

# SAYAK BANERJEE

Email: sayakbanerjee022@gmail.com

Phone Number: (412) 844-1988

[LinkedIn/sayak-banerjee/](#)

[Google-Scholar/sayak-banerjee/](#)

## EDUCATION

### Carnegie Mellon University

Master of Computational Data Science (MCDS) | QPA – 4.00/4.00

**Coursework** – Large Language Models, Introduction to Machine Learning, Search Engines, Cloud Computing, Foundations of Computational Data Science, Machine Learning in Production (MLOps)

Pittsburgh, PA

Dec 2025

### Vellore Institute of Technology

B. Tech, Electronics and Communication Engineering | CGPA – 9.41/10

**Relevant Coursework** – Data structures, Algorithms, Operating Systems, Database Systems, Statistics, Applied Linear Algebra

**Recognition** – Departmental Top 10 and Best Outgoing Student from the School of Electronics Engineering (SENSE).

Vellore, India

May 2022

## SKILLS

**Languages & Tools:** Python, C++, PySpark, MySQL, Snowflake, Git, Apache Airflow

**Libraries and Frameworks:** PyTorch, Huggingface, Tensorflow, NumPy, Pandas, SciPy, Scikit Learn, Matplotlib, SpaCy

**Technologies:** Machine Learning, Large Language Models (LLM), Natural Language Processing, Foundation Models, Retrieval-Augmented Generation (RAG), LLM Fine Tuning, LLM In Context Learning, Deep Learning & Neural Networks, Data Engineering/Warehousing, Extract Transform Load (ETL) Pipelines, Azure ML Studio

## EXPERIENCE

### Acuity Knowledge Partners

Associate (Data Engineer)

Senior Analyst

Gurugram, India

Dec 2022 – July 2024

July 2022 – Nov 2022

- Orchestrated and analyzed **alternative financial data** for a US-based Hedge Fund covering various investment portfolios including consumer, healthcare, technology, and e-commerce.
- Spearheaded the integration of Apache Airflow with existing ETL pipelines – creating event-based and time-based triggers, reducing manual efforts to monitor jobs **by more than 60%**.
- Designed and implemented a dynamic data warehouse to store aggregated KPI data from multiple datasets in a SQL database, empowering PMs **to generate financial models by automating the extraction of over 1 million data points**.
- Developed and maintained **robust ETL data pipelines** utilizing Python, PySpark, Pandas, Airflow, AWS, SQL, and Snowflake for seamless data integration.
- Implemented a full-fledged analytics pipeline involving product tagging and KPI visualizations using PySpark, Pandas, and Dash, **achieving ~40% reduction in time taken to generate KPI reports**.

### Bajaj Finance Limited

Data Science and Engineering Intern

Pune, India

Jan 2022 – June 2022

- Developed various customer segmentation and propensity models using machine-learning algorithms for the health insurance business. **Accomplished a ~6% overall increase in effective lead generation of prospective customers**.
- Enhanced search recommendations for health insurance products listed on the website leveraging NLP algorithms **lowering the bounce rate by ~12%**.

### HoloWorld

SDE and ML Intern

Mysuru, India

May 2021 – Oct 2021

- Developed a 3D Pose Animation Pipeline in Unity Engine involving animation of a Y-bot, shadowing movements based on inputs from a pre-recorded or a live-streamed video.
- Collaborated with senior software engineers to develop an Artificial Intelligence-based Yoga Application for Android.

## PROJECTS

### Carnegie Mellon University

Text2Utility – Talk with your Mobile Apps

Pittsburgh, PA

Dec, 2024

- Developed a LLM-based AI agent for end-to-end task automation from natural language inputs.
- It involves intent classification, structured code generation (JSON) followed by LLM tool-use to help user's **complete tasks such as booking an Uber, finding a restaurant table or getting weather updates**.
- Performed full finetuning and PEFT (LoRA) along with in-context few-shot learning to evaluate the performance of various foundation models from Hugging Face.

### GPT Tiny - Decoder only Transformer Architecture

Sept, 2024

- Developed a decoder only transformer model from scratch using Python and PyTorch. Trained the model on the Deep Learning AMI GPU PyTorch 2.0.1 Ubuntu Server using AWS with 16M parameters within a compute budget of 1e+17 FLOPS and integrated it with the Hugging Face API.
- Achieved a validation loss of 3.83 and a validation perplexity of 46.11 on generating new tokens using the model.**

### Vellore Institute of Technology

Real-Time Priority Management of Ambulances at Traffic Intersections

Vellore, India

Dec 2021

- Developed an IOT and CV-based solution to monitor ambulances at traffic intersections, enabling rapid movement and optimized rerouting to reduce wait times by integrating the **YOLO-v3** object-detection with audio tagging algorithms.
- Findings presented at – '2023 IEEE 2nd International Conference on Data, Decision and Systems.' at NITK Surathkal.