Hieu (Henry) Ngo, Ph.D.

Union City, NJ • ngohieu241@gmail.com • 508-542-2191
Profiles: LinkedIn • Google Scholar

PROFESSIONAL SUMMARY

Applied Researcher with Ph.D. in Computer & Information Science and a postdoc at Yeshiva University. Experience delivering privacy-preserving ML systems and data pipelines across NIH/NSF projects. Strengths in modeling, evaluation, optimization, and deployment for high-dimensional, real-world data (digital health, IoT/CPS, V2X). Passionate about translating state-of-the-art AI into reliable, human-centered products.

CORE SKILLS

Machine Learning • Deep Learning • Applied Research • PyTorch • Hugging Face • Lightning • Vector Databases • Retrieval/Evaluation • AWS (EC2, S3, Batch, ECR) • Containers (Docker) • Experiment Tracking • Self-Supervised Learning • Robustness/Explainability • RLHF (exposure) • Training Optimization (profiling, mixed precision) • Distributed Training (DDP) • Data Privacy/Security • Federated/Distributed Learning • SQL/Python Data Engineering • Reproducible ML Pipelines

TECHNICAL STACK

Languages/Tools: Python, SQL, MATLAB, R; PyTorch, scikit-learn, Lightning, Hugging Face Transformers/Datasets; pandas/NumPy; matplotlib/plotly; Docker; Git.
Cloud: AWS (EC2, S3, IAM, Batch, ECR); Linux.
Data/Systems: Federated learning, privacy-preserving analytics, vector indexing & retrieval, streaming-friendly evaluation.
ML Topics: SSL, temporal modeling, clustering/representation learning, robustness, explainability, model validation.

APPLIED RESEARCH EXPERIENCE

Postdoctoral Researcher — Yeshiva University, Katz School of Science & Health (NYC)

Nov 2025 – Present

- Lead applied research on trustworthy ML: privacy-preserving learning, federated evaluation, and visualization-aided validation on longitudinal datasets.
- Prototype PyTorch pipelines with Lightning for training/evaluation; package datasets with Hugging Face Datasets; automate evaluation on AWS.
- Design retrieval/evaluation workflows with VectorDB-style embeddings; author reports and present to stakeholders.

Graduate Researcher — University of Massachusetts Dartmouth

2019 – 2025

- NIH privacy-aware analytics for multi-site digital health: federated learning and fuzzy clustering with multiple imputation; peer-reviewed publications.
- NSF cooperative perception (V2X): compact perception representations for bandwidth/latency constraints; PyTorch-based training and evaluation; IEEE publications.
- mmWave WBAN beamforming; IEEE ComSoc MCTC Best Paper Award (2021).

SELECTED PROJECTS

- Federated Longitudinal Modeling: PyTorch + Lightning pipeline for multi-site time-series; privacy-preserving training/evaluation; uncertainty reporting; AWS Batch orchestration.
- Representation Learning for Retrieval: embedding-based similarity search over heterogeneous signals; vector indexing for cohorting/audit; reproducible evaluation harness.
- Robustness & Explainability Toolkit: perturbation tests, ablations, SHAP-style attribution for incomplete/irregular data; reporting templates aligned to product goals.

SELECTED PUBLICATIONS

- Ngo, H.; Fang, H.; Wang, H. (2023). *Federated fuzzy clustering for decentralized incomplete longitudinal behavioral data*. **IEEE Internet of Things Journal**. doi:10.1109/JIOT.2023.3343719
- Ngo, H.; Fang, H.; Wang, H. (2023). *Cooperative perception with V2V communication for autonomous vehicles*. **IEEE Transactions on Vehicular Technology**. doi:10.1109/TVT.2023.3264020
- Ngo, H.X.; Fang, H.; Wang, H. (2021). *Beamforming and scalable image processing in vehicle-to-vehicle networks*. **Journal of Signal Processing Systems**. doi:10.1007/s11265-021-01696-6

EDUCATION

Ph.D., Computer & Information Science — University of Massachusetts Dartmouth

Dissertation: iMIF2V2: Intelligent Multiple Imputation Federated Fuzzy Clustering with Visualization and Validation for Longitudinal Digital Trials

B.S., Computer Science — University of Massachusetts Dartmouth

AWARDS & HONORS

- CIS Graduate Research Award, UMass Dartmouth (2023)
- NSF Student Travel Grant (2022)
- IEEE ComSoc MCTC Best Paper Award (2021)

KEYWORDS

PyTorch; Hugging Face; Lightning; AWS; Vector Databases; Foundation Models; LLMs (evaluation/finetuning exposure); Self-Supervised Learning; Robustness; Explainability; RLHF; Training Optimization; Distributed Training; Quantization/Sparsity (familiarity); Model Compression (familiarity); Data Privacy; Federated Learning; Reproducibility; Experiment Tracking; ML Engineering; Applied Research; Cross-functional Collaboration; Productization