David Yang

+1 (424) 394-3236 dongjun.yang@emory.edu dong-jun-david-yang-a97825228 /david0530

Degree	Institute	GPA
Computer Science	Emory University, 2022-2026	3.85

Skills

Languages: Python, Java, C++, SQL
Domains: Machine Learning, Deep Learning, Artificial Intelligence, LLM
Platforms: Linux, Git, CI/CD, RESTful APIs, Google Cloud Platform
Frameworks: PyTorch, Scikit-Learn, Pandas, numpy, Matlab

Experience

ML Research Assistant | Emory University @ MAIX — Jan 2025 - Current
Mentor: Dr. Ran Xiao

- Engineered **feature extraction** and pre-processing on infant growth scale **EEG data** for 245 visits ,totaling 142GB with 4217 segment, transforming **time-series** signals into FSD and DE features while constructing electrode distance adjacency matrices
- Deployed implemented and evaluated multiple Machine Learning models (**SVR, MLP, XGboost, CNN, GCN, GAT, SparseDGCNN, DCGNN, RGNN, HET**) for **regression** tasks on growth scale establishing performance benchmarks for infant developmental prediction

NSF REU ML Research Intern | UNC Greensboro @ GraLNA — May 2024 - Nov 2024
Mentor: Dr. Minjeong Kim

- Engineered **Pre-processing** on Parkinson's and Alzheimer's **fMRI** data for 259 patients with 4 different label status, generating node features, edge features, and node and edge **functional connectivity matrices**
- **Designed domain specific co-embedding GNN** model more simulating a human brain integrating both edge and node connectivity matrices with heterogeneous dimensions during aggregation
- Deployed traditional brain disease classification approaches as a **benchmark** including **CNN, GCN, CRGNN, and MGNN**. Our model achieved superior diagnostic accuracy of **7%** improvement compared to latest GNN model with multi **classification**. It have been accepted and pending publication in **IEEE ISBI**

ML Research Assistant | Emory University @ EGM — January 2024 - May 2024
Mentor: Dr. Carl Yang

- Deployed **Hypergraph Neural Network** model (hyEHR) on **Electronic Health Records** for type-2 diabetes prediction, achieving competitive performance against established benchmarks
- **Enhanced** model performance(**2%** acc increase) by implementing multiple self-supervised pretraining strategies for **GNN**
- Constructed comprehensive benchmarking framework for MIMIC-3 and type-2 diabetes dataset using diverse machine learning approaches (**SVM, Logistic Regression, Random Forest, XGBoost, MLP)** for healthcare outcome prediction

Software Engineering Intern | METY Technology — 2023 - 2024

- Engineered **CICD** pipeline for **Git** repositories, detecting and classifying **100+ vulnerabilities**. The system generated comprehensive reports including contributor identification, vulnerability categorization, file location mapping, and severity assessment, leveraging OpenAI's API for intelligent classification.

AI Intern | Skytain Capital — 2023

- **Developed and deployed** an OpenAI-powered **chat AI** system for housing rent inquiries, integrating the solution across web and social platforms (WhatsApp, Instagram, Facebook)

Publication & Presentation

1. **Yang, David**; Abdelmegeed, Mostafa; Modl, John; Kim, Minjeong. "Edge-Boosted Graph Learning for Functional Brain Connectivity Analysis", IEEE International Symposium on Biomedical Imaging (ISBI) 4page paper, Houston, TX, April 14-17, 2025.
2. **Yang, David**. "Supervised Data Mining classification on Dry Bean dataset", Annual Oxford Research Scholars Spring Symposium; April 2024; Oxford, GA.

AWARDS

- Received a Travel Grant from Computing Research Association (CRA) 2025

PROJECTS & ACTIVITIES

ML implementation on Iris dataset | JAVA 2023
Oxford Research Scholar, Oxford College of Emory University Project Link
Mentor: Dr. Hai Le

- Implemented Naive Bayes and Decision Tree algorithms from scratch in Java, achieving 96% accuracy on Iris dataset and 90% on Dry Bean dataset, demonstrating performance comparable to SVM benchmarks without using external ML libraries

Morenmost Fashion Brand Website | HTML, CSS 2019 - 2021
Co-founder, Morenmost, South Korea Project Link

- Founded and developed e-commerce platform for fashion brand Morenmost, implementing custom website design and digital marketing strategies that achieved 2,000+ monthly visitors

Relevant Coursework: Foundations of CS, Machine Learning Applications, Database Systems, DS & A, Linear Algebra, Analysis of Algorithm, Machine Level Programming, Deep Learning, Data Mining