**Madhu Bagamma Gari** madhu.bagamma@dtskill.com

## **Professional Summary**

Python & Generative AI Developer with 7 months of experience in AI-powered applications, specializing in LLMs, chatbot development, and AI agents. Skilled in Python, Django, FastAPI, and AI frameworks like LangChain. Passionate about exploring LLM fine-tuning, quantization, and GPU-based optimizations.

## **Work Experience**

### **DT Skill**

**Role:** Python & Generative AI Developer  
 **Duration:** 7 months (Currently working)

* Developing AI agents with predefined inputs like name, role, goal, and backstory.
* Implementing real-time data-fetching tools.
* Creating persona builders with predefined prompt templates.
* Implementing authentication, organization, and authorization features.
* Working on an AI-powered chatbot with private/public chat and enterprise mode.

## **Skills**

### **Programming Languages & Frameworks**

* **Languages:** Python, JavaScript
* **Frameworks:** Django, Django Rest Framework, FastAPI, React, Next.js, Tailwind CSS

### **AI & Machine Learning**

* LLMs (Large Language Models), Retrieval-Augmented Generation (RAG)
* LangChain, LangGraph, OpenAI, Hugging Face
* Data Analysis & Feature Engineering
* SpaCy, NLTK, Scikit-learn, TensorFlow

### **Databases & Cloud**

* PostgreSQL
* Docker
* Milvus

### **Other Tools**

* WebSockets, Pandas, NumPy, RabbitMQ

## **Projects**

### **Company Projects (DT Skill)**

* **GenE AI Agents:** AI-driven agents with predefined attributes and tool integration.
* **Persona Builder:** Feature that allows users to create and modify predefined prompt templates.
* **Chatbot System:** Ongoing development of a chatbot with private/public chat and enterprise features.

### **Personal Projects**

* **Cosmology AI Assistant:**
  + LLM-based assistant limited to answering cosmology-related questions.
  + Modes: Kids Mode, Casual Mode, Scientific Mode, Story Mode.
  + Automatic note-taking for each conversation.
  + Upcoming Features: Voice Mode, Vector DB Integration.
* **Machine Learning Projects:**
  + House Price Prediction
  + Salary Prediction
  + Heart Disease Prediction
  + Movie Recommendation System
  + Text Emotion Detection
* **Deep Learning Projects:**
  + Predicting Temperature on a Specific Date
  + WhatsApp Chat Data Analysis (Active users, message trends, etc.)

## **Education**

* **BSc Computer Science** - Rayalaseema University (2021)

## **Interests & Exploration**

* Fine-tuning LLM models
* Quantization of LLM models
* Running LLM models on GPU and multi-GPU setups
* Techniques like Flash Attention, Offloading
* DeepSpeed, vLLM, TGI
* Voice AI tools like DeepFilterNet, NVIDIA Riva, OpenVINO