PRATHMESH GAJANAN RAUT

Mumbai,Maharastra

📞 +91-750683144 ✉ prathmeshraut150904@gmail.com 💼 LinkedIn ⭕ GitHub Ⓗ HackeRank

EDUCATION

Dwarkadas J Sanghvi College of Engineering, Mumbai Oct 2022 – May 2026
*B.Tech - **CGPA - 9.20 (Top 1%)*** *Mumbai, India*

S.K Somaiya College Sep 2020 – Aug 2022
*Maharashtra State Board Exam - Higher Secondary - **Percentage - 85%*** *Mumbai, India*

TECHNICAL SKILLS

Languages: Python, Java, C, C++, JavaScript
Machine Learning Tools: Tensorflow ,PyTorch , Scikit-Learn ,Open-CV, Mediapipe ,Transformers ,Stable Diffusion, Yolov8/7,Hugging Face, gym(Open ai Gym) , Amazon Sage Maker,Pillow
Natural Language Processing: Langchain ,Llama Index,nltk,Scapy,Llama ,Gemini(pro,pro-vision,flash),Eleven Labs,DeepGram,Crew AI,Groq,Replicate,Assembly AI,Open AI(API)
Data Science: Tableau,Sqlite,BeautifulSoup, Selenium ,Numpy ,Pandas,Matplotlib,Seaborn
Database: LanceDB(Multi Modal Retrieval),Neoj(Graph Vectore Databse),FAISS DB,PineCone,ChromaDB,VectorDB,MySQL
Web Development: ReactJS,NodeJS,Express,MongoDB,Flask,Tailwind CSS,Bootstrap,Firebase,SASS

Experience

AI Intern ChumsAI↗

Created an AI Human which used 3D avatar to respond in Real Time *Mumbai, India*

- Developed a 3D avatar system for real-time interaction, enhancing user engagement and experience.
- Implemented a Retrieval Augmented Generation (RAG) pipeline, integrating advanced AI to improve response accuracy and relevance.
- Created vision capabilities for Large Language Models (LLMs), enabling avatars to visually perceive and interact with their environment.
- Addressed and reduced LLM hallucinations, increasing the reliability and trustworthiness of AI-generated content.
- Successfully developed and delivered three major client projects, leading to strategic partnerships and client meetings.

Championed 10 Hackathons; secured victories in 2 by collaborating with Companies work↗

Hackathon Participant / Winner *Mumbai,*

- Demonstrated strong technical and collaborative skills across multiple competitive hackathon events.
- **Achievements: Winner of *HackHeaven National Level Hackathon* and *M Indicator Hackathon*.**
- **Additional Hackathons Participated**: KnowCode (36hr),Aeravat S.P.I.T(24hr), Recursion 5.0(24hr), ACE of Hacks (24hr), Quasar 2.0(24hr),Data 2 Knowledge (6hr),Codiessaance (24hr),HakXite(24hr)

Projects

Coded Image Generation From Scratch ↗ | CLIP,UNET,VAE,PyTorch,DDPM June 2024

∗ Developed a Stable Diffusion-based text-to-image model, integrating a UNet generator with self-attention and cross-attention mechanisms for detailed image synthesis. Enhanced the model with CLIP for accurate text-image alignment and a Variational Autoencoder (VAE) using ELBO loss to optimize image diversity and quality. Incorporated upsampling techniques and temporal embeddings to improve image resolution and temporal consistency. Managed the full project lifecycle using Python and PyTorch, focusing on scalable and maintainable code.

Self Driving Car using RL ↗ | PyTorch, Deep Q Learning, AI Agents June 2023

∗ Developed a 2D self-driving car simulator using Deep Q-Learning for autonomous navigation. The system incorporates a neural network for decision-making, Experience Replay for efficient learning, and a softmax policy for optimal action selection. Employed Kivy for UI development, enhancing the interactive visualization of the car's performance.Also Tried to work on Convo Q Learning

Anamoly Detection [↗] | Masked R-CNN, Transfer Learning,Open CV   **Nov 2023**

* Created a real-time violence detection system using the Masked R-CNN framework, an interesting choice given its strengths in instance segmentation which can provide detailed context by differentiating individual objects or persons in an image. By not relying on YOLO, which is typically faster but less precise in distinguishing overlapping objects, you've prioritized accuracy in identifying actions and participants involved in violence.

Video Format Story Generator [↗] | Stable Diffusion, NLP   **Jan 2024**

* Architected and implemented a cutting-edge model that dynamically generates videos from user-provided prompts and crafts narratives through text inputs like character names, types, and themes. Enhanced processing speed by 20% through optimization on a T4 GPU with CUDA, ensuring rapid content delivery and enhanced performance.

Research Paper Studied

Segment Anything By Meta AI [↗] | Image Encoder,Prompt Encoder,Masking,Decoder   **June 2024**

* Studied the use of Masked Vision Transformer encoder which is use as Image encoder to SAM models,Pompt Encoder for text,boxes and points,Converts the points in vectors of 256 dimension and which acts as an input of Prompt encoder in combination with output tokens which is similar to class token of BERT models.for the text Part in Prompt Encoder it used CLIP.It consist of 2 loss function one is dice loss and other is IOU loss

META AI Llama 2 [↗] | Rotary embeddings,RV Cache,Grouped Query Attention   **June 2024**

* Understand the use of RV cache to store he dot product of RV for decreasing computation,Importance of CLS token,Group Query Attention,Use of SwiGLU activation function,RMSnorm and the concept of Rotary Positional Encoding which are added for absolute relative distance.

LongNet: Scaling Transformers to 1,000,000,000 Tokens [↗] | **June 2024**

* Reduced complexity using Sliding Windows and Sparse Attention.
* Developed Efficient Transformer Variants for faster processing.
* Tackled Scaling Challenge for sequences up to 1 billion tokens.
* Introduced Dilated Attention for short and long-range dependencies.
* Implemented Mixture of Dilated Attentions with Dynamic Weights.

Quantization-Aware Training [↗] | Stable Diffusion, NLP   **June 2023**

* Explored Quantization techniques to enhance neural network efficiency, beginning with an introduction to numerical representation of integers and floating points. Delved into key concepts like Asymmetric and Symmetric Quantization, and Quantization Range. Focused on Dynamic and Static methods, particularly Post-Training Quantization and Quantization-Aware Training, achieving significant performance improvements.

EXTRACURRICULAR

DJS Nova Tech Head [↗]   **Jul 2024 – Present**
AI Head   *Mumbai*

* Conduct Research with the space project to apply machine learning on space research papers

DJS Computee [↗]   **Jan 2024 – Present**
Co Commitee Member   *Mumbai*

* Work on Machine Learning Projects for fulfilling client order and their Request

CERTIFICATIONS

• Hackathon Achievement Certificate [↗]   • Convolution in Depth Course - Udemy[↗]