

CHETAN SAI BORRA

979 575 9026 | chetansai2003@tamu.edu | LinkedIn/chetan-sai | github/Chetansai11 | Collage Station, TX

Technical Skills

Programming: Python (Proficient), C++, Java, SQL, MATLAB.

Machine Learning & AI: PyTorch, TensorFlow, Scikit-learn, Keras, XGBoost, OpenCV, Large Language Models (LLMs).

LLM & GenAI: LangChain, LangGraph, LlamaIndex, RAG Pipelines, Prompt Engineering.

MLOps & Deployment: Docker, MLflow, Git, CI/CD (GitHub Actions), CUDA, Linux, Model Optimization.

Perception & Computer Vision: OpenCV, YOLO, Object Detection & Segmentation, SLAM, Sensor Fusion.

Robotics Frameworks: ROS 2, Gazebo, Rviz2, SLAM Toolkit, MATLAB/Simulink, Reinforcement Learning.

Data Engineering: Pandas, NumPy, AWS, MySQL, PostgreSQL, Vector Databases (FAISS, Pinecone).

Frameworks & Environments: ROS 2, Anaconda, VS Code, MATLAB/Simulink.

Experience

ML Engineer – Student, Utilities and Energy Services, Texas A&M University

Feb 2025 – Present

- Designed scalable time-series pipelines in Python, Pandas, and SQL to automate anomaly detection and forecasting for HVAC and electricity across 200+ campus buildings, cutting fault response time by 40%.
- Deployed custom ML models (ARIMA, Isolation Forests) for real-time energy anomaly detection, reducing undetected system faults by 30%.
- Orchestrated end-to-end ETL workflows integrating live IoT sensor feeds with historical data, enabling automated meter diagnostics and predictive maintenance alerts.
- Optimized SQL and Python validation layers to identify consumption anomalies with 95% precision, minimizing manual audits and elevating operational efficiency.

Research Intern – Controls, DRDO – DRDL

May 2023 – Jul 2023

- Engineered advanced aircraft pitch control systems using PID controllers in MATLAB/Simulink, improving system stability by 11% and precision of flight dynamics.
- Applied machine learning-based optimization to refine controller parameters, cutting overshoot by 10% and improving response time by 10% for robust autonomous performance.
- Collaborated with senior scientists (including Shri. Murali Mohan Gade, Scientist 'F') to design adaptive control algorithms for aerospace platforms, contributing to next-gen robotics and autonomous navigation systems.
- Validated control architectures through simulation-driven testing, ensuring resilience under dynamic operating conditions relevant to robotic and aerospace applications.

Research Assistant – Deep Learning, Vellore Institute of Technology

Aug 2023 – May 2024

- Implemented U-Net and W-Net architectures for brain tumor segmentation on BraTS 2020 using PyTorch, reaching 90% accuracy and 72% Mean IoU with transfer learning from VGG16, ResNet, and DenseNet.
- Refined model performance via cross-dataset validation, hyperparameter tuning, and advanced OpenCV preprocessing, producing high-fidelity clinical inputs and stronger generalization.

Projects

Telecom KPI Anomaly Detection Using AI Agents | LangGraph, MCP, LLMs

- Architected an AI agent integrating LangGraph and MCP with anomaly detection models (DWT-MLEAD, Isolation Forest, Ensemble Voting) and NVIDIA LLMs to track more than 10 KPIs across 100 plus cellular sites in real time.
- Elevated anomaly detection reliability by 30% via ensemble fusion of ML models and LLM reasoning, ensuring anomaly rates stayed within 0.5–5% thresholds while minimizing false positives.
- Launched a full-stack platform with Gradio conversational UI, LLM (Llama, GPT-4o) reasoning, and Tavily Search, supporting 50+ advanced telecom queries (e.g., SINR drops, throughput spikes, KPI co-anomalies).

Retrieval Augmented Generation from YouTube for Long-Form QA | LangChain, LLMs, Python

- Developed a RAG pipeline transforming YouTube videos into searchable knowledge bases by transcribing with OpenAI Whisper, refining with GPT, and embedding with LLaMA.
- Indexed contextual transcript chunks into FAISS and Pinecone for high-speed similarity search and ranked retrieval across multi-hour video datasets.
- Delivered an end-to-end QA interface, processed natural language queries and generated grounded answers via LLM reasoning, reducing hallucinations and improving factual accuracy.

Face Tracking Robot | Computer Vision, Arduino, Robotics

- Structured a face-tracking robot with a pan-tilt servo mechanism, achieving two degrees of freedom for precise facial alignment and continuous tracking.
- Implemented Python-based computer vision algorithms for real-time face detection and tracking, ensuring robust operation under dynamic lighting and motion conditions.
- Built and integrated a 4-wheeled mobile base with microcontroller, enabling autonomous navigation and consistent person following behavior.

- Demonstrated perception-to-control integration by synchronizing camera input with actuation commands, showcasing end-to-end robotic autonomy.

Multi-Robot Coordination in Smart Warehouses | Multi-Agent Systems, MARL, Decentralized Planning

- Conducted a literature review and analysis of multi-robot warehouse coordination strategies, emphasizing decentralized planning, MARL-based task allocation, and environment-aware optimization.
- Examined coordination frameworks such as Decentralized Prioritized Planning, Multiagent Rollout with Reshuffling (MARR), and MARL with QMIX extensions, highlighting scalability and resilience under robot failures.
- Identified research gaps in perception-driven coordination, simulation-to-real transfer, and integration of tactile/visual/RFID sensing for robust warehouse autonomy.

Robust Reinforcement Learning for Mixed Autonomy Traffic Systems | PPO, TRPO, SUMO

- Designed and tuned PPO and TRPO agents with KL annealing, entropy regularization, and return normalization, stabilizing policy learning in multi agent partially observable Markov decision process (POMDP) settings.
- Scaled training by orchestrating 40+ parallel SUMO simulations with Python multiprocessing, accelerating throughput and improving cross-scenario generalization.
- Achieved a 20% increase in traffic throughput and zero safety violations compared to rule-based systems, validated through trajectory and time-space analytics.

Meal Nutrition Analysis via Multi-Modal Deep Learning | CNNs, LSTMs, Regression

- Built a multi-modal pipeline integrating CNNs (meal images), LSTMs (glucose logs), and demographic embeddings, boosting prediction accuracy by 34% over benchmarks.
- Evaluated feature importance using regression metrics and correlation heat maps, reducing test loss to 0.34 and identifying top predictors of calorie absorption.

Education

Texas A&M University, College Station, TX

Aug 2024 – May 2026

Master of Science in Computer Engineering, **GPA - 4**

College Station, TX

Coursework: Machine Learning, Parallel Computing, Reinforcement Learning, Computer Vision & Robot Perception.