

# YASH KUMAR SAIN

yashkumarsain66@gmail.com | +919783181471 | linkedin.com/in/sainsama

## EDUCATION

| Year | Degree/Exam                         | Institute     | CGPA/Marks |
|------|-------------------------------------|---------------|------------|
| 2023 | MTech (Dual Degree 5Y) <sup>1</sup> | IIT Kharagpur | 7.89/10    |
| 2023 | BTech (Dual Degree 5Y) <sup>1</sup> | IIT Kharagpur | 7.89/10    |

## EXPERIENCE

**Frontend Tech Lead, Hiccup Dating App** Feb'25 – Oct'25

- Architected production-grade cross-platform Flutter app, gRPC, and protobuf with fully async, event-driven communication
- Built auth with refresh tokens, persistent sessions, offline caching, and seamless recovery across lifecycle
- Integrated Riverpod state management ensuring SOLID compliance and scalable async architecture
- Maintained >90% crash-free rate via Firebase Crashlytics, performance monitoring, 23 analytics events
- Designed post upload with background compression, retry-backed Dio queueing, network-aware state handling
- Built bidirectional encrypted chat via gRPC streams with live delivery/read receipts and offline sync
- Engineered TikTok-style feed with optimized batch querying, lazy media loading, swipe-to-profile transitions
- Implemented Razorpay payments with efficient data reuse across tabs minimizing redundant queries
- Enabled unified profile editing with optimistic updates and single-call synchronization
- Ensured GDPR/CCPA compliance for secure data handling, consent tracking, analytics privacy

**Software Engineer / Full-Stack Developer, Aachaarya AI** Dec'25 – Feb'26

- Architected multi-tenant SaaS with Next.js 14, TypeScript, Supabase, RLS across 13 tables, JSONB schema
- Engineered AI pipeline with Gemini Batch API, Function Calling for structured outputs, context caching (1M+ tokens, 24h TTL) reducing costs 75%
- Implemented RAG with vector embeddings, pgvector similarity search, and document chunking
- Built serverless PDF pipeline with Puppeteer solving Devanagari rendering, bilingual layouts, KaTeX math
- Designed structured output with Gemini Function Calling, TypeScript schemas eliminating JSON parsing failures
- Optimized token economics via context caching (1000-word prompts, 24h), Batch API queueing (50% cost savings)
- Architected vector retrieval with 500-token chunks (50-token overlap), Gemini embeddings, cosine similarity, hybrid filtering
- Built analytics with token logging (cached vs. non-cached), batch monitoring, embedding costs, per-operation attribution

**Software Engineer, Optym** Jul'23 – Feb'25

**Project 1 – WheelsUp Flights Scheduling**

- Optimized MySQL queries in PostgreSQL database reducing redundant multi-endpoint calls for faster response
- Designed complete ERD for documentation and comprehensive knowledge transfer of database structure
- Resolved production bugs across PostgreSQL database, Python backend, MySQL query layer ensuring data accuracy

**Project 2 – GateAI | Southwest Airlines**

- Built backend managing aircraft gate assignments optimizing cost, passenger/crew transfer time, gate prioritization
- Designed scalable APIs with Node.js/TypeScript for event-driven architecture processing ~1M flight events daily
- Integrated AWS Lambda, EventBridge, RDS, SQS, SNS, SES for async workflows and fault-tolerant ingestion
- Developed event models for flight/crew/commodity streams with real-time transformation aligned to DB schema
- Implemented algorithms connecting passengers/crew/commodities with transit calculations and flight connection validations
- Refactored modules adopting CQRS pattern, SOLID principles, functional programming improving maintainability
- Engineered API layer for optimized data retrieval enabling Gate Optimization Engine for real-time scheduling insights

## SKILLS

|                                   |  |  |
|-----------------------------------|--|--|
| Flutter & Production Architecture | gRPC & Protocol Buffers                  | Clean Architecture & SOLID                 |
| Riverpod State Management         | Cross-Platform Development (iOS/Android) | Backend (Node.js, AWS, PostgreSQL)         |
| Async Event-Driven Systems        | Auth & Session Management                | Payment Integration (Razorpay, Stripe)     |
| Realtime Chat & Messaging         | Performance Monitoring (Firebase)        | Data Privacy (GDPR, CCPA)                  |
| UI/UX & Animations                | API Design & Data Modeling               | Database Architecture & Query Optimization |

<sup>1</sup> - Completed a 5-year Dual Degree program at IIT Kharagpur, earning both B.Tech and M.Tech degrees upon graduation in 2023