

ANSH SINGH

B.Tech. Industrial Internet of Things, 2026 | CGPA 8.45 | New Delhi, India
+91-7525993400 | singhanshz172@gmail.com | linkedin.com/in/anshz | github.com/Martian172

SUMMARY

B.Tech. graduate (Industrial Internet of Things, 2026, CGPA 8.45) seeking AI/ML and Generative AI engineering roles. Takes ML systems from model to production: trained a PyTorch autoencoder to 96% accuracy on ECG anomaly detection and served it through a FastAPI backend with a Gemini RAG chatbot. Engineers LLM tooling from scratch — multi-provider pipelines with token and cost tracking and statistical prompt A/B testing — alongside production monitoring with data-drift detection and alerting. Data Analyst Intern at IBM. Authored an IEEE conference paper accepted at ICIC3S-2026. Available immediately for full-time roles.

TECHNICAL SKILLS

Languages: Python, SQL, C++, Java

Generative AI: LangChain, RAG, Prompt Engineering, LLM Evaluation, Prompt A/B Testing, Token & Cost Optimization, OpenAI API, Anthropic API, Gemini API, FAISS (Vector Search)

Machine Learning: PyTorch, TensorFlow, Scikit-learn, Deep Learning, CNNs, Autoencoders, Anomaly Detection, Reinforcement Learning

Backend & MLOps: FastAPI, REST APIs, WebSockets, JWT Authentication, Docker, Docker Compose, GitHub Actions (CI/CD), Git

Data & Analytics: PostgreSQL, MySQL, Pandas, NumPy, Matplotlib, Seaborn, Power BI, Exploratory Data Analysis, Statistical Testing

EXPERIENCE

IBM — Data Analyst Intern, Core Engineering

Jun 2025 – Aug 2025

- Automated end-to-end data cleaning and exploratory data analysis pipelines in Python and SQL across three business datasets, reducing data inconsistencies by 30%.
- Trained and validated a heart-rate anomaly detection model in Scikit-learn, achieving a 91% F1-score on held-out test data.
- Built Power BI dashboards adopted for stakeholder reporting, cutting manual reporting time by 40%.

PROJECTS

CardioSense AI — Real-Time Cardiac Anomaly Detection Platform

github.com/Martian172/cardiosense_ai

- Trained a 1D convolutional autoencoder in PyTorch to detect cardiac anomalies in ECG signals, reaching 96% accuracy on held-out test data.
- Built a FastAPI backend with JWT authentication and WebSocket streaming for live ECG inference, paired with a React 18 frontend.
- Integrated a Gemini-powered RAG chatbot that explains model predictions in plain language to non-technical users.
- Containerized the full stack with Docker Compose and automated build and deployment through GitHub Actions CI/CD.

LLM Chain Forge — LLM Pipeline Framework with Cost Tracking

github.com/Martian172/llm-chain-forge

- Built a framework from scratch for multi-step LLM pipelines with pluggable OpenAI and Anthropic providers, LRU/TTL response caching, and per-call token and cost tracking.
- Shipped a statistical prompt A/B testing harness using the Mann-Whitney U test to compare prompt variants on measurable output quality.
- Included a web playground and CLI; a mock provider makes the full suite offline-testable at zero API cost in CI.

MLOps Sentinel — Production Model Monitoring & Alerting Toolkit

github.com/Martian172/mlops-sentinel

- Implemented data-drift detection using the Kolmogorov-Smirnov test and Population Stability Index, with Slack and email alerting on threshold breach.
- Exposed Prometheus metrics and a live WebSocket dashboard; packaged as a one-command Docker Compose stack with Grafana.

RESEARCH

Physics-Informed Neural Networks with Federated Learning for Microgrid Management

Accepted

ICIC3S-2026, IIT Una — IEEE

- Proposed a physics-constrained federated learning framework enforcing Kirchhoff's laws directly in the loss function, enabling edge training without sharing raw data.
- Delivered a 10x inference speedup (0.019s to 0.002s) and 50%+ faster reinforcement learning convergence while preserving full data privacy.

EDUCATION

B.Tech., Industrial Internet of Things — USAR, GGSIPU, New Delhi | CGPA: 8.45 / 10

Nov 2022 – Jun 2026

Relevant Coursework: Data Structures & Algorithms, Machine Learning, Deep Learning, Reinforcement Learning, DBMS

LEADERSHIP

- Core Organiser, ICAPIE Research Conference — coordinated speakers and paper submissions across 15+ institutions.
- Class Representative, USAR (2023–2026) — sole liaison between 120+ students and faculty on academic scheduling, exam logistics, and grievance escalation.