AMAN SHARMA

Bengaluru, India | +91-9955306911 / +91 6204984854| developeras1999@gmail.com | amansharma.in | LinkedIn | leetcode | GitHub

PROFESSIONAL EXPERIENCE

Salesken.ai – Machine Learning Intern (On-site) | Bengaluru, India *March 2025 – Present*

- ❖ **Fine-tuned** the Mistralai/Mistral-7B model to give better responses format and increased goal detection accuracy by 10%.
- ❖ **Optimized ASR service**, enhancing memory efficiency by **up to 25%** and improving overall processing performance by **15%**.
- ❖ **Integrated a noise cancellation feature** into the core product, leading to a **20% improvement** in audio clarity for users.
- ❖ Conducted **multiple proofs-of-concept (POCs)** for **10-15 LLM models**, evaluating their suitability and performance
- ❖ POCs on **quantized LLM models** within frameworks like Ollama, identifying potential compared to unquantized alternatives.
- ❖ Managed **LLM inference servers** using cutting-edge tools like VLLM and Triton, supporting **high-throughput inference** for critical applications.

Marathon Trends Advisory Private Limited – Software Developer & Research Analyst | Mumbai, India *May 2024 – March 2025*

- ❖ Revamped **backtesting algorithms**, reduced processing time by 67% (from over 20 hours to under 7 hours) through implementation of **memoization, dynamic programming, and caching**.
- ❖ **Spearheaded** the design and deployment of a **full-stack application** for research analysts using **FastAPI and React.js**, to a 50% reduction in manual data processing.
- ❖ Developed a **finance-specific chatbot** based on **RAG models with OpenAI embeddings, ChromaDB, and LangChain**, 40% better response times.
- ❖ Developed a sentiment analyzer service that leveraged real-time scraped data from sources like Google Finance to deliver actionable market insights with 90% accuracy
- ❖ Regularly Constructed a **technical financial database** to manage over 3,000 + stocks data containing essential metrics (e.g., PE ratios, PAT, MC), optimizing data processing by 30% using **Pandas**.

GS Designe- Full Stack Developer | Remote , May 2023 - April 2024

- ❖ Worked on front and backend of global e-commerce clients like(squarepeg) (custoking).
- ❖ Designed and developed frontends for 5+ e-commerce apps using React js and Tailwind Css.
- ❖ Integrated APIs with cross-functional teams, .

EDUCATION

Maulana Abul Kalam Azad University of Technology | West Bengal, India **Bachelor of Technology (B.Tech) in Computer Science and Engineering** *Graduated: 2023 | CGPA- 8.9*

Courses- [**Mathematics for Machine Learning: Multivariate Calculus(coursera)**](#) , [**A Crash Course in Data Science(coursera)**](#),[**AWS Fundamentals: Going Cloud-Native,**](#) [**Cybersecurity and Its Ten Domains(coursera)**](#), [**Interactivity with JavaScript(coursera)**](#)

PROJECTS

- ❖ [**Interview Bot:**](#)Created an **AI-powered speech-to-speech interview bot** using Llama(LLM), PlayAI(TTS), and Whisper(ASR), allowing users to practice interviews for selected roles and receive immediate feedback.
- ❖ **Brain Tumor Detection Application:**Engineered a **Deep Learning application** for brain tumor detection utilizing **Convolutional Neural Networks (CNN)**, achieving an impressive 92% accuracy on MRI scans.(collage project)
- ❖ **OpenChat** :Realtime group chatting application ([link](#))

SKILLS

- ❖ **Programming Languages:** Python, JavaScript, C++
- ❖ **AI & ML Tools:** VLLM, Ollama, Transformers, Hugging Face, Pydub, ChromaDB, LangChain, TensorFlow, Keras
- ❖ **Frontend Development:** React.js, Next Js,
- ❖ **Backend Development:** FastAPI, Django
- ❖ **Deployment Platforms:** Vercel, Netlify, Firebase, VPS(Hostinger), Streamlit, Docker ,Kubernetes
- ❖ **Other Tools:** Kafka , Prometheus ,Grafana ,Pandas, NumPy

ACHIEVEMENTS

- ❖ **First Prize RoboMac 2020:** Created an intelligent interactive dustbin system.
- ❖ **LeetCode Enthusiast:** Solved **600+ coding problems** with over 130,000 profile views, showcasing strong expertise in Data Structures & Algorithms (DSA).