

# Shreyans Soni

+919827067491 • sonishreyans01@gmail.com • Website • LinkedIn • GitHub

## Education

**Birla Institute of Technology & Science Pilani, Hyderabad Campus**

**Nov 2020 – Jul 2024**

*Bachelor of Engineering in Electronics and Instrumentation*

*Hyderabad, Telangana*

**Relevant Coursework:** Artificial Intelligence, Computer Programming, Operating Systems, Digital Design, Object Oriented Programming, Cryptography, Internet of Things, Microprocessors, Probability & Statistics

## Skills

**Languages:** Python, C, C++, HTML/CSS, JavaScript, SQL

**Developer Tools:** VS Code, AWS SageMaker, Docker, Autotrain, Redis, CUDA

**Technologies/Frameworks:** React, Next, Node, MongoDB, Linux, Multithreading, Django, Flask, MQTT, Neural Networks, GitHub, Deep Learning, Machine Learning, L NP, Parallel Processing, Fine tuning, LangChain, RAG systems, WebSockets, HTTP, OpenCV, PyTorch

## Experience

**ArgenBright Innovation Labs**

**Jan 2024 - Present**

*SDE Intern*

*Bengaluru, Karnataka*

- Implemented real-time face recognition and pose estimation functionalities for the security platform
- Optimized backend performance, achieving a 7x improvement in output frame rate by leveraging parallel processing and batching technique
- Developed a complete Conversational AI system utilizing web sockets and a MongoDB database

**Arbunize Digital Media**

**May 2022 - Jul 2022**

*Web Development Intern*

*Remote*

- Implemented UI updates for the VideoWiki Class website using React and NodeJS
- Resolved 14+ website issues in course of a month, including bug fixes and enhancements to the React UI

## Projects

**Multilingual Speech Integrated RAG Conversational AI | Server Optimization, Full Stack**

**Apr 2024 - May 2024**

- Optimized Backend with Python and Socket.IO for real-time Communication using Socket.IO and reduced server latency and average inference time by 85%
- Implemented Scalable Microservice Architecture with MQTT Broker and Workers to handle each micro service efficiently for multiple requests
- Deployed a ReAct Agent using Llama Index for the conversational engine
- Implemented multilingual speech (TTS & STT), in 53 languages, to a LangChain RAG system, using Whisper

**Personal Portfolio and Interactive Resume Chat | MERN Stack, Artificial Intelligence**

**Feb 2024 - May 2024**

- Built a full-stack personal website using Next.js for a dynamic user experience. Leveraged MongoDB and Sanity.io for efficient data management
- Successfully deployed the website on Vercel and achieved a high performance score (RES: 92) through code optimization
- Implemented a conversational resume chat using the Gemini API and a vector database, enabling users to interact with resume information

**Intelligent Task Scheduler with LLM Integration | Web Development**

**Mar 2024 - May 2024**

- Designed and developed a web-based task scheduler utilizing Llama Index agents and Gemini for intelligent task management
- Implemented Redis caching to optimize task retrieval and updates, leading to faster response times
- Built the backend application with Flask and integrated a MongoDB database for scalable data storage

**Distributed Graph Database System Simulation | Operating Systems, Databases**

**Nov 2023 - Dec 2023**

- Developed a POSIX-compliant multithreaded program simulating a load balancer, clients, and servers for efficient graph database access
- Implemented inter-process communication mechanisms, including message queues and shared memory with semaphores, with multi threaded algorithms for Depth-First Search (DFS) and Breadth-First Search (BFS) traversals

## Leadership / Extracurricular

**Entrepreneurship Cell, BPHC**

**Jun 2022 - Jun 2023**

*Vice Chairman*

*Launchpad*

- Headed one of the largest Entrepreneurial Summit in South India, involving over 500 participants from various cities, renowned speakers, internship drive, and national level competitions.