

Ramagiri Srujan

☎ 9390983535 | ✉ ramagirisrujan123@gmail.com | 💻 [linkedin.com/in/srujan](https://www.linkedin.com/in/srujan) | 🐙 github.com/srujan3421

PROFESSIONAL SUMMARY

Results-driven Data Scientist strong hands-on experience in Python-based application development, data handling, and Generative AI solutions. Experienced in building web applications using Streamlit, integrating LLMs via LangChain and Gemini APIs, and applying machine learning concepts to real-world datasets. Familiar with low-code platforms (Pega – academic exposure), Pega Constellation concepts, and GenAI-driven workflows. A self-driven learner with strong problem-solving, analytical thinking, and communication skills, eager to contribute in collaborative engineering environments.

EDUCATION

Vaagdevi College of Engineering
B.Tech in Computer Science , **CGPA: 8.07**

Warangal, Telangana
Nov 2021, July 2025

SKILLS

- **Languages:** Python, SQL, Java, C, HTML, CSS.
- **Developer Tools:** Excel, MySQL, PostgreSQL, SQL, VS Code, PyCharm, Jupyter Notebook.
- **Data Handling:** Data Cleaning, Feature Engineering, Data Modeling, Structured & Unstructured Data
- **Technologies / Frameworks:**
 - **Artificial Intelligence, Machine Learning & Deep Learning:** Regression (Linear & Non-linear), Classification, K-Means Clustering, Hierarchical Clustering, TensorFlow, Keras, Scikit-learn, PyTorch, XGBoost
 - **Statistics:** Descriptive Statistics, Probability, Hypothesis Testing, Model Assumptions
 - **NLP Libraries:** NLTK, SpaCy
 - **Data Analytics:** NumPy, Pandas, Matplotlib, Seaborn, Power BI, MS Excel, PowerPoint
 - **Cloud & APIs:** AWS, MongoDB Atlas, REST APIs, Flask, Streamlit
 - **Version Control:** Git, GitHub
- **Technologies / Frameworks:**
 - **Transformer Architectures:** BERT, GPT, LLaMA, T5
 - **LLMs & APIs:** OpenAI, Gemini, LLaMA, Groq (Mixtral)
 - **Frameworks:** Hugging Face (Transformers, Datasets, Model Hub), LangChain, LangGraph
 - **RAG Systems:** Chroma, FAISS, Embeddings, Semantic Search
 - Prompt Engineering, Fine-Tuning, Model Evaluation, Agentic & Multi-Agent Workflows
- **Other Skills:** Problem Solving, Analytical Thinking, Time Management, Team Collaboration, Communication

Work Experience

Data Science Intern

April 2025- June 2025

AI-Powered Insurance Claim Assistant :

- Built an **AI-powered Insurance Claim Assistant** using Streamlit and Groq Llama/Mixtral models with multi-model selection.
- Implemented **context-aware claim guidance** with structured prompts and user profiling.
- Created **insurance-specific document checklists** for auto, bike, home, health, travel, and life claims.
- Added **conversation memory, clean chat UI, avatars, and test scenarios** for a smooth user experience.

PROJECTS

AI Study Assistant

- Engineered **RAG pipeline with ChromaDB** and sentence transformers, **achieving 92% retrieval accuracy and 0.87 cosine** similarity across 500+ document chunks for semantic search
- Automated document preprocessing using PyPDF2 and python-docx, reducing manual processing time by 85% while implementing vector embeddings with all-MiniLM-L6-v2 model
- Developed AI-powered quiz system via Groq API integration, improving knowledge **retention by 40%** and optimizing query processing to **reduce latency by 60% with 94% relevance accuracy**.

Next Words Prediction using NLP

- Preprocessed large text corpus with cleaning, tokenization, and n-gram sequence generation, improving data quality by 35%.
- Built and optimized an LSTM-based language model, **achieving 92% next-word** prediction accuracy and reducing overfitting by 28% using dropout and early stopping.
- Developed a fast prediction tool capable of generating next words in <0.3 seconds, increasing text generation **efficiency by 50%**.

Real-Time Automated License Plate Recognition (ALPR) System

- Developed an automated car number plate detection system in Python using **OpenCV and advanced image processing** for real-time vehicle identification.
- Engineered a robust pipeline with grayscale conversion, edge detection, and contour extraction, achieving accurate localization under diverse conditions.
- Integrated **Pytesseract OCR** for high-precision license plate text extraction, enhancing automation and workflow efficiency.
- Visualized results and detection steps with Matplotlib and imutils, demonstrating end-to-end skills in computer vision, OCR, and image analytics.

CERTIFICATIONS

- Data Science Professional (Oracle) - [Link](#)
- Python Essentials (Cisco) - [Link](#)
- Data Analytics and Visualization (Accenture) - [Link](#)