Shrey Ganatra

Research Fellow, Microsoft Research India
Indian Institute of Technology Bombay
ganatrashrey2002@gmail.com – (+91)9106238918 – www.linkedin.com/in/shreyganatra

Education

Indian Institute of Technology Bombay *[2020-2025]*
Master of Technology in *Artificial Intelligence and Data Science* Cummulative GPA: 9.0/10.0
Bachelor of Technology in *Electrical Engineering* with Minor in *Computer Science and Engineering*

Achievements

- Ranked **9th** in **Optiver Trade-a-thon** at IIT Bombay | **14th**(India)-**232th**(Global) in **IMC Prosperity** *[2024]*
- Placed **5th** in **AlphaVerse Premier League** | **4th** in **International Quant Championship** (University) *[2023]*
- One among the **top 30** students in the institute to secure **IDDDP** in the department of **CMInDS**, IITB *[2023]*
- JEE Main: **99.82 %tile** among **1.2M** candidates | JEE Advanced - **98.75 %tile** among **0.16M** aspirants *[2020]*
- Among the top **1%** in **NSEC** out of **49K** candidates and successfully qualified for **INChO** examination *[2019]*
- Secured **West Zone Rank 2** in **Vidyarthi Vigyan Manthan**, largest Talent Search for Future India *[2018]*

Publications

HalluDetect: Detecting, Mitigating, and Benchmarking Hallucinations in Conversational Systems in the Legal Domain Spandan Anaokar, Shrey Ganatra et. al. (EMNLP'25) *[2025]*
Jago Grahak Jago: Consumer Grievance Redressal through Large Langugae Models
Shrey Ganatra et. al. (JUST-NLP Workshop, AACL'25) *[2025]*
Timing Matters: Enhancing User Experience through Temporal Prediction in Smart Homes
Shrey Ganatra*, Spandan Anaokar*, Pushpak Bhattacharyya. (***Equal Contribution**) *[2024]*

Research Experience

Microsoft Research India | Research Fellow *[Jul '25 - Present]*

- Working with the Retrieval team to improve ad relevance, user engagement, and ranking quality at large-scale
- Designing efficient cross-encoders for long-document understanding, focusing on latency, accuracy and scalablity

Consumer Grievance Redressal Chatbot *[Mar '24 - Jun '25]*

*M.Tech Project | Prof. Pushpak Bhattacharyya | Collaboration with **Meta, NLSIU***

- Deployed a chatbot using **LangChain** & **HuggingFace** LLMs helping **1Bn+** people to address consumer claims
- Used non-persistent **FAISS** vector store for **Retrieval-Augmented Generation** & **instructor-xl embedding**
- Reused chat history using **ConversationBufferMemory** in the conversation-chain using **Meta's Llama-3** model
- Exploring techniques like Parameter-Efficient Fine-Tuning **(PEFT)**, **RLHF** and **RAGAS** for model evaluation

Action and Time Recommendation in Smart Homes *[Aug '23 - Mar '24]*

*B.Tech Project | Prof. Pushpak Bhattacharyya | Collaboration with **LG Software India***

- Implemented two-stage **transformer-encoder** architecture with **80%** accuracy in sequential recommendation tasks
- Designed a new task for better insights into human behavior and improved **10%** accuracy over baselines
- Incorporating **Large Language Models (LLMs)** into the framework to leverage their advanced capabilities

Professional Experience

Floworks.ai | NLP Engineer Intern *[Apr '24 - Jul '24]*

*Awarded **Pre-Placement Offer** for exceptional dedication and performance*

- Developed synthetic data pipeline using **GPT-4** and fine-tuned **GPT-3.5** models for next-step prediction task
- Engineered **Chain-of-Thought** and **Few-shot** prompting methods to automate meeting slot extraction from emails
- Integrated email preview functionality into the campaign setup workflow using **TypeScript** enhancing usability

Futures First | Quantitative Analyst Intern *[May '23 - Jul '23]*

*Awarded **Pre-Placement Offer** based on persistence, insights and analytical skills*

- Generated **alphas** for **Energy Market products (CL, HO, NG)** by analysing market **seasonality** and patterns
- Researched various futures trading strategies such as **Spreads, Flies** and **D-flies** to backtest and generate alphas
- Automated the strategy by generating trading signals in Python and sending it **Trading Terminal** for execution

Technical Projects

Automatic Voice Activity and Word Recognition *[Nov '23]*

Course Project | Speech Processing | Prof. Preeti Rao

- Integrated **speech end-pointing** using energy thresholds and applied **pre-emphasis** to enhance higher frequencies
- Extracted Mel-Frequency Cepstral Coefficients **(MFCC)** and trained **GMM-HMM** for voiced activity detection
- Achieved **83.75%** accuracy on clean data and **70.86%** on noisy data by augmenting the dataset with diverse noise

Classification for Patch-Gastric pathological images *[Apr '24]*

Course Project | ML in Healthcare | Prof. Kshitij Jadhav

- Enhanced **few-shot** pathological image classification by text-guided model adaptation through prompt-tuning
- Utilized pre-trained **ViT CLIP** and medical LM **BioLinkBert** to reduce distance in image and text embeddings
- Boosted accuracy levels from **40%** to **60%** when compared to base architecture using **tunable prompt** and text

Hindi-English Speech Recognition *[Apr '24]*

Course Project | Automatic Speech Recognition | Prof. Preethi Jyothi

- Adapted OpenAI's **Whisper**-small **transformer** model to recognize the **code-switched** Hindi-English speech
- Enhanced standard **beam search** decoding for the constraint of ensuring at least one English token in the output
- Achieved Word Error Rate (**WER**) of **0.67** using Beam Search decoding strategy on **0.5 hours** of test utterances

Entity-Level Factual Consistency of Text Summarization *[Apr '23]*

Course Project | Deep Learning for NLP | Prof. Pushpak Bhattacharyya

- Reduced entity-level **hallucination** using a data filtering approach and evaluated using a Question-Answering model
- Optimized the **BART-large** model by fine-tuning it on the filtered **CNN/DailyMail** article-summary dataset.
- Improved the precision score to **0.989** on the filtered dataset as outperforming the baseline score of **0.973**

Deep Learning: Theory to Applications *[Sep '22 - Oct '22]*

***TCS Research** | SysCon, IIT Bombay | Prof. Mayank Baranwal*

- Gained an understanding of the nuances on several aspects of **Deep Learning** and their implementation in **PyTorch**
- Achieved proficiency in **RNNs**, **GANs**, **Neural-ODEs**, **Physics-Informed NNs**, and **Reinforcement Learning**
- Applied algorithms to complex systems like **double pendulum motion**, **generating faces**, house price predictions

Bridging GNNs and MLPs *[Dec '23]*

Course Project | Deep Learning | Prof. P. Balamurugan

- Compared the **generalization** and **representation** capabilities of GNNs vs MLPs using **Propagational MLPs**
- Concluded that **GNNs** outperform **MLPs** due to superior generalization rather than better representation
- Extended PMLPs by adding **parametric message passing** in GNNs enhancing speed in complex graph tasks

Billards Game using Reinforcement Learning *[Dec '23]*

Course Project | Foundations of Intelligent and Learning Agents | Prof. Shivaram Kalyankrishnan

- Implemented multiple regret minimizing algorithms such as **UCB**, **KL-UCB** and **Thompson Sampling**
- Applied **Value Iteration**, **Policy Iteration**, and **Linear Programming** to compute an optimal policy
- Developed RL based algorithm to **play the best action** based on the current state of **billiards** pool table

Positions of Responsibility

Team Leader | Inter-IIT Tech | Brain Quantitative Analyst Challenge *[Nov '23 - Dec '23]*

- Headed a team of **10** students representing IIT Bombay, resulting in a **4th** place finish out of **21** participating IITs.
- Submitted **35** alphas, **20k+** simulations and explored **100+** datasets, automating through **genetic algorithm**

Interview Coordinator | Institute Placement Team *[Dec' 21]*

*Part of the team of **250+** students coordinating for placements of **1800+** students*

- Entrusted and assisted in the task of securing placements for final year students in top global firms/universities
- Ensured smooth conduction by being the **Point-Of-Contact** between students and respective firms for 300+ students

Technical Skills

Programming Languages	Python, C++, VHDL, Assembly, Embedded C
Web Development	HTML, CSS, JavaScript, React, Redux, Next.js, Typescript, Git and GitHub
Data Science Libraries	Numpy, Pandas, Matplotlib, Seaborn, OpenCV, Scikit-Learn, spaCy, PyTorch, Tensorflow, HuggingFace, LangChain, Streamlit
Relevant Courses	Natural Language Processing, Automatic Speech Recognition, Speech Processing, Reinforcement Learning, ML in Healthcare, Optimization, Image Processing

Extracurricular Activities

- One of the founding member of **Quant Community** engaging **2000+** students in research & hackathons *[2024]*
- Actively trading in Indian **Equity Markets** with **20%** profit on **INR 50K** capital in less than 6 months *[2024]*
- One of **18** Teaching Assistants for **Natural Language Processing** course with over **400** students *[2024]*
- Volunteered at **Electronics and Robotics Club** to conduct workshops & events catering **1000+** students *[2022]*
- Successfully completed two **Himalayan** treks: **Bhrigu Lake (14000 ft)** and **Brahmatal Trek (12,250 ft)** *[2023]*