

TUAN ANH NGO

ngotuananh10906@gmail.com

| Waterville, ME

| [Linkedin](#)

| [GitHub](#)

| [Website](#)

EDUCATION

Colby College, Maine, USA

Expected Graduation: 5/2028

- Intended Major: Computer Science
- Related Courses: CS231 - Data Structures & Algorithms, MA160 - Series and Multi-Variable Calculus

VNUHCM - High School for The Gifted, Ho Chi Minh City, Vietnam

8/2021 - 5/2024

- Specialized Math GPA: 10.0/10.0
- Overall GPA: 9.8 / 10 (**Top 1%**)
- Dean's List (2022 & 2023)

International Qualifications:

- IELTS: 8.5 / 9.0 7/2023 | SAT: 1560 / 1600 8/2023

AWARDS

- Student Best Poster Presentation, Division of AI & Data Science for Health, BME10 Conference 7/2024
- Global Finalist, Top 10/2545, High School Academic Research Competition (SARC) 5/2024
- 7% distinction, Runner-up, Immerse Education Essay Competition, Topic: Artificial Intelligence 10/2023

EXPERIENCE

STEMulate Research Program, CS & AI Research Scholar

6/2024 - Present

- Contributed to current understanding by investigating the influence of data modality and the feature importance of demographic, socioeconomic, neurobiological, and cognitive indices in Machine Learning for dementia detection
- Utilized cutting-edge Python libraries (numpy, pandas, seaborn) on the adapted dataset from the OASIS-2 dataset, containing 373 data instances of 150 subjects, to conduct Early Data Analysis, namely normalizing, feature selection, and visualization
- Conducted a comparative analysis of 4 techniques - Random Forest, AdaBoost, Support Vector Machine, and Naive Naves - to identify an optimal model for dementia classification on multimodal data
- Achieved 87.50% accuracy in classifying the Clinical Dementia Rating of 112 test subjects using the Random Forest Classifier
- Employed eXplainable AI (XAI) to highlight the modalities' feature importance and alignment with current knowledge on dementia risks, improving the transparency and reliability of Machine Learning models

Jane Street's WiSE Program, Participant

7/2024

- Selected as one of the 41 pre-college students to attend a fully-covered multi-day program for at Jane Street Hong Kong
- Engaged in sessions on the application of math, computer science, and probability concepts in quantitative trading
- Networked with Jane Street professionals, gaining insights into transitioning from STEM studies to finance careers
- Participated in interactive problem-solving activities and social events for fostering collaboration and professional growth

Brain Health Lab, International University, Research Fellow, Team Representative

6/2023 - 6/2024

Brain Health Lab focuses on developing novel diagnostic tools and interventions for mental health and brain health.

- Researched the application of Artificial Intelligence in Alzheimer's Disease (AD) early detection through narrative speech
- Utilized data-centered Python libraries (numpy, pandas, re) to analyze the ADDreSS Challenge 2020 dataset comprising 54 Alzheimer's Disease patients and 54 healthy control
- Constructed 4 AI language models BERT, ROBERTA, distilBERT, and ALBERT using Object-Oriented Programming, fine-tuning, and Learning Rate Scheduler for dynamic hyperparameter adjustment during training
- Yielded the highest results at 0.82 precision, 0.81 recall, 0.84 f1-score, and 83% overall accuracy in detecting AD patients
- Accepted for publication & poster presentation at the BME10 Conference in Phan Thiet City, Vietnam

Publication: Khoi, N., Anh, N., Thu, N. and Phuc, P. (2023). A Comparative Study of Transfer Learning Techniques for Alzheimer's Disease Early Detection

PROJECTS

Molecule Semantic Search

- Developed a Molecule Semantic Search Engine using text embeddings from NomicAI API to retrieve relevant chemical compounds from a dataset of 50,000 compounds using text-based queries
- Performed filtering, truncating, and dropping to clean and lighten the dataset while retaining relevant targetted information
- Implemented NomicAI setup for embedding and searching descriptions, including mapping data, creating an interactive map for visualization, and retrieving search results
- Designed and implemented a vector search function that embeds text-based queries and retrieves 10 most relevant compounds

Traffic Signs Detection

- Developed a Traffic Sign Classifier for Autonomous Vehicles using Convolutional Neural Networks (CNNs) with the German Traffic Sign Recognition Benchmark (GTSRB) dataset
- Constructed a deep learning model with 3 convolutional layers, dropout regularization, and 2 fully connected layers, achieving an 88.56% accuracy on the test set.
- Optimized model performance using CrossEntropyLoss and Adam optimizer, fine-tuning hyperparameters: learning rate of 0.0005, batch size of 64, and 24 epochs, yielding a significant reduction in training loss from 3.34 to 0.026

SKILLS

Languages: Vietnamese (native), English (fluent), German (A1 - beginner)

Computer Skills: Machine Learning, Deep Learning, Software Development

Research Skills: Journal club presentation, literature review, research proposal composing, data visualization