ASHWIN RAJA SHANMUGA RAJA

Data Scientist

ashraja941@gmail.com · (469) 604 4434 · https://www.linkedin.com/in/ashwinraja941 · github.com/ashraja941

Computer Science graduate with expertise in Machine Learning, Deep Learning, and AI model deployment with 1.5+ years of experience in Data Science and Machine Learning

EDUCATION

Texas A&M **May '25**
Master of Science in Computer Science **CGPA: 4.0/4.0**
SRM Institute of Technology **May '23**
B.Tech CSE specialization in Artificial Intelligence and Machine **CGPA: 9.68/10**

PROFESSIONAL EXPERIENCE

Data Scientist, Texas A&M **Jan '24 - Present**

- Optimized machine learning models **(XGBoost, Random Forest, Logistic Regression)** for predicting Sudden Hypoxemia, achieving 15% higher accuracy through **Recursive Feature Elimination and Lasso Regression.**
- Designed Deep Learning models for analyzing ECG spectrograms using **ResNet-18, CNNs, and Bayesian optimization**, improving shunt blockage prediction accuracy by 10%.
- Built scalable data cleaning and preprocessing pipelines for **500,000+ real-world patient records** from Texas Children's Hospital and Baylor College of Medicine, improving model training efficiency by 35%.
- Built **group-based trajectory analysis** for DKA prediction in R and Python, extracting maximum insights from a limited dataset to enhance early detection strategies and improved **Logistic Regression** performance by 12%.

Python Intern, The Arm Academy **Jan '23 - May '23**

- Programmed an **Interactive dashboard using Flask backend and JavaScript frontend** to display and **analyze real-time data streams** from solar panels, to **predict trends**, improving monitoring accuracy and efficiency.
- Migrated Arduino workflows to Raspberry Pi with Python, utilizing its **40x faster processing power**
- Authored a comprehensive technical manual to facilitate knowledge transfer for new engineers using IoT systems.

Research Intern, Samsung Prism **Jan '22 - Jul '22**

- Designed an open-source header compression tool using **ROHC based header-compression and decompression** achieving 93% of Samsung's proprietary algorithm efficiency.
- Delivered a cross-platform frontend to test compression-decompression metrics using JavaScript and Python.

PROJECTS

Automated Personal Assistant using LangChain Texas A&M, 2025

- Developed a chatbot-driven personal assistant capable of scheduling meetings and summarizing emails. Utilized **multi-agent RAG** and **prompt-engineered LLM agents** for efficient tool invocation and task execution.

Outfit recommendation system Texas A&M, 2024

- Built a fashion recommendation system that used **ResNet-50** for features extraction as well as visual semantic embeddings, **using BERT.** Increased the accuracy of the model by 23% by incorporating transformers.

Meal Plan Assistant Texas A&M, 2025

- Created an assistant capable of creating custom meal plans and recipes based on user preferences using **LangGraph** and **RAGs**.

Gemma-3B Fine-Tuning for Text-to-SQL Retrieval 2025

- Fine-tuned **Gemma-3B** with Unsloth on **100,000+** text-to-SQL data. Built an instruction-tuned model for accurate, schema-aware **SQL query generation**, improving performance on complex queries.

Aggieland Art Trail web application Texas A&M, 2023

- Built a **web app for the Visual Art Society,** enabling an interactive Art Trail with stamps and achievements using Python, Firebase for the Backend and React, and JavaScript for the Frontend.

Medium posts auto tagger SRMIST, 2023

- Developed a **Chrome extension**, available on the Edge Extension Store, that automatically tags Medium posts using Natural Language Processing (NLP) Designed and implemented both the backend (Python) and frontend (JavaScript) for real-time data processing.

Healthcare Treatment Scheduling Optimization SRMIST, 2023

- Proposed a **genetic algorithm** to optimize the scheduling for fixed-time medical treatments and achieved a 15% jump in performance compared to the Firefly Algorithm. Published on **IEEE Xplore Digital Library.**

Reinforcement learning for Autonomous car driving in simulation.

- Developed and fine-tuned a PPO-based reinforcement learning agent for autonomous driving in a Rocket League simulation, outperforming A2C by 2× in goal efficiency. Engineered custom reward functions and training pipelines in PyTorch, optimizing policy stability and accelerating convergence.

SKILLS

- **Programming Languages** : Python, C++, JavaScript, Kotlin, C#, Ruby
- **Frameworks and Tools** : TensorFlow, PyTorch, Keras, Transformer Models, LangChain, LangGraph, SQL, Firebase, MongoDB, REST API, Git, NumPy, Pandas, Data Science, SQL, Git, Flask, Node.js
- **Core Competencies** : Machine Learning, Deep Learning, Natural Language Processing, Computer Vision, Agile Methodology, IoT, Operating Systems, Reinforcement Learning, LLMs, Generative AI, ETL

ACHIEVEMENTS

- **Published Paper :** Presented at **ESCIIOIT 2023** and published on **IEEE Xplore Digital Library**
- **Best paper :** won best paper award at **ICIoT 2022**
- **First Place :** Defense Services **Hackathon**