Sreevaatsav Bavana

Hyderabad, Telangana, India

+91 9177554960 | bavanasreevaatsav1@gmail.com | SreevaatsavB | sreevaatsav-bavana | sreevaatsavbavana | Sreevaatsav Bavana |

Portfolio website

Personal Profile

Gold medalist in B.Tech Artificial Intelligence at Mahindra University, currently working as Data Scientist. Strong expertise in machine learning and generative AI. Passionate about developing multi-modal AI techniques to solve real-world challenges.

Education

Mahindra University, Hyderabad, India *Aug 2020 - Aug 2024*

Bachelor of Technology in Artificial Intelligence

- **Gold medalist**
- **Cumulative GPA** : 8.76
- **Relevant Courses:** Data Structures and Algorithms, Linear Algebra, Probability and Statistics, Optimization Techniques in AI, DBMS, Machine and Deep Learning, Natural Language Processing, Reinforcement learning

Work Experience

Associate Data Scientist *Jun 2024 - May 2025*

Awone AI, Hyderabad, India

- Developing a multi-agent tool for automating workflows related to medical insurances.
- Modeled client-specific ML methods to classify insurance claims on the basis of denial rate, along with regular monitoring of metrics and re-training
- Improved the inference speed of domain specific-LLMs by 15% by quantization and developing speculative decoding n-gram head.
- Enhanced customer chatbot solutions through custom LLM routing, dynamic memory updates, and finetuned reranking models, decreasing the token cost by 30%.

Data Science Intern *Jan 2024 - Jun 2024*

Awone AI, Hyderabad, India

- Developed a strong foundation in LLMs and transformers by coding state-of-the-art models like Llama-2 and Mistral from scratch, focusing on architectures, attention mechanisms, and optimizations.
- Experimented with several LLM optimization and quantization techniques, helping to reduce latency on developement platform.
- Finetuned task-specific SLMs for medical-insurance related tasks.

Research intern *Jul 2023 - Sept 2023*

Mahindra University, Hyderabad, India

- Collaborated with a four-person team to conduct a thorough research survey on video summarization techniques.
- Co-authored a survey paper published in the ACMMM Workshop 2023.

Projects

An Explainable Visual Question-Answering system for chest X-rays *Aug 2024 - Jan 2025*

Supervisors: Dr. Nidhi Goyal, Assistant Professor, Mahindra University

- Curated a new dataset with 892,364 question-answer pairs with grounding on chest X-rays.
- Developed baseline VLMs with grounding capabilities, which outperformed baselines by over 10% (Rouge-L metric) on the validation split.
- Currently under review for SIGKDD-2025.

Transformers and LLMs reconstruction from scratch *April 2024 - May 2024*

Self-learning

- Coded the fundamental modules such as transformers and language models after thorough conceptual understanding
- Matched the intermediate outputs of each model with the corresponding model from huggingface for a input
- Provided the code corresponding to the same in a published book

AI-Assisted Learning for NVIDIA SDKs and Toolkits *Aug 2023 - Sept 2023*

Mahindra University

- Developed a LLM that assists users in understanding and effectively using various NVIDIA SDKs.
- Fine-tuned Falcon-7B using QLoRA with data scraped from the official Nvidia SDK documentation.

Enhancing video summarization by enhancing text-to-image module *Aug 2023 - Dec 2023*

Supervisors: Dr. Nidhi Goyal, Assistant Professor, Mahindra University

- Improved text-to-image projection by augmenting it with a learnable projection layer.
- Finetuned the selected baseline's projection module using the QVHighlights dataset.

Comprehensive analysis of posts/tweets about mass layoffs *Dec 2022 - Jun 2023*

Supervisors: Dr. Sanatan Sukhija, Assistant Professor, Mahindra University

- Collected and analyzed data about the mass layoffs that started in the last quarter of 2022 from social media platforms.
- Several reactions from users, trends, and patterns from discovered using language modeling and hashtag analysis.

Publications

Conference Proceedings

GIV-CXR: Densely Grounded, Visually Interpretable, Chest X-Ray Question Answering Dataset
Sreevaatsav Bavana, Adit Rushil Potta, Krishi Raju Vysyaraju, Sai Amrit Patnaik, Nidhi Goyal
Under review at SIGKDD, 2025

A Systematic Study on Video Summarization: Approaches, Challenges, and Future Directions
Kajal Kansal, Nikita Kansal, Sreevaatsav Bavana, Bodla Krishna Vamshi, Nidhi Goyal
Proceedings of the 2nd Workshop on User-centric Narrative Summarization of Long Videos, 2023

Research Interests

- Multi modal model research & frameworks development.
- Explainable AI techniques in the field of conversational AI and multi-modal architectures.
- Efficient fine-tuning and compression techniques for domain-specific small LMs/VLMs.

Achievements

2024 **4 years consecutive merit scholarship (2020-2024)** for securing a rank in the top 5% of the branch

2023 **3rd position** in ICETCI Hackathon (2023), Mahindra University

2023 **Top 150 teams (Rank 111)** in Amazon ML Challange 2023, a total of 4000+ teams participated

Skills

Languages & Libraries Python, C++, SQL, Pandas, NumPy, Scikit-learn, HuggingFace, PyTorch, Tensorflow

GenAI & Infra MCP Protocol, Pydantic AI, VLLM, MLFlow, Nvidia-SLRUM, Ray

Miscellaneous AWS, Docker, Async programming, Web Sockets, Linux, FastAPI, Computer networks, GitHub Actions, Javascript

Extra-Curricular

Power-lifting Volunteer for AERO Coordinated events at Aero 2023 Sports Fest

College Gym committee member (MU) Oversaw equipment and mentored new students

Yoga club member and volunteer (MU) Volunteered for the University's celebration of International Yoga Day.