B. Mahesh Raj

+91-8921662922
bmaheshraj23@gmail.com
cb.en.u4cse22006@cb.students.amrita.edu
GitHub : B. Mahesh Raj
LinkedIn : B. Mahesh Raj

Bachelor of Technology
Computer Science and Engineering
Amrita Vishwa Vidyapeetham, Coimbatore

About

I am a Computer Science undergraduate with a strong foundation in software development and artificial intelligence. My interests span deep learning, computer vision, generative models, and data-driven problem solving. I am passionate about applying technology to build impactful solutions and continuously expanding my knowledge across emerging areas in computer science.

Education

Degree	Institute	Board / University	CGPA/Percentage	Year
B.Tech CSE	Amrita School of Computing	Amrita Vishwa Vidyapeetham, Coimbatore	8.41 (Till 6th Sem)	2022-2026
Senior Secondary	St. Peter's Senior Secondary School	CBSE	86.2%	2022
Higher Secondary	St. Peter's Senior Secondary School	CBSE	89.8%	2020

Experience

- **Amazon** *May 2024 – July 2024*
  SDE Intern — Finance Automation Bangalore, India

  – **Finance-Document VQA and Field-Value Extraction POC**
  Tech: Python, PyTorch, Amazon Bedrock, SageMaker, Textract, LoRA/QLoRA, Knowledge Distillation, Quantization
    * Built a prototype for **Visual Question Answering (VQA)** and **field-value extraction** on tax documents using LLMs/VLMs.
    * Boosted accuracy from **80–85%** to **99% overall**, with **97%** success on critical fields.
    * Gained hands-on experience in **prompt engineering** (zero-/few-shot, chain-of-thought), **parameter-efficient fine-tuning with LoRA/QLoRA**, **knowledge distillation** for compressing large teacher models into lightweight students, and **quantization**.

  – **Data Pipeline Monitoring System**
  Tech: AWS Lambda, Athena, DynamoDB, S3, EventBridge, CloudWatch, Secrets Manager, SQS/SNS, Docker
    * Designed a scheduled monitoring service that queried multiple data sources to check for missing/late records.
    * Implemented automated alerting to pipeline owners on discrepancies using **CloudWatch metrics, dashboards, and alarms**.
    * Ensured reliability and data integrity across finance automation pipelines.

Research Experience

- **Estimating Soil Moisture from Satellite Data** *March 2024 – July 2025*
  ***Tech**: PyTorch, TensorFlow, Excel, Pandas, Matplotlib* Amrita Vishwa Vidyapeetham, Coimbatore
  – Collaborated with INRAE, France under the guidance of Dr. Amit Agarwal, Professor, TIFAC-CORE in Cybersecurity.
  – Engaged in research involving remote sensing, agriculture, and machine learning.
  – Leveraged advanced satellite data processing and ML techniques to address challenges in soil moisture prediction and plant life cycles.

- **Text-Prompted 4D Mesh Character Animation using GNNs and Diffusion Models** *November 2024 – August 2025*
  ***Tech**: PyTorch, PyTorch Geometric, Trimesh* Amrita Vishwa Vidyapeetham, Coimbatore
  – Contributed to research on 4D mesh generation using Graph Neural Networks (GNNs) and diffusion models as part of a professional research elective.
  – Developed a latent graph diffusion model to overcome limitations in handling meshes with varying topologies.
  – Designed a pipeline integrating GNN autoencoders with diffusion models for text-prompted 4D mesh generation.
  – Explored applications in animation and game development by creating a versatile, generalizable approach to dynamic mesh generation.

- **Lightweight Student Network for NNUNet** *September 2025 – Present*
  ***Tech**: PyTorch, nnU-Net, Knowledge Distillation, Deep Supervision* Amrita Vishwa Vidyapeetham, Coimbatore
  – Developing a multi-stage **model compression pipeline** for NNUNet aimed at reducing parameter count, memory footprint, and inference time while maintaining performance.
  – Employing a **multi-phase knowledge distillation** strategy: first at the **feature level**, then at the **soft-label level**, guided by deep supervision.

– Designing preprocessing mechanisms to emphasize **contrast-enhanced regions in DCE-MRI scans**, narrowing the search space and simplifying learning.
– Ultimate goal: deliver a compact, efficient NNUNet variant suitable for clinical deployment in low-resource environments without significant accuracy trade-offs.

- **Final Year Project — Self-Driving Cars with Small Language Models** *September 2025 – Present*
  ***Tech**: Qwen-0.5B LLM, Multimodal Encoders (LiDAR + Multi-axis Camera), Edge AI* Amrita Vishwa Vidyapeetham, Coimbatore
  – Experimenting with **edge-focused autonomous driving solutions** using a lightweight Qwen-0.5B decoder paired with encoders for multimodal inputs.
  – Designed pipelines for real-time **waypoint prediction, scene understanding, and object detection (bounding box prediction)** from LiDAR and camera data.
  – Exploring two architectures:
    * **Parallel multi-encoder design** — separate encoders for LiDAR and multi-axis camera, integrated via the decoder.
    * **Single fusion encoder** — fusing LiDAR and camera into a unified encoder trained on RGB–point cloud data, reducing compute overhead and improving inference speed.
  – Implementing a **safety system** to predict future states of surrounding objects, ensuring generated waypoints are safe to execute — addressing gaps in many current systems.
  – Entire solution optimized for **local, real-time inference** to enable closed-loop autonomous driving on resource-constrained edge hardware.

Projects

- **Deep Fake Detection** *Jan 2024 – Apr 2024*
  *Tech: **PyTorch, OpenCV*** **GitHub**
  – Developed a video deepfake detection system leveraging multiple detectors with unique methodologies.
  – Built an ensemble framework where results were intelligently combined by a meta-model that assigned weights based on each detector's historical performance on a custom dataset.

- **DDPM Image Generation — Deep Learning Coursework** *2024*
  *Tech: **PyTorch, Denoising Diffusion Probabilistic Models (DDPMs)*** **Demo Repository**
  – Implemented a **DDPM-based generative model** for image synthesis as part of deep learning coursework.
  – Demonstrated the underlying techniques behind **deepfake generation**, enabling fine-tuning on a handful of images of a target person to produce new, realistic samples.
  – Showcased diffusion-based generative modeling and its applications in media synthesis and AI ethics demonstrations.

- **Adobe India Hackathon — Team Starks (Connecting the Dots)** *Jan 2025*
  *Hackathon Project / Tech: **Qwen2.5-0.5b (Int8, llama.cpp), YOLOv8n*** **GitHub**
  – Built an intelligent, lightweight, **CPU-only offline system** to transform static PDFs into dynamic, structured, persona-aware knowledge artifacts.
  – Designed a **layout-aware Small Language Model (SLM)** using **Qwen2.5-0.5b (Int8 quantized)** on **llama.cpp** for efficient low-resource inference.
  – Integrated **2× YOLOv8n models** (distilled with PP-DocLayout-L + a custom outline detector), **SentenceTransformers** for semantic search, and **K-means clustering** for hierarchical structuring (H1, H2, H3).
  – Enabled **semantic retrieval + summarization** by ranking the top 5 relevant sections in embedding space, then summarizing them via the SLM.
  – Optimized the pipeline to meet hackathon size limits (**200 MB in Round 1A**, **1 GB in Round 1B**), achieving high portability and efficiency. The system processed 10–15 documents of 50 pages each in under one minute.

- **Fire Fighting Drone for Early Forest Fire Detection and Extinguishment** *2018*
  *School Project / Tech: **Drones, Sensors, Fire Suppression Systems***
  – Developed a drone capable of early forest fire detection and suppression, with integrated surveillance and rapid response mechanisms.
  – **Achievements:** CBSE Science Fair State Finalist, PPTIA Innovation Award National Finalist (Top 10), First Prize at Shastra Science Fair.

Technical Skills

- **Programming Languages**: Python, Java, C++, C
- **DeepLearning Frameworks**: Pytorch, Pytorch3D, TensorFlow, Scikit-Learn,
- **Data Analytics Tools**: NumPy, Pandas, Matplotlib, Seaborn
- **Image processing Libraries**: MediaPipe, OpenCV
- **Cloud Platforms**: AWS (Lambda, S3, CloudWatch, Athena, DynamoDB, Bedrock, Sagemaker)
- **LLM Techniques**: Prompt Engineering, Model Fine-tuning, Quantization, Knowledge Distillation
- **Accelerated Computing**: CUDA programming for GPU-accelerated deep learning and high-performance computing

Certifications

- **The Joy Of Computing Using Python || NPTEL IITM || ELITE**

Achievements

- **1st place in Nestria Jan Built Hackathon** ,Built a deep fake detection website, **Link** *April 2024*