Sai Raghavendra Maddula
Email: raghudo888@gmail.com
Phone: 940-843-3560
LinkedIn: https://www.linkedin.com/in/sairaghavendramaddula/
Blog: https://myblogspott.com

Professional Summary

Dynamic and results-driven **Senior Research Engineer** with expertise in **Machine Learning, Artificial Intelligence, and Large Language Models (LLMs)**. Adept at architecting **scalable AI/ML solutions**, integrating **neural network models** into production, and optimizing **ML infrastructure** for high-performance applications. Proven ability to drive **breakthrough innovations** in AI applications across **Legal, Tax, Compliance, and News industries** by leveraging cutting-edge research and real-world implementations. Passionate about pushing the boundaries of **AI-driven automation, intelligent data processing, and next-generation model deployment** to solve complex industry challenges.

Technical Skills

- **Programming:** Python, Java, Rust, Scala, TypeScript
- **Machine Learning:** TensorFlow, PyTorch, Scikit-Learn, Hugging Face Transformers
- **NLP & LLMs:** Named Entity Recognition (NER), Information Extraction, Retrieval-based AI
- **MLOps & Model Deployment:** ModelOps, Docker, Kubernetes, MLflow, CI/CD pipelines
- **Cloud Platforms:** AWS, GCP, Azure
- **Data Engineering:** Spark, Hadoop, SQL/NoSQL (PostgreSQL, MongoDB)
- **Software Development:** Agile, DevOps, SDLC, Test-Driven Development
- **Monitoring & Logging:** Prometheus, Grafana, ELK Stack

Professional Experience

Research Engineer

GR Simplot | Jun 2019 - July 2020 | **Hyderabad, India**

- Developed and optimized **AI-powered legal research tools**, improving document processing efficiency by 40%.
- Built **scalable ML pipelines** to process millions of legal, tax, and compliance documents.
- Integrated **LLMs** for **information retrieval, Named Entity Recognition (NER), and intelligent data extraction**.
- Designed **cloud-based AI solutions** using AWS/GCP, ensuring high availability and low-latency responses.
- Led **ModelOps & MLOps initiatives**, automating AI/ML model deployment and monitoring.

- Collaborated with **cross-functional teams** to translate AI research into real-world applications.

Machine Learning Engineer
Microsoft | Jun 2020 - Jul 2021 | **Hyderabad, India**

- Developed **state-of-the-art NLP models** for document classification and semantic search.
- Implemented **AI-powered fraud detection algorithms**, reducing anomalies by 30%.
- Designed **RESTful APIs** for AI-driven legal and tax compliance tools.
- Built **CI/CD pipelines** for ML model deployment, improving automation and scalability.
- Applied **Agile & DevOps methodologies** for rapid AI product development and deployment.

AI Research Engineer
GM Financials | Sept 2021 – August 2022 | **Hyderabad, India**

- Conducted research on **transformer-based models** for improving legal document search and summarization.
- Developed **multi-modal AI solutions**, combining text and image processing for legal case analysis.
- Worked with **cross-functional teams** to bridge the gap between AI research and enterprise applications.
- Led the **development of explainable AI models**, improving transparency and trust in AI-driven compliance tools.

Education
Master's in Computer Science Christian Brothers University, 2024

Certifications

- Certified Kubernetes Administrator (CKA)
- Google Cloud Professional Cloud Architect
- AWS Certified Solutions Architect

Projects & Achievements

- Developed an **AI-driven compliance tool** that reduced document processing time by 50%.
- Led the **integration of NLP models** into enterprise search applications for legal & tax research.
- Built an **LLM-based chatbot** for automated legal consultation, enhancing client support.

- Implemented **self-healing ML pipelines**, reducing model drift and improving predictions.
- Developed and optimized a **CI/CD pipeline** that reduced deployment time by 30%.
- Led the implementation of a **security-enhanced OpenShift platform**, ensuring compliance with enterprise security policies.
- Built a **scalable logging and monitoring solution** using **Prometheus and Grafana**, improving system observability.
- Designed and deployed **microservices handling millions of requests per day**, achieving high availability and fault tolerance.
- **Scorify:** Led the development of an **AI-powered sports analytics platform**, providing real-time player performance insights using ML-driven predictive modeling.
- **Zeez AI:** Engineered an **AI-driven customer service assistant**, leveraging **LLMs and NLP models** to enhance automated responses and improve user experience.
