

Sachin Loddiya Karthik

[in LinkedIn](#) | [517-243-9938](#) | [Portfolio](#) | sachinlkece@gmail.com | [GitHub](#) | USA

Skills

- Python | R | SQL | Java | C++ | Machine Learning | PyTorch | TensorFlow | Scikit-Learn | XGBoost | OpenCV | MLflow | NLTK
- LLMs | LangChain | FAISS | OpenAI API | Gemini | NLP | Conversational AI | RAG Systems | Text Embeddings | Generative AI
- Azure | AWS | GCP | Databricks | Apache Spark | Kafka | Airflow | Docker | Kubernetes | CI/CD | Jenkins | Docker | CI/CD
- PostgreSQL | MongoDB | Power BI | Tableau | Streamlit | FastAPI | Git | Jupyter | Statistical Analysis

Experience

- AI Engineer** Goodie Bag **Remote, USA** **06/2025 – Present**
- Developed an **AI-powered Customer Marketing Agent** to generate branded copy across channels (social media, email, website, in-app content, and physical ads), ensuring consistent brand voice.
 - Orchestrated **RAG (Retrieval-Augmented Generation) pipelines** using brand guidelines, historical campaign data, and tone-of-voice documents to improve message consistency.
 - Designed and deployed a stateful conversational AI with **FastAPI** to handle **7+** customer support use cases, reducing manual ticket handling by **30%**.
 - Engineered a high-performance, non-blocking API processing **500+** daily transactions for real-time refunds and ticket creation.
 - Built a **PostHog-powered analytics dashboard** by writing custom **HogQL** queries to track active users across multiple cities, enabling dynamic date filtering and actionable insights for growth teams.

- Data Scientist** WMU **Kalamazoo, USA** **01/2025 – 04/2025**
- Created an automated class scheduling optimizer using **Google OR-Tools**, reducing manual scheduling effort by 85% for 200+ course assignments
 - Built a Streamlit interface for real-time CSV ingestion/validation, **minimizing data errors** by 90%.
 - Applied operations research and **constraint optimization** to generate 100% conflict-free schedules for 5 academic departments, accommodating 20+ instructor preferences and 15+ classroom resource constraints

- Data Scientist** Green Expectations LLC **Remote, USA** **01/2024 – 04/2024**
- Formulated a rule-based Home Sustainability **Scoring model** leveraging 1,000+ data points, enhancing recommendation accuracy by 35%.
 - Refined **NLP chatbot** workflows, reducing latency by 40% and improving user engagement by 25%.
 - Directed **feature engineering** for sustainability predictions across varied user profiles, improving model accuracy by 15%.

- Data Engineer** Accenture Solutions **Remote, India** **07/2021 – 07/2023**
- Architected **ETL** pipelines using **Azure Data Factory** and **Databricks**, improving data processing efficiency by 40% and accuracy by 25%.
 - Enhanced batch transformation workflows on **Data Lake Gen2**, halving processing times while maintaining schema consistency across diverse datasets.
 - Constructed **Power BI** dashboards and integrated **CI/CD pipelines**, accelerating delivery speed by 30%.

- Data Engineer** Claritrics India Pvt Ltd **Remote, India** **11/2020 – 05/2021**
- Implemented **OCR-based** document processing systems via **OpenCV** and Python, boosting text extraction accuracy by 30%
 - Streamlined recurring workflows by automating manual processes, reducing **processing time by 40%** and boosting team efficiency.

Education

- Master of Science** WESTERN MICHIGAN UNIVERSITY **MI, USA** **08/2023 – 04/2025**
- Data Science GPA: 3.86

Project

Ask My PDF – LLM-Powered Document Q&A System (LangChain, FAISS, BM25, PyMuPDF):

- Engineered a scalable, multilingual PDF Q&A system that combines LLMs with hybrid retrieval (FAISS + BM25) to deliver context-aware answers over long, unstructured documents.
- Launched a Streamlit-based web app supporting multi-PDF uploads and memory-aware chat sessions, delivering real-time AI-powered responses that reduced user query time by 40%

Enhanced ATS Resume Expert Pro (Streamlit, Gemini Pro, PyMuPDF, scikit-learn, Plotly):

- Engineered a scalable, multilingual PDF Q&A system that combines LLMs with hybrid retrieval (FAISS + BM25) to deliver context-aware answers over long, unstructured documents.
- Integrated PDF/DOCX parsing, TF-IDF + cosine similarity, and skills gap detection to provide actionable insights, boosting resume relevance scores by up to 30% for users.