Srishti Garg

✉ srishti.19.garg@gmail.com — 📞 +91-9310206967 — 🔗 LinkedIn — GitHub

Education

IIIT Lucknow
Master of Data Science
Expected Graduation: 2026
CGPA: 8.9

Daulat Ram College, Delhi University
B.Sc. in Mathematics
Graduated: 2023
CGPA: 8.122

Internships

Data Science Intern, Climate Resilience Observatory (CRO), UP Government *Sept 2024 - June 2024*

- **Flood Prediction Project:** Spearheaded the creation of **ML** flood forecasting models with the CRO team, achieving a peak accuracy of **77%** using **XGBoost** for locations along the Ghaghra, Rapti, Saryu, and Gandhak rivers. Addressed data quality issues by translating Hindi datasets and implemented **SMOTE** to tackle class imbalance to predict inundations using flood level data.
- **UP Lightning Tracker: Real-time Monitoring and Future Prediction System:** Designed an automated data pipeline for real-time lightning strike monitoring, extracting JSON data via **Python**, storing in **SQL Server**, and visualizing with **Folium** and **Plotly**. Enabled data-driven insights into lightning patterns, supporting the CRO's goal of an **8-minute** lightning alert system and significant fatality reduction

Data Science Intern, Dawaa Dost *June – August 2024*

- **Optimized Inventory:** Analyzed **79+ CSVs**, optimized stock (20-30% reduction in stockouts).
- **Customer Insights:** Analyzed **79+ datasets**, identified **top 2.5K customers** using **sales, recency, and engagement** metrics.
- **Cohort Analysis & Retention:** Segmented **10M+ customers** into cohorts, analyzed **5-year retention**, and identified **20%+** drop-offs using **SQL, R, and data visualization**.
- **Web Scraping & Automation:** Extracted structured data via **Selenium & Python**, refined **JSON files**, and converted data into **PDF reports** for seamless documentation.

Academic Projects

LSTM Next Word Prediction System 🔗

Tech Stack: TensorFlow (2.15.0), **LSTM**, Pandas, NumPy, scikit-learn, TensorBoard, Matplotlib, Streamlit, scikeras, Pickle

- Developed LSTM model for next-word prediction using **30,000+ words** and **10,000+** training sequences
- Built end-to-end ETL pipeline for data cleaning, **NLP tokenization**, and feature extraction
- Implemented MLOps practices: Early stopping, model checkpointing (.h5), and tokenizer serialization (.pickle)
- Deployed Streamlit web app with sub-second inference and **TensorBoard** visualization

Remote Sensing Data Analysis with STAC API 🔗

- Conducted an in-depth analysis of **Landsat** satellite imagery using the **STAC API**.
- Automated selection of high-quality satellite images with less than 10% cloud cover.
- Used Python libraries like **Rasterio** and **Geopandas** to generate insights into land use and environmental patterns.

Drowsiness Detection System 🔗

Tech Stack: Deep Learning, YOLO, OpenCV, dlib

- Designed and implemented a **drowsiness detection system** using **YOLO**, **OpenCV**, and **dlib**.
- Achieved **94%** training accuracy, **0.97** precision, and **0.975** recall.
- Developed logic to **differentiate between blinking and closing eyes** for real-time **drowsiness alerts**.
- Compared **YOLO** with traditional computer vision methods and documented results.
- Implemented **alert system** using the **mixer** library.

Volunteer Experience

SEEDS NGO *Sep 2024 – Apr 2025*

Tech Stack: Python, Pandas, SQL, Power BI

- Automated data cleaning and **preprocessing workflows**, including **normalization** and validation for disaster relief datasets.
- Built interactive dashboards in **Power BI** for real-time insights.
- Applied **statistical analysis** to support data-driven decision-making.

Position of Responsibility

Placement Coordinator, MSc Data Science

Organized placement drives, coordinated with recruiters, and guided students, demonstrating strong **leadership** and **communication** skills.

Certifications

JAM (AIR 754), Udemy Python, Coursera Data Analytics, R Programming A-Z, Math Summer Program IIITD

Skills

Technical Skills

- **Programming:** Python, R, SQL, C, HTML, LaTeX
- **Data Science:** ML, NLP, Deep Learning, Predictive Modeling, Web Scraping
- **Survey & Field:** CAPI tools, Data Quality (MoSPI), Field Ops, Tech Troubleshooting
- **Reporting & BI:** Report Writing, Dashboarding (Power BI, Tableau), Data Viz, Excel
- **Cloud & Big Data:** AWS, Azure, Spark, Hadoop, Hive, Airflow, Docker, ETL
- **Databases:** MySQL, MS SQL Server
- **Tools:** Jupyter, MLFlow, Git, VS Code, Flask, Mathematica

Leadership and Soft Skills:

Joint Secretary, BSc Mathematics Dept, Vice Head, Finance & Investment Cell, Research Team Member, Vaarta Club, Teaching Assistant