Sandipan Paul

+91-7892710285 | psandipan20@gmail.com | LinkedIn | GitHub | Portfolio | LeetCode | Kaggle

EDUCATION

Dr. Ambedkar Institute Of Technology — Bengaluru, India
Bachelor of Engineering - Electrical & Electronics; CGPA: 8.13 — *Dec 2022 – Present*

EXPERIENCE

AI/ML Intern — Jun 2025 – Present
Strydden Technologies LLP. — *Ahmedabad, India — Remote*

- Built a **FastAPI-based clinical AI triage service** using **RAG pipelines and Pinecone vector search** for symptom assessment.
- Improved symptom retrieval relevance using **semantic search and ontology-driven symptom taxonomy**.
- Implemented **stateful session management with MongoDB** to maintain multi-turn triage conversations.
- Integrated **WATI WhatsApp webhooks** to enable real-time AI triage interactions.
- Built a **regulatory compliance assistant using RAG** to answer and compare frameworks such as **HIPAA, CCPA, and SOC2**.
- Implemented **tamper-proof** medical report signing workflows using **SHA-256** hashing and **ECDSA digital signatures.**
- Built a **25**-case automated security test suite for a **medical e-signature pipeline**, achieving **68%** overall code coverage and **95–100%** coverage across cryptographic signing, key rotation, and tamper detection modules.

Robotics Intern — Jul 2024 – May 2025
Just Robotics — *Bengaluru, India*

- Engineered an RC controller using websockets/serial communication; optimized interface latency by **30%**.
- Implemented a lightweight **Flask** backend to process real-time robotic control signals via WebSockets.

PROJECTS

FrameForge | *Python, VideoDB, Redis* | Web — Jan 2026

- Built an **AI** video processing pipeline that converts long-form footage into short clips using **multimodal analysis** and automated scene segmentation.
- Engineered a **fault-tolerant ingestion pipeline** with advanced anti-bot detection and multi-strategy fallback mechanisms using yt-dlp, achieving **90%** successful video downloads
- Implemented intelligent content moderation with automated safety filtering and virality scoring **(0-100)** using **Groq Llama 3**, ensuring brand-safe content generation.
- Deployed **serverless HLS streaming** via VideoDB with **semantic video indexing**, enabling instant playback.
- Instrumented **LLM workflows** using **LangSmith** for tracing, debugging, and evaluation of prompt responses.

Observa | *Python, Gradio, DVC, MLFlow, Grafana* | Code — Oct 2025

- Developed an **end-to-end MLOps pipeline** integrating FastAPI, DVC, MLflow, and Grafana for **real-time NLP monitoring**
- Fine-tuned **DistilBERT** for news classification with custom tracking of inference latency and request metrics.
- Containerized workflow for observability, visualizing latency and confidence trends via **Grafana** dashboards.

Law-Vector | *Python, Flask, Pinecone, Supabase* | Code | Demo — Jun 2025

- Built an AI legal doc search tool using **transformer-based vector embeddings** and Pinecone retrieval
- Implemented user **authentication** with Supabase, ensuring **90% uptime** and securing access.
- Designed a modular vector pipeline using sentence-transformers, achieving **95% retrieval precision**.
- Refactored the **RAG** pipeline to use **Groq-hosted Llama-3.3-70B** instead of Perplexity Sonar, improving cost efficiency and model control.

SKILLS

Languages: Python, C, C++, JavaScript, HTML/CSS
Libraries & Frameworks: Flask, FastAPI, Streamlit, Gradio, TensorFlow, PyTorch, OpenCV, NLTK, LangChain, LangGraph, LangSmith, Mem0 , Postman, SQLModel
Cloud/DevOps: AWS (S3, EC2), GCP, Docker, Kubernetes, Git, GitHub Actions, Airflow, DVC, Dagshub, Redis
Databases: PostgreSQL, MongoDB, SQLite, Pinecone, Qdrant, Firebase
MLOps & Data Tools: MLflow, Grafana, Prometheus, Jupyter Notebook, VideoDB

ACHIEVEMENTS

- Won the **Best Innovation** award for developing a Drone surveillance system using CV and AI (Tech Nexus 2025).
- Won the **Innovator's Edge** award for developing an interactive Braille script system (Hack Colossus 2024).