Deepak Kumar Soni

+91 9179687775 | dks91@case.edu | https://glitchdoescode.github.io/myportfolio/

Career Objective

An innovative AI Engineer and Systems Architect with proven experience in leading projects from ideation to deployment. Eager to leverage expertise in multi-agent systems, product architecture, and client-facing roles to build scalable, intelligent solutions in a dynamic and growth-oriented environment.

Technical Skills

Languages: Python, C, C++, Bash/Shell Scripting, JavaScript, Rust
AI/ML & GenAI: Multi-Agent Systems, LangGraph, RAG, NLP, LLMs, Model Fine-tuning, Prompt Engineering, TensorFlow, PyTorch, LangChain, OpenCV
Backend & Frameworks: FastAPI, Flask, Node.js, Django, Streamlit
Cloud, DevOps & Infra: GCP, AWS, Docker, Kubernetes, CI/CD, Terraform, Nginx, gVisor, Celery, HashiCorp Vault
Databases & Vector Stores: PostgreSQL, MongoDB, Redis, Elasticsearch, Supabase, Pinecone, ChromaDB
Systems & Tools: Linux/Unix, Raspberry Pi, Socket Programming, Network Protocols, Git, Wireshark, Neovim

Work Experience

SLOTHPAYS (powersmy.biz) Remote, Bengaluru, Karnataka, India
Founding Engineer *Oct 2024 – Present*

- Architected and managed the development of **SmartPlaybook OS**, a multi-agent decision-making platform using LangGraph, deterministic analytics, and human-in-the-loop governance to automate business workflows.
- Designed the architecture for **Instant-Filtro**, an AI application factory that transforms natural language requirements into fully deployed applications, significantly reducing development time.
- Led the full-cycle technical recruitment process; personally screened over 300 candidates and hired skilled developers to expand the core engineering team.
- Acted as a key technical point of contact in client and advisory meetings, confidently communicating complex architectural concepts and project roadmaps.
- Drove the end-to-end development of an AI-driven interview scheduling automation system, reducing meeting coordination time by 85-90% **(Demo Link)**.

SHUNYA.EK Remote, Goa, India
Software Development Engineer Intern ***(Certificate Link)*** *May 2024 – Sept 2024*

- Spearheaded the creation of a command-line workflow automation tool using Bash, Python, and JavaScript, eliminating 85-95% of manual tasks and boosting engineering team productivity.
- Deployed a Raspberry Pi-based web server using Nginx and Ngrok to enable secure remote SSH tunneling, significantly improving system accessibility and reliability.
- Designed and implemented an automated disk imaging utility in Bash for USB drive flashing with custom configurations, increasing deployment speed and reducing error rates **(Github Link)**.
- Pioneered a real-time video segmentation and motion tracking proof-of-concept using Python and Mediapipa, enabling advanced analytics and visualizations **(Demo Link)**.
- Fine-tuned GPT-3.5-turbo with JSONL datasets and OpenAI APIs, elevating the platform's conversational capabilities for specialized NLP tasks.
- Implemented robust CI/CD pipelines using Git and Terraform for automated model deployments, reducing downtime and accelerating updates.

Volunteer Experience

Red Hen Labs Distributed / Online
Volunteer Researcher/Contributor *February 2025 – Present*

- Co-authored and submitted a research paper, **"Frame Blending Is Essential For Advancing Creative Generative AI,"** to the NeurIPS 2025 Position Paper Track in collaboration with Prof. Mark Turner.
- Developed **TalentScout**, a proof-of-concept multi-agent system using LangGraph and RAG to intelligently match candidates with internship opportunities based on deep contextual understanding.
- Initiated and implemented architectural improvements to the Frame Blender framework, enhancing multi-modal data processing performance and scalability for a distributed research team.

IIT Hyderabad Hyderabad, India

Volunteer Researcher/Contributor *Jan 2025 – Present*

- Prototyping a warzone-focused object detection pipeline with Detectron2 and a ConvNeXt backbone, targeting challenges such as camouflage, partial occlusion, and changing environmental conditions **(Colab Link)**.
- Experimenting with advanced data augmentation (dust, smoke, low-light) to improve model generalization in early-stage R&D, currently limited by a small dataset of a few hundred images.
- Investigating methods for scaling to thousands of images and optimizing real-time inference performance, contributing to ongoing research and development efforts.

Projects

Real-Time Sign Language Detection | *Python, TensorFlow, OpenCV, Mediapipa* GitHub

- Spearheaded the design and implementation of a real-time sign language recognition system utilizing TensorFlow and OpenCV, achieving 95% accuracy in complex gesture detection. Integrated Mediapipa for efficient hand tracking, significantly improving user experience.
- Constructed robust data preprocessing pipelines to augment training data, increasing model accuracy and reducing false positives.

TCP Client-Server Implementation | *C, Socket Programming* GitHub

- Orchestrated the design and construction of a robust TCP client-server model in C based on Beej's Guide to Network Programming, ensuring stable data transmission and fault tolerance.
- Optimized network protocols to enhance data throughput and minimize latency, aligning with demanding high-frequency requirements.

9-Axis Spatial Orientation Tracker | *Python, WebSocket, ADB, Sensor Fusion* Github

- Built a real-time phone orientation tracking system leveraging IMU sensors (accelerometer, gyroscope) and magnetometer data to support advanced spatial analytics.
- Integrated Android Debug Bridge (ADB) with WebSocket communication for live data streaming and near-real-time visualization.
- Implemented sensor fusion techniques to accurately interpret 3D spatial orientation, enhancing reliability and performance.

Certifications

CCNA (Cisco Certified Network Associate) – Mastered fundamental networking concepts including IP connectivity, security, and automation. **(Certificate Link)**

Python Essentials – Covered foundational Python programming from basic syntax to object-oriented principles, as evidenced by Python Essentials 1 and 2 course completions. **(Certificate Link)**

Linux Essentials – Acquired core skills in Linux command-line operations, system administration, and scripting. **(Certificate Link)**

Introduction to Cybersecurity – Gained foundational knowledge of cybersecurity principles, common threats, and defense strategies. **(Certificate Link)**

Achievements

Co-authored and submitted a research paper to the **NeurIPS 2025** Position Paper Track, collaborating with renowned cognitive scientist Prof. Mark Turner. **(Link)**

Maintained consistent development activity with **935+ GitHub contributions** in the last year, demonstrating sustained commitment to software development and open-source contributions. **(Profile Link)**

Selected for the Smart India Hackathon (SIH) 2024, contributing to the development of a Non-Invasive Intraocular Pressure (IOP) Measurement Device.

Participated in the LLM Agents MOOC Hackathon, hosted by UC Berkeley RDI, contributing to the development of 'ScheduleBot', an AI-driven scheduling agent application **(Project Link)**.

Secured a position among the Top 10 teams at The Great India Hackathon - Jabalpur, with a project focused on innovative skin disease detection and prediction.

Advanced to the second round of HackOn with Amazon - Season 3, showcasing strong proficiency in Data Structures and Algorithms.