Saife Shuhaib Md.

LinkedIn | +8801766217329 | enansaifme33@gmail.com | saifeshuhaib.com | GitHub

Skills

Python | JavaScript | PyTorch | Tensorflow | LangChain | Django | Prompt Engineering | RAG | Supabase | ONNXRuntime | FastAPI | Docker | Git | Generative AI | Natural Language Processing (NLP) | Deep Learning | Computer Vision | Google Cloud Platform (Vertex AI, BigQuery) | SQL | PostgreSQL | REST APIs

Experience

Software Engineer **REEA Digital Limited** **01/2024 - Current**

● Developed a SaaS application for generating meeting transcripts and file notes. Enabled querying of meeting insights with accurate citations via Perplexity. Improved query performance, reducing response time from 1.5 minutes to 30 seconds. Optimized vector query performance by 30%, through query augmentation.
● Developed a custom AI chatbot leveraging GPT-4 and Retrieval Augmented Generation (RAG) using LangChain, FastAPI, OpenAI, and Google Cloud Platform, improving document query response time by 50%.
● Led research into optimizing LLM prompts, integrating OpenAI's models with Google Cloud Platform to create efficient, scalable solutions.
● Utilized ONNXRuntime and advanced image processing for a real-time comic book text translation system serving 50k users, reducing translation errors by 85%.

Software Engineer **BJIT Limited** **11/2021 - 12/2023**

● Developed a proprietary AI automation tool using computer vision and machine learning techniques, which reduced manual QA efforts by 80%.
● Developed an AI-based document extraction solution using FastAPI and Nanonets, automating data extraction from PDFs, reducing manual work by 90%.
● Led the design and development of a face analysis application providing real-time detection and landmark information. This internal tool, now used by 700 employees, saw a 40% performance boost through parallel processing and frame skipping techniques. Built using OpenCV, and Intel OpenVino.
● Developed an Android face recognition app with fully homomorphic encryption. Improved performance by reducing file size by 93% and achieving inference times under 1 second, using C++, Microsoft SEAL, and Dlib.
● Created a Food Detection API using a custom-trained Yolov5 model to accurately identify 13 types of Japanese food. Utilized Python, Flask, OpenCV, with image processing times under 200 ms.

Education

Bachelor of Science **Chittagong University of Engineering and Technology** **02/2016 - 06/2021**

● Electronics and Telecommunication Engineering | CGPA: 3.58/4.00

Projects

● Chess AI web app: Developed a Chess AI web app using python-chess, chessboard.js, pytorch, and Django, featuring AI opponents powered by the Classic MiniMax Algorithm and Convolutional Neural Network (CNN).
● Snake AI: A snake bot developed using PyTorch and reinforcement learning, utilizing a fully connected neural network.

Achievements & Leadership

● Ranked among the Top 5 teams in Robi Datathon 2.0
● Vice President, CUET Photographic Society