# NARRA VENKATA RAGHU CHARAN

## Professional Summary

Data & AI Engineer with experience delivering scalable Machine Learning, Deep Learning, and AI solutions. Skilled in Retrieval-Augmented Generation, Computer Vision, and Optimization Tools, with Strong Proficiency in Python, PyTorch, and Cloud Platforms (AWS, GCP). Proven ability to take projects from concept to deployment.

#### EXPERIENCE

# Data & AI Engineer

Sep 2024 - Present

Freshbus, Hyderabad

- Developed an automated customer service bot using a **RAG Agent** (OpenAI, ChromaDB, Vanna), reducing ticket resolution time by 40%.
- Built a scheduling software for dynamic fleet expansion in many stations, creating dashboards and improving decision-making speed by 30%.

SDE-1

Apr 2024 - Aug 2024

MoveInSync, Bangalore

- Built a facial recognition attendance system with OpenCV and InsightFace for real-time employee tracking with 95% + precision.
- Integrated cloud-based pipelines for secure, real-time attendance with in-trip level.

Software Intern

May 2024 - Jun 2024

IIT Madras (remote)

- Designed a CNN model achieving 96% accuracy on MNIST using IBM's AI Hardware Kit (aihwkit), eliminating traditional DL libraries.
- Benchmarked neuromorphic vs. traditional inference, achieving a 15% less computation.

## **EDUCATION**

# B.Tech in Computer Science & Engineering (Data Science)

2020 - 2024

Gayatri Vidya Parishad College of Engineering (A)

CGPA: 9.1/10

#### Projects

#### RAG Agent Chatbot

GitHub

- $\bullet \ \ Domain-specific\ Retrieval-Augmented\ Generation\ chatbot\ using\ Vanna,\ ChromaDB,\ and\ OpenAI.$
- Reduced query retrieval latency by 20% via optimized embeddings and easy access to DB via NLP

#### Synthetic Image Generation Using GANs

GitHub

- PyTorch pipeline for dataset augmentation and GAN-based synthetic image creation.
- Improved model generalization and accuracy in low-data vision tasks using synthetic samples by 10%.

#### **EV Bus Scheduling Optimization**

GitHub

- Optimization system using OR-Tools for EV bus scheduling, maximizing occupancy or given weights while adhering to operational constraints.
- Generated HTML tables and JSON outputs for visualization and integration.

#### SKILLS

# Languages & Programming

FastAPI, Flask, HTML, CSS, PyTorch, TensorFlow, LangChain, OpenCV, FAISS, Scikit-learn, Streamlit, Postman, AWS, GCP, Docker, Git, OR-Tools, EDA, Weka, An-

droid Studio

Databases
PostgreSQL, MongoDB, Redis, SQLite

# CERTIFICATIONS & OPEN SOURCE

Python, C/C++, Java, JavaScript, R (Basics), SQL

#### Certifications:

- Google Cloud Career Practitioner
- AWS Cloud Foundations
- $\bullet\,$  Python (HackerRank Gold Badge)
- Android Development with Kotlin Google Developers

#### Open Source & Volunteering:

Frameworks, Tools & Platforms

- DeepFace: Added Buffalo-L model
- InsightFace: Added Face Recognition script
- Mentor GSSOC
- DevOps Lead GDSC GVPCE
- Core Member GDG Vizag