Dae Yon Hwang

• Address: 150 Logan Ave, Toronto, ON, Canada • Cell: 1-437-345-3631 • Email: eoduself@gmail.com
Linkedin: www.linkedin.com/in/dae-yon-hwang-a39076153/
Personal Website: eoduself.github.io/daeyonhwang/

EDUCATION

University of Toronto	Ph.D. in Electrical & Computer Engineering, **GPA**: 4.0/4.0	Nov 2022
Texas A&M University	Master of Science in Electrical Engineering, **GPA**: 4.0/4.0	May 2016
Hanyang University	B.S. in Electronic Engineering, **GPA**: 3.56/4.0 ***(Cum Laude)***	Feb 2014

WORK EXPERIENCE

Layer 6 AI @TD, Trustworthy AI - Senior Research Machine Learning Scientist — May 2025 - Present

- **Sequential Monte Carlo (SMC)-based Agentic Framework**
  - Designed SMC-driven agentic systems with solid theoretical grounding and strong empirical performance
  - Built an SLM-based critic integrated with a black-box policy LLM for improved generalizability and scalability
  - Tested across diverse agentic scenarios, showing competitive performance at lower computational costs
- **Content Safety for Frontier Conversational Model**
  - Designed agentic red-teaming workflows to stress-test customer-facing conversational models
  - Deployed a three-layer safety guardrail system, significantly improving model reliability and trustworthiness
  - Built a model-agnostic, geometry-based blue-teaming framework scalable across content-safety methods
- **Customer-Facing Chatbot for a Finance Application**
  - Customized RAG-based system for high-stake finance domain, balancing model capability with trustworthiness
  - Designed query classification, rewriting, and abstention logic to ensure accurate and policy-aligned responses
  - Coordinated human-in-the loop testing to refine conversational tone, professionalism, and user experience

Amazon Science, Artificial General Intelligence - Applied Scientist II — Sep 2022 - Apr 2025

- **Post-Training of Multimodal Foundation Models for General and Reasoning Applications**
  - Generated synthetic data with multi-step reasoning traces for post-training
  - Applied LLM-as-a-Judge techniques for data filtering and quality assurance
  - Conducted fine-tuning and preference optimization focused on domain-specific knowledge and reasoning
  - Utilized chain-of-thought prompting to improve reliability and explainability
  - Contributed to the launch of **Amazon Nova** and improved performance to achieve higher leaderboard ranking
- **Recommendation System with Foundation Models**
  - Customized foundation model with pre-training, instruction fine-tuning, and preference optimization
  - Adapted recommendation system for online settings with continuous user feedback updates
  - Integrated RAG to maintain accuracy and up-to-date recommendation without retraining
  - Optimized prompts with chain-of-thoughts and proposed novel online evaluation metrics
  - Deployed a customer-facing recommendation system after positive online testing
- **Information Retrieval Model for Alexa Devices**
  - Generated diverse synthetic data using LLM-based methods and fine-tuned models for zero-shot scenario
  - Developed a novel GAN-based augmentation approach to produce high quality synthetic data
  - Customized traditional retrieval systems, such as ElasticSearch, for specific usage scenarios
  - Advanced representation learning to build transferrable and generalizable data representation
  - Designed a novel dimensionality reduction to reduce computational costs and boost the generalizability

University of Toronto, Biometrics Security Lab - Research Assistant — Sep 2018 - Sep 2022

- **User Verification System using Heart Signal (CNN, RNN, GAN, VAE)**
  - Applied time- and frequency-domain signal processing techniques to construct datasets
  - Identified time-stable and unique features in cardiac signals for robust user verification
  - Benchmarked traditional ML models and deep learning models to determine optimal methods

- Developed a system robust to adversarial attacks and security threats
- Collected physiological signals from 170 participants, creating the largest public dataset in this domain

• **Human Activity Recognition with Wearable Device**
- Used inertial and physiological sensors to build a robust multimodal activity recognition system
- Designed hierarchical deep learning models to classify the diverse human activities

Hyundai MOBIS, DAS Control Engineering - Research Engineer Jul 2016 - Feb 2018
- Evaluated the recognition rate of multi-function camera across varied environments (urban, rural, highway)
- Designed and optimized the driver attention warning logic
- Conducted on-road vehicle testing in challenging scenarios to diagnose and resolve new-vehicle issues

Texas A&M University, Laboratory for Optical Diagnosis and Imaging - Research Assistant Sep 2014 - May 2016
- Applied image-processing methods (deconvolution, filtering) to enhance biomedical image quality
- Performed feature selection experiments on large scale datasets
- Optimized ML classifiers (primarily Gaussian-kernel SVM) to achieve lower error rate

RECENT PUBLICATIONS (Full list covered in Google Scholar)

The Amazon Nova Family of Models: Technical Report and Model Card
Amazon Artificial General Intelligence (Contributor: DY Hwang) Dec 2024

Empirical Methods in Natural Language Processing (EMNLP) 2024
Link, Synthesize, Retrieve: Universal Document Linking for Zero-Shot Information Retrieval Nov 2024
DY Hwang, B Taha, H Pande, Y Nechaev

The 4th Workshop on Multilingual Representation Learning 2024 @ EMNLP 2024
Unsupervised Text Representation Learning via Instruction-Tuning for Zero-Shot Dense Retrieval Nov 2024
Q Zeng, Z Qiu, DY Hwang, X He, WM. Campbell

International Conference on Natural Language Generation (INLG) 2023
GAN-LM: Generative Adversarial Network using Language Models for Downstream Applications Sep 2023
DY Hwang, Y Nechaev, CD Lichy, R Zhang

Association for Computational Linguistics (ACL) 2023
EmbedTextNet: Dimension Reduction with Weighted Reconstruction and Correlation Losses for Efficient Text Embedding Jul 2023
DY Hwang, B Taha, Y Nechaev

2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)
Eeg Emotion Recognition Via Ensemble Learning Representations Jun 2023
B Taha, DY Hwang, D Hatzinakos

IEEE Journal of Selected Topics in Signal Processing
EyeDrive: A Deep Learning Model for Continuous Driver Authentication Jan 2023
B Taha, SNA Seha, DY Hwang, D Hatzinakos

2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)
Hierarchical Deep Learning Model with Inertial and Physiological Sensors Fusion for Wearable-based Human Activity Recognition May 2022
DY Hwang, PC Ng, Y Yu, Y Wang, P Spachos, D Hatzinakos, KN. Plataniotis

Journal of Signal Processing Systems (Invited paper)
A New Score Level Fusion Approach for Stable User Verification System Using the PPG Signal Mar 2022
DY Hwang, B Taha, D Hatzinakos

IEEE Transactions on Information, Forensics and Security
PBGAN: Learning PPG Representations from GAN for Time-Stable and Unique Verification System Oct 2021
DY Hwang, B Taha, D Hatzinakos

2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Variation-Stable Fusion for PPG-based Biometric System Jun 2021
DY Hwang, B Taha, D Hatzinakos
IEEE Transactions on Information, Forensics and Security
Evaluation of the Time Stability and Uniqueness in PPG based Biometric System Jul 2020
DY Hwang, B Taha, DS Lee, D Hatzinakos

HONORS

SGS Conference Grant - Outstanding student who do conference presentation May 2019
Hanyang International Scholarship - Outstanding student who is studying abroad Sep 2014 - May 2016
Full National Science & Engineering Scholarship - Outstanding engineering student: 5 times Sep 2009 - Sep 2013
Full Grade Scholarship - Top student in major (Rank in 1/215) Mar 2009

PROFESSIONAL SERVICE

Reviewer - EMNLP 2023-2025**,** ACL 2023-2025, ACL Rolling Review, RepL4NLP @ ACL 2024, SyntheticData4ML @ NeurIPS 2023, IEEE Journal of Biomedical and Health Informatics, IEEE Transactions on Information, Forensics and Security
Program Committee - EMNLP 2023 Industry Track
Talks - Career Guidance Seminar @ Incheon National University (Dec 2023), GAN with LM @ ML for Healthcare Roundtable in Amazon (Oct 2023)

SKILLS

Technical Skills - C, C++, Python (including TensorFlow, PyTorch), MATLAB (including Stateflow), AWS
Technical Areas - Signal Processing, Computer Vision, Natural Language Processing, Machine Learning, Deep Learning
Foreign Language - Native in Korean, Fluent in English

REFERRERS

At Amazon - Collaborated closely with **Sr. Applied Scientist Harshit Pande** harshit.pande@mail.mcgill.ca
At Amazon - Collaborated closely with **Sr. Applied Scientist Yaroslav Nechaev** remper@me.com
During Ph.D. degree - Supervised by **Prof. Dimitrios Hatzinakos** dimitris@comm.utoronto.ca
During Master degree - Supervised by **Prof. Javier A. Jo** javierjo@ou.edu