Vijayabharathi Murugan

vmurugan@cs.stonybrook.edu | +1-934-246-0713 |

Education

Stony Brook University **Aug 2024 – May 2026**

Master of Science in Computer Science GPA: 3.82/4.0 *Stony Brook, NY*

- *Coursework: Vision Language Models, Intro to Computer Vision, Natural Language Processing, Distributed Systems*

Indian Institute of Technology Madras **Aug 2016 – Jul 2020**

Bachelor of Technology in Chemical Engineering GPA: 3.48/4.0 *Chennai, India*

- *Coursework: Multivariate Data Analysis for Process Modeling, Process Optimization, Integer Optimization*

Technical Skills

Programming: Python (PyTorch, TensorFlow, Scikit-learn, Pandas, NumPy), SQL, Golang
Computer Vision: DINOv2, CLIP, Stable Diffusion, LLaVA, OpenCV, Pose Estimation
AI/ML: Hugging Face, Transformers, LLMs, RAG / LangChain, Fine-tuning (LoRA), Bayesian Optimization, XGBoost
Tools: Docker, Git, FastAPI, CI/CD, pytest, MLflow, Wandb, Excel, PowerBI

Research Experience

Research Assistant — Master's Thesis **Jan 2025 – May 2026**

Advisor: Prof. Zhaozheng Yin, Stony Brook University *Stony Brook, NY*

- Developed a training-free pipeline for zero-shot 6DoF object pose estimation by fusing Stable Diffusion (DIFT) and DINOv2 features into 3D point descriptors from multi-view CAD renders, with pose recovered via Kabsch+RANSAC
- Outperformed DINOv2-only baseline by **+6.8 Mean AR** over all 7 BOP benchmark datasets; surpassed DZOP diffusion baseline on LM-O and YCB-V
- Calibrated feature configuration through systematic ablations on fusion weight, timestep, and decoder layer selection to maximize descriptor quality

Professional Experience

Anheuser-Busch InBev, Bangalore **Aug 2020 – Jul 2024**

- Fast-tracked career at AB InBev with **2 promotions in 3 years**; recognized with 4 awards for outstanding performance

Senior Data Scientist *Oct 2023 – Jul 2024*

- **D2C Sales Decomposition**: Built a deep learning sales attribution model in TensorFlow with custom domain-specific layers and hierarchical parameter sharing across brands; automated retraining pipeline via GitHub Actions with DVC-managed data pulls
- Achieved **88.4% R^2** on sales prediction; measured impact of **$49M+** in media spend across 60+ brands

Data Scientist *Sep 2021 – Sep 2023*

- **E-Retail ROI**: Developed a Linear Mixed Effects model to measure impact of $1.5M+ promotion spend, achieving **82.5% R^2**; automated hyperparameter tuning via Bayesian Optimization
- **Generative AI**: Built a RAG pipeline using LangChain and Chroma with chain-of-thought prompting to surface insights from internal databases; demonstrated directly to ABI's CTO
- Engineered a GPT-4-powered summarization tool with hierarchical temporal aggregation of social signals across hourly, daily and weekly horizons; developed in collaboration with Bain & Company

Associate Data Scientist *Aug 2020 – Aug 2021*

- **Media Effectiveness**: Developed an XGBoost model with SHAP-based feature attribution to explain platform-level drivers of media ROI; achieved **87.3% R^2** with top-K features capturing **>75%** cumulative SHAP attribution
- **MMM Lead, Canada**: Led end-to-end modeling and stakeholder presentations for the Canada market across a $17M+ marketing budget; recommendations drove an estimated **$4M in incremental net revenue**

Competitions & Projects

Whisper-Accent **Feb 2026**

- Designed accent-robust ASR by conditioning Whisper's decoder with learned accent embeddings via Adaptive Layer Normalization, supporting 23 accents within a single shared architecture while **training under 5% of total parameters**
- Achieved **3.2 pp** WER reduction over the fine-tuned Whisper Medium baseline; accent classifier reached **95.7%** accuracy
- Deployed a Dockerized FastAPI + Gradio inference service with GPU/CPU fallback, async audio handling, and a public HuggingFace model release

ABI Global Analytics Hackathon — 7th Place Globally (150+ Teams) **Sep 2023**

- Designed a promotion spend optimization engine modeling discount elasticity as learned SKU-level coefficients to allocate budget across brand, pack, and segment levels; achieved **MAPE < 10%** on sales prediction and **+0.7 ROI** on optimized spend

Byzantine Fault-Tolerant Banking Ledger **Nov 2025**

- Implemented a Byzantine fault-tolerant distributed transaction processing system in Golang across 7 replicas communicating via gRPC
- Optimized consensus by linearizing PBFT and implementing SBFT with BLS threshold signatures and checkpointing to minimize communication overhead and bound memory growth
- Benchmarked under YCSB workloads achieving **36.2 OPS** throughput with **17.5ms** read and **44.3ms** write latency