

# Dhruvraj Singh Rathore

Bryan, Texas | [dhruvrajrathore2011@gmail.com](mailto:dhruvrajrathore2011@gmail.com) | [dhruvrajsinghrathore.github.io/Portfolio/](https://dhruvrajsinghrathore.github.io/Portfolio/) | [linkedin.com/in/dhruvrajsingh/](https://linkedin.com/in/dhruvrajsingh/) | 7372061179

## SUMMARY

AI Engineer with 2.5 years of professional experience, including multi-agent code generation and RAG pipelines. Delivered internal AI tools adopted by 10+ engineers, automating 60% of development workflows and accelerating client PoC delivery.

## EDUCATION

**MS in Data Science**, Texas A&M University, College Station, Texas, USA | GPA: 4.0

Aug 2024 - Dec 2025

**BS in Computer Science**, SRM Institute of Science and Technology, India | GPA: 3.8

Jul 2018 - May 2022

## TECHNICAL SKILLS

**Programming & Databases:** Python, SQL, PostgreSQL, NoSQL, Vector DBs (FAISS, ChromaDB).

**LLM & ML:** LangChain, LangGraph, LlamaIndex, RAG, Agentic Workflows, MCP, Fine-tuning (LoRA), PyTorch, scikit-learn.

**Data Engineering & Cloud:** Spark, Airflow, AWS (Lambda, S3), FastAPI, Docker, Git, GitHub Actions, CI/CD.

## WORK EXPERIENCE

**AI Engineer Intern**, TechSur Solutions, Reston, Virginia

Jun 2025 - Dec 2025

- Built **multi-agent LLM system** using CrewAI and structured prompt engineering to **generate full-stack code from Jira stories**, including business logic, APIs, and UI components.
- Integrated **safety guardrails** with syntax validation, unit tests, and LLM review, **achieving 90% validation pass rate**.
- Introduced LLM driven diagnostic pipeline in collaboration with backend and test engineers to analyze build failures and apply corrective fixes, **improving success rate from 50% to 80%**.
- **Productionized the system** as FastAPI microservices with GitHub Actions CI/CD, **streamlining 65% of development tasks for 10 developers and accelerating client PoC delivery**.

**Data Engineer**, Draup Business Solutions, Bangalore, India

Dec 2022 - Jun 2024

- Owned Spark big data pipelines, **cutting job runtime by 30%** through efficient partitioning and resource allocation.
- Established data quality and monitoring workflows orchestrated via Apache Airflow, validating 200M+ daily records, **decreasing production issues by 45%**.
- Partnered with consultants and analysts to deploy serverless data pipelines on AWS to transform raw client data into query-ready datasets, **resulting in 50% faster responses to custom data requests**.

**Data Scientist**, HighRadius Corporation, Hyderabad, India

Jul 2022 - Nov 2022

- Developed gradient boosting models to predict expected payment dates and late payment risk, allowing collections teams to prioritize accounts and **enhance cash flow planning**.
- Optimized legacy payment prediction models through GridSearchCV hyperparameter tuning and L2 regularization, **increasing accuracy from 58% to 78%**.

## PROJECTS

**Real-Time IoT Analytics Platform with LLM Query Interface** | [GitHub-Link](#)

Aug 2025 - Dec 2025

- Built dashboard with React/TypeScript frontend and FastAPI backend to track live lab equipment sensor data, **cutting manual monitoring effort by 60% for 10+ researchers**.
- Developed **agentic AI assistant** enabling natural language queries over 1M+ sensor readings to analyze historical trends, equipment performance, and anomalies without coding knowledge.
- Created Model Context Protocol (**MCP**) **server** exposing database schemas to AI agents, enabling generation of accurate SQL queries including aggregations and time-based filtering, **reducing query latency from 3 minutes to 20 seconds**.

**AI System for Employee Policy Violation Detection - Capstone** | [GitHub-Link](#)

Aug 2025 - Nov 2025

- Delivered hierarchical **RAG** system using FAISS to retrieve and rank relevant company policy text for workplace violation classification using AI, **reducing HR manual review time by 50%**.
- Evaluated zero shot and few shot **prompting strategies** using a prompt only **LLM** with a RAG pipeline to establish a baseline for policy violation classification, **achieving 89.1% accuracy**.
- Optimized SaulLM-7B model using LoRA **fine-tuning** on synthetic workplace scenarios with feedback-driven retraining loop, **reaching 92% accuracy**.

**Travel Genie - AI Powered Travel Planning and Recommendation System** | [GitHub-Link](#)

Mar 2025 - May 2025

- Designed intelligent travel itinerary planner leveraging parallel REST APIs to fetch real-time flight, hotel, and attraction data, **saving users hours of trip research time**.
- Integrated BM25 + embedding retrieval to improve relevance and generate personalized, budget-aware itineraries.
- Deployed **RAG**-based LLM workflow generating entire itineraries with cost breakdowns in 30 seconds vs hours manually.