# Aviroop Mitra

aviroopmitra5@gmail.com | +91 79809 58951 | India | LinkedIn | GitHub

## PROFESSIONAL EXPERIENCE

**Neurologic-ai | Machine Learning Intern**                                                     *Nov, 2023 – Jul, 2025*

**Intelligent Claim Orchestration in Guideware** | Skan AI                                        *May – Jun, 2025*

Automated claim creation and assignment in **Guidewire ClaimCenter** with a vision-to-reasoning pipeline that pairs **GPT-4 Vision/Text** (**async Python**, prompt-engineered with **token-cost guardrails**) and a **hierarchical scorer** operating on a **30-event window**. Delivered **80% confidence** (**+23 pts over benchmark**) and **halved inference latency**.

**Document Intelligence Platform** | Carelon Health                                              *Mar – May, 2025*

**Streamlit**-based interface drives a PDF-to-image flow where a **layout-transformer** isolates semantic zones; cropped sections are passed to async **GPT-4 Vision** (backed by **Mistral OCR** and **Python imaging**) and rendered with **Pandas**. The generalized architecture now ingests diverse healthcare documents and keeps end-to-end latency under **3s** without sacrificing accuracy.

**PwC India | Advisory Intern**                                                                  *Jun – Jul, 2024*

**Full-Stack Ticket Platform**                                                                   *Jun – Jul, 2024*

**Angular** + **Spring Boot** + **MongoDB** stack with **JWT**-secured role-based access. Provides a real-time ticket dashboard, template creation/search, appraisal & query threads, and an animated UI—while keeping CRUD latency under **50 ms**. An **API Gateway** and **client-side service discovery** route traffic to horizontally duplicated ticket and auth services for effortless scale-out.

## EDUCATION

**Indian Institute of Science (IISc)** | **Master of Technology - MTech in Computer science and engineering**     *Aug, 2025 – Jul, 2027*

**Jadavpur University** | **Bachelor of Engineering - BE (Hons.) in Computer Science and Engineering** | *GPA: 8.25 (76.3%)*     *Dec, 2021 – Jul, 2025*

**Burdwan Municipal High School** | **Higher Secondary in Science** | *GPA: O (93.4%)*            *Jul, 2019 – Mar, 2021*

## PROJECTS

**CVPipeline** | GitHub                                                                          *Jul, 2024 – Aug, 2025*

- Implemented a **6-step** modular pipeline using **linkedin-api**, **openai** and **Playwright** to fetch, transform, AI-enhance, validate URLs, and generate **HTML/PDF**.
- Developed async URL validation with **httpx** achieving sub-**1s** per link across bulk runs and cascading **Google Knowledge Graph**, **LinkedIn**, profile fallback.
- Integrated **SQLite**-based API response caching with **24-hour TTL** to store and auto-clean entries for **LinkedIn**, **PyGithub**, **Google Knowledge Graph** and job search calls.
- Created configurable **HTML** generator and **Playwright**-based PDF generator with centralized styling config for fonts, colors, and responsive layout in scripts/config.py.

**VARC_APP** | GitHub                                                                           *Apr, 2025*

- Developed a **Python Streamlit** application with **Web scraping** modules for The Hindu and The Telegraph and **caching** in data directory.
- Implemented a topic-based probability distribution of **50%** science, **20%** business, **20%** art, and **10%** philosophy to ensure consistent daily article selection.
- Structured code into **modular architecture**: **app/components** for UI, **app/scrapers** for source-specific scrapers, **app/utils** for utilities, and **app/main.py** as the entrypoint.

**distributed-p2p-network** | GitHub                                                            *Mar, 2025*

- Developed distributed text processing platform in **Python 3.8+** with **5 Main Nodes** for coordination, **10 Data Nodes** for chunk-based storage, and **10 Compute Nodes** for embedding and similarity search.
- Implemented **data replication**, **persistence**, **load balancing**, and **real-time health monitoring** in the **Python**-based cluster across Data Nodes (ports **9000-9999**) and Main Nodes (ports **8000-8999**).
- Automated **node orchestration** and **Python virtual environment** setup via **scripts**, and provided frontend at localhost:8501 with API documentation at localhost:8000/docs.

## SKILLS

**Programming Languages:** Python, C++, JavaScript, Java, C

**Databases & Data Storage:** MySQL, MongoDB, SQL

**Frameworks & Platforms:** Streamlit, FastAPI, Spring Boot, Angular, Node.js

**APIs & Libraries:** OpenAI API, API Gateways, REST APIs, pytest, SQLAlchemy, PIL

**Data Science & Machine Learning:** Prompt Engineering, Scikit-Learn, Pandas (Software), NumPy

**Tools & Environments:** Git, Jupyter

## AWARDS & ACHIEVEMENTS

**GATE CS, 2025** | *AIR 65, 99.96 percentile*                                                   *Feb, 2025*

**Senior Scholar** | **Jagadish Bose National Science Talent Search** | *98.98 percentile*        *Jul, 2022*

**Junior Scholar** | **Jagadish Bose National Science Talent Search** | *98.5 percentile*         *Mar, 2020*