PRASAD P R

+919740146523 | shenoy.prasad20@gmail.com | Bengaluru, Karnataka, India | LinkedIn | GitHub

SUMMARY

Innovative AI Engineer with 8+ years of experience in developing AI applications using Python, OpenVINO, Pytorch, and TensorFlow. Skilled in optimizing models and building scalable, production-ready AI systems for customer-facing products. Proficient in Machine Learning Methodologies, Text, Predictive Modelling, and Data Analysis. Experienced in Generative AI systems, RAG and Langchain based solutions with a proven track record in delivering impactful solutions.

PROFESSIONAL EXPERIENCE

AI Software Development Engineer, **Intel Corporation**   *Jun '22 — Present*

- Model Benchmarking and Optimization Achieved a 10x performance enhancement on Intel GPU by optimizing transformer models, Whisper and M2M100, on Intel platforms.
- Implemented an automated pipeline for inference and benchmarking across multiple Intel platforms, reducing the benchmarking time per machine from 5 hours to 30 minutes.
- Optimized model performance by enhancing command line Report Generation with ONNX runtime and OpenVINO tools, reducing analysis time by 50% from days to minutes 3D Based Camera Calibration and Mapping.
- Developed a Real-Time 3D Camera Calibration strategy, reducing inference time by 80% using Python, PyTorch, Open3D, and Docker.
- Enhanced AI models, boosting training and inference speeds by 30%, and incorporated the solution into an IoT-enabled device with features for MQTT and video feed operations.
- Created a real-time visualization application and an efficient inference pipeline for segmentation and labelling of point clouds utilizing Python, Intel RealSense Sensor, PyTorch, and Open3D, leading to a more than 50% decrease in inference time.

System Software Development Engineer, **Intel Corporation**   *Aug '18 --- Jun '22*

- Led the development of a MERN stack Solution Delivery System, that garnered over 1000+ downloads.
- Improved access time by over 30% through remodelling the architecture from monolithic to microservices based-architecture.
- Guided the development of a management portal, integrating CICD tools, reducing maintenance time by 80%.

INTERNSHIPS

Software Engineering Intern, **Intel Corporation**   *Oct '17 --- Aug '18*

- Contributed code to 2 repositories in the GitHub Archives program as a part of Hyperledger Sawtooth Enterprise Blockchain.
- Developed a UI on Angular JS for POC concept on visualization of blockchain nodes and their transactions.
- Built a POC to gather data for a Proof of Concept (POC) based on Stereo Vision and Computer vision.

EDUCATION

Liverpool John Moores University - **Liverpool** & upGrad   *Apr '25 -- Present*
Master of Science in Artificial Intelligence and Machine Learning

International Institute of Information Technology - **Bangalore** & upGrad   *Feb '24 — Apr '25*
Executive Post Graduation in Artificial Intelligence and Machine Learning   *3.63/4 CGPA*

JSS Academy of Technical Education, Bangalore   *Aug '14 — Jul '18*
Bachelor of Engineering (B.E.), Computer Science & Engineering   Grade – Distinction

KEY PROJECTS (Beyond Work Contributions)

Find Me A Cube (Intel Internal tool)
Objective: No cost solution for identifying empty cubes in a Hybrid workspace environment. Also provide cube usage analytics. **Solution:** Cloud hosted solution using Django REST API as a backend, React JS based frontend and Mongo DB database. **Result:** Faster, hassle-free cube identification for employees with a clean frontend. Analytics that help the building management plan and act on any issues associated with the cube management.

Automated Smart Reports using Generative AI (Intel Internal Tool)

Objective: Speed up report collation, reduce manual report generation based on multiple mails. **Solution:** Using Gen AI & prompt engineering, extract required information from multiple reports and combine the information into a single status report. **Result:** Reduced the managers time spent on report generation from 5 days to 30 minutes.

Smart Hiring and Resume Screening using Generative AI (Intel Internal Tool)

Objective: Help hiring manager screen suitable resumes based on the job description required for the role. **Solution:** Using LLM models, understanding the data of the resume and then trying to perform a similarity mapping with the job description to obtain the top resumes that are suitable for the job. **Result:** Helped managers reduce the resume screening time from hours to minutes.

Finance and Stock Markets (Personal Project)

Objective: Fetch the financial details of a company based on their ticker and perform the fundamental analysis. **Solution:** Using Agentic AI, fetch the company data and using Python pandas perform fundamental analysis of the company to suggest if the company stock can be purchased or not. All this is performed using phidata module and **agentic AI**

AWARDS & CERTIFICATIONS

Divisional Recognition Award (DRA)	Intel
Innovation Star	Intel
Deep Learning With Pytorch Image Segmentation	Coursera
LangChain- Develop LLM powered applications with LangChain	Udemy
Generative AI with Large Language Models	Coursera

SKILLS

Programming Languages: Python, JavaScript
Machine Learning Techniques: Deep Learning, Generative AI, LLM, NLP
Frameworks and Libraries: Pytorch, TensorFlow, Huggingface, OpenVINO, ONNX
Data Processing and Analysis: Computer Vision, Image and Video Processing, Quantitative Analysis, Data Visualization Techniques, NoSQL
Software Development: Application Design, Microservices, Git, Docker, Kubernetes