Anuj Gupta

Final Year Undergraduate
B. Tech in Electrical Engineering
Indian Institute of Technology Bombay, India

21d070014@iitb.ac.in
guptaanuj166@gmail.com
+91 8690567330
anujgupta
guptaanuj166

Research Interests

Deep Learning, Computer Vision, NLP, Applied Probability, Optimization, Theoretical Machine Learning

Education

Indian Institute of Technology Bombay *[Nov 2021 – May 2025]*
Bachelor of Technology in Electrical Engineering with Honors
*GPA: 9.49/10 (Department Rank **8th** amongst 200+ students)*
Minor in Artificial Intelligence and Data Science
Minor in Industrial Engineering and Operations Research

Publications

Gouranga Bala*, **Anuj Gupta***, Subrat Kumar Behera, Amit Sethi (2024), "Mitigating Instance-Dependent Label Noise: Integrating Self-Supervised Pretraining with Pseudo-Label Refinement" *arXiv preprint*, arXiv:2412.04898
*Equal contribution as first authors

Research Experience

Noisy Labels Training Improvement *[Jan '24-Present]*
Guide: Prof. Amit Sethi, EE, IITB *Research Project*

- **Bachelor's Thesis Project**
  - Developed an end-to-end pipeline for training deep learning models on datasets with **Intance-dependent Noise (IDN)**
  - Incorporated **contrastive learning** for pre-training, a **stage-wise filtering mechanism** to identify low-loss samples and **consensus-based labelling** for pseudo labels assignment, outperforming many existing methods significantly
  - Implemented a combination of **Self-supervised Learning (SSL)** & **Semi-supervised Learning** to construct a feature extractor and **Supervised Learning along with Data augmentation** to train on the extracted features
  - Achieved **93%** accuracy on **CIFAR-10** with **20% IDN** using **ResNet-18** following **DivideMix** and **SimCLR** for SSL
- **Exploratory Research**
  - Conducted a comprehensive **literature survey** on **noisy labels training**, analysing existing methods and limitations
  - Implemented a **hybrid** approach with **loss correction and sample selection** aimed to improve robustness by **10.5%**

Urban Building Energy Modelling *[Dec '22-Apr '23]*
Guide: Prof. Chirag Deb, Building Science, The University of Sydney *Research Project*

- Analyzed the **energy demand** and **renewable potential** of IIT Bombay campus using **City Energy Analyst** software
- Extracted **building footprints** from satellite images using image segmentation techniques such as **thresholding, edge detection, and clustering**, along with QGIS plugins like **MapFlow and MagicWand**, to analyse the energy potential

Professional Experience

Goldman Sachs | Internship *[May '24-July '24]*
Received a Pre-Placement Offer for exceptional performance and contribution during the Quantitative Analyst internship

- Modelled the **interest rate forward curve** using **Gaussian Process** with **Ornstein-Uhlenbeck** covariance function
- Derived expressions for **basis functions** of the interest rate forward curves, spanning a spectrum from **piecewise constant (infinite degree) to piecewise quadratic** forms by varying a hyperparameter in the covariance function
- Calculated **dollar price and risk impacts** of the new covariance function on various **live and synthetic securities**

Scholastic Achievements

- Awarded a **Department Change** to **Electrical Engineering** (B.Tech) among 20 out of 900+ students *['22]*
- Achieved **All India Rank 607** in **JEE Advanced 2021** examination among **0.22 million** registered candidates *['21]*
- Secured **99.86** percentile in **JEE Main 2021** examination among **1.2 million** registered candidates *['21]*
- Bagged an All India Rank **689** in **KVPY** and awarded with the prestigious scholarship and fellowship by GoI *['20]*
- Scored **100 percent** marks in **Mathematics** and **Physical Education** in class 12th CBSE Boards Examinations *['21]*

Key Technical Projects

Hinglish Sentiment Analysis *[Oct '24-Nov '24]*
Guide: Prof. Pushpak Bhattacharyya, CSE, IITB *Course Project (Speech, NLP and the Web)*

- Developed classifier for **sentiment analysis** on code-mixed social media text using unsupervised multlingual embeddings
- Benchmarked on cross-lingual word embeddings like **VecMap and mBERT**, and monolingual embeddings like **Word2Vec and GloVe** using neural models such as **Long Short-Term Memory (LSTM)** and **Gated Recurrent Unit (GRU)**
- Achieved **62% accuracy** in classifying Hinglish tweets as **positive, neutral, or negative** using LSTM with GloVe

Animal Classification: CNN Implementation *[Mar '23-Apr '23]*
Guide: Prof. Biplab Banerjee, C-MInDS, IITB *Course Project (Introduction to Machine Learning)*

- Achieved **98%** accuracy by implementing a **Convolutional Neural Network** in **tensorflow** to classify animal species
- Constructed end-to-end pipeline for **preprocessing, visualization & prediction** using python libraries like **OpenCV**
- Deployed **7 layers** with multiple filters, **dropout regularization**, **Adam optimizer** and **ReLU activation function**

Undersampled Image Reconstrucion *[Mar '24-Apr '24]*
Guide: Prof. Ajit Rajwade, CSE, IITB *Course Project (Advanced Image Processing)*

- Achieved RMSE of **0.23** in reconstructing the **medical images** from its undersampled **projections** using **MATLAB**
- Solved a **constrained** l_1 **minimization problem** based on a **compressed sensing algorithm**, employing **discrete gradient transforms** as the sparsifying basis and using a **Filtered Backprojection image** as a prior for reconstruction
- Utilized **algebraic reconstruction technique** along with **steepest gradient descent** iteratively to obtain the images

Image Quilting: Texture Synthesis and Transfer *[Oct '23-Nov '23]*
Guide: Prof. Ajit Rajwade, CSE, IITB *Course Project (Digital Image Processing)*

- Implemented a **texture synthesis** algorithm in **MATLAB**, yielding quality textures by stitching small base textures
- Extended the algorithm for **texture transfer**, enabling objects to be rendered with **textures from a source image**
- Transformed **RGB images to YCbCr** for detailed alignment, enhanced correlation through **histogram equalization**, and applied **Image Quilting** to integrate matching source texture blocks by minimizing error across overlapping region

Intelligent Loan Repayment Forecasting *[Dec '22-Jan '23]*
Python Deep Learning Project *Self Project*

- Curated **Kaggle LendingClub dataset**, encompassing the personal, professional and financial details of its members
- Preprocessed and feature engineered the data, **handled null values** and **transformed categorical data** for ML model
- Achieved **94%** accuracy in predicting repayment status using **ANN** of **four dense layers & dropout regularization**

Housing Rent Valuation and Analysis *[Oct '22-Nov '22]*
Guide: Prof. Amit Sethi, EE, IITB *Course Project (Programming for Data Science)*

- Scraped and collected data about various factors affecting housing rent from various resources available on the web
- Performed EDA on various features and compared the rent using **linear, ridge and lasso regression** on our dataset

Computer Vision Project *[May '23-July '23]*
Maths and Physics Club, IITB *Summer of Science*

- Studied the theory for vision problems including the relevance of **hyperparameters, padding & pooling operations**
- Explored the use of Convolutional Neural Networks in **object detection & localisation** and **image classification**

Other Projects

Portfolio Optimisation *[Oct '24-Nov '24]*
Guide: Prof. Vishnu Narayanan, IEOR, IITB *Course Project (Mathematical Optimisation Techniques)*

- Employed **QP, DE, and SLSQP** to optimize **asset allocation** and create efficient portfolios under complex constraints
- Implemented **VaR and CVaR** to assess portfolio risk, focusing on downside and tail risks for comprehensive management
- Applied **Ledoit-Wolf and OAS shrinkage** methods to improve covariance matrix stability in portfolio optimization

Miniature DC Current Controller *[Jan '24-Apr '24]*
Guide: Prof. Siddharth Tallur, EE, IITB *Course Project (Electronic Design Lab)*

- Designed a **high-current source** with **0.1%** precision and control algorithms for stable regulation & spike protection
- Received **Best Project Award** for **easy user accessibility** and designing an effective **consumer deliverable product**

RISC Processor *[Oct '22-Dec '22]*
Guide: Prof. Virendra Singh, EE, IITB *Course Project (Digital Systems)*

- Modelled a **16-bit** processor with **8 register** in **VHDL** using **structural and behavioural modelling** in Quartus
- Designed **hardware flowchart** & implemented components like **Datapath & Controller** for its optimal functionality
- Simulated the controllers' state transition process using a **self-prepared testbench** by running the **RTL Simulation**

Pipelined Processor *[Apr '23-May '23]*
Guide: Prof. Virendra Singh, EE, IITB *Course Project (Microprocessors)*

- Modelled a **16-bit pipelined processor** incorporating 6 stages & 8 registers in **VHDL** using **behavioral modelling**
- Streamlined **hazard mitigation** by implementing **Forwarding mechanism**, eliminating structural and data hazards
- Simulated the designed processor by running **RTL Simulation** of **Modelsim**, validating its operational integrity

Bubble Shooting Game *[Nov '21-Mar '22]*
Guide: Prof. Parag Chaudhuri, CSE, IITB *Course Project (Computer Programming and Utilization)*

- Modelled a **Bubble Shooting Game** using **Codeblocks** in **C++**, implementing **game programming** using a recursive algorithm incorporating concepts of **object oriented programming like classes, loops and structures**
- Simulated the multi-feature game play by utilizing the graphics and canvas capabilities of **Simplecpp graphics**
- Implemented multiple features like **parabolic motion of bubbles, score, health & bullet triggered bubbles split**

Technical Skills

Programming Languages	C++, Python, Embedded C, SQL, VHDL, Assembly
Python Libraries	Tensorflow, Keras, Pytorch, NumPy, Pandas, Matplotlib, Seaborn, Plotly, Cufflinks, Scikit-Learn, OpenCV, NLTK, SciPy, Pyomo
Other Softwares	Git/GitHub, LaTeX, MATLAB, Quartus, Keil, QGIS, NGSpice, GNU Radio, AutoCAD

Key Courses Undertaken

Computer Science	Computer Programming and Utilization, Programming for Data Science, Introduction to Machine Learning, Digital Image Processing, Advanced Image Processing, Speech and Natural Language Processing and the Web*
Probability & Optimisation	Probability & Random Processes, Advanced Probability, Markov Chains & Queuing Systems, Optimization Models*, Mathematical Optimization Techniques*
Mathematics & Physics	Calculus I & II, Linear Algebra, Differential Equations I & II, Complex Analysis, Number Theory and Cryptography, Applied Linear Algebra*, Quantum Physics and Application, Basics of Electricity and Magnetism
Electrical Engineering	Analog Circuits, Digital Systems, Microprocessors, Control Systems, Communication Systems, Power Engineering, EM Waves, Signal Processing, Electronic Devices, Communication Networks, Electronic Design Lab, VLSI Design

*to be completed by Nov'24

Mentorship Experience

PROJECT MENTOR | SUMMER OF SCIENCE *[May '23-July '23]*
Topic: Data Science *Maths and Physics Club, IITB*

- Exercised strong mentorship capabilities while guiding and instructing a team of **6 students** for a **2 months** project
- Fostered a profound and comprehensive understanding of core concepts in **Data Science & Machine Learning** to them
- Guided them in formulating the **Plan of Action** and compiling the **final report** for successful project completion

Extra-curricular Activities

- Volunteered **80+** hours of social service to **Green Campus** initiative by National Service Scheme at IIT Bombay *['21]*
- Participated in an **Electrical based General Championship** as a representative of Hostel 3 in a team of four *['23]*
- Awarded a **certificate** on completing a **Python for data science & machine learning** course issued by Udemy *['23]*