







Tuan Nguyen Van Anh

 (+84)972865664
 nvatuan3@gmail.com
 <https://tuanio.github.io/>

Work Experience

- May 2021 – Aug 2021  **WeAi - EduplaX**, Data Engineer Intern.
- Study Data Engineer concepts, build a small tool for internal use, and do data preprocessing and data ingesting steps.
- Aug 2021 – Present  **WeAi - EduplaX**, Machine Learning Researcher.
- Successfully apply the research's results to build a new feature for the product.
 - Research Computerized Adaptive Testing, Automatic Speech Recognition.

Education

- 2019 – Present  **Student, Industrial University of Ho Chi Minh City (IUH), Faculty of Information of Technology**, Data Science Department - GPA: 3.35/4.
Thesis title: Semi-Supervised Learning approach for Automatic Speech Verification.
(Progress: 60%)
- Supervisor: Associate Professor [Linh Nguyen](#).

Research Publications


Conference Proceedings

- 1 Trinh, T. B. B., Dinh, H. H. D., Phan, L. H. V., Tran, N. B., Nguyen, T. T. H., **Nguyen, V. A. T.**, ... Nguyen, C. K. (2021). Using machine learning to detect abnormal in cyber security. In *Young Science Conference - Industrial University of Ho Chi Minh City (ysc-IUH 2021)* (pp. 221–231). Ho Chi Minh city, Vietnam. Retrieved from <https://bit.ly/3SNYTTt>
- 2 **Nguyen, V. A. T.***, Ha, B. A., Trinh, T. B. B., Nguyen, T. T. H., Pham, T. S., Pham, H. T. T., ... Nguyen, T. A. N. (2020). Applying machine learning to determine the level of air pollution basing on historical data. In *Young Science Conference - Industrial University of Ho Chi Minh City (ysc-IUH 2020)* (pp. 256–256). Ho Chi Minh city, Vietnam. Retrieved from <https://bit.ly/3yvyPEt>

In Press

- 1 **Nguyen, V. A. T.***, Nguyen, T. T. H., Nguyen, T. D., Pham, M. T., Dao, D. T., & Dang, T. P. (2022). A novel approach for Vietnamese Speech Recognition using Conformer. (Accepted as short paper in *The 9th International Conference on Future Data and Security Engineering - (FDSE 2022)*), Vietnam.
- 2 **Nguyen, V. A. T.***, Pham, H. T. T., Pham, T. S., & Dang, T. P. (2022). Vietnamese voice recognition using deep learning model. (Accepted, to be published in end of October 2022 in *Young Science Conference - Industrial University of Ho Chi Minh City (ysc - IUH 2022)*), Ho Chi Minh city, Vietnam.
- 3 **Nguyen, V. A. T.***, Trinh, T. B. B., Nguyen, T. T. H., Phan, L. H. V., Tran, N. B., Dinh, H. H. D., ... Nguyen, C. K. (2022). Using Bayesian and Neyman-Pearson hypothesis testing for Autoencoder to detect anomalies in network security. (Accepted, to be published in May 2023 in *Journal of Science and Technology - Industrial University of Ho Chi Minh City (jst - IUH 2022)*), Ho Chi Minh city, Vietnam.

Skills

Languages  English (Strong reading, listening and Intermediate writing, speaking).

Skills (continued)

Coding	Python (PyTorch, Tensorflow, Numpy), C/C++, R, \LaTeX .
Databases	SQL, PostgreSQL, SQLite, MongoDB.
Web Dev	HTML, CSS, JavaScript, ReactJS, Flask, Django.
Misc.	Academic research, teaching, teaching assistant, training, consultation.

Research Experience

- Nov 2021 – Present **Automatic Speech Recognition (ASR).**
- **Topic detail:** Research how a computer precisely recognizes what a human says. Languages are Vietnamese and English.
 - **Skills gained:** Deeply understand the pipeline of Hybrid ASR Systems and End-to-end ASR Systems. Successfully applied state-of-the-art Deep Learning ASR Architecture for English dataset on Vietnamese dataset (Accepted paper). Build a toolkit for experiment End-to-end ASR System: [See my GitHub](#).
- Nov 2020 – Jan 2022 **Anomaly Detection.**
- **Topic detail:** Research how to build Intrusion Detection System (IDS) for Network Security. Experiment on *NSL-KDD* Dataset.
 - **Skills gained:** Understand the pipeline of IDS, and know how to use hypothesis testing for the Unsupervised Learning model and Supervised Learning model for IDS (Both papers have been accepted). [See my GitHub](#).

Research Interest

My current research focuses mainly on Deep Learning techniques to solve End-to-end Automatic Speech Recognition (ASR). i.e., I try to make the computer system exactly recognize what humans say. The languages that I research are both Vietnamese and English. Till now, I have successfully applied state-of-the-art Deep Learning architecture on ASR with English datasets to the Vietnamese datasets. In the near future, I want to research both Natural Language Processing and Speech Processing.

Awards and Achievements

- 2022 **Third Prize**, Euréka at [IUH](#).
- 2021 **Top 20**, startup contest [Inno Greenlife IUH](#).
 Second Prize, [TDMU - Entropy Data Analytics](#) contest.
- 2020 **Consolation Prize**, [Olympic Informatics for Vietnamese students](#) (specialize board).
- 2019 **Consolation Prize**, [ACM/ICPC Asia region](#).

References