

HARSHITH GUJJETI

Tampa, FL | +1 813-786-8772 | Harshithgujjeti@icloud.com | [LinkedIn](#) | [Github](#)

SUMMARY

MS in AI & Business Analytics student focused on machine learning and data science, experienced with production ML model development, generative AI applications, and healthcare analytics using Python, R, SQL, and TensorFlow.

EDUCATION

University of South Florida

Master of Science, Artificial Intelligence and Business Analytics (GPA: 4.00)

Aug 2025 - May 2027

Tampa, FL

Vignana Bharathi Institute of Technology

Bachelor of Science, Computer Science with Artificial Intelligence focus

Nov 2021 - Jul 2025

India

SKILLS

- **Machine Learning & AI:** Python, scikit-learn, XGBoost, PyTorch, statistical analysis, hyperparameter tuning
- **Generative AI & LLMs:** OpenAI API, prompt engineering, LLM fine-tuning, chatbot development
- **Data Science:** Python, pandas, NumPy, StatsModels, R, SQL, exploratory analysis, healthcare analytics, JavaScript, Java
- **Computer Vision:** CNN, SRGAN, NeRF, volumetric rendering, image processing, Deep Neural Networks
- **Tools & Infrastructure:** Jupyter, Git, Docker, SQL Server, Snowflake, MySQL, MongoDB, Tableau, Power BI
- **BI & Visualization:** Tableau, Power BI, Dashboards, Data Storytelling

PROJECT EXPERIENCE

Healthcare Research Agent - Multi-Agent AI System | <https://github.com/Harshxth/Healthcare-Research-Agent>

- Built autonomous 4-agent system (Researcher/Analyzer/Validator/Reporter) using CrewAI that conducts end-to-end PubMed research and generates executive reports with charts
- Agents autonomously search literature, validate sources, extract insights; processed 50+ healthcare queries (hospital readmissions, treatment trends) with 95% accuracy vs manual review

Hospital Length of Stay Prediction - Machine Learning | <https://github.com/Harshxth/Hospital-LOS-Prediction>

- Developed a predictive model for hospital length of stay using 100,000 patient records, achieving a mean absolute error (MAE) of 0.30 days (± 7.1 hours) with XGBoost.
- Systematically evaluated 12+ machine learning algorithms and 230+ hyperparameter configurations, demonstrating 5-6x improvement over published research benchmarks.
- Delivered actionable insights with potential cost savings of \$1.7-2.6M annually per 400-bed hospital by optimizing hospital length of stay predictions.

Local PDF RAG Chatbot

- Developed fully-local RAG app with LangChain, Ollama Llama 3.1, and ChromaDB; enables natural-language Q&A over uploaded PDFs with source citations
- Implemented document chunking, semantic retrieval, and grounded generation; achieved <2s query response on 50-page docs, eliminating hallucinations
- Deployed interactive Streamlit UI for real-time chatting; open-sourced with eval benchmarks showing 90%+ faithfulness score

WORK EXPERIENCE

Centre of Transportation and Research | *OPS Student Assistant – Computer Vision Video Analytics* Feb 2026 - May 2026

- Deployed production YOLOv8 + ByteTrack pipeline processing 100+ hours of traffic video; tracked 1,200+ unique motorcycles across Tampa Bay intersections for injury prevention studies
- Automated manual counting workflow for FDOT safety research; delivered 50+ annotated videos, CSV datasets, and batch processing scripts supporting SS4A grant deliverables

Muma College of Business (University of South Florida) | *Research Assistant (Artificial Intelligence)* Sep 2025 - May 2026

- Developed machine-learning models for healthcare billing analysis using scikit-learn, hyperparameter tuning, and SQL, which lowered coding errors and streamlined diagnostic workflows
- Applied statistical analysis and Python to validate ML models, reducing coding errors and improving diagnostic decision support for hospital operations

ISRO (Indian Space Research Organisation) | *Artificial Intelligence Research Internship* Jan 2025 - Jul 2025

- Designed and implemented a super-resolution pipeline for satellite imagery using Python, TensorFlow, a CNN baseline and an SRGAN variant, which improved spatial resolution while preserving radiometric consistency and produced clearer images for downstream analysis
- Performed 4x upscaling from 10 m to 2.5 m ground sampling distance, enabling finer spatial detail.

Venkusa Technologies | *AI Developer Internship*

May 2024 - Nov 2024

- Built an AI code-completion tool using the OpenAI API and PHP; created API wrappers and performed prompt/LLM tuning, which improved suggestion relevance and reduced developers' coding time
- Achieved ~8% efficiency improvement versus GitHub Copilot benchmarks and enhanced overall usability for travel bookings.