

Sayan Banerjee

sayan112207@gmail.com | 9475979608 | sayan112207.github.io | linkedin.com/in/sayan-18 | github.com/sayan112207

Summary

Computer Science graduate with hands-on experience building scalable AI-driven web and data solutions. Optimized 80+ ETL pipelines, implemented LLM-based RAG systems, and deployed monitoring tools that boost engagement by 15%. Skilled in Python, SQL, and Web technologies to deliver globally reusable solutions that enhance client support and operational efficiency.

Skills

- **Programming Languages:** Python, C++, SQL, JavaScript, HTML, CSS
- **Web & Data Technologies:** REST APIs, Streamlit, Flask, Git, MySQL
- **Cloud & Platforms:** SageMaker, Azure Studio, GitHub Actions, VS Code, Snowflake, Redshift, MongoDB Atlas, SQLite
- **Frameworks & Libraries:** React, Bootstrap, FAISS, OpenVINO, PaddleOCR

Education

- KIIT University**, B.Tech in Computer Science 2021 – 2025
- **CGPA:** 8.88/10.0
 - **Coursework:** Data Structures, DBMS, Algorithm Design, Machine Learning, Cloud Computing, Distributed Systems
- Sri Chaitanya Techno School**, CBSE 2020 – 2021
- **Percentage:** 92.8%
- St. Patrick's High School**, ICSE 2018 – 2019
- **Percentage:** 92.3%

Experience

- Research Analyst Intern**, Hevo Data – Bangalore Jun 2024 – Dec 2024
- Administered **80+ ETL/ELT pipelines** with real-time workflows and connector development via REST APIs and SSH/SSL.
 - Led the design and implementation of a Snowflake Pricing Calculator, increasing MQLs by **25%**.
 - Developed a context-sensitive support chatbot which is **deployed on 50+ blogs**, boosting user retention by 15%.
- Research and Development Intern**, Samsung – Remote Nov 2023 – May 2024
- Achieved **86.3% accuracy** in anomaly detection by analyzing patterns in app policy violations using RoBERTa-Large.
 - Built monitoring systems for **10K+** user interactions, **reducing false positives by 40%**.
 - Implemented a **RAG-based** architecture using LLMs such as **Llama 3.2 1B** to develop scalable systems.

Projects

- Perplexa**, KIIT Jan 2025 – April 2025
- Integrated Google OAuth with MongoDB-backed authentication and Captcha verification, ensuring **99.9% uptime**.
 - Built a FAISS-based RAG with real-time web context, **reducing hallucination by 7%** and maintaining **<150ms latency**.
 - Deployed via SSL-secured GitHub webhooks to Streamlit Cloud, scaling to **500 concurrent queries/min**.
- Vehicle Movement Analysis and Insight Generation**, Intel Unnati May 2024 – Jul 2024
- Leveraged YOLOv8 for real-time vehicle detection and tracking, achieving **<100ms inference latency** and **95% accuracy**.
 - Used PaddleOCR for license plate recognition, processing 60+ frames/min with **92% accuracy** under variable conditions.
 - Optimized edge inference using OpenVINO, realizing a **3x speed boost** for efficient traffic monitoring.
- Text-to-SQL**, KIIT Dec 2023 – Feb 2024
- Fine-tuned Star-Coder2-3b on SQL data using PEFT LoRA, achieving a **28% improvement** in query-to-code accuracy.
 - Leveraged vectorization and token embedding to cut **input processing time by 35%**, in a resource-constrained environment.
 - Applied 4-bit quantization using bitsandbytes, **lowering memory usage by over 20%**.

Certifications

- Intel Unnati Industrial Training**, Intel **IBM Data Science**, Coursera
Machine Learning Specialization, Deeplearning.ai **Certified Data Science Professional**, Oracle

Publications

- Precision Agriculture: Digital Twins with Advanced Crop Recommendation**, IEEE ICOCT 2025
Sayan Banerjee, Aniruddha Mukherjee, Suket Kamboj, DOI: 10.48550/arXiv.2502.04054
- Efficient Waste Collection and Filtration using IOT**, IJSREM 2023
Sayan Banerjee, Rahul Naugariya, Shubham Patel, Shubham Kumar, DOI: 10.55041/IJSREM17403