

- **Data Scientist** with hands-on experience in **Machine Learning**, **Deep Learning**, and data-driven automation. Skilled in building **end-to-end ML pipelines** using **Python**, **Scikit-learn**, **TensorFlow**, and **PyTorch**. Experienced in **OCR**, **NLP**, and **LLM-powered systems** for intelligent data extraction and automation. Developed a **multi-agent email automation system** and contributed to **MCP** and **A2A** research. Passionate about applying **AI**, **predictive modelling**, and **MLOps** to deliver scalable, real-world solutions with business impact.

**EDUCATION****Prestige (PIEMR), RGPV, Indore, MP, India** | Bachelor of Technology (Computer Science & Engineering)

2021-2025

- **Courses:** Object-oriented programming and design principles (OOPS), Database, Data Structures, Software Design Patterns.

**WORK EXPERIENCE****Soothsayer Analytics, Hyderabad - Junior Data Scientist**

Feb 2025 – Current

- Created and delivered **RESTful APIs** for seamless integration between frontend and backend services, enabling data flow and automation.
- Collaborated and optimised **MySQL** database schemas and queries, reducing query execution time by **30%** and ensuring data integrity.
- Managed backend development, including data ingestion, processing, and storage with **OCR** and **LLM-powered extraction pipelines**.
- Collaborated with internal **stakeholders** and participated in code reviews with cross-functional team members, demonstrating strong communication skills to build a secure and scalable platform.

**AvniPride Technology, India - Python Intern**

April 2024 – May 2024

- **Revitalised** a legacy press management system using Python, which reduced previous system errors by **20%**
- **Engineered and enhanced backend solutions** with **MySQL integration**, improving **database performance** and reducing query latency.
- Designed and integrated automated data processing workflows, boosting system efficiency and reducing manual intervention by **40%**.

**KEY PROJECTS****AUTOMATED VOICE CALL HANDLING AGENT** | **ElevenLab API**, **Twilio**, **Flask**, **Railway**, **SQLite**, **FastAPI** | [LINK](#)

- Integrated **Twilio Voice API** for real-time call management and **ElevenLabs TTS/STT APIs** to enable natural, human-like speech communication.
- Achieved a **75% reduction in manual call-handling time**, showcasing potential for scaling AI-driven customer support solutions.

**Automated Invoice Extraction** | **N8N**, **Python**, **LlamaIndex**, **DeepSeek** | [LINK](#)

- The system **automatically reads invoice emails**, **extracts PDF attachments**, parses complex invoice layouts, **converts unstructured data into a structured schema**, and stores the results in Google Sheets — **all without manual intervention**.
- Performed comprehensive **text preprocessing** (tokenisation, stop-word removal) and leveraged **Count Vectorizer** for robust feature extraction, improving model **interpretability, efficiency, and reliability**.

**AUTOMATED BILL EXTRACTION SYSTEM** | **OCR**, **Google Vision API**, **Streamlit**, **LangChain**, **LLMs**, **Python** | [LINK](#)

- Planned and executed a Streamlit-based application to **automate expense bill processing**, **accurately extracting key details**(date, amount, vendor) and **reducing manual workload by 70%**
- **Integrated and Advanced** Google OCR API to improve text recognition accuracy by **30%**, ensuring faster, error-free data extraction.
- Leveraged **LLM-powered** entity recognition (via Hugging Face API) to categorise extracted data with **85% accuracy**

**AUTOMATED RAG-AGENT COMMUNICATION SYSTEM** | [LINK](#)

- Building a system where AI agents can communicate with each other autonomously using an **A2A (agent-to-agent) workflow**.
- Exploring how agents can share context, delegate tasks, and collaborate to produce better decisions and outputs.
- Designing flexible modules for intent routing, message planning, and reasoning to make **agent communication** smoother and more efficient.

**TECHNICAL SKILLS**

- **Programming Languages** : C, C++, Python (expertise in backend & AI development)
- **AI & ML** : Deep Learning, Natural Language Processing (NLP), Computer Vision, Generative AI (LLMs)
- **Cloud Platforms & Deployment**: AWS (EC2, S3, Lambda), Google Cloud Platform, Docker, CI/CD with GitHub Actions
- **Databases** : MySQL, complex SQL queries (Joins, Subqueries, CTEs), database administration.
- **Algorithmic Techniques** : Data Structures, algorithm optimisation, efficient coding practices.

**PUBLICATION****A Study on the Potential of AI in the Healthcare Sector**

Feb 2024

- **Metszet Journal** | **Key Skills: Artificial Intelligence, Healthcare, Uses of AI in Healthcare**
- Conducted in-depth analysis of various AI technologies and their impact on the healthcare industry. [metzet 4303 R.pdf](#)