

# RISHABH SHARMA

AI ML Engineer

+91 9131494753

rishabhksharma19@gmail.com

www.linkedin.com/in/rishabhssharma

---

## PROFESSIONAL SUMMARY

Oracle Certified GenAI Engineer with around 2 years of experience building end-to-end intelligent systems across Agentic AI, GenAI, Deep Learning, NLP, Computer Vision, IoT, and predictive analytics. Skilled in developing scalable AI/ML solutions using Python, Transformers, Azure, GCP, and modern MLOps workflows. Experienced in training and fine-tuning advanced deep learning models—including LLMs, multimodal architectures, and vision models—for real-world applications such as image/video analysis and autonomous AI agents. Proven track record of delivering high-impact AI solutions in both independent roles and cross-functional team environments.

---

## TECHNICAL SKILLS

1. **AI/ML & Data Science:** Machine Learning, Deep Learning, Transformers, Time Series, Predictive Modeling, Data Visualization, Computer Vision, CNN, DCNN, RNN/LSTM
  2. **GenAI & Agentic AI:** RAG, Graph RAG, LLM Fine-Tuning, LangGraph, CrewAI, Azure AI Foundry, Prompt Engineering
  3. **NLP & Text Analytics:** Text Classification, NER, Similarity Search, STT/TTS, Document Understanding
  4. **Cloud & MLOps:** Azure (AI Search, ML, VMs), GCP (Gemini), AWS SageMaker, MLflow, Docker
  5. **Databases:** SQL, PostgreSQL, AuraDB, MongoDB, Pinecone, Neo4j (Knowledge Graphs)
  6. **Programming & Frameworks:** Python, FastAPI, PyTorch, JAVA (J2SE) , HTML CSS
  7. **Tools & Platforms:** Git, Bitbucket, Jira, Postman, Power BI
- 

## SOFT SKILLS

Problem Solving, Research & Development, Project Leadership, Cross-Functional Collaboration, Strong Communication, Documentation, Mentoring

---

## CERTIFICATIONS

1. Oracle Certified GenAI Professional
  2. Oracle Certified Vector Database and Search Professional
  3. Oracle Certified Data Science Professional
  4. Oracle Certified AI foundation Associate
  5. AWS Cloud Essentials
  6. TCS iON Career Edge
  7. Microsoft Applied Skills : Accelerate app development using Github Copilot
  8. Microsoft Applied Skills : Build a natural language solution using Azure AI Language
  9. AWS Cloud Practitioner Essentials
  10. Career Essentials in Data Analytics By Microsoft and Linkdein
  11. Microsoft - Computer Vision Solution with Azure AI Vision
  12. Linkdein Learning - Github Professional Certificate
  13. Microsoft - Vision Models with Azure AI custom vision
  14. Microsoft - AZ400 - Development for enterprise DevOps
  15. AWS - Building a Machine Learning ready Organization
- and other 90 + certifications
-

## **PROFESSIONAL EXPERIENCE**

- **AI ML Engineer**

2024 - Present

### **InfoBeans Technologies**

- Led and contributed to large-scale data science projects, working with datasets containing millions of rows to generate actionable insights.
- Built multimodal machine learning solutions for high-volume data, delivering impactful analytics for supply chain management, healthcare management, and related domains.
- Engineered computer vision systems for the infrastructure industry, including advanced image and video processing pipelines.
- Delivered a fully independent solution integrating Azure AI Search, vector search (RAG), Neo4j Knowledge Graphs, and OData metadata filtering, enabling intelligent contextual retrieval.
- Contributed as both a team collaborator on multiple client projects and a project lead for several high-impact in-house AI initiatives.
- Developed end-to-end ML pipelines using Azure MLOps, covering model development, deployment, monitoring, and automation.
- Fine-tuned multiple LLMs using Azure and Ollama, optimizing them for domain-specific tasks and production usage.
- Trained and fine-tuned multiple YOLO and video-based models for object detection, tracking, and real-time inference using Deep learning.
- Delivered agentic AI systems using CrewAI, LangGraph, n8n, and Azure AI Foundry, enabling autonomous workflows and intelligent multi-agent orchestration.
- Built and deployed data science workflows on AWS SageMaker for scalable training and experimentation.
- Individually contributed to AI-driven IoT projects, integrating ML models with edge devices and real-time sensor data.

---

- **AI ML Intern**

Feb 2024 - Jun 2024

### **InfoBeans Technologies**

- Assisted in designing and developing machine learning and deep learning models for real-world business use cases across NLP, computer vision, and predictive analytics.
- Worked as a shadow contributor on production-grade AI/ML projects, gaining hands-on exposure to industry best practices, scalable architectures, and client-facing workflows.
- Supported senior engineers in building data pipelines, performing exploratory data analysis, and preparing large datasets for model training and evaluation.
- Contributed to the development of LLM-based and GenAI prototypes, including text classification, summarization, and retrieval-augmented workflows.
- Implemented and tested computer vision modules, including object detection, image preprocessing, and video analytics, using OpenCV, TensorFlow, and PyTorch.
- Built small-scale AI/ML solutions that delivered measurable impact, including automation scripts, prediction models, and proof-of-concept demos.
- Assisted in training, fine-tuning, and validating models, performing hyperparameter tuning, and optimizing performance metrics.

- Supported cloud-based development using Azure / AWS, including model deployment, notebook experimentation, and data storage operations.
  - Collaborated with cross-functional teams to document workflows, prepare reports, and present insights to mentors and project leads.
  - Gained experience with version control, CI/CD practices, and MLOps fundamentals while supporting end-to-end lifecycle tasks.
- 

- **Student Intern**

2024 - Present

**InfoBeans Foundation**

- Gained hands-on experience in full-stack web application development using Java, MERN stack (MongoDB, Express.js, React, Node.js), and modern web technologies.
- Built multiple web applications applying core concepts of HTML, CSS, JavaScript, responsive design, and UI/UX fundamentals.
- Developed backend services using Java, Node.js, and REST API principles, including routing, authentication, and database integration.
- Worked extensively with MongoDB and SQL databases (MySQL/PostgreSQL) for data modeling, schema design, CRUD operations, joins, indexing, and query optimization.
- Practiced modular coding, clean architecture, and industry best practices for scalable and maintainable web development.
- Collaborated in team-based mini-projects, following Git/GitHub workflows for version control and code reviews.
- Strengthened problem-solving, debugging, and deployment skills through guided assignments and real-world project simulations.
- Gained foundational exposure to software engineering principles, SDLC methodologies, and agile teamwork practices.

---

- **Projects**

1. **AI Powered Cognitive Search System (RAG+GraphRAG+ OmniRAG)**

CSA Group

**Tools & Technologies**

Agentic AI, GenAI, Azure Cloud, Azure Vector Search, Knowledge Graph (Neo4j), Advanced RAG, Graph RAG, OData Metadata Filtering, Document Analysis, LangGraph, Databricks MLflow, Azure App Insights

**Description**

Individually Built an AI-powered cognitive search platform processing 1M+ multimodal documents, integrating Azure Vector Search, Traditional RAG + Graph RAG, Neo4j Knowledge Graph, and OData metadata filtering for high-accuracy retrieval. Enabled natural-language and structured search with advanced indexing and LLM-driven query understanding. Delivered seamless integration with Salesforce and MuleSoft, and utilized Databricks MLflow and Azure App Insights for tracking and monitoring.

---

## 2. AI-Driven Audit Management System

CSA Group

### **Tools & Technologies**

AI Modeling, Machine Learning, Deep Learning, Embeddings, Streaming Pipelines, Workflow Automation, RAG, Graph RAG, OpenAI, Agentic AI, NLP, Knowledge Graphs (KG), Azure AI Search, Azure Blob Storage, Azure App Insights, Audit Analysis, Vector Search, Document Intelligence, Python, REST APIs

### **Description**

Delivered the intelligent assessor assignment module for the CSA Group's large-scale MDR-based audit management system. Designed dynamic matching algorithms using agent-based reasoning, embeddings, NLP, and ML models to evaluate audit requirements, assessor competencies, availability, historical performance patterns, and interpersonal compatibility through sentiment and behavior analysis. Implemented RAG + Graph RAG pipelines, enterprise Knowledge Graphs, and Azure AI Search to enable context-aware retrieval and smarter audit decision-making. Utilized Azure Blob Storage for scalable data ingestion and Azure App Insights for end-to-end system monitoring and telemetry. The solution enabled optimized, fair, and accurate audit team formation, significantly enhancing operational efficiency and decision-making across the CSA Group's audit lifecycle.

---

## 3. AI-Based Tender & RFP Management System

International SOS

### **Tools & Technologies**

Automation, Data Extraction, NLP, OpenAI, Pinecone Vector DB, Azure Cloud, Azure AI Services, Docker, Document Intelligence, ML Pipelines

### **Description**

Developed an end-to-end AI-powered Tender & RFP Management System to automate the evaluation and response creation workflow. Built intelligent parsing pipelines to extract eligibility criteria, critical clauses, financial requirements, and deadlines from large tender/RFP documents using NLP and document extraction techniques. Integrated Pinecone vector search and OpenAI-based retrieval workflows to match extracted requirements with internal business capabilities, generating real-time eligibility scores and risk insights.

Implemented automated form population and generated human-review-ready draft proposals, reducing turnaround time from 2–3 days to just a few hours. Containerized the solution using Docker and deployed scalable components on Azure, ensuring high availability and seamless processing for large volumes of tender documents.

---

## 4. AI-Driven Automated Interview Bot System

InfoBeans Technologies

### **Tools & Technologies**

11Labs (TTS), Google Speech-to-Text (STT), Gemini AI Models, Python, FastAPI, Azure Virtual Machines, WebSockets, NLP, Prompt Engineering, Automation, Scheduling APIs

### **Description**

Built an AI-driven automated interview bot that analyzes job descriptions, schedules interviews, and autonomously joins meetings to conduct real-time candidate evaluations. Integrated Google STT, 11Labs TTS, and Gemini AI for adaptive questioning and live conversation. Developed using Python, FastAPI, and Azure VMs, the system auto-generates detailed candidate feedback for HR. Achieved ~90% reduction in developer interview time, significantly accelerating the hiring process.

---

## **5. AI-Based Manufacturing Warehouse Safety & Disaster Prevention System**

ISOS

### **Tools & Technologies**

Computer Vision, CNN, DCNN, YOLO-based models, 3D CNNs, Optical flow CNNs, CCTV Streaming, Real-Time Object & Motion Detection, Python, OpenCV, Edge Deployment, Alert Automation

### **Description**

Developed a computer-vision–driven warehouse safety and disaster prevention system using a fine-tuned YOLOX model for real-time object detection, hazard identification, and motion tracking. Integrated live CCTV video ingestion to monitor worker safety, detect unsafe activities, and trigger instant alerts for high-risk events. Enabled proactive safety enforcement and early disaster prevention across manufacturing warehouse environments.

---

## **6. Infobeans.ai Intelligent Chatbot (RAG + Graph RAG)**

InfoBeans Technologies

### **Tools & Technologies**

RAG, Graph RAG, Azure AI Search, Neo4j Knowledge Graph, Vector Embeddings, NLP, Python, FastAPI, Web Integration, Metadata Filtering

### **Description**

Individually developed a RAG + Graph RAG–powered intelligent chatbot for the Infobeans.ai homepage. The system delivers accurate, real-time information about the company, projects, services, and internal processes by combining vector search with knowledge graph reasoning. Implemented intelligent metadata filtering to interpret vague or ambiguous user queries and return precise, context-aware responses. Successfully deployed as a production-grade assistant enhancing user engagement and information discovery.

---

## **7. Expona – Multimodal RAG-Based Document Intelligence System**

InfoBeans Technologies

### **Tools & Technologies**

RAG, GPT-4, GPT-4o Mini, CNN/ViT, Pinecone Vector DB, Computer Vision, Azure Cloud, Docker, Embeddings, NLP

### **Description**

Designed and implemented the visual content response module for Expona, an in-house enterprise AI assistant built for querying large, complex documents. Enabled image-inclusive responses by integrating GPT models with computer-vision pipelines for extracting, matching, and rendering document visuals (graphs, diagrams, tables, screenshots). Built multi-modal retrieval using RAG + CV to align semantic queries with the most relevant textual and visual content. The system delivered superior enterprise QnA performance, outperforming industry tools like NotebookLM in handling visual-aware, context-rich insights.

---

## **8. HR Expona**

### **Tools & Technologies**

RAG, GPT-4, GPT-4o Mini, Pinecone Vector DB, NLP, Embeddings, Azure Cloud, Docker, Python, Automation Pipelines

### **Description**

Built the intelligent resume screening and talent matching module for HR Expona, an AI-driven hiring assistant designed for automated candidate shortlisting. The system ingests the Job Description (JD), extracts key skills and requirements, and automatically retrieves bulk resumes from internal and external repositories. Implemented vector-based matching using Pinecone and LLM-driven relevance scoring to identify the most suitable candidates.

Delivered a fully automated scoring pipeline that ranks resumes, highlights matching evidence, and generates recruiter-ready insights—reducing Talent Acquisition (TA) team effort by over 50%. Enabled faster, more consistent, and bias-free shortlisting for high-volume hiring workflows.

---

## EDUCATION

- **Bachelors of Science in Biotechnology Lifesciences Chemistry**

2019- 2022

Maharaja Ranjeet Singh College

- **Higher Secondary Certificate Examination**

Shree Neema Vidya Niketan

- **High School Examination**

Shree Vaishnav Academy

---

## OTHERS

1. President of InfoBeans Foundation Alumni Council - leading approx 400 students.
  2. AI ML Trainer at InfoBeans Foundation
  3. Provided Guest lectures on AI in MU20 University ,Sage University.
  4. AI Blog writer for InfoBeans Technologies.
-