Suman Mandava

Buffalo, New York | +1 716-319-4581 | sumanm2003@gmail.com | linkedin | github | Portfolio

Summary

Software Engineer with entry-level experience in computer science, specializing in AI research and full-stack development. Developed an LLM model for emotion recognition, achieving 78% test accuracy, and built responsive HMI applications that improved system responsiveness by 30%. Seeking to apply technical expertise to deliver high-impact software engineering solutions.

Education

University at Buffalo, The State University of New York **Aug 2024 - Dec 2025**

Master of Science, Computer Science and Engineering

- **GPA:** 3.8
- **Coursework:** Data Structures & Algorithms, Database Systems, Object-Oriented Programming, Operating Systems, Machine Learning, Deep Learning

St.joseph's institute of technology **Mar 2020 - Apr 2024**

Bachelor of Engineering, computer science

- **GPA:** 3.6
- **Coursework:** Operating Systems, Computer Networks, Microprocessors and Microcontrollers, Data Structures and Algorithms, Design and Analysis of Algorithms, Database Management System

Experience

Kornia (Computer Vision Library) **Jan 2026**

Open Source Contributor *Remote*

- Implemented core components of the Qwen2.5-VL Vision Encoder in Kornia's main codebase
- Built modular, type-safe PyTorch components optimized for high-dimensional tensors (B, N, 1280).
- Resolved CI/CD issues (Ruff, Mypy) and added pytest coverage to ensure CPU/GPU stability.

UB research, University at Buffalo **Oct 2024 - Dec 2025**

Research Assistant *University at Buffalo*

- Developed a multi-head Transformer model for multi-label classification of cognitive-affective states (engagement, boredom, confusion, frustration) using Action Unit (AU) and Valence-Arousal (VA) features from the DAISEE dataset, which later achieved 78.08% test accuracy
- Reached 78.08% test accuracy via late fusion of parallel Transformers for emotion recognition.
- Integrated Qwen3 (LLM) + LoRA + TRL's SFT for efficient cross-modal fine-tuning, enabling the model to interpret structured AU/VA descriptors and generate multi-head emotion predictions via instruction-style prompts.

Centum T&S **Apr 2024 - Jun 2024**

Software Engineer Intern *bengaluru, India*

- Built responsive HMI applications for BMRCL and DMRC metro systems using protobuf, ZeroMQ, and web sockets, achieving a 30% improvement in system responsiveness.
- Designed intuitive UI components (fault info screen, settings page) with Handlebars.js and Node.js, enhancing driver control over console and audio configurations.
- Translated business requirements into detailed technical specifications, ensuring accurate implementation and timely delivery of HMI features.

HCLTech **Jan 2024 - Mar 2024**

Software Engineer Intern *Chennai, India*

- Built a full-stack Online Video Platform using React.js, Node.js, and MongoDB, improving system performance by 25% and reducing authentication latency by 15%
- Engineered core features like JWT authentication, role-based user interfaces, and video streaming, boosting user engagement by 20% and admin efficiency by 30%.
- Deployed the Online Video Platform using Docker containers on Google Cloud Platform (GCP) to ensure scalability

Projects

(Qwen + LoRA) - Video Affective State Analyzer | huggingface.co

- Built an end-to-end video affective state analyzer that predicts Engagement, Boredom, Confusion, and Frustration from short face videos using Action Units (AUs) and an LLM-generated explanation.
- Fine-tuned **Qwen (LoRA adapters)** as four separate label heads (one per affect) to produce calibrated **0–3** scores with probability distributions.
- Designed a **Gradio + Hugging Face Spaces** demo with an "example case" workflow and interpretable outputs (top contributing AUs + confidence per label).

Trip Planner CrewAI (Gemini + Serper) | Github

- Designed and shipped a multi-agent travel planner with CrewAI-agents (city selector, local expert, concierge) collaborate to produce a personalized 7-day itinerary.
- Integrated Gemini LLMs via CrewAI's LLM and Serper web search; added DuckDuckGo fallback and a website reader tool for robust, real-world data gathering.
- Implemented a config-driven architecture using .env (model/key switching, timeouts) with Markdown/PDF export; hardened reliability with retry/backoff for 429/503, provider fallbacks, and version-compatible tools.

Visual Question Answering | Github

University at Buffalo

- Developed a Visual Question Answering (VQA) system for real-world images from the VizWiz dataset, addressing challenges such as blur and poor lighting.
- Integrated pre-trained CLIP (ViT-B/32) vision-language encoder with a lightweight multi-head classifier, achieving 70% test accuracy in textual answer prediction.

Road Lane Line Detection | Github

- Computer vision system for detecting road lane lines with 93.8% accuracy using OpenCV and Python. Implemented advanced image processing techniques for real-time lane detection.

Technologies

- **Languages**: Java, C++, Python, HTML, CSS, JavaScript
- **Frameworks**: React.js, Node.js, Express.js, Handlebars, Bootstrap, Material UI, Context API, JSON Web Token, CrewAI
- **Databases & Libraries**: MySQL, MongoDB, PostgreSQL, ReportLab, python-dotenv, Requests
- **DevOps/Cloud**: Git, GitHub, Firebase, GCP, Docker
- **Tools & Technologies**: REST APIs, Postman, WebSockets, OAuth, Ganache, Nmap, Wireshark, Tableau, Google Dialogflow, crewai-tools, Serper API, Google AI Studio, DuckDuckGo Search
- **AI & ML**: Linear/Logistic Regression, SVM, k-Means, CNNs, RNNs, Transformers, VAEs, GANs, Reinforcement Learning, scikit-learn, TensorFlow, PyTorch, NumPy, NLP, Agentic AI, LLMs