JITEN BHALAVAT

✉ jbha0504@umd.edu | 📞 (240) 481-5543 | 🔗 LinkedIn | ⌂ GitHub | 🌐 jitenbhalavat.com

Education

University of Maryland, College Park   **Aug 2024 - Expected May 2026**
Master of Science in Applied Machine Learning   *College Park, Maryland*

Charotar University of Science and Technology (CHARUSAT)   **Sept 2020 - May 2024**
Bachelor of Technology in Information Technology   *Gujarat, India*

Experience

Plutomen Technologies Pvt. Ltd.   **Sept 2023 - April 2024**
Machine Learning Engineer Intern   *Gujarat, India*
- Engineered a high-performance **RAG chatbot** using **LlamaIndex**, **LangChain**, and HuggingFace embeddings, boosting semantic search relevance by **30%** through **advanced indexing** with Qdrant and Pinecone, demonstrating proficiency in API integration and innovative model experimentation
- Optimized **multi-model pipeline** by **evaluating** LLMs, embeddings, and re-rankers, refining generation parameters to enhance output quality and consistency, aligning with generative AI tool experimentation
- Developed a **scalable Flask API** to connect UI with retrieval backends, facilitating seamless real-time interactions and smooth integration with client-side apps
- Identified and fixed model failures in RAG chatbot testing, boosting robustness via refined prompts and streamlined flow

NXON Pvt. Ltd.   **May 2022 - June 2022**
Machine Learning Engineer Intern   *Gujarat, India*
- Designed a CNN achieving over **99%** test accuracy on MNIST by **fine-tuning** convolutional and pooling layers
- Applied dropout regularization techniques, improving model generalization by **7%** and reducing overfitting risks
- Refined architecture to boost classification reliability, streamlining deployment in production-ready ML pipelines

Projects

ClassTopper: AI-Powered Exam Preparation Platform
- Collaborated in a cross-functional team to build an exam prep platform with advanced document and video processing, enabling interactive learning from uploaded materials and **improving study efficiency** by **40%**
- Engineered a mind-mapping feature to visualize topic relationships, enhancing **content retention by 60%** and enabling personalized knowledge structures; contributed to design through collaborative brainstorming and iterative refinement.
- Developed an **intelligent quiz system** through cross-functional collaboration and analytical problem-solving, creating tailored assessments that identify knowledge gaps and focus on key concepts, **improving performance by 27%**

AI Voice Assistant
- Engineered a real-time **AI Voice Assistant** pipeline integrating Speech-to-Text (Whisper.cpp, AssemblyAI), Text-to-Text (Groq LLMs), and Text-to-Speech modules, optimizing end-to-end **latency** for seamless voice interactions
- Optimized large language model (LLM) performance by **reducing Time to First Token (TTFT) and inference latency**, enabling **sub-800ms response time** using speculative decoding and low-latency architectures
- Spearheaded the implementation of **WebRTC**-based voice framework with integrated Voice Activity Detection (**VAD**), End-of-Utterance (**EOU**) detection, and real-time Agent Turn Detection for dynamic multi-speaker handling

Brain tumor segmentation in MRI images using U-Net
- Constructed a **U-Net** model for pixel-wise segmentation on clinical MRI datasets, achieving **91%** segmentation accuracy
- Preprocessed DICOM images with intensity normalization and resizing to enhance model input quality
- Benchmarked model using IoU metrics, achieving a **15%** improvement over baseline CNN architectures

Technical Skills

Languages and Databases: Python, C++, Java, SQL, MySQL, MongoDB, Pinecone, Qdrant, Chroma
Frameworks and Libraries: PyTorch, TensorFlow, Keras, HuggingFace, Scikit-learn, NumPy, Pandas, Transformers
Generative AI: Large Language Models (LLMs), AI Agents, RAG, Prompt Engineering, Fine-tuning
Tools & DevOps: Git, GitHub, Bitbucket, Docker, Kubernetes, AWS ECS, GCP, FastAPI, Flask, Uvicorn
Data Science: Machine Learning, Deep Learning, Computer Vision, Natural Language Processing

Certifications and Achievements

Google Cloud Computing - Google Cloud Educational Badges (9 Badges)
AWS Academy Graduate - AWS Academy Machine Learning Foundations and AWS Academy Cloud Developing
Academic Achievement - Secured **2nd rank (Silver Medal) among 180 students** in Bachelor of Technology
Team Lead, Data Science Club, CHARUSAT - Organized events and mentored junior students on data science careers, and **managed a successful 6-day AWS Machine Learning workshop with 100+ participants**
Hackathon Achievement - Earned **2nd rank (Silver Medal)** in the 2023 SolutionChallenge Hackathon at CHARUSAT