Saumya Thakur

AI Architect, Accenture Japan Ltd.

Google Scholar | Homepage | LinkedIn | Email

Education

Indian Institute of Technology Bombay (*IITB*) — *Mumbai, India*
Integrated Master's (BS + MSc) in Chemistry — *(Jul '18 - May '23)*
Cumulative GPA: ***8.25***/10.0

Research Experience

DEEP LEARNING OF KINETIC MODELS
University of California, San Diego — *(May'21 - Feb'22)*
Advisors: Anand Ojha, Prof. Rommie Amaro

- Employed a **Variational Markov Process Net** to sample statistically relevant **($\sim 10^{-3}$ s)** structural transitions
- Obtained free-energy plots with an **increased coverage** using conformations from above net as input to the WE
- Achieved **improved rate constants** than the conventional WE by performing extensive experiments on 2 systems
- **Co-authored** a research paper, wrote detailed documentation for the codes and designed a cover art for the project

DEVISING A NOVEL GENERALITY METRIC
University of British Columbia — *(May'22 - Jul'22)*
Advisors: Isaiah Betinol, Prof. Jolene P. Reid
*Selected for the prestigious and highly competitive Globalink Research Internship fully funded by **MITACS**, Canada*

- Developed a novel **Catalyst Generality metric** and validated it using a biased dataset of 365 reaction datapoints
- Performed **Virtual augmentation** of the dataset by fitting **Random Forest** models obtaining a test R^2 of **0.96**
- **Clustered** the substrate space and structured **5** different end-to-end generality frameworks for different catalysts
- Optimised the obtained Generality values using **MLR** models and validated the metric using **rediscovery** studies

OPTICAL PROPERTY OPTIMISATION
Massachusetts Institute of Technology — *(Jul'22 - Dec '22)*
Advisors: Akshay Subramanian, Prof. Rafael G. Bombarelli

- Developed an algorithm using **RDKit** to obtain exhaustive symmetric products for an improved synthesizability
- Experimented with various **decoder uncertainty cutoffs** to a JT-VAE to improve the band gap optimisation
- Worked on implementing a **Latent Space-based Diffusion model** to perform efficient optimisation of molecules

Professional Experience

Accenture Japan || AI Architect — *(Nov '23 - Present)*

- Integrated **Google Cloud Speech-to-Text** and **Gemini AI APIs** into an internal transcription tool's backend
- Gained extensive end-to-end experience with **AWS** core services **(like SQS, S3, EC2, DynamoDB, Lambda)**
- Optimized the time performance by implementing **multi-threading** and **asynchronous** programming techniques
- Developed a comprehensive framework using **OpenAI APIs** to merge content and style from two distinct webpages
- Created an interactive UI with a user-controlled Chatbot using **Gradio** to demonstrate the real-time style transfer
- Implemented admin APIs for a **Java Springboot** based Railway Ticketing System and performed **JUnit Testing**

Publications and Preprints

- *Ojha, A.A.,* ***Thakur, S.****, Ahn, S.H. and Amaro, R.E.*, 2023. DeepWEST: Deep Learning of Kinetic Models with the Weighted Ensemble Simulation Toolkit for Enhanced Sampling. **Journal of Chemical Theory and Computation**, 19(4), pp.1342-1359.
- *Betinol, I., Lai, J.,* ***Thakur, S.*** *and Reid, J.*, 2023. A Data-Driven Workflow for Assigning and Predicting Generality in Asymmetric Catalysis. **Journal of American Chemical Society**, 145, 23, 12870–12883
- *Karmakar, S.,* ***Thakur, S.****, Jain, A.*, 2024. Can classical mechanics sense conical intersection?. **The Journal of Chemical Physics**. 160(12).

Scholastic Achievements

- Recipient of the prestigious **Inspire-SHE** scholarship by the Department of Science & Technology, GOI *('18)*
- Attended the reputed **Vijyoshi Science Camp** conducted by the Indian Institute of Science, Bangalore *('18)*
- Designed a **supplementary art cover** accepted in the Journal of Chemical Theory & Computation, ACS *('21)*
- Awarded Certificate of **Quantum Excellence** by QiSkit Global Summer School on Quantum Computing *('21)*
- Selected to receive the **MITACS Globalink Research Fellowship** at the University of British Columbia *('22)*

Key Projects

IMAGE GENERATION FROM SCENE GRAPHS | *Advanced ML* *(Jan'22 - Apr'22)*
Prof. Sunita Sarawagi, Department of Computer Science *Course Project*

- Achieved labelled object layout prediction from scene graphs using **Graph Neural Network** (GNN) embeddings
- Employed a **Generative Adversarial Network** for converting the obtained layouts to **photorealistic images**
- Improved **Google's sg2im** architecture by utilising **NVIDIA's** Spatially Adaptive normalisation **(SPADE)** layer

CONICAL INTERSECTIONS: A CLASSICAL INVESTIGATION | *Master's Project* *(Jul'22 - Apr'23)*
Prof. Amber Jain, Department of Chemistry *Course Project*

- Investigated the effects of classical molecular dynamics on a system containing a vibrational **Conical Intersection**
- Contrasted with the state populations observed from semi-classical **Surface Hopping** and pure Quantum Dynamics
- Co-authored a research article on the performed experiments and demonstrated a classical-quantum correspondence

TRANSFER LEARNED MOLECULAR GAN | *Research Project I* *(Jul'21 - Dec'21)*
Prof. Raghavan B. Sunoj, Department of Chemistry *Course Project*

- Implemented a **graph-based** molecular generative adversarial architecture capable of generating novel molecules
- **Fine-tuned** the above network by freezing the lower layers of discriminator followed by training on smaller data
- Successfully achieved **structurally biased molecular generation** from the above transfer-learned GAN network

PERSON DETECTION IN LOW-LIGHT IMAGES | *Introduction to ML* *(Mar'21 - Apr'21)*
Prof. Amit Sethi, Department of Electrical Engineering *Course Project*

- Performed literature review regarding available benchmarks for object detection networks in low-light/NIR condition
- Compared the performance of **R101-FPN** and **X101-FPN** object detection models from FAIR's **Detectron 2**
- Attained an accuracy score of **98.54%** on test dataset from X101-FPN using F1 score evaluation metrics with **IoU**

LEARNING TO RANK IN CASCADE MODEL | *Online Learning* *(Mar'21 - May'21)*
Prof. M. K. Hanawal, Industrial Engineering and Operations *Course Project*

- Formulated the ranking assignment of a cascade model as a stochastic **combinatorial partial monitoring** problem
- Derived gap-dependent upper bounds on the regret obtained for **CascadeUCB** and **CascadeKL-UCB** algorithms
- Evaluated that CascadeKL-UCB outperforms CascadeUCB with a regret value **3 times** smaller than that of latter

MICROSTRUCTURE IMAGE SEGMENTATION | *Materials Informatics* *(Mar'22 - Apr'22)*
Prof. Alankar, Department of Mechanical Engineering *Course Project*

- Executed extensive Gabor and 1st statistic feature extraction on the images followed by **t-SNE** & **PCA** visualisation
- Deployed **K-Means** & **Bayesian Gaussian Mixture Model** clustering algorithms for the structure segmentation
- Performed hyperparameter tuning of BGMM to attain an **entropy** value of **6.12** and validated it on Tin images

YIELD-PREDICTION WITH GRAPHS | *Research Project II* *(Jan'22 - Apr'22)*
Prof. Raghavan B. Sunoj, Department of Chemistry *Course Project*

- Benchmarked the various molecular representations: **String, Graph & Fingerprint**, for prediction of yield values
- Evaluated a **Message Passing Neural Network** for the prediction of yield and obtained a **test R^2** of **0.924**

QUANTUM MACHINE LEARNING | *Institute Technical Council* *(May'21 - Jul'21)*
Maths and Physics Club, IIT Bombay *Summer of Science Project*

- Studied the fundamental operators and algorithms of Quantum Computing like **Deutsch-Jozsa** & **Grover's algo**
- Implemented **Variational Quantum Eigensolver** using QiSkit for calculating the ground state energy of He atom

SOLVING SCHRODINGER'S EQUATION FOR A 1-D POTENTIAL *(May '19 - June '19)*
Advisor: Prof. Amber Jain, Department of Chemistry | IIT Bombay *Summer Project*

- Developed a C++ code using Eigen library to solve the **Schrodinger's equation** for a **1-D protonic potential**
- Implemented **Discrete Variable Representation** and visualised potentials with eigenfunctions using **Matplotlib**

Industrial Internships

GENERATIVE DESIGN MODULE | Palpx Technologies Pvt. Ltd. *(Dec'19)*
Mentor: Mr. Subhash Kunnath *Winter Intern*

- Conducted literature and applicative survey of neural transfer algorithms like **Style Transfer** and **Deep Dream**
- Analysed and developed mathematically defined parametric visualisation of **Julia and Mandelbrot Fractals**
- Studied, assessed and implemented Generative Architectures - **CycleGAN** and **DCGAN** on benchmark datasets

Technical Strengths

Languages	Python, C++, Javascript, Bash, HTML, MATLAB, Fortran, Markdown, Git, & LaTeX
Frameworks	PyTorch, Keras, Tensorflow, RDKit, OpenCV, Pandas, Scikit-learn, SciPy, & Gradio
Softwares	Anaconda, GitHub, Jupyter, WESTPA, SimuLink, AutoCAD, & Cluster Computing
Interests	Generative Architectures, Accelerated Drug Discovery, RL, & Quantum Mechanics

Positions of Responsibility

Institute Student Mentor - *Student Mentor Program | IIT Bombay* *(Jun'21 - Jul'22)*

- Mentored a group of **12** freshmen girls on their institute onboarding procedures, academics and personal growth
- Part of a **130-membered team** playing a facilitative role in the transition of freshmen students to the institute
- Assisted in developing the centralised and department academic blogs to help students make informed choices

Institute Sports Convener - *Institute Sports Council* *(Apr'19 - Apr'20)*

- Part of a **3-tier council** catering to execution and publicity of all the sports activities of IITB to **10k+** students
- Furnished **100+** pre-event and post-event write-ups for FB page of **IIT Bombay Sports** having **16.5k+ likes** and **16.6k+ followers** and ideated **Freshmen Orientation** targeting at an audience of **900+**
- Conceptualized and co-authored news and articles for the biannual Sports magazine **KREEDA** of **IIT Bombay**

Techfest Coordinator - *Asia's Largest Technical Fest* *(Apr'19 - Apr'20)*

- Worked under **Publicity** department in organising and coordinating Techfest '19 with an overall footfall of **1,75,000+**
- Publicized and solicited students for **College Ambassador Program, Techfest** by contacting various institutes
- Identified suitable companies for TechFest merchandise and contacted to deal with them for the fest sponsorships

Department Academic Mentor - *Department of Chemistry* *(June'20 - May'21)*

- Part of a 16-membered intra-departmental team, selected on the basis of overall balanced personality
- Mentoring a pool of **5 sophomores** on various areas pertaining to their academics and general well-being

Courses Undertaken

Chemistry Courses	Programming Techniques in Chemistry, Data Analysis for Chemists, Quantum Chemistry, Computational Chemistry, Electrochemistry, Molecular Spectroscopy
Computer Science	Computer Programming and Utilisation, Online Learning, Advanced Machine Learning, Linear Algebra, Introduction to Machine Learning, Calculus
Other Certifications	Practical High-Performance Computing (edX MIT), Quantum Machine Learning (QiSkit GSS), Building Basic Generative Adversarial Networks (Coursera)
Interdisciplinary	Differential Equations, Engineering Drawing, Economics, Material Informatics

Extracurriculars

- Presented a poster on my publication at the **Student Academic Conference** of the Inter IIT Tech Meet *('23)*
- Represented 2-Mah Engg Regiment of **National Cadet Corps** contingent in the **Republic Day Parade** *('19)*
- Procured **1st** position in the **Product Management** General Championship for a Creddit App Statement *('21)*
- Performed in the Broadway Category in **Annual Insync's Dance Show** in the Convocation Hall, IITB *('20)*
- Counseled BMC school students under a **Career Counseling Campaign** organised by Abhyuday & NCC *('19)*
- Represented the hostel as part of the dramatics team of **Performing Arts Festival-1.0** in the OAT IITB *('19)*
- Volunteered in the **Dialogue With Stakeholders** (DWS) event organised by the SINE & DS Center IITB *('19)*