

# Shreerama D S

+918217731857

[Mail ID](#)

[Linked IN](#)

[Github](#)

[My Portfolio](#)

## Summary

**AI/ML Engineer** with a **B.Tech in Artificial Intelligence and Machine Learning** (CGPA: 8.6) from **Dayananda Sagar Institutions**, experienced in **LLMs, RAG pipelines, and Generative AI**. Interned at **Hindustan Aeronautics Limited(HAL)** and **Bulk Beings(Startup)**, building real-time AI solutions using **Python, PyTorch, and TensorFlow**. Published two **IEEE research papers** and contributed as a **peer reviewer** for **NetACT 2025**. Skilled in **NLP, image processing, and end-to-end AI deployment**.

## Education

### Bachelor of Technology in Artificial Intelligence

Dayananda Sagar University

Expected Graduation: June 2025

CGPA: 8.6

Jan 2021 – June 2025

Bangalore, India

## Skills

- **Programming Languages:** Python, HTML, CSS, SQL, C, JavaScript
- **Machine Learning:** Supervised & Unsupervised Learning, Classification, Clustering
- **Deep Learning:** NLP, CNN, RNN, LSTM, GAN, Autoencoders
- **AI Frameworks:** LLM, Generative AI, RAG, PyTorch, TensorFlow
- **Tools & Technologies:** GPT, BERT, APIs, Git, Jupyter, PyCharm, Replit
- **Data Analysis:** Pandas, NumPy, EDA, Data Visualization
- **Other:** Image Processing, Computer Vision, Problem-Solving

## Experience

### AI Engineer, Bulk Beings (Startup)

Jan 2025 – June 2025

- Built a real-time **NOTAM Dashboard** using **LLMs, RAG, and Qwen model** for aviation alerts.
- Integrated multi-source data pipelines and implemented **citation-backed responses**.
- Optimized **prompt logic** and deployed scalable **RESTful APIs**.

### AI Research Intern, Hindustan Aeronautics Limited (HAL)

July 2024 – Aug 2024

- Improved the **HiFiC model** for high-quality, **lossless** image compression.
- Enhanced **deep learning** performance through **model tuning** and **error resolution**.

## Projects

### 1. [Cysinfo AI](#): Fine-tuned LLaMA 3.1 using LORA

- **Description:** Fine-tuned **LLaMA 3.1** using **LoRA** and domain-specific datasets to achieve high contextual accuracy for cybersecurity and ethical hacking queries, improving response accuracy for Kali Linux commands by **20%**. Developed web interface, integrating advanced NLP APIs, deployed the model on **Ollama server**, optimizing query processing latency by **30%**
- **Skills Highlighted:** LLMs, Generative AI, AI Development

### 2. [DocuSense](#): Designed RAG-based assistant with Langchain

- **Description:** Designed an AI-powered documentation assistant utilizing advanced **Retrieval-Augmented Generation (RAG)** with **LangChain, LLaMA 2, and Nomic Embed Text** to enable real-time query resolution and contextual awareness. Developed and deployed an interactive application using Streamlit, integrated with the Ollama server.
- **Skills Highlighted:** RAG, LLaMA 2, LangChain, Generative AI, AI Development, Streamlit, NLP.

### 3. [Voice Genius](#) : It's Echoes of Intelligence

- **Description:** Developed an AI-driven voice assistant using **OpenAI API** and **SpeechRecognition**. The application processes spoken questions and provides answers in both voice and text formats.
- **Skills Highlighted:** NLP, Speech Recognition, API Integration, Deployment

### 4. [Beyond Skin-Deep](#): Melanoma Detection

- **Description:** Implemented and optimized **deep learning models** for melanoma detection, focusing on **hyperparameter tuning** and performance evaluation.
- **Skills Highlighted:** Deep Learning, Image Classification, Hyperparameter Tuning

## Publications

1. **[“Voice Genius: It's Echoes of Intelligence”](#)** — Published in 15th IEEE ICCNC Conference 2024.
2. **[“CysinfoAI: Fine-Tuning LLaMA for Unrestricted Cybersecurity Insights”](#)** — Published in 3rd International Conference NETACT 2025.

## Certifications

### [Machine Learning \(Coursera\)](#)

June 2023

- Completed **Andrew Ng's Machine Learning** course, gaining foundational skills in supervised and unsupervised learning, offered by **Stanford University**