Amitesh Badkul

📱 +1 551 404 3996 | @ amitesh.badkul@gmail.com | 🔗 LinkedIn | GitHub | 🌐 Portfolio | 📍 New York City, NY

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 architectures in drug discovery. (*Supervisor:* Dr. Lei Xie)

EXPERIENCE

Graduate Center, CUNY — New York City, NY

Visiting Researcher (Supervisor: **Dr. Lei Xie**) — *Jun 2022 – Present*

- Utilized BLAST to analyze protein sequence similarities for large dataset (used in PortalCG), providing valuable insights into out-of-distribution generalization of PortalCG in predicting chemical protein interactions.
- Successfully implemented a deep learning-based residue-residue contact classification with a BERT-based protein large language model (LLM) DISAE and achieved an accuracy of **98.81%**.
- Implemented a comprehensive pipeline leveraging diverse deep learning algorithms, including Transformers, Bi-directional LSTMs, Attention Mechanisms, and Deep CNNs, to predict protein binding affinity. Integrated LLMs BERT-based DISAE and ESM-2 models into a deep learning architecture for protein binding affinity prediction.
- Modeled Mutli-Task learning-based deep regression algorithms for prediction of chemical protein binding affinity and achieved mean absolute error of **0.43**, Pearson correlation of **0.86**, and Spearman correlation of **0.88**.

Arizona State University — Remote

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

- Performed extensive analysis on large chemical datasets (ChEMBL, ZINC, GDB-13) to identify the optimal dataset for training a deep generative model, with a specific focus on compounds exhibiting hydrogen bonding.
- Designed a generative model leveraging Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures, enhanced with hydrogen-bonding groups, for the generation of self-healing compounds.
- Successfully trained the deep generative model, achieving exceptional results with 98.43% validity and 99.2% uniqueness of the generated compounds. Evaluated the generated compounds using a wide range of metrics, including properties such as physicochemical properties, mechanical properties, and synthesizability, ensuring their applicability and potential for further optimization.

Birla Institute of Technology and Science Pilani — Hyderabad, India

Undergraduate Research Assistant (Supervisor: **Dr. Srinivas Prasad K**) — *Jun 2021 – Jun 2022*

- Designed and developed image and video processing techniques to accurately track and measure the pupil of rat eyes in images and videos using canny edge detection and contour detection.
- Implemented deep learning models, with backbones as MobileNet, EfficientNet, and ResNet, for precise tracking and measurement of the rat eye pupil.
- Performed extensive evaluations and comparisons of a deep learning model's performance, surpassing state-of-the-art image processing methods and achieving superior correlation with human ground truth.

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

- Modelled a water box cube and carried out Molecular Dynamic (MD) simulation for 20 nanoseconds. Performed statistical analysis on the output position, velocity, and energy data obtained.
- Developed algorithm for calculating the Mean Square Displacement and Diffusion Coefficient of the water system. Explored the Reaction Center of Rhodobacter Sphaeroides. [GitHub] [Blog 1] [Blog 2]

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

- Fine-tuned various pretrained models such as ResNet, MobileNet, Xception, and VGG for multi-class classification of Chest X-Ray Scan (CXR) and achieved an accuracy of **97%**. Developed CXR enhancement algorithm using image processing, and improved the accuracy of the previously fine-tuned model by **2%**.
- Took the lead in conceptualizing and implementing a project focused on creation of a novel dataset by extraction of various statistical descriptors after performing wavelet transform on the CXR. Attained an accuracy of **97.46%** on the novel dataset using 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*

- Made significant contributions to the refactoring and cleaning of existing JavaScript code for a website, leading to increased efficiency and usability.
- Successfully developed and enhanced educational materials for for over **20+** teachers, resulting in improved 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, TensorFlow, Keras, OpenCV, RDkit, Scikit-Learn, Biopython, Numpy, Pandas, Matplotlib, Seaborn, DeepLabCut, Networkx, DGL, PyTorch Geometric, Pywt, Deepchem

PUBLICATIONS

1. Cai, T., Xie, L., Zhang, S., Chen, M., He, D., **Badkul, A**., ... and Xie, L. (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]

2. Vyshnavi S K, **Amitesh Badkul**, Vamsi Inturi, Sudha Radhika (2023)., "Comparative Study of DCNN and Image Processing based classification of Chest X-rays for identification of COVID-19 patients using Transfer Learning", **Expert Systems With Applications**. [Status: Under Review]

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.
- Implemented machine learning algorithms like logisitic regression, random forests classifier, and XGBoost classifier on the the curated datasets and obtained 77% accuracy.

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 and for calculation of various properties - Mean Square Displacement, Radius of Gyration, Position Autocorrelation Function and more.

RESEARCH INTERESTS

Computational Sciences: Bioinformatics, Cheminformatics, Molecular Dynamic Simulation

Image Processing: Image Segmentation, Image Enhancement, Computer Vision