York City, NY
*Visiting Researcher (Supervisor: **Dr. Lei Xie**)* — *Aug 2022 – Present*

- **Developed** a deep learning framework for predicting chemical-protein binding affinity, incorporating uncertainty. **Achieved** a Pearson correlation of **0.92** and MAE of **0.25** in OOD settings, surpassing state-of-the-art methods.
- **Engineered** and **Integrated** Multi-Task learning-based algorithms with Protein Language Models (DISAE and ESM-2) for enhanced chemical protein binding affinity prediction. Utilized deep learning techniques such as Transformers, RNNs, and CNNs, **resulting in** a Pearson correlation of **0.81** and MAE of **0.56**. [Report (Thesis)]

Arizona State University — Remote
*Summer Research Intern (Supervisor: **Dr. Ashif Iquebal**)* — *Jun 2022 – Aug 2022*

- **Analyzed** large chemical datasets to **identify** the optimal dataset emphasizing compounds with hydrogen bonding.
- **Optimized** a generative model using LSTM and GRU architectures for self-healing compound generation.
- **Trained** a generative model yielding 98.43% valid compounds, with a higher ease of synthesis on average. [Poster]

Birla Institute of Technology and Science Pilani — Hyderabad, India
*Undergraduate Research Assistant (Supervisor: **Dr. Srinivas Prasad K**)* — *Jun 2021 – Jun 2022*

- **Employed** architectures such as MobileNet, ResNet, and EfficientNet to achieve accurate rat pupil measurements.
- **Achieved** minimal deviation from ground truth, outperforming other state-of-the-art image processing algorithms.

*Undergraduate Research Assistant (Supervisor: **Dr. Durba Roy**)* — *Aug 2021 - May 2022*

- **Simulated** a water box cube using Molecular Dynamics for 20 nsec and **analyzed** positional and energy data.
- **Devised** algorithms for Mean Square Displacement (MSD) and Diffusion Coefficient in water systems; **investigated** Rhodobacter Sphaeroides' Reaction Center. [GitHub] [Blog 1] [Blog 2].

*Undergraduate Research Assistant (Supervisor: **Dr. Sudha Radhika**)* — *Feb 2021 - April 2022*

- **Optimized** pretrained models (ResNet, MobileNet, etc) for CXR classification, **achieving 97%** accuracy. **Enhanced** accuracy by **2%** with a CXR enhancement algorithm.
- **Developed** a novel dataset from CXR using statistical descriptors post wavelet transform. **Achieved 97.46%** accuracy with XGBoost and Random Forest models.

Sensordrops Networks, IIT Kharagpur — Kharagpur, India
*Research Intern (Supervisor: **Dr. Sudip Misra**)* — *Dec 2020 – Feb 2021*

- Developed a Graph Neural Networks (GNNs) based algorithm for Contact Tracing of COVID-19 patients.
- Created a novel Twitter dataset for training and testing. Used Twitter metadata as features and deployed the GNN model. Obtained accuracy of **92.31%**.

Million Sparks Foundation — Noida, India
Summer Intern — *Jun 2020 – Aug 2020*

- **Refactored** and **refined** of existing JavaScript code leading to increased efficiency and usability.
- **Developed** and **enhanced** educational content, **benefiting over 20+ teachers** and **elevating** learning outcomes.

Skills

Programming Languages: Python, MATLAB, R, Verilog, LaTeX, C, Bash, Java, Javascript, HTML, CSS

Software Skills: BLAST, NAMD, VMD, OpenCV, EMU8086, LTSpice, Microsoft Office Suite, Adobe Suite, AutoCAD

Languages: Hindi (Native), English (Professional)

Libraries: PyTorch, PyTorch Geometric, TensorFlow, Keras, OpenCV, RDkit, Scikit-Learn, Biopython, Numpy, Pandas, Matplotlib, Seaborn, DeepLabCut, Networkx, DGL, Pywt, Deepchem

Publications

- Tian Cai, Li Xie, Shuo Zhang, Muge Chen, Di He, ***Amitesh Badkul***., ... and Lei Xie. (2022)., "End-to-end Sequence-Structure-Function Meta-learning Predicts Genome-Wide Chemical-Protein Interactions for Dark Proteins", **PLoS Computational Biology** [DOI: 10.1371/journal.pcbi.1010851]
- ***Amitesh Badkul***, Sonakshi Mishra, and Srinivasa Prasad Kommajosyula (2023)., "A Comparitive Study of DeepLabCut and Other Open-Source Pupillometry Data Analysis Algorithms". [Status: Under Review]
- Vyshnavi S K, ***Amitesh Badkul***, Vamsi Inturi, Sudha Radhika., "Comparative Study of DCNN and Image Processing based classification of Chest X-rays for identification of COVID-19 patients using Transfer Learning". [Status: In preparation]

Service

- **Reviewer**, International Journal of Computational Biology and Drug Design (*2023*)
- **Student Member**, IEEE Organization, BITS-Pilani, Hyderabad (*2019-2022*)

Awards & Achievements

- **Scholarship for Higher Education (SHE):** Recipient of Scholarship for Higher Education provided by the Indian Government for excellence in academics Higher Secondary School Board examination, given to the top 1% of students, held in month of March 2018 in India.
- **Scholarship for Practice School - 1 (held in the summer of 2020):** industry exposure program, given to the students with excellent performance (highest grade - 'A') in the industry provided by BITS Pilani.

Projects

CYP3A4 Inhibition Classification | *GitHub*

- **Curated** and **cleaned** the datasets for improved accuracy of machine learning models.
- **Deployed** machine learning algorithms such as Random Forest, and XGBoost. **Achieved** an accuracy of 77%.

CXR Bit Plane Classification | *GitHub*

- **Implemented** CXR classification on bit plane sliced CXRs using MobileNet model.
- **Obtained** a highest accuracy of 95% validating the fact that all layers of a CXR are equally important.

Cdh23EC1 Analysis | *Blog*

- **Conducted** thorough and **extensive** analysis of the data obtained from the MD Simulation of Cdh23EC1 protein.
- **Programmed** functionalities for calculation of various properties - MSD, Radius of Gyration and more.

Research Interests

Computational Sciences: Bioinformatics, Cheminformatics, Molecular Dynamic Simulation

Image Processing: Image Segmentation, Image Enhancement, Computer Vision