SAMRIDHI GUPTA

Phone:	Mail:	LinkedIn:	Github:	Portfolio:
7839454003	samridhiii.guptaa@gmail.com	linkedin.com/in/samridhiii-gupta	github.com/SamridhiiiGupta	samridhiiigupta.netlify.app

Professional Summary

Machine Learning and **Software Engineering** intern with hands-on experience building end-to-end **ML systems** and **full-stack applications**. Proficient in developing scalable backend APIs with **Flask**, executing **data preprocessing pipelines**, and deploying **NLP**-driven applications using **Hugging Face Transformers**. Strong foundation in **data structures**, **algorithmic thinking**, and **production-level code** practices, with a focus on building performant, real-world solutions.

Education

Pranveer Singh Institute of Technology — Kanpur, India

Bachelor of Technology in Artificial Intelligence & Data Science — *Dec 2022 – Jul 2026*

- **CGPA:** 7.68/10 | **Coursework:** Machine Learning, Deep Learning, NLP, Data Structures & Algorithms, Cloud Computing, Computer Vision, DBMS

Technical Skills

- **Languages:** Python, SQL, C, C++, Java
- **Frameworks/Libraries:** Flask, scikit-learn, pandas, NumPy, matplotlib, seaborn, NLTK, BeautifulSoup, Hugging Face Transformers
- **Tools/Technologies:** Git, GitHub, VS Code, Jupyter Notebook, REST APIs, Linux/Command Line, AWS (SageMaker, Cloud Practitioner fundamentals)
- **Databases:** MySQL
- **AI/ML Technologies:** Machine Learning, Data Preprocessing & Feature Engineering, NLP (Text Summarization & Sentiment Analysis), Model Evaluation, Earth Similarity Index (ESI) Modeling, Semantic Segmentation (AWS SageMaker)
- **Other:** HTML, CSS, Bootstrap, Full-Stack Application Development, Data Visualization, API Integration

Experience

Machine Learning Intern – Certificate of Completion | **ExoHabitAI** | **Live** | **GitHub** | **Video Demo** — Remote

Infosys Springboard — *Feb 2026 – Mar 2026*

- Engineered an end-to-end machine learning pipeline on NASA Exoplanet dataset (39,386 records, 289 features), reducing manual habitability analysis from hours to millisecond-level inference using Python, Pandas, and scikit-learn
- Constructed a custom habitability classification system using astrophysical constraints and physics-based feature engineering (ESI, Stellar Flux), improving model interpretability and achieving 94.6% accuracy, 93.0% F1-score, and 0.971 ROC-AUC with Tuned XGBoost (12.2% gain over baseline)
- Deployed a production-grade ML system via Flask REST API (4 endpoints) with real-time prediction ($<$500ms latency), structured logging, and a Three.js-powered SPA dashboard, enabling scalable ranking of 6,000+ exoplanets.

Projects

MailForge | *Python, FastAPI, LangChain, Groq, LLaMA 3.1, JavaScript* | **Live** | **GitHub** — *Jan 2026 – Feb 2026*

- Produced a full-stack AI application using FastAPI, LangChain, and LLaMA 3.1 (Groq), automating end-to-end cold email generation from job URLs and reducing manual outreach time by $\sim$85%
- Designed a multi-stage AI pipeline (web scraping $\rightarrow$ structured JSON extraction $\rightarrow$ portfolio matching $\rightarrow$ email generation), improving personalization accuracy and enabling context-aware email generation with deterministic LLM outputs (temperature=0)
- Implemented a resilient 3-strategy scraping system (requests, Selenium, WebBaseLoader) with fallback handling and deployed a scalable split architecture (Vercel + Render), ensuring high availability across client-rendered and JS-heavy websites

Recallio | *Python, FastAPI, React, Groq, LLaMA, SQLite, SM-2 Algorithm* | **Live** | **GitHub** — *Dec 2025 – Jan 2026*

- Launched a full-stack learning platform using React, FastAPI, and Groq LLM that converts PDFs into 15–40 high-quality flashcards within 30 seconds, accelerating study content creation by $\sim$90%
- Integrated the SM-2 spaced repetition algorithm with interval-based scheduling (ease factor, interval, retention modeling), improving long-term retention through adaptive review cycles
- Architected a production-grade system with PDF parsing (PyMuPDF), semantic chunking (800–1200 tokens), and LLM-based card generation pipeline, supporting real-time analytics (heatmaps, retention curves) and interactive study sessions

Certifications

- **Semantic Segmentation with Amazon SageMaker** – Optimized image segmentation workflows on SageMaker using computer vision, data labeling, and deployment pipelines.
- **AWS Cloud Practitioner Essentials** – Completed cloud fundamentals training, demonstrating proficiency in cloud architecture, IAM, EC2, S3, and scalable infrastructure design.
- **Career Essentials in Data Analysis by Microsoft and LinkedIn** – Applied data analysis techniques covering cleaning, visualization, and insight generation using Excel and BI tools.
- **Career Essentials in Software Development by Microsoft and LinkedIn** – Built software engineering skills in OOP, version control, debugging, and development workflows.

Achievements

- **Software Engineering Job Simulation** – Simulated real-world tasks applying system design, feature implementation, and deployment in an industry-aligned environment.
- **GenAI: LLMs & Hugging Face** – Developed GenAI solutions with LLMs and Hugging Face — NLP pipelines, prompt engineering, and model integration.
- **Adobe India Hackathon** – Competed in Adobe India Hackathon, delivering innovative solutions under time constraints with a focus on UI/UX.
- **Cybersecurity Job Simulation** – Executed cybersecurity tasks identifying vulnerabilities, applying threat analysis, and implementing secure coding with risk mitigation.