

# Vuthtyra (Teera) Yong

Ithaca NY 14850 | (607) 229 | vuthtyrayong@gmail.com | <https://ty3340.github.io> |  
<https://www.linkedin.com/in/vuyong/>  
Permanent Resident (Green Card Holder)

## EDUCATION

### Syracuse University, School of Information Studies

Syracuse, NY

M.S. in Applied Data Science, GPA: 4.0/4.0

December 2025

**Relevant Coursework:** Applied Machine Learning, Natural Language Processing, Deep Learning, Advanced Big Data Management, Big Data Analytics, Data Warehousing, Cloud Management, Business Analytics

## TECHNICAL SKILLS

**Languages:** Python, SQL, R, C++

**Machine Learning:** PyTorch, TensorFlow, Hugging Face, LangChain, RAG, supervised/unsupervised learning

**Data & Big Data:** Apache Spark, Kafka, MongoDB, Redis, Cassandra, Hadoop, Neo4j, Elasticsearch, dbt

**Cloud & Dev Tools:** Azure, AWS, Snowflake, Docker, Kubernetes, Git, Linux

**Visualization & Web:** Tableau, Power BI, Excel, Flask, FastAPI, Streamlit, React

## WORK EXPERIENCE

### Syracuse University

Syracuse, New York

AI Developer

April 2025 – April 2026

- Built a full-stack AI research alert system using FastAPI, React/Vite, Azure Functions, Cosmos DB, Azure AI Search, Azure OpenAI, and Azure Communication Services Email.
- Developed an automated recommendation pipeline that ingests arXiv preprints, enriches researcher profiles with ORCID/OpenAlex data, generates embeddings, performs vector-based matching, and sends daily email digests.
- Designed persistence, deduplication, alert logging, and 24-hour temporary vector cleanup to reduce duplicate recommendations and control Azure AI Search storage growth.

### DENSO (Cambodia) Co., Ltd.

Phnom Penh, Cambodia

Machine Learning Engineer (Internship)

May 2025 – August 2025

- Fine-tuning pretrained YOLO and Faster R-CNN models on manufacturing image datasets for automated defect detection and quality inspection use cases.
- Benchmarked model performance using precision, recall, mAP, and inference speed to identify the most practical solution for real-time inspection workflows.
- Prototyped a containerized inference service using Docker and FastAPI to simulate real-time automated inspection.

## PROJECT

### Multimodal Product Search & Retrieval | Python, SigLIP, PyTorch, FAISS

- Fine-tuned a SigLIP dual-encoder on a curated subset of the Amazon Berkeley Objects dataset (11K+ products) to align image and text embeddings for semantic search.
- Built an end-to-end retrieval pipeline with FAISS indexes for text-to-image and image-to-text search, achieving **Recall@3 = 0.95** and **MRR = 0.86** on the test set.
- Analyzed learned embeddings with t-SNE visualizations and qualitative search examples to validate stronger cross-modal alignment after fine-tuning.

## CERTIFICATIONS & BADGES

IBM AI Engineering Professional Certificate

April 2026

Microsoft Certified: Azure AI Engineer Associate (AI-102)

March 2026

AWS Certified Cloud Practitioner (CLF-C02)

February 2025

Microsoft Certified: Azure Fundamentals (AZ-900)

February 2025

Snowflake Hands-on Essentials: Data Warehouse

January 2025