# Kartik Gupta

kartikgupta1474@gmail.com | https://www.linkedin.com/in/kartik-gupta14 | https://github.com/Kkartik14

## EDUCATION

**University of Petroleum and Energy Studies – Dehradun, Uttarakhand** *August 2022 - June 2026*
Bachelor of Technology, Computer Science (Spl. AI&ML)

## EXPERIENCE

**Swiggy (CREW)**
*Bangalore, Onsite* *Jan 2026 – Present*

- Built backend services for AI-agent workflows and an internal alternative to Langfuse, enabling versioning, debugging, and evaluation; used SQS, Kafka, and Redis for async execution and caching.
- Identified redundant LLM invocations and inefficient multi-turn flows, and refactored orchestration logic to reduce unnecessary turns and optimize context usage, lowering execution cost by ~50%
- Improved response times by restructuring async task execution (introducing worker pools and better concurrency control), reducing latency and speeding up complex request turnaround time.

**Upsurge Labs (Bhindi AI)** **Product Link: bhindi.io**
*Bangalore, Onsite* *July 2025 – Dec 2025*

- Engineered and maintained **150+ AI agents**, building evaluation loops and prompt-refinement pipelines to improve factual accuracy and reduce hallucinations.
- Built and maintained internal dashboards and CI/CD pipelines using Next.js**,** TypeScript**,** and REST APIs, providing real-time visibility into agent performance, analytics, and deployment health across environments.

## PROJECTS

**Tably |** Golang, Next.js, Typescript, MongoDB, chart.js

- Built a formula compilation pipeline supporting 200+ Excel-compatible functions using a custom recursive-descent parser and type-safe evaluator.
- Built an advanced **topological sorting system** that tracks formula dependencies and optimizes recalculation, making it 43x faster when compared to Microsoft excel

**WritePad |** GoLang, React, tiptap, Y.js

- Built a custom WebTransport sync engine using QUIC streams + datagrams to eliminate head-of-line blocking.
- Reduced bandwidth by 82%, achieving 99.8% delivery rate and 17.75ms P50 latency vs WebSockets on high-latency networks.

**VeriStream |** Python, Apache Kafka, Apache Spark, Google-gemini, langchain, Explainable-AI, Knowledge Graph

- Developed a scalable deepfake detection system by fine-tuning DinoV2 and using Whisper AI for transcription and optimized deep learning models, achieving 95% accuracy in identifying manipulated media. Integrated sentiment analysis, fact-checking, and emotional trigger detection by fine-tuning different variants for Bert for the same to provide comprehensive video credibility assessment.
- Leveraged Apache Spark for distributed video processing (10x faster) and Apache Kafka for real-time data streaming (1000+ frames/sec). Designed the system for distributed processing and real-time streaming, with a focus on scalability, fault tolerance, and operational reliability.

## ACHIEVEMENTS

- Research Article on **A royalty framework for copyright protection and accountability in AI-generated art** published at Springer Nature (https://link.springer.com/article/10.1007/s10791-025-09766-0)
- Won Best Research Paper Award for **Reconstructing Bezier Curve Segments in SVG Graphics** (https://ieeexplore.ieee.org/document/11070764 )
- Gold medalist in IBM ICE day for assistive technology for autism
- Winner In Truth Tell Hackathon (World Audio Visual and Entertainment Summit)
- Smart India Hackathon 2023 Finalist