Madhav Kataria

Jodhpur, Rajasthan 342030 | b23ch1025@iitj.ac.in | Website | Linkedin | Github

Education

Indian Institute of Technology Jodhpur , B.Tech in Chemical Engineering 2023 - 2027

- CGPA: 7.76/10.0
- **Coursework:** Introduction to Computer Science, Introduction To Machine Learning,Data Structure and Algorithm, Probability, Statistics & Stochastic Process

U.S.M. Sr. Sec. Public School, Senior Secondary 2023

OP Jindal Modern School, Secondary 2021

Industry Experience

SWE Intern Belongg AI

June 2024 – July 2024

- Optimized systems across Edtech, Fintech, and Healthtech using advanced algorithms and NLP, boosting performance by 30%.
- Debugged 1500+ lines of code daily, reducing bugs by 40% and ensuring robust data validation.
- ERevamped UI with interactive dashboards, driving a 15% increase in customer satisfaction and product insight.

Academic Experience

VisFake: Explainable Vision-Language Model for Fake Image Detection IAB Lab(IITJ)

September 2024 – Ongoing

Advisor: Dr. Mayank Vatsa

- Developed an advanced Vision-Language Model (VLM) for high-accuracy fake image detection with interpretable reasoning.
- **Dataset**: Curated and synthesized a diverse dataset using outputs from 6-7 different VLMs, ensuring robust training and generalization.
- Enhanced image authenticity verification by leveraging the combined strengths of multiple VLM architectures, improving reliability and interpretability.
- **Tools**: Python, VLM's, Fast API

Projects

Dynamic Agentic RAG RAID (AI Club IITJ)

September 2024 – December 2024

- Developed a real-time RAG pipeline with dynamic crewAI agents ingesting streaming Google Drive documents, delivering analysis in < 40 s/query.
- Integrated hybrid retrieval (Pathway vector store), Jina AI fallback and ethical guardrails to achieve 95%+ answer relevance and +30% system robustness.
- **Tools**: Docker, FastAPI, Pathway, crewAI, OpenAI API, Gradio, Google Drive Connectors.

Personalized Image Editing RAID (AI Club IITJ)

September 2024 – December 2024

- Designed and deployed a text-prompt-driven generative AI system using GANs, VAEs, and Diffusion Models, enabling intuitive image edits (e.g., converting buildings into royal palaces, adding dynamic elements like "cats holding secret meetings") without traditional software.
- Built an end-to-end pipeline in PyTorch, including data curation (self-collected datasets), model training, and inference deployment, reducing manual editing time by 70% for complex tasks.
- Enhanced image authenticity verification by leveraging the combined strengths of multiple VLM architectures, improving reliability and interpretability.
- **Tools**: GANs, VAEs, Diffusion Models, NumPy, Matplotlib, Git, Fast API.

Intelligent Football Agents Using Deep Q-Networks RAID (AI Club IITJ)
January 2024 – July 2024

- Implemented Deep Q-Networks (DQN) to train two agents to play football from scratch.
- **Achievements**:
  - Designed and implemented a reinforcement learning system, empowering agents to independently learn game dynamics and strategies, leading to a 40% enhancement in strategic decision-making accuracy and gameplay efficiency.
  - Validated the effectiveness of DQN by achieving a 85% success rate in training intelligent agents to master complex football strategies.
  - Presented 3D simulation using the UNITY engine, providing a visual representation of agent behaviors and strategies.
- **Tools**: Unity Engine, PyTorch, TensorFlow, DQN, PPO

Competitions

Adobe Challenge – Image Classification & Artifact Identification Inter IIT Tech Meet 13.0
November 2024 – December 2024

- Developed a multi-head CNN model for detecting AI-generated images (85% accuracy on 32×32 low-resolution images) and fine-tuned a vision-language model (VLM) to detect artifacts while providing reasoning for image authenticity.
- **Dataset**: Curated a 500,000+ image dataset from BigGAN, Stable Diffusion, StyleGAN, and CIFAKE. Used GPT-4o to generate artifact-based reasoning for fine-tuning the VLM with LoRA adapters.
- Implemented adversarial training (FGSM & PGD) to enhance model robustness against perturbations.
- **Tools**: Python, Vision-Language Models (VLMs), GANs, KANs, VAEs, LoRA, Knowledge Distillation.

AI Agent for Domain-Specific QA Inter IIT Tech Meet 12.0
December 2023 – March 2024

- Developed multi-task learning frameworks with GPT-4 and advanced prompt engineering; boosted model efficiency by and achieved reduction in training times across diverse data sets
- Designed prompts and used Chain of Thought (CoT) strategy to improve AI capabilities, resulting in a 40% reduction in support ticket resolution times.
- Forecasted annual savings of over $2.2 million through increased efficiency.
- **Tools**: GPT-4, prompt engineering, CoT, multi-task learning, model distillation.

Additional Experience And Awards

- **10th in Tech Meet**: Awarded 10th place in Inter IIT Tech Meet 13.0
- **BCCA UG Rep**: Managed all technical clubs under BCCA
- **Robotics Core Member**
- **13th in KGP Data Science**: Ranked 13th out of 9000 participants in KGP Data Science Hackathon

Technologies

Languages: C++, C, TypeScript, Swift, JavaScript, Python

Tools & OS: Unity engine, Git, Jupyter Notebook, Google Colab, Linux, macOS, Windows, MS Office, PostgreSQL, ROS

Software: NextJS, React, Fast API, Streamlit, XCode, Git, Flask

Embedded Systems: NVIDIA Jetson, Arduino, NodeMCU, Raspberry Pi

Machine learning, Deep learning and Data Science: Pandas, NumPy, OpenCV, Neural Networks, Computer Vision, Reinforcement Learning, PyTorch, Keras, TensorFlow, Matplotlib, Weights and Biases, NLP