

SAIF UR REHMAN

Senior Backend Engineer | Python | AI/ML | Cloud Architecture

Toronto, ON, Canada | +1 (647) 539-0940

saif.rehman2498@gmail.com | linkedin.com/in/saif-ur-rehman-1913a7212 | github.com/saifrehman100

Permanent Resident – Authorized to work in Canada

PROFESSIONAL SUMMARY

Senior Backend Engineer (Python/FastAPI) with 3+ years of experience designing and developing cloud-native microservices, recommendation engines, and AI-powered backend systems on Google Cloud Platform (GCP) and AWS. Specialized in building scalable FastAPI/Flask services, event-driven architectures, containerized applications (Docker + Kubernetes), and CI/CD pipelines (Jenkins, GitLab CI, GitHub Actions). Delivered production-grade ML systems including recommendation engines and RAG applications, optimized high-volume APIs, and architected platforms used by thousands of users with 99% uptime.

TECHNICAL SKILLS

Languages: Python, JavaScript, SQL, Bash

Frameworks: FastAPI, Flask, Node.js, SQLAlchemy, Pydantic

Cloud/DevOps: GCP (GKE, Compute Engine, Cloud Storage, BigQuery, Pub/Sub, Cloud Functions, Vertex AI), AWS (Lambda, SQS, SNS, RDS), Docker, Kubernetes, CI/CD (Jenkins, GitLab, GitHub Actions)

Databases: PostgreSQL, MySQL, MongoDB, Redis, DynamoDB

AI/ML: OpenAI, Anthropic Claude, Google Gemini, LangChain, RAG, ChromaDB, YOLO, OCR

Architecture: Microservices, Event-Driven, RESTful APIs, Async Processing, Message Queues, WebSocket

PROFESSIONAL EXPERIENCE

Senior Software Engineer | 10Pearls Pakistan | Jan 2024 – Oct 2025

AI-Powered Appraisal Analysis System (GCP + FastAPI + ML)

- Developed and deployed large-scale FastAPI microservices powering an ML-driven document analysis workflow

- Built GCP architecture using Vertex AI, Compute Engine, Cloud Storage, and Pub/Sub
- Improved processing throughput by 70% via async processing, caching, and concurrency
- Reduced monthly cloud/ops costs by \$4,000 and cut manual processing time by 60–70%
- Ensured 99% uptime and production-grade reliability for high-volume pipelines

Enterprise Integration & Microservices Platform (Python + FastAPI)

- Built a microservices platform integrating 20+ enterprise systems with 50+ REST APIs
- Reduced integration failures by 75% using strict validation, retries, and Redis caching
- Designed high-throughput APIs with connection pooling, async workers, and structured logging
- Maintained 99% uptime under load using Kubernetes-based scaling policies

Event-Driven Workflow Automation (Redis + Celery + Pub/Sub)

- Implemented event-driven pipelines using Redis queues and Celery workers
- Increased system throughput by 250% and eliminated 90% of manual workflows

Database Optimization & Search Systems

- Optimized PostgreSQL indexes and heavy queries using SQLAlchemy and native SQL
 - Designed scalable full-text search APIs powering millions of records
-

Software Engineer | 10Pearls Pakistan | Jan 2022 – Dec 2023

IoT Backend Platform (Python + AWS Serverless)

- Built FastAPI-powered backend for 1000+ IoT devices using Lambda, SQS, SNS, DynamoDB
- Reduced infrastructure costs by 35% through serverless scaling patterns
- Added real-time monitoring with CloudWatch metrics and automated alerting

Real-Time Collaboration Platform (WebSockets + PostgreSQL)

- Engineered multi-tenant WebSocket backend supporting real-time collaboration
 - Implemented RBAC, JWT auth, and optimized SQL schemas for high concurrency
-

FEATURED PROJECTS

Pharmacy Microservices (GCP + Kubernetes)

github.com/saifrehman100/pharmacy-microservices-gcp

- Cloud-native microservices deployed on GKE for Product, Order, Inventory & API Gateway
- Event-driven messaging with Pub/Sub, CI/CD via Cloud Build
- Stack: FastAPI, Kubernetes, GCP (GKE, Cloud SQL), PostgreSQL, Docker

Smart Research Assistant (RAG + Gemini)

github.com/saifrehman100/smart-research-assistant

- ML-based RAG Q&A system using Google Gemini 2.0
 - Semantic search with ChromaDB, async workers with Celery/Redis
 - Deployed using Cloud Run, PostgreSQL, and FastAPI
-

EDUCATION & CERTIFICATIONS

Bachelor of Computer Science | Bahria University | 2018 – 2022

GPA: 3.74/4.0 (Cum Laude)

Certifications:

- Database Design & Query Optimization (10Pearls, 2022)
- Python Data Structures (Coursera, 2020)

Awards: Employee of the Month (2023) • Code Quality Award • AI Hackathon Team Lead

KEY ACHIEVEMENTS

- Built Python/FastAPI microservices at scale on GCP and Kubernetes
- Architected ML and recommendation-style systems with async processing
- Delivered 4+ enterprise AI/ML systems with 99% uptime
- Reduced operational and cloud costs by \$4,000+ monthly through architectural optimizations
- Scaled services to thousands of users and millions of data records