Vishal Raj

Senior Undergraduate, IIT Kanpur
Department of Chemical Engineering

vishalr23@iitk.ac.in
vishalkaay
Vishal Raj

Academic Qualifications

Year	Degree/Certificate	Institute	CPI/%
2023–Present	B.Tech	Indian Institute of Technology Kanpur	**7.1/10**
2022	CBSE (XII)	SKP Vidya Vihar, Bhagalpur	**95.4%**
2020	ICSE (X)	St. Teresa's School, Bhagalpur	**96.4%**

Work Experience

IoT Security Auditing Framework | Internship, C3iHub, IIT Kanpur | Prof. S Venkatesan *Sep'26–Oct'26*

- Designing and setting up a **secure IoT Security Testbed** to emulate real-world distributed and connected IoT systems and networks.
- Performing **hardware security assessment** through interface identification, **UART/SPI Flash** testing, firmware extraction, and bootloaders.
- Developing automated **reconnaissance and scanning** for Wi-Fi networks, covering discovery, device identification, and security assessment.
- Performing **Wi-Fi security testing**, including deauthentication attacks and Aircrack-ng-based analysis in controlled environments.
- Developing an integrated framework for **device scanning, hardware interface analysis, firmware extraction, and Wi-Fi security auditing**.

Key Projects

Pipelined AES-128 Hardware Implementation | Course Project | Prof. Urbi Chatterjee — CS666 | *Sep'25–Nov'25*

- Designed and implemented a high-throughput **AES-128 encryption core** using an **11-stage outer-round pipelined** architecture on FPGA.
- Enhanced overall performance by **parallelizing key expansion with cipher rounds**, removing critical timing delays and improving throughput.
- Functionally verified the design using standard NIST test vectors through detailed simulation and waveform-level timing analysis results.
- Validated on PYNQ-Z2 (Zynq-7000 SoC) via AXI4-Lite, delivering complete RTL source code, testbenches, and Vivado project files.

Fusion Power Prediction using PINNs | Course Project | Prof. Salman Khan — CHE657 | *Sep'25–Nov'25*

- Developed and compared machine learning models, including a **Physics-Informed Neural Network** and standard neural networks.
- Designed a **custom physics-based loss function** to enforce steady-state fusion power laws for physically meaningful, consistent predictions.
- Preprocessed and normalized **10,000+ experimental runs**, using scaling to manage extreme values and ensure numerical stability.
- Evaluated model performance using error-based metrics, highlighting the **PINN's advantage** in efficient, physically grounded prediction.

Gene Expression Prediction from Histone Modification Signals | Course Project | Prof. Sana Khanum — CHE629 | *Feb'26–Apr'26*

- Predicted binary gene expression (ON/OFF) from **5 histone modification marks** across **100 spatial bins** spanning ±5,000 bp around the transcription start site, on **19,799 genes** of the **REMC E003** human embryonic stem cell dataset.
- Designed and benchmarked a **CNN** against a **hierarchical BiLSTM with dual attention**, using bin-level **spatial attention** and mark-level **combinatorial attention** to model both long-range spatial dependencies and inter-mark correlations.
- Achieved **0.788 test AUC** with the attention model — matching the CNN using **11× fewer parameters** (59K vs 644K) and **5× higher minority-class F1** — closing in on the published **DeepChrome** and **AttentiveChrome** baselines.
- Countered **severe class imbalance** (79.2% inactive genes) with **Focal Loss**, LR scheduling on validation AUC, gradient clipping and early stopping; showed learned **attention weights independently re-discover known epigenetic logic** (TSS focus, H3K4me3 for active, H3K27me3 for repressed) without manual annotation.

2D Poisson Solver: GPU Acceleration | Course Project | Prof. Mahendra Kumar Verma — IDC606 *Jan'26–Apr'26*

- Engineered a high-performance **spectral 2D Poisson equation solver** on an NVIDIA RTX PRO 6000 GPU using CuPy and NVIDIA Warp.
- Achieved an **8,000× runtime speedup** (28.9s to 0.36ms for a 1024×1024 grid) using a **GPU-accelerated Fast Fourier Transform (FFT)**.
- Replaced the vectorized Red-Black SOR baseline with an $O(N^2 \log N)$ **spectral FFT algorithm**, eliminating iterative convergence bottlenecks.
- Integrated **DLPack zero-copy interoperability** for a complete on-GPU data path, achieving **machine-precision accuracy** ($L2 \sim 10^{-16}$).

GIFSleuth: GIF Forensics & Steganography Scanner | Self Project | *Apr'25*

- Built a **zero-dependency Python** GIF security scanner that parses **GIF87a/89a** byte streams and runs **7 detectors** on non-rendered regions.
- Detected **GIFAR** , trailer-appended payloads, content-sniffing **XSS polyglots**, embedded web shells, and **gifshuffle steganography**.
- Combined **19 embedded file signatures** with **Shannon entropy** analysis to detect encrypted or packed payloads and report offsets for verification.
- Reduced false positives by scanning renderer-ignored carrier regions and stripping power-of-two palette padding from LZW image data.
- Designed a **self-registering plugin architecture** with batch scanning, **JSON/CSV/ANSI** reporting, and contractual exit codes for CI pipelines.
- Wrote **143 stdlib unittest cases** with byte-level GIF fixtures and a **GitHub Actions** pipeline across **Python 3.11–3.13**.

Machine Learning Foundations | Mini-Projects | Prof. Rajesh M. Hegde — EE708 *May'25-July'25*

- Completed a semester-long series of weekly assignment mini-projects covering the **full data-science pipeline** — statistical inference, **regression**, **clustering**, **Gaussian Mixture Models**, **decision trees**, and **deep learning** (**CNNs**, **RNNs**, **autoencoders**).
- Implemented core algorithms **from scratch in NumPy**, including a **perceptron** trained by gradient descent with decision-boundary visualisation and a **feedforward neural network** trained via **SGD**; derived **backpropagation** weight updates for **ReLU** MLPs by hand.
- Applied **Pandas**, **Matplotlib** and **scikit-learn** to real datasets for feature engineering, normalisation and binning, least-squares fitting with **residual diagnostics**, **hierarchical clustering** with dendrograms, **vector quantisation**, and **Gini/SSE**-based tree induction.

ML Kickstart | Association for Computing Activities, IIT Kanpur | *May'25–Jun'25*

- Built a machine learning model to classify songs by genre using the **Spotify Songs and their Characteristics dataset**.
- Implemented ML algorithms including linear regression, classification models, and neural networks using **scikit-learn** and **TensorFlow**.
- Performed **exploratory data analysis (EDA)**, feature engineering, and thorough data preprocessing on real-world datasets.
- Gained foundational experience in model training, validation, and interpreting results for **regression and classification** tasks.

Mesoscale and Continuum Simulation | Course Project | Prof. Indranil Saha Dalal — CHE643 *Jun'26–Jul'26*

- Built **NumPy simulators** from scratch; benchmarked **explicit Euler** on **Stokes settling** showing accuracy collapse as $\Delta t \to \tau$ with large timesteps.
- Extracted diffusivity from **multiple-time-origin MSD** fits ($D = \text{slope}/6$) on 10^5-step **3D Langevin** trajectories for reliable transport estimates.
- Handled **periodic boundaries** by tracking wrapped and unwrapped coordinates in parallel, keeping box crossings out of the MSD calculation.

- Ran a 10^7**-step FENE dumbbell** simulation ($\nu = 500$): RMS end-to-end **22.74** vs. theory $\sqrt{\nu} =$ **22.36**, showing strong agreement.
- Mapped **shear-thinning** across **8 shear rates spanning 3.5 decades** via the **Kramers stress tensor** under steady shear flow.
- Implemented **DPD** (81 beads, $\rho = 3$) with **fluctuation–dissipation**-coupled noise; MSD resolved **ballistic-to-diffusive** crossover across timescales.
- Grafted a **5-bead chain into 76 solvent beads**: RMS end-to-end **2.494** vs. Spenley's **2.525** (**1.2%** error), closely matching theory.

MediQ | Mentor, Society of Civil Engineers, IIT Kanpur | *Dec'25–Jan'26*

- **Mentored** ∼**25 students** through end-to-end ML workflow, guiding them from data preprocessing to a deployed disease-diagnosis web application.
- Guided design of a **multi-class disease classification pipeline** using a **Random Forest ensemble (300 estimators)** with balanced class weighting and `sqrt` feature selection for symptom-based diagnosis.
- Directed the ML pipeline build, covering **symptom feature normalization, duplicate-column handling, label encoding, stratified 80:20 train-test splitting**, and automated model evaluation using accuracy and classification reports.
- Oversaw implementation of a **probabilistic inference engine** using `predict_proba()` to rank and visualize **Top-5 differential diagnoses**, with confidence scores and symptom-severity weighted risk scoring.
- Supervised development of a **Flask REST API** for real-time inference, integrated with a structured knowledge base of disease descriptions, precautions, diets, medications, and lifestyle recommendations.

Algorithmic Problem Solving with Advanced Graph Techniques | Self Project | *May'25–Jun'25*

- Implemented graph algorithms including BFS/DFS, shortest paths using Dijkstra, Bellman-Ford, Floyd-Warshall, and minimum spanning trees.
- Applied **Kosaraju's and Tarjan's algorithms** for strongly connected component detection in complex large-scale networks.
- Developed and optimized solvers for Sudoku, maze navigation, and Rubik's Cube using **A* search** and other pathfinding strategies.
- Evaluated and compared algorithm performance through detailed complexity analysis and extensive hands-on benchmarking experiments.

Integrated Single-Train Ammonia Plant | Course Project | Prof. Harshvardhan Katkar — CHE251 *Oct'24–Nov'24*

- Designed a **1500 t/d integrated single-train ammonia plant** flowsheet covering reforming, CO_2 removal, methanation, and synthesis.
- Modeled the process in **Aspen Plus**, computing reaction extents for primary/secondary reforming, shift, and Haber-Bosch synthesis.
- Performed detailed mole and component balances across all **9 process streams**, validating simulation outputs against given plant data.
- Analyzed hydrogen recycle, purge control, and heat integration strategies to improve ammonia yield and overall energy efficiency.

Fluroxypyr Production EHS Analysis | Course Project | CHE261 | *Feb'25–Mar'25*

- Conducted an **Environmental, Health and Safety** analysis for Fluroxypyr production, focusing on hazardous waste streams and safety risks.
- Designed a complete **Zero Liquid Discharge (ZLD)** treatment system incorporating catalyst recovery, evaporation, and crystallization stages.
- Calculated sustainability metrics: **Atom Economy (61.58%)**, **E-Factor (0.624)**, and **EQ-Factor (1.56)** to evaluate green process efficiency.
- Studied and documented exposure limits and disposal regulations for chemicals including **Toluene, NaOH, Ethanol, and TBAI**.
- Co-authored the final technical report including mass balance, regulatory compliance, and safety measures for lab and industrial scale-up.

LSTM Neural Network for Time Series Prediction | Self Project | *May'26-July'26*

- Studied and reproduced a **stacked LSTM architecture in Keras/TensorFlow** for univariate and multivariate time-series forecasting, training on both synthetic sine-wave signals and historical stock-market data across varying forecast horizons.
- Implemented and compared three forecasting strategies—**point-by-point prediction, full-sequence prediction, and multi-sequence (windowed) prediction**—to systematically evaluate trade-offs between short-term accuracy and long-horizon forecast stability under different settings.
- Built the data pipeline with **sliding-window sequence generation and per-window normalization** to handle non-stationary financial price data and prepare multi-dimensional input tensors for robust network training and evaluation.
- Investigated **recursive multi-step forecasting**, identifying and visualizing **prediction drift** as forecast errors accumulated progressively over longer prediction horizons compared against corresponding ground-truth sequences.
- Tuned **network depth, dropout regularization, and sequence length** hyperparameters, and benchmarked training/inference performance across multiple configurations to understand their effect on convergence, generalization, and forecast quality.

Technical Skills

Programming Languages	Frameworks & Libraries	Tools & Softwares
C, C++, Python, Bash, MATLAB, SQL, HTML, CSS, Verilog	NumPy, Pandas, TensorFlow, OpenCV, Matplotlib, scikit-learn	Git, Linux, Jupyter Notebooks, Vim, Xilinx Vivado, LaTeX, Aspen Plus, Raspberry Pi, Arduino

Relevant Courses

(*: Ongoing)

Intelligent Pattern Recognition*	Hardware Security for Internet of Things	Computer Arithmetic on FPGA*
Fundamentals of Computing-I	Machine Learning in Chemical Engineering	Introduction to Electronics
Bayesian Models and Data Analysis	Fundamentals of Data Science and Machine Intelligence	Multivariable Calculus
Computer Methods for Engineers	Mesoscale and Continuum Simulations In Chemical Engineering	Data Structure & Algorithm
AI in Systems Biology	High Performance Computing with Applications	Linear Algebra & ODE

Positions of Responsibility

Senior Executive | Udghosh'24, IIT Kanpur *Oct'24*

- Managed logistics and coordination for **100+ football participants** from multiple colleges, ensuring seamless event operations.
- Oversaw team registrations, match scheduling, and **ground-level execution** to ensure smooth conduct of the tournament.
- Acted as the primary **point of contact** between the organizing committee and team captains for football events and related activities.

Extra-Curricular Activities

- Mentored **20 students** on the MediQ project, leading hands-on sessions in machine learning and web development for an ML web application.
- Served as a cadet in the **National Cadet Corps (NCC)** under the **2 UP Composite Technical Regiment (2UPCTR)**.
- Volunteered at **Antaragni'23**, assisting in the planning, organization, and execution of events during IIT Kanpur's annual cultural fest.
- Built and deployed a static **portfolio website** on GitHub Pages and learned responsive design principles and deployment workflows.
- Passionately pursue music as a hobby by recording and uploading vocal covers of Bollywood songs on a personal **YouTube channel**.