

Krishna Kaushik

Data Scientist

Personal Info

Phone

+91-6900400923

E-mail

k.kaushik25@gmail.com

Links

LinkedIn

<https://www.linkedin.com/in/kaushik-2523/>

AI Projects

<https://github.com/krish-nakaushik25/AI-MiniLabs/>

Portfolio

<https://krishnakaushik25.github.io/>

Technical Skills

Languages & Frameworks: Python

(Advanced), SQL, Bash, PyTorch, FastAPI, Flask, AsyncIO

Backend & APIs: Microservices

Architecture, REST APIs, GraphQL, Asynchronous Programming

Databases: PostgreSQL, Redis, MongoDB, Vector Databases

Cloud & DevOps: Docker, Kubernetes, GCP, AWS, CI/CD Pipelines, Git, Unit Testing

Data Engineering: ETL Pipelines, Spark, Kafka, Large-scale Data Processing, Feature Engineering

AI/ML: LLM Fine-tuning (RLHF/LoRA), RAG Systems, Transformers, Distributed Training

Financial Systems: Real-time Data Pipelines, Performance Optimization, High-throughput Systems

Python Developer with 3.5+ years building high-performance backend systems, data pipelines, and AI/ML applications. Strong experience with REST/GraphQL APIs, asynchronous programming, and microservices development. Delivered \$90M+ business impact through scalable data systems and ML pipelines. Proficient in PostgreSQL, Docker, cloud platforms (GCP/AWS), and distributed computing. Thrives in fast-paced startup environments requiring rapid iteration and production-grade code.

Experience

AI/ML Engineer - Freelance Consultant

Independent Projects & Startup Collaboration / 06/2025 - Present

- **TableRAG Document Intelligence:** Built hybrid RAG system for heterogeneous document processing achieving 85% accuracy improvement over baseline approaches, implemented Python async services with SQL execution engine for multi-hop queries
- **Product Catalog System:** Led entity resolution pipeline canonicalizing 26K ambiguous products using PostgreSQL, reduced duplicate listings by 70% through deterministic matching with sub 100ms latency using small LLMs

Data Scientist III

Walmart Global Tech / 06/2024 - 06/2025

- **Pricing Optimization Framework:** Developed ML system using advanced statistical methods and dietary compliance classifiers, generated \$28M incremental revenue and improved model accuracy from 10% to 60% processing 10M+ GMV
- **LLM Compliance System:** Deployed multi-model dietary classification pipeline with REST APIs, expanded coverage by 80% and reduced tagging errors by 60%, generated \$10M GMV impact
- **Intelligent QA Automation:** Engineered automated testing modules for Smart Creatives platform, reduced manual QA cycles by 60% and improved banner production efficiency by 80%

Data Scientist II

Walmart Global Tech / 07/2022 - 06/2024

- **Full-Stack ML Pipeline:** Architected substitution system with SLMs increasing GMV by \$40M, achieved 250 bps coverage improvement, deployed Python microservices handling 100K+ requests/hour
- **Inference Optimization:** Applied LoRA fine-tuning and dynamic thresholding achieving 70% performance improvement, deployed GCP infrastructure with Docker/Kubernetes serving production traffic
- **Customer Choice Interface:** Launched substitution interface boosting adoption rates by 90%, created high-confidence SKU models with optimized thresholds generating \$10M additional GMV

Education

B.Tech

IIT Guwahati / 07/2018 - 05/2022

- **Relevant Coursework:** Data Structures & Algorithms, Database Systems, Software Engineering
- Top 5% in Amazon ML Challenge with ensemble NLP models

Internships

Software Developer Intern

Publicis Sapient / 05/2021 - 07/2021

- Delivered an inventory analytics app using React, Node.js, and GraphQL, which increased data visibility and reduced restocking latency by 50%.