

Yash Maheshwari

yashmaheshwari8983@gmail.com | linkedin.com/in/yash-maheshwari-3b891a307 | github.com/ApexYash11 | yashmaheshwari.is-a.dev | +91-8128564549

EXPERIENCE

Orydle

Jan 2026 – Present

Backend Engineer Intern

Remote

- Built AI-powered cloud IDE with **LLM-based code completion** using **OpenAI API** and few-shot **prompt engineering**, enabling real-time semantic code understanding and context-aware suggestions.
- Implemented full **RAG pipeline** — chunking, embeddings, vector indexing, reranking, citation grounding — for semantic code search with **hallucination mitigation**.
- Integrated **GitHub API** for repository ingestion and branch-level code indexing, enabling the IDE to reason over entire codebases with automated context retrieval.
- Architected async **FastAPI + PostgreSQL** backend with distributed workflow execution and **WebSocket streaming**; implemented checkpointing and fallback paths for agent reliability.
- Improved backend throughput by **30%** via optimized async design; reduced LLM token cost by **25%** through structured output schemas and prompt optimization.

DevX

Dec 2025 – Jan 2026

AI/ML Intern

Remote

- Developed production **ML forecasting engine** using SARIMAX and Scikit-learn, improving prediction accuracy by **70%** through feature engineering and model optimization.
- Built and evaluated **LLM text generation and summarization** prototypes using **Hugging Face Transformers** for automated business report generation and document understanding.
- Designed **prompt engineering** workflows for GPT-style models; iterated on prompt templates for domain-specific tasks, documenting experiments and evaluation results clearly.
- Built offline **eval sets and regression tests** to catch model drift before deployment; established quality scoring per workflow ensuring reliable **structured outputs**.
- Built scalable **FastAPI** inference pipelines with async endpoints; implemented **Pandas + NumPy** preprocessing reducing data preparation time by **40%**.

PROJECTS

ClaimWise | Python, FastAPI, LangChain, OpenAI API, RAG, Vector DB

GitHub

- **LLM-powered insurance assistant** with full **RAG stack** (chunking, embeddings, reranking, citation grounding); **document AI** pipeline parses PDF policies into structured coverage logic; reduced latency by **60%**.
- Engineered **prompt engineering system** with function calling and golden eval sets for regression testing; supports text extraction, classification, and **summarization** over insurance documents.

Jasper | Python, FastAPI, LangGraph, OpenAI API, Prompt Engineering

GitHub

- Production-grade **autonomous research agent** with multi-step **LLM pipelines**, tool calling, state machines, human-in-the-loop checkpoints, and audit logging; **100%** deterministic execution across test cases.

StockIQ | Python, SARIMAX, FastAPI, Pandas, NumPy

GitHub

- Inventory forecasting with **probabilistic SARIMAX (P10/P50/P90)**, anomaly/fraud detection, multi-warehouse allocation, and explainable **AI decision support**; reduced logistics waste by **20%**.

Wealthify | Python, FastAPI, PostgreSQL, Docker, Scikit-learn

GitHub

- AI-driven budget forecasting achieving **90% accuracy** with anomaly detection; optimized DB queries improving sync performance by **35%**; deployed via Docker with CI/CD.

TECHNICAL SKILLS

Languages: Python, Java, C, SQL

GenAI & LLMs: OpenAI API, Hugging Face Transformers, LangChain, LangGraph, Prompt Engineering, RAG, Text Generation, Summarization, Chatbot Development, Fine-tuning, Structured Outputs, Hallucination Mitigation

ML & Data: TensorFlow, Scikit-learn, SARIMAX, Pandas, NumPy, Feature Engineering, Model Evaluation, Anomaly Detection, Document AI, Data Preprocessing

Backend: FastAPI, RESTful API Design, Async Programming, Microservices, WebSocket Streaming, Distributed Systems

Databases: PostgreSQL, Supabase, Vector Databases (FAISS, Chroma), Query Optimization, Embeddings

Developer Tools: Git, GitHub, GitHub API, Docker, AWS EC2, CI/CD, VS Code, Postman, Logging, Debugging

EDUCATION

Manipal University Jaipur

Jaipur, Rajasthan

B.Tech in Computer Science and Engineering (Data Science)

Expected 2027