

SUPRAJA MADALA

+1 (475) 202 1437 | Supraja2224@gmail.com | www.linkedin.com/in/suprajamadala New Haven, CT- 06511

OBJECTIVE

Data Scientist with expertise in AI/ML, NLP, deep learning, and predictive analytics. Strong experience in building RAG systems, developing end-to-end ML pipelines, automating data workflows, and delivering scalable intelligence for real-world operations.

SUMMARY OF TECHNICALSKILLS

Large Language Models (LLMs): Llama, GPT models, RAG systems, Prompt Engineering, Fine-tuning (LoRA/QLoRA), Embeddings (FAISS, Chroma), Vector Databases, Tokenization, Retrieval Pipelines, Model Quantization (GGUF), API Integration

Machine Learning: TensorFlow, PyTorch, Scikit-learn, XGBoost, CNNs, NLP, Computer Vision

Programming: Python, SQL, R, JavaScript

Data Engineering: ETL, Airflow, Kafka, Docker, AWS (S3, EC2), Batch & Streaming Pipelines

Visualization: Power BI, Tableau, Matplotlib

Web/Dev: Flask, REST APIs, Salesforce Dev

Tools: Git, Jira, Selenium, JMeter

Soft Skills: Strong presentation, time management, organizational skills, teamwork, communication

EDUCATION

Master of Science, Data Science | University of New Haven, West Haven, CT

December 2025

Academic Accomplishments: CGPA: 4.0/4.0

Bachelor of Technology, Computer Science and Engineering | St. Peter's Engineering College, Hyderabad

May2023

Academic Accomplishments: First Class with Distinction.

EXPERIENCE

Data Scientist | Sperry Rail, CT

Aug 2025- present

- Training Elmer-based machine learning models for vision, Eddy Current, Joint Bar, and MagFluxdata to improve defect detection accuracy across rail-end inspections.
- Developed deep learning models for signal recognition, MagFlux anomaly classification, and eddy current pattern analysis, improving detection precision and reducing false positives.
- Building automated data preprocessing pipelines processing 10,000+ sensor frames per inspection, enabling scalable real-time analysis.
- Collaborating with engineering teams to integrate model output into predictive maintenance workflows.

Research Assistant | University of New Haven, CT

Jan 2024 - present

- Developed an advanced Risk Assessment and Guidance (RAG) system for construction inspections utilizing Llama (Large Language Model Meta AI).
- The project aimed to streamline and enhance the inspection process by leveraging AI to analyze construction site data, identify potential risks, and provide actionable insights.
- Implemented a robust framework for real-time data processing and risk assessment, which significantly improved the accuracy and efficiency of inspections.
- The system integrated natural language understanding capabilities to interpret and generate comprehensive inspection reports, facilitating better decision-making and compliance with safety regulations.

Data Engineer | Larsen & Toubro Limited, Hyderabad

Jan2023-Jan2024

- Built predictive models for crew scheduling that improved routing efficiency by 28% and reduced station-level delays by 15%.
- Automated feature engineering and ETL pipelines, cutting manual processing time by 60% and improving data quality by 40%.
- Performed large-scale EDA on 5M+ operational records, identifying bottlenecks that increased resource utilization.
- Developed Kafka-based real-time data streams and analytics dashboards, accelerating decision-making speed by 3× across metro operations.

Salesforce Developer Intern / Cloudely, Hyderabad

Sep2022-Jan2023

- Developed complex Salesforce applications using Apex, Visual Force and Lightning Components. Assisted in designing custom objects, fields, formulas, validation rules, workflow rules. Created and maintained reports and dashboards in Salesforce to track key metrics.
- Utilized Salesforce APIs for integration with external systems. Participated in requirements gathering sessions with stakeholders to identify business needs and develop solutions accordingly. Tested software before delivering it to users.

Human Resources Analyst Intern / Lernx, Hyderabad

Sep 2022-Sep2022

- Prepared onboarding documents for new hires, such as contracts and employee handbooks.
- Analyzed employee data and maintained secure, organized personnel records in the database to ensure compliance and confidentiality.
- Assisted in streamlining the recruitment process by using data analytics to screen resumes, assess candidate qualifications, and schedule interviews.
- Generated reports on recruitment trends, candidate performance, and onboarding efficiency to optimize HR processes.

Intern | Girl Summer Script Of Code , Hyderabad

March2022-May2022

- Implemented a voice assistant named "Kavi" using Natural Language Processing (NLP) techniques to enable voice- driven interactions for users.
- Developed speech recognition and synthesis modules, integrating them with a chatbot framework to perform tasks through voice commands, enhancing user experience and productivity.
- Contributed to various open-source projects, leveraging collaborative tools to improve and refine machine learning models and NLP algorithms.
- Worked on improving the voice assistant's accuracy and responsiveness through data preprocessing, feature engineering, and model optimization.

COURSES AND CERTIFICATIONS

Google Data Analytics With R| Coursera

July 2024

Google Data Analytics | Coursera

Nov 2023

Lean Six Sigma White Belt| M&S Institute

Nov 2023

AWS Academy Certification | AWS

July 2022

Python For Beginners | Tap Academy

April2022

ACADEMIC PROJECTS

- **Advertisement Recommendation System (99% Accuracy):** Implemented Gradient Boosting, Random Forest, and Decision Trees to recommend ads, improving targeting accuracy by 35% and engagement prediction by 42%. Optimized hyperparameters and reduced inference latency by 20%.
- **Job Market Analysis Pipeline:** Automated ETL using Airflow, scraping 20,000+ job listings daily and storing in AWS S3. Built dashboards tracking hiring trends across 50 states, improving visibility for workforce planning.
- **Brain Tumor Detection (CNN):** Developed CNN achieving 92% classification accuracy for MRI tumor detection. Applied augmentation and KNN refinement to reduce false negatives by 18%.
- **OCR Text Extraction System:** Built OpenCV + CNN model boosting text extraction accuracy by 27% on low-quality images. Implemented preprocessing pipeline reducing noise by 40%.
- **KAVI Voice Assistant:** Integrated NLP pipeline capable of answering 10+ categories of user queries. Reduced speech-recognition error rate by 30%.

EXTRA-CURRICULAR ACTIVITIES

Vice President| St. Peter's Engineering College| E-cell CSE

Aug 2022-Aug 2023

Radio Jockey, President | St. Peter's Engineering College| Radio club

Aug 2019- May 2023

Club Secretary | St. Peter's Engineering College| Student Council

Aug 2019- May 2023

Associate Marketing lead | Entrepreneur Club

Jan 2024 -Present

President| University of New Haven| Girls Who Code

Sep 2024 -Present

Member| Indian Student Council

Jan 2024 -Present

RESEARCH PUBLICATIONS

Title: Automatic Segmenting of Brain Tumours within MRI Images

Journal: ODJ Journal, Volume 29, Issue 06, June 2023

- Published a research paper focused on glioma segmentation using deep learning techniques, aiming to improve diagnostic accuracy and treatment planning for brain tumour detection in MRI images.
- Proposed a novel Depth Normalizing technique to address information heterogeneity across different MRI acquisition settings, ensuring consistency in analysis across multiple sites and scanners.
- Explored the use of compact convolutional architectures for precise segmentation, demonstrating the effectiveness of deep learning in medical imaging applications.