# Sonu Kumar

+91 7595020599   sonuhits047@gmail.com   LinkedIn   GitHub   Portfolio

## SUMMARY

AI Engineer with hands-on experience in building LLM-based agents, fine-tuning computer vision models like YOLOv8, and backend development using Java Spring Boot. Skilled in LangChain, LangGraph, OCR, RAG, and modern ML/NLP stacks with a proven ability to deliver scalable and secure solutions.

## EDUCATION

| | |
|---|---|
| **Masai School** | **Bengaluru, India** |
| *Full Stack Development* | *June 2023 – Nov 2023* |
| **Dream Institute of Technology** | **Kolkata, India** |
| *B.Tech in ECE* | *July 2019 – July 2022* |
| **Holy Mary Institute of Technology and Science** | **Hyderabad, India** |
| *Diploma in EEE* | *July 2014 – June 2017* |

## TECHNICAL SKILLS

**Languages:** Python, Java, JavaScript

**AI/ML:** LangChain, LangGraph, OpenAI, Sklearn, Pandas, Numpy, OpenCV

**ML Architectures:** CNN, YOLOv8, BERT, LSTM

**Backend:** Spring Boot, Flask, REST APIs

**Databases:** MySQL, PostgreSQL, MongoDB, Vector DB

**Cloud/Tools:** AWS (S3), Git

## WORK EXPERIENCE

### Software Engineer (AI/ML)
*Monocept Consulting Pvt. Ltd., Gurugram, India*                    *April 2024 – Present*

– Developed AI-powered solutions using Python and ML frameworks to automate business workflows.
– Built LLM-based agents using LangChain, LangGraph, and OpenAI for question answering and automation.
– Implemented RAG pipelines for knowledge retrieval and chatbot integration.
– Fine-tuned YOLOv8 model to detect Aadhaar info in documents with high accuracy.
– Developed backend services using Java and Spring Boot following microservices architecture.
– Led migration of databases from Oracle to PostgreSQL for Care Health Insurance.

### Associate Software Engineer
*SuperSeva Global Services Pvt. Ltd., Bengaluru, India*                    *July 2022 – May 2023*

– Contributed to backend development using Java, Python and Spring Boot, Flask.
– Implemented REST APIs and optimized code for service reliability.
– Worked with teams to automate internal workflows and enhance service delivery.

## PROJECTS

### Major Projects

- **PolicyAssist – Monocept:** Designed and developed an intelligent insurance assistant that enables users to **upload policy-related documents** and interact with them through natural language queries. By leveraging **RAG (Retrieval Augmented Generation)** pipelines and **LangChain**, the system delivers highly specific answers directly from the uploaded documents, eliminating the need to manually read through lengthy policies for small details. This significantly **improves customer experience, reduces query resolution time, and increases overall efficiency in insurance workflows.**

- **Mono-DB-Agent – Monocept:** Engineered an LLM agent to translate natural language into MySQL queries using **LangChain** and **OpenAI**. This significantly improved data access for non-technical users by automating complex data retrieval workflows and reducing dependency on database administrators, leading to **increased operational efficiency and faster data insights**.

- **Mono-Aadhar-Masking – Monocept:** Developed a secure and automated pipeline for redacting sensitive Aadhaar numbers from uploaded documents. Utilized **Optical Character Recognition (OCR)** for text extraction and fine-tuned **YOLOv8** for precise object detection. This solution drastically **reduced manual effort and potential human error**, significantly enhancing data security and compliance with privacy regulations.

- **mPro – MaxLife Insurance:** Contributed to the backend development and maintenance of a web application for insurance agents to capture user details and process policy purchases. Developed and integrated new **Spring Boot APIs** and ensured secure storage with **AWS S3**, leading to **improved system stability and enhanced user experience** by resolving performance bottlenecks and adding new features.
- **DB Migration – Care Health Insurance:** Led a critical database migration from **Oracle** to **PostgreSQL**. This involved comprehensive planning, data mapping, and rigorously testing data consistency. Crucially, I refactored and adapted existing **Java Spring-based project queries** to be compatible with PostgreSQL, ensuring **zero-downtime** and significant improvements in database performance, scalability, and cost-efficiency, laying a robust foundation for future data-driven initiatives.
- **PikmyKid – SuperSeva:** Handled and resolved customer queries via email and phone, acting as a primary point of contact. Ensured timely and accurate responses, improving customer satisfaction and maintaining quality support standards.

## Data Science & AI/ML Development Projects

- **MoviesBot:** Developed a responsive chatbot application using **HTML, CSS, JavaScript, Python (Flask)**, and **OpenAI**. This bot leverages a **Retrieval Augmented Generation (RAG)** pipeline with **Pinecone** as the vector database to allow users to ask movie-related questions and provides relevant answers from a pre-defined database, enhancing user engagement and information retrieval efficiency.
- **SpringAIBot:** Created an AI-powered chatbot using **HTML, CSS, JavaScript, Java (Spring Boot)**, and **OpenAI**. This chatbot enables users to inquire about Spring, Spring Boot, and Java topics, and offers a feedback mechanism, improving user support and knowledge dissemination.

## Java-Oriented Projects

- **Electricity Bill Management System:** Built a Java-based application for electricity bill management. This system allows customers to create accounts, manage bills and payments, and check bill status, streamlining billing operations and improving customer service.
- **EasyRentHub:** Developed a housing system application using **Java, Hibernate**, and **MySQL**. This platform enables users to view and manage property rentals, enhancing the efficiency of property management and rental processes.
- **Trip Management System:** Designed a web application for online trip management using **Java, Spring Boot, MySQL, HTML, CSS**, and **JavaScript**. This system provides a convenient platform for managing tours and travels, improving organization and accessibility for travel planning.

## General Web Development Projects

- **Purplle Clone:** Contributed to the development of a prominent online beauty and personal care platform using **HTML, CSS**, and **JavaScript**. This project involved implementing cutting-edge technologies and strategic improvements to enhance the user experience and platform functionality.

## CERTIFICATIONS & COURSES

- Full Stack Development – [View Certificate](#)
- Complete Data Science Bootcamp – [View Certificate](#)
- LLM Engineering – [View Certificate](#)