Varsha Balaji

Chicago, IL | +14022372373 | vbala7@uic.edu
linkedin.com/in/varsha-balaji | github.com/B-Varsha | b-varsha.github.io

Education

University of Illinois at Chicago Aug. 2024 – May 2026
Masters of Science in Computer Science *GPA: 4.00/4.00*
Coursework: *Data Science, Computer Algorithms, Neural Networks, Natural Language Processing, AI, Big Data Mining, Information Retrieval*

SSN College of Engineering Sep. 2020 – July 2024
Bachelor of Engineering in Computer Science and Engineering *GPA: 3.65/4.00*

Technical Skills

Programming Languages: Python, Java, C++, C, SQL, JavaScript, HTML, CSS, LaTeX
Frameworks & Libraries: FastAPI, Flask, Matplotlib, Node.js, NumPy, OpenCV, Pandas, PIL, Pinecone DB, PyTorch, ReactJS, Scikit-learn, Seaborn, Streamlit, TensorFlow
Cloud & Development Tools: AWS, Docker, Git, Google Cloud Platform (GCP), GKE, Kubernetes, MLflow, Vertex AI
Other: Agentic AI, Agile Methodologies, CI/CD, Data Cleaning, Ensemble Learning, Exploratory Data Analysis (EDA), Explainable AI, REST API

Experience

Machine Learning Engineer Intern Jun 2025 – Aug 2025
Vosyn AI *Chicago, IL*
- Engineered and fine-tuned a custom **Text-to-Speech (TTS)** deep learning model for multilingual audio synthesis using **semantic-to-acoustic tokenization**.
- Designed a data pipeline deployed on GCP (Vertex AI) that processed 800K+ audio files using dynamic **in-memory batching** and concurrent uploads, achieving **5× throughput improvement**.
- Integrated **Vertex AI Experiment Tracking** and executed distributed model training via **Custom Jobs**, improving experiment reproducibility by **98%** and reducing training iteration setup time by **75%**, streamlining the workflow.
- Implemented **automated checkpointing**, error recovery to ensure fault-tolerant batch inference at scale.

Data Science Intern July. 2022 – Aug 2022
Synergy Maritime *Chennai, India*
- Engineered **ML-powered virtual sensors** for ”Scrubber Management,” applying advanced **regression techniques** to improve accuracy from **60% to 85%**, enhancing environmental compliance for over 30 maritime vessels.
- Streamlined data pipelines, reducing preprocessing time by **30%**, and implemented hyperparameter tuning to optimize model performance, achieving a **25%** accuracy improvement.

Projects

AI-Stylish Fashion Agent - GKE Hackathon | *RAG, GKE, Cloud Run, Streamlit* Aug. 2025 – Sep 2025
- Engineered and deployed a containerized microservice on **Google Kubernetes Engine (GKE)** to power a vibe-based fashion recommendation system using **Gemini 2.5 Flash** and a **Pinecone vector database**.
- Designed a data ingestion pipeline with **gRPC** for real-time embedding generation, enabling a live **Retrieval-Augmented Generation (RAG)** system.

Progressive In-Context Alignment for LLMs | *LLMs, Gen AI, HuggingFace, NLP* Feb. 2025 – May 2025
- Implemented the Progressive In-Context Alignment (PICA) inference method, leveraging **Mistral** and **LLaMA** models to improve instruction alignment by encoding few-shot demonstrations.
- Designed a multi-stage inference pipeline to extract **In-Context Learning (ICL) vectors** and inject them to steer targeted, query-only generation.
- Evaluated PICA performance against zero-shot and few-shot baselines using **BLEU, ROUGE-L, and cosine similarity**, confirming consistent performance gains across models.

Facial Antispoofing | *Transformer Networks, Flask* Sept. 2023 – Apr. 2024
- Developed an advanced facial antispoofing system using **Vision Transformers (ViTs)** and **Exploratory Data Analysis (EDA)** to clean and optimize training data.
- Created proprietary **SPAN (Cross-Layer Relation)** and **MSWF (Feature Fusion)** techniques, which significantly improved model accuracy and robustness for real-time use.
- Designed and deployed a **Flask-based web application** with **WebRTC** for real-time image capture and dynamic client-server interactions.

Publications

- P. Ravisankar, V. Balaji, et al. Deep learning-based renal stone detection: A comprehensive research study and performance analysis. Applied Computer Systems. | Paper