

KHUSWANT RAJPUROHIT

☎ +91-7852045031 ✉ khuswantrajpurohit18@gmail.com 🔗 linkedin 🐙 github 📁 LeetCode 🧑🏻 GeeksForGeeks

Summary

Experienced in building end-to-end AI workflows spanning data ingestion, preprocessing, inference, and deployment. Skilled in LLM systems, OCR + VLM pipelines, and integrating AI into scalable applications with React, Next.js, and Node.js. Proficient in evaluation loops and vector search using PostgreSQL pgvector, with strong problem-solving and production experience.

Internship Experience

AI Intern Nov 25

Kunishe Software Private Limited

- Build a **production OCR + Vision-Language Model (VLM) pipeline** for structured data extraction from complex, varied document layouts handling preprocessing, inference, and post-processing as a unified system.
- Improved extraction accuracy through **iterative data labelling and model fine-tuning**, making the pipeline robust to real-world document noise, skew, and format inconsistencies.
- Integrating the inference pipeline into a **scalable backend service**, bridging the gap between ML model output and production-ready structured data delivery.

Junior Developer Intern June 25 – Aug 25

Ardent Financial Solutions

- Delivered **production features end-to-end** for a live financial platform from design to deployment using **React, Next.js, Node.js, Supabase, and PostgreSQL**, directly serving real users.
- Set up and maintained **CI/CD pipelines** to automate deployments, cutting release friction and resolving production edge cases under real-world load.

Open Source Contributor 2026 – Present

Kubescape (CNCF)

- Contributing to **Kubescape**, a CNCF Kubernetes security project, across multiple repositories including **Kubescape, Kubescape Operator, and Regolibrary**, focusing on security controls, bug fixes, testing, and tooling.
- Serving as a **Triager**, reviewing and categorizing issues, identifying actionable problems, and collaborating with maintainers on issue and PR workflows.
- Developing and debugging **Go** components and **Rego** security policies while working with Kubernetes, Docker, CI, and security scanning infrastructure.

Projects

PDF Invoice Extraction | *PaddleOCR, Flask, Qdrant, React, RapidFuzz* [GitHub] 2025

- Developed a **fully offline invoice parsing pipeline** using PaddleOCR + a custom **fuzzy-matching engine (RapidFuzz)** that maps 200+ OCR-noisy label variants (*qty/qnty/nos/pcs*) to 30+ structured fields.
- Designed a **3-strategy line-item extractor** (multicolumn OCR recovery, regex pattern matching, HY-product fallback) to accurately reconstruct complex invoice tables from noisy OCR output.

AstroAgentAI | *LangGraph, FastAPI, Groq Llama-3.3-70B, ChromaDB, React, TypeScript* [GitHub] 2026

- Engineered a **production-grade agentic AI assistant** using **LangGraph** with a multi-stage workflow for intent classification, enabling reliable multi-step reasoning with up to **6 tool invocations per conversation**.
- Implemented a **Retrieval-Augmented Generation (RAG) pipeline** using **ChromaDB** and curated astrological knowledge, integrated with **Swiss Ephemeris**-based natal chart and a streaming **FastAPI** backend.

Skills

Programming Languages: JavaScript (ES6+), TypeScript, Python, SQL, C++

Frontend: React.js, Next.js, Tailwind CSS, Bootstrap

Backend & Databases: Node.js, Express.js, PostgreSQL, MongoDB, MySQL, SQLite, Prisma, Supabase

AI & Tools: LangChain, LangGraph, Hugging Face, LangSmith, GenAI, NumPy

Additional: Data Structures, Docker, Kubernetes

Achievements & Certifications

- Active open-source contributor in **Hacktoberfest** and **GSSoC (GirlScript Summer of Code)**.
- Mumbai Hacks – GenAI Hackathon (2025): **Cleared Round 1**.
- SIH Internal Hackathon (2024): **Team Lead**, delivered working prototype.

Education

B.Tech in Artificial Intelligence 2024 – 2028 (Ongoing)

Newton School of Technology, Rishihood University

Grade: 7.13 / 10