HARJEET SINGH CHAHAL

+1 732-558-3249 | hc1242@scarletmail.rutgers.edu | linkedin.com/in/harjeet-singh-chahal/
github.com/harjeet-chahal | leetcode.com/u/harjeetchahal/

Education

Master of Science in Computer Science — **Aug 2025 – May 2027**
Rutgers University–New Brunswick, New Jersey, USA — GPA: TBD

Bachelor of Technology in Computer Science and Engineering (AI) — **Nov 2020 – June 2024**
Netaji Subhas University of Technology, New Delhi, India — GPA: 3.74 / 4.00

Technical Skills

Programming Languages: Java, Python, C, C++, JavaScript, HTML, CSS, SQL, R
Frameworks & Tools: PyTorch, TensorFlow, Keras, Scikit-Learn, NLTK, Flask, OpenCV, React.js, Node.js, Express.js, FastAPI, Next.js, MongoDB, MySQL, Pandas, NumPy, Matplotlib, Seaborn, SciPy, Jupyter, Google Colab, Git/GitHub, Docker, Kubernetes, GCP, Data Structures & Algorithms, OOP

Work Experience

Associate Business Analyst (Full Time), Dhruv Research, Gurugram, India — Jan 2025 – May 2025

- Processed and analyzed large-scale election survey datasets (India's 2024 national elections and multiple state elections) using SQL and Python to identify key voter behavior trends.
- Automated data cleaning and transformation pipelines, reducing manual effort and turnaround time by 30%.
- Designed dashboards and visual reports in Google Sheets to monitor gender differential and sampling accuracy across districts in several states.
- Collaborated directly with senior analysts, the CEO, and the COO to deliver insights that informed client-facing election forecasts.

Intern – Analytics Team, Dhruv Research, Gurugram, India — Sept 2024 – Dec 2024

- Supported survey data analysis for state assembly elections, working with datasets of 200,000+ responses per state.
- Wrote SQL queries and Python scripts to clean, standardize, and validate survey responses.
- Conducted demographic-based breakdowns (age, gender, ethnicity) to detect and correct sampling distortions at the local level.
- Built automated Google Sheets trackers for state-level vote share estimates, improving team efficiency during field surveys.

Projects

- **PneumoXAI – Trustworthy Medical Diagnostics** — PyTorch, OpenCV, Explainable AI
  - Built a ResNet50 pneumonia classifier with 0.978 AUC and 93.4% sensitivity using patient-wise splits to prevent leakage.
  - Added Grad-CAM, LIME, and SHAP explanations for clinician-facing visual validation of model focus.
  - Measured uncertainty via Monte Carlo Dropout and calibrated outputs with temperature scaling to flag risky predictions.
  - Improved preprocessing (CLAHE) and class-imbalance handling, reducing false negatives by 15%.
- **VideoQuery – Temporal Alignment for Video RAG** — PyTorch, Transformers, LLMs
  - Designed a video RAG pipeline to handle the 10–30s "semantic gap" between events and narration.
  - Implemented timestamp injection to produce grounded citations (e.g., `[[24.5s]]`) for precise retrieval.
  - Aligned Whisper transcripts with CLIP embeddings; ablations showed a 3-chunk sliding window improves reasoning QA.
- **CommentAnalysis – Multi-Label Toxicity Detection** — RoBERTa, FastAPI, Streamlit, Docker
  - Trained and compared baselines vs Bi-LSTM vs RoBERTa for multi-label toxicity detection (0.98 ROC–AUC).
  - Built robustness tests (noise/length shifts) and mitigations to reduce a 10% drop under typos.
  - Improved rare-label detection via loss weighting + per-label threshold tuning ("Threats" +40% F1).
  - Deployed with FastAPI + Streamlit, containerized via Docker Compose.

Academic & Coding Achievements

- **Test Scores:** GRE: 315/340; JEE Main 2020: All India Rank 1513 (among ∼1,000,000 test-takers); JEE Advanced 2020: AIR 5677 (among ∼200,000); JEE Advanced 2019: AIR 6540.
- **Coding Proficiency:** Solved 1300+ problems on LeetCode with an active streak of 1000+ days.