Dakshesh Gusain

dgusain@buffalo.edu , 716-398-1720, https://github.com/dgusain , https://linkedin.com/in/dgusain04
Portfolio: https://dgusain.github.io/

Graduating with MS in AI in May 2025, seeking research, job opportunities in LLM, ML/ AI applications, Data Science

Profile

Machine Learning Engineer with hands-on experience in NLP, deep learning, and transformer architectures. Skilled in developing, training, and evaluating state-of-the-art models for real-time applications, including content moderation, semantic search, and chatbot assistants. Proficient in Python, PyTorch, and scikit-learn, with a track record of delivering scalable, production-ready systems. Adept at collaborating with cross-functional teams to innovate and enhance model performance, driving impactful AI solutions.

Key Skills

Programming	Python, PySpark, MATLAB, C/C++, C#, Java, Linux, SQL, Apache Solr, Scala, Apache Spark
Software Expertise	Web development (React, Node.js, Flask), building RESTful APIs, Git version control, Database management (SQL, NoSQL), Distributed Systems (Apache Kafka, Hadoop HDFS, Kubernetes, Amazon S3), SSH, SSL/TLS, Shell scripting, Cloud Computing, Cloud Migration, Apache Hive
ML Expertise	Distributed Computing, Data Parallel, Proximal Policy Optimization, Time Series Forecasting, Predictive Analytics, Prescriptive Analytics, Feature Engineering, Clustering, Statistical modeling, Vision Machine learning, chatbots
Libraries	PyTorch, TensorFlow, scikit-learn, OpenMim, Gymnasium, NumPy, Transformers, Tokenizers, NLTK, OpenCV, Docker, MuJoCo, GNU, CUDA, Robotic Operating System (ROS), Internet of Things (IOT), Pandas, NumPy,
Engineering	Raspberry Pi, Arduino, Blender, CAD, CFD, Embedded Systems, Audio-visual systems, User-interaction, Autonomy
Frameworks	Git, Jupyter Notebook, Azure, AWS, Databricks, Spark-SQL, Airflow, Google Cloud Platform (GCP)

Experience

University at Buffalo, New York:

Research Assistant, Center for Unified Biometrics and Sensors (CUBS/CEDAR) **Sept 2024 – Present**

- Quantized and fine-tuned **multimodal vision** large language models (LLMs) for speech-language therapy, integrated into the **embedded system** of MISTY 2 **social robot**, enabling real-time interactive sessions with **5-second** latency.
- Experimenting with **Llama 3.2** and **Llama 3.1** to enhance the robot's multimodal capabilities, enabling more human-like conversations through multi-GPU support, RESTful APIs, AV streaming, supporting **robotic autonomy**.

Research aide, SUNY Research Foundation, High Performance Computing lab **Aug 2024 – Oct 2024**

- Developed a comprehensive natural language processing pipeline to identify toddlers as late talkers or typically developing, based on parent-child audio conversations. Implemented audio transcription and **diarization** using OpenAI Whisper and Pyannote, introducing novel part-of-speech tags, and applying causal modeling to analyze word usage in vocabulary.
- Annotated a dataset of **2.5 million** words and **40K** vocabulary across new linguistic categories Shape/Non-shape nouns and result/manner verbs using GPT-4o-mini and **LLaMA 3.1 405B** models via prompt engineering. Fine-tuned **RoBERTa-base**, achieving **94%** accuracy for Shape/Non-shape noun classification and **97%** for result/manner verb classification.

Graduate Student Assistant, State University of New York **Sept 2023 – Oct 2024**

- Managed Noldus software for human behavioral observation ensuring smooth operation across **6** psychological laboratories; executed reliability studies and administrative duties.
- Developed and deployed cost effective **$300 autonomous** vape smoking device using **Raspberry Pi** for **20-member** biomedical team across **3** universities: increasing efficiency by **200%**, experimental time reduced by **66%**, with operational performance equivalent to **$45000** smoke machine at Roswell Park Cancer Institute(Github).

Projects

Automobile Inspector: AI-Powered Product for Car Damage detection and Chatbot, (Github) **July 2024 – Aug 2024**

- Deployed an AI-driven **web application** integrating computer vision and NLP to deliver comprehensive car repair solutions, including visual damage assessment and a **RAG**-based **conversational chatbot** to provide econometrics.
- Engineered a segmentation pipeline using **Mask R-CNN** with ResNet-101 backbone and Deformable Convolution Networks (DCN) to accurately identify and categorize vehicle damages like dents, scratches, broken lamps, glass shatters, flat tires, and cracks.
- Developed CarBot, a **context-aware chatbot** leveraging chat history and Ollama to enhance prompt accuracy; implemented Facebook FAISS vector store and LLaMA models for efficient **document retrieval** and context-specific responses.
- Architected a user-centric **Flask** application enabling damage detection, cost estimation, and repair requests within five clicks.

Education

University at Buffalo	Master of Science	**Artificial Intelligence**	Aug 2023 – May 2025	3.46 / 4.0
Amrita Vishwa Vidyapeetham	Bachelor of Technology	Aerospace Engineering	July 2019 - June 2023	3.78/4.0

Coursework: Computer Vision & Image Processing, Data Intensive Computing, Pattern Recognition, Reinforcement Learning, Analysis of Algorithms, Robotic Algorithms, Information Retrieval, Deep Learning, Orbital Mechanics

Certifications

- Generative AI with Large Language Models, Deep-Learning.AI, AWS.