Madhav Kataria

b23ch1025@iitj.ac.in | 9034701178 | madhavkataria.com | madhavkataria
madhavkataria1010

Education

Indian Institute of Technology, Jodhpur — **Jodhpur, Rajasthan**
B.Tech – Chemical Engineering — *Expected Graduation: June 2027*

Relevant Coursework: Data Structures and Algorithms, Introduction To Machine Learning, Deep Learning, Linear Algebra, Probability, Statistics, and Stochastic Processes

Experiences

Full Stack ML Developer, SACOM Technologies — May 2025 – Ongoing

- Developed production-grade CI/CD pipelines for medical systems, enabling real-time data processing, seamless ML model deployment, and faster iteration cycles.
- Integrated interactive front-end dashboards with backend ML inference services to support clinical decision-making.
- Collaborated cross-functionally with data scientists and clinicians to ensure seamless deployment of AI-driven tools.

Data Automation Internship, GreenTree — May 2025 – Ongoing

- Collaborated cross-functionally to analyze business operations and identify automation opportunities, leading to streamlined workflows using Power Automate.
- Developed and deployed data automation solutions leveraging SQL and low-code platforms, enhancing reporting efficiency and reducing manual effort.
- Utilized Power Automate, LLM's, Fast API, SQL, and Azure ML to create automated data processing workflows, which improved data accuracy and reduced processing time

Research Experiences

Undergrad Researcher, IAB Lab(IITJ) — September 2024 – May 2025
Advisor: Dr. Mayank Vatsa

- Developed an advanced Vision-Language Model (VLM) for fake image detection enhancing accuracy and providing interpretable reasoning which improved detection reliability.
- Curated and synthesized a diverse dataset using outputs from 6-7 different VLMs, ensuring robust training and improved model generalization.
- Enhanced image authenticity verification by leveraging the combined strengths of multiple VLM architectures which improved reliability and interpretability of the verification process
- Utilized Python, VLMs, and Fast API to streamline the development process, contributing to the overall success of the project

Publications

- A.Baranwal, **M. Kataria**, N. Agrawal, S. Vyas, and Y. S. Rawat,
  "Re:Verse—Can Your VLM Read a Manga?"
  In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2025.
  Accepted (Oral Presentation, Top 5% of papers).

Projects

NeuroLearn – AI-Driven Course Recommendation — Github
RAID (AI Club IITJ) — *January 2025 – April 2025*

- Designed and implemented NeuroLearn using a fine-tuned BGE-M3 model for semantic search and integrated Qwen2.5-14B-Instruct for natural language explanations, enhancing recommendation transparency.

- Integrated hybrid retrieval (Pathway vector store), Jina AI fallback and ethical guardrails to achieve 95%+ answer relevance and +30% system robustness.
- **Tools**: Docker, FastAPI, React.js, Bert, LLM, Fine-Tuning, PostgreSQL.

Dynamic Agentic RAG Github
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.

RealmWeaver Github
***Advisor**: Suchetana Chakraborty* *November 2024 – February 2025*

- Procedurally generated a vast 2D open world using Gemini API and recursive tree graphs for infinite exploration and dynamic quests, while optimizing quest assignment and player navigation via advanced graph algorithms and priority queues to significantly boost traversal efficiency and player engagement
- Integrated DSA-driven mini-games (Maze, Connect 4 with Minimax and – pruning, 2048 logic) to enhance gameplay variety and demonstrate highly optimized algorithmic performance
- **Tools**: C++, CMake, Gemini API, Graph Algorithms, DSA

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

- Developed a multi-head CNN (85% accuracy on 32×32 images) for detecting AI-generated content and fine-tuned a Vision-Language Model (VLM) with LoRA adapters for artifact detection and reasoning on image authenticity.
- **Dataset**: Curated a 500K+ image dataset (BigGAN, StyleGAN, Stable Diffusion, CIFAKE) and used GPT-4o for artifact-based reasoning; incorporated adversarial training (FGSM PGD) to boost model robustness.
- **Tools**: Python, Vision-Language Models (VLMs), GANs, KANs, VAEs, LoRA, Knowledge Distillation.

AI Agent for Domain-Specific QA Github
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

- **14th in Tech Meet**: Awarded 14th place in Inter IIT Tech Meet 13.0
- **BCCA UG Rep**: Managed all technical clubs under BCCA
- **RAID , Devlups & Robotics Core Member**
- **13th in KGP Data Science**: Ranked 13th out of 9000 participants in KGP Data Science Hackathon

Technologies

Programming: C/C++, TypeScript, Swift, Java, JavaScript, Python, Rust, Shell, HTML/CSS

Tools & OS: Unity engine, Git, Jupyter Notebook, Google Colab, UNIX systems, macOS, Windows, PostgreSQL, ROS, Figma, SQL, Github

Development: Django, NextJS, React, Fast API, Streamlit, XCode, Git, Flask

Machine learning, Deep learning and Data Science: Pandas, NumPy, OpenCV, Neural Networks, Computer Vision, Reinforcement Learning, PyTorch, Keras, TensorFlow, Matplotlib, Weights and Biases, NLP, Graph Neural Nets, Unslothm, llama.cpp, diffusers, accelerator