Tarun R Jain

jain.tarun7501@gmail.com | +919986197355
github.com/lucifertrj | tarun.netlify.app | linkedin.com/in/jaintarun75

Work Experience

AI Planet, Luxembourg April 2023 - Present

Data Scientist and Head of DevRel

- I have worked on deploying state of art model at AI Marketplace and further worked on **Large Language models** and **Agents** for customer POC projects in **Education**, **Banking**, **Finance**, and **CPG** industries.
- Regarding the Open Source LLMs contribution I have built **Panda Coder 13B** which has over **25K+** overall downloads on **HuggingFace** and also **AI research project lead** for **Buddhi-128K chat model** that is one among the first open-source **128K context LLM** with over **3K+ overall downloads**.
- I am the lead maintainer, contributor, and reviewer for **2** of our **Open Source** repositories: **BeyondLLM**: An open-source library that enables you to build advanced RAG pipelines, embeddings, and **LLM evaluations** in just **5-8 lines of code** and **OpenAGI**: An **Autonomous Multi-Agent framework** with **advanced long-term memory integration**.
- AI Planet is a global community of 300K+, I maintain **8200+** community on AI Planet Discord Server, moderated **30+ live sessions**, collaborated with **IIT Guwahati** & **IIT Bombay** for Hackathons and hosted **2 Generative AI Bootcamps**.

Google Summer of Code 24, Red Hen Lab May 2024 - August 2024

Contributor

- I developed a multilingual news dataset for fine-tuning LLMs, encompassing over **64,000** question-and-answer pairs across **6** languages, including English, Spanish, French, German, and Portuguese as per **Self Instruct framework**.
- I also optimized the fine-tuning and inference processes for multilingual LLM using **PEFT** and **Flash-Attention 2**.
- Our 3B LLM model built during GSoC at RedHen Lab outperforms **Gemma-7B**, **Mistral-7B**, **Llama-3-8B**, and other models on the **HuggingFace** OpenLLM leaderboard. Read full article: news-reporter-3b benchmarks.
- Upon the completion of the model training, we inference the LLM on upcoming news data using the **Retrieval Augmented Generation** pipeline along with guardrails implementation and **RAG evaluation**.

CCCIR, Bangalore August 2022 - January 2023

Research Intern

- Worked on a project related to **Hyperspectral Satellite Images** on **FPGA** for the **ISRO-funded scheme**.
- I wrote a paper on how to interface **Intel's OneAPI** for **Object and Text Detection** on **Raspberry PI** and improved the performance throughput **8x** compared to CPU.
- Contributed to various **AI/ML projects** that are based on Object Detection. One project where I was lead consisted of **Crop Tracking** using **DeepSort** and **TensorFlow** and **AI based License plate detection and Violence notification automation**.

Education

Cambridge Institute of Technology, Bangalore Aug 2019 - Jun 2023

B.E. in Computer Science and Engineering ***CGPA: 8.9/10***

Relevant Coursework: Majors in Image Processing and FPGA VLSI. I have been awarded **Best project presentation** for two consecutive years at **Intuit** competition and won the **Best outgoing student award** for 2019-2023 batch

Project Work

- **OpenAGI** OpenAGI helps you build Autonomous Multi-Agent architecture. I have added the support for the autonomous task creation and task execution of the Agents via ReACT prompting. OpenAGI consists of 10+ custom action tools integration implemented by me. You can build advanced Agentic Workflows using 5+ different LLMs such as **Mistral**, **Gemini**, **Azure**, **OpenAI**, **Ollama** and so on.
- **BeyondLLM** BeyondLLM helps you build an **advanced RAG pipeline** and LLM evaluation with 8-19 lines of code. I have added 7+ custom LLM inference implementations, evaluating the retriever pipeline metrics i.e., **hit rate** and **MRR** and maintaining the entire repo.

Awards and Recognition

- **Google Developer Expert** in **AI/ML** Helping students and working professionals to get better at AI, Agents, RAG and LLMs. Delievered **40+ talks** including **7 International APAC talks** and **35+ talks** within India.
- Content Creator with **2.47K+** subscribers on YouTube channel: AI with Tarun. Mostly Agents and RAG videos.
- **Won 6+ Hackathons** including one international hackathon in **Generative AI** space.
- **Visiting faculty** at **MS Ramaiah** - **Top #1 Engineering** in **Bangalore** to conduct **15 hours** of course material Deep learning using **TensorFlow** and **PyTorch**.