(+84) 972865664
Ho Chi Minh City, Vietnam
nvatuan3@gmail.com

Nguyen Van Anh Tuan
Junior Data Scientist

Personal blog: tuanio.github.io
github.com/tuanio
linkedin.com/in/tuanio
kaggle.com/tuannguyenvananh

I have 1+ year of experience working on data science-related tasks like data processing, visualization, and modeling. I also have the skills to research Anomaly Detection (3+ months) and Automatic Speech Recognition (9+ months). Those things come with good programming skills, and I am seeking a job as a Data Scientist.

SKILLS

Programming	Python (PyTorch, Tensorflow, Numpy, Pandas), C/C++, LaTeX.
Communication	English (Strong reading, listening and Intermediate writing, speaking)
Misc.	Academic research, teaching, teaching assistant, training, consultation.

TECHNICAL EXPERIENCE

Machine Learning Researcher / EduplaX   **Aug 2021 — Present**
WeAi   *Ho Chi Minh City, Vietnam*

- Successfully apply the research's results to build a new feature for the product.
- Research Computerized Adaptive Testing, Automatic Speech Recognition.

PROJECTS

Image to Latex   **Aug 2022 — Oct 2022**

- Successfully built a codebase for experimenting with various Encoder-Decoder models for Image to Latex task (GitHub).
- Achieve BLEU score of 77% on im2latex-100k Dataset.
- Team leader.

Vietnamese Automatic Speech Recognition (ASR)   **Dec 2021 — Present**

- Successfully built a toolkit for experimenting with some state-of-the-art ASR models on English dataset to Vietnamese dataset (GitHub).
- Achieve WER 20% on Vivos Test Dataset, training on 115 hours of data (VLSP 2020 and Vivos Train Dataset).
- Have accepted paper to local university YSC - IUH 2022 and accepted as a short paper at FDSE - 2022.
- Team leader.

Anomaly Detection   **Dec 2021 - Mar 2022**

- Successfully applied Hypothesis Testing on an Unsupervised Autoencoder for an Anomaly Detection task.
- Experiment on NSL-KDD Dataset. Achieved accuracy: 96%. GitHub.
- Team leader.

Recommendation System   **Dec 2021 - Mar 2022**

- Successfully apply Content-based Collaborative Filtering in the Android app, Book Recommendation app (Mobile source, API source).
- Team leader.

EDUCATION

Bachelor of Science in Data Science, *Industrial University of Ho Chi Minh City, Vietnam*   Aug 2019 — Present

- GPA: 3.35/4.0
- Thesis title: *Semi-Supervised Learning approach for Automatic Speech Verification.* (Progress: 60%)

REWARDS & ACHIVEMENTS

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

ACTIVITIES

Administrator of Problem Based Learning, the Data Science playground for IUH students.	May 2021 — Aug 2021.
Teaching Assistant, including subjects: Programming Stats and Application, Programming basic, Programming Techniques, Introduction to Programming	2020, 2021