Yuvraj Singh

Github Website Linkedin yuvraj.mist@gmail.com HuggingFace
+919354672378

Experience

IISER, Kolkata | *Summer Research intern* May 2024 – Present

- Working under Prof. Kripabandhu Ghosh on exploring Fine-tuning LLMs and Using GenAI Models on Medical and Finance Data for QnA

AIISC | *Research intern* March. 2024 – July 2024

- Worked with Prof. Amitava Das on Prevention of Hallucination in LLMs
- Part of the data collection team- Performed web scraping of news articles and posts. Generated synthetic data using various open-source LLMs.
- Tagged LLM-generated synthetic data with entities using GLiNER and other open-source models.

Clinical AI Assistance | *Research intern* Dec. 2023 – March 2024

- Worked closely with the Research Team for the main project, building dataset and testing the model's generated output with other Text Generative tools like GPT etc.
- Used HuggingFace API to use various LLMs to test our generated response to the original ones
- Data Gathering and cleaning and performing Data Analysis on the generated output to view model's efficiency

Skills

Languages: C++, Python
Frameworks: Pytorch, Keras, Flutter, Flask
GenAI Tools Langchain, LlamaIndex, Chromadb, Pinecone, FAISS, HuggingFace
NLP: GRU, RNN, LSTM, Bi-LSTM, Transformers, LLMs
Computer Vision: CNN, OpenCV, YOLOv8, GANs, ViTs
Tools: Streamlit, Git, Github, Firebase, Docker, GCP

Courses

Machine Learning Specialization by Andrew Ng: Coursera
Generative Adversarial Networks (GANs) Specialization: Coursera
Deep Learning Specialization by Andrew Ng: Coursera

Projects

Paper/Architecture Replication from Scratch using PyTorch | *GitHub Repository* 2024

∗ **Tools**: Pytorch, HuggingFace Spaces (for deployment of smaller versions of replicated models), Git, Github

FarmGenie (GeoHack 2024) | *Empowering farmers with real-time insights and expert guidance via AI-driven space* 2024

∗ **Role**: Backend (AI/ML Integration)
∗ **Vision/Goal**: In the ever-evolving agricultural landscape, farmers often face challenges in accessing up-to-date knowledge and resources to improve their farming practices. To address this, we have developed a comprehensive product that leverages the power of LLMs as Experts and Agents to create an interactive platform for farmers.

* **Solution**: Our platform utilizes LLMs and a Mixture of Expert (MoE) approaches to provide precise guidance on soil management, plant disease identification, and irrigation techniques. Built as a scalable web application with a Next.js frontend and backend, and supported by a Redis queue and multiple worker nodes, FarmGenie ensures robust performance. The system's multilingual support, interactive community forum, and up-to-date knowledge base facilitate seamless, expert-driven assistance for both new and experienced farmers.
* **Tools**: Tailwind CSS, TypeScript, Next.js, Flask, Python, Langchain, GeminiPro API, peft , bitsandbytes, transformers, PostgreSQL, FAISS, Pinecone, MySQL, Unsloth, GCP, Docker, Vercel, DigitalOcean

Insight-Ed (HackNITR 5.0) | *EdTech Platform for Student and Teacher* 2024

* **Role**: Backend (AI/ML+Server)
* **Vision/Goal**: Our solution tries to bridge the knowledge gap between a teacher and a student in online classes by leveraging the power of AI to provide the teacher and student with insights to better prepare themselves for the course/lecture.
* **Solution**: Implemented features such as- Detection of Emotion and Attentiveness of a Student for better understanding of student's understanding of a topic, Topic Modelling to know exactly where he/she lost attention, QnA with uploaded PDFs (Handwritten/Printed) along with any Videos/Tutorials and Reverse Video Search to quickly look for any concepts in the same.
* **Tools**: Flask, Python, Gemini API, Whisper API, HuggingFace Open-Source Models (EmotionLLM), MTCNN, LangChain, Cloud Run, GCP(VM)

MoviesMania | *A Reverse Search based Movies Recommendation System(Geek-o-thon @IIIT-BH)* 2023

* **Role**: Team Lead
* **Vision/Goal**: To provide an interface to users to find similar movies/web-series recommendations based on an uploaded video clip/YT Short.
* **Solution**: To make use of the various faces of actors in the provided clip and the details provided(title, genre, plot) for the prediction of the movie's title (if available in the dataset) or similar movies/web series.
* **Result**: Prediction of movie's title(if available in the dataset) with 78% accuracy and similar movies with 85% accuracy.
* **Tools**: Keras, Tensorflow, Word2Vec (Word Embeddings), MTCNN, NLTK, Spacy, VGGFace, OpenCV, Streamlit, Render

Achievements

2nd place @GeoHack '24 Finale 2024
Secured 2nd position at finals of GeoHack 2024 with our project- FarmGenie organised by IEEE GRSS Kolkata and SAADRI.

Finalist @YESIST12 '24- Special Track 2024
Led the winning project PlogPayouts at YESIST12 '24 under Special Track category

Geek-o-thon (D3 @IIIT-BH) 2023
Led the winning project MovieMania at D3 2023, inter-college hackathon in AI/ML domain

Education

International Institute of Information Technology Bhubaneswar 2023- 2027
BTech Computer Science Engineering

Delhi Public School 2022-23
CBSE Grade 12 *Percentage: 91*

Amity International School 2021-22
CBSE Grade 10 *Percentage: 96*