Nguyen Huu **Manh**

manhnh594@gmail.com | +84 868-601-594
https://www.linkedin.com/in/manhitv/ | https://manhitv.github.io/

SUMMARY

I am pursuing my PhD at Applied Artificial Intelligence Institute (A2I2), Deakin University, Australia. My research focuses on explainability of large language models, including hallucinations and fine-tuning techniques.

Before that, I have spent nearly 5 years applying Machine Learning to solve Data Science problems in the industrial fields such as Martech, Fintech, Natural Language Processing and Computer Vision. I am skilled in both data engineering, applying Machine Learning and Deep Learning models. Furthermore, I am certified by Google Cloud as a Professional Data Engineer and Machine Learning Engineer.

In addition to my full-time job, I also spend time teaching and mentoring students, and working on side projects in Data Science.

EDUCATION

Hanoi University of Science and Technology **Hanoi, Vietnam**
Bachelor of Science (B.S.) in Electronics and Communications Engineering *Aug. 2012 - Jun. 2017*

- Honors Program
- GPA: **3.28/4** (convertible to 10-scale: **8.2**) - Class rank: **3rd/22**
- Thesis: "Precoding techniques and optimization for capacity enhancement in visible light communications (VLC) systems".
  – Advisor: Associate Professor Chuyen T. Nguyen
  – Researched and built a mathematics model to simulate multi-user broadcast channels VLC system using zero-forcing precoding technique.
  – Solved optimization problems to find proper parameters of the precoder.

PUBLICATIONS

Luan Tran, **Manh Nguyen**, Cyrus Shahabi. *Representation learning for early sepsis prediction*. *2019*

PROFESSIONAL EXPERIENCE

Applied Artificial Intelligence Institute (A2I2, Deakin University) **Geelong (Victoria), Australia**
PhD Student *Oct 2024 - Current*
My research focuses on explainbility of large language models, including hallucinations and fine-tuning techniques.

XPON Technologies Group (ASX:XPN, Repensentative Office) **Hanoi, Vietnam**
Data Scientist *May 2022 - Sep 2024*
My work focused on developing the Data Science portion of Wondaris (a Customer Data Platform product).

- *Identity Resolution.* Research, design and implement solutions for incremental and batch deterministic identity matching.
- *End-to-end Machine Learning solutions.* Leverage Google Cloud to build end-to-end solutions (from processing data, feature engineering to training, evaluating, predicting and monitoring ML models) for Machine Learning problems such as propensity models, churn prediction model, lifetime value prediction, customer segmentation... based on unified CDP data.
- *Automate Feature Suggestions.* Build a solution that automatically suggests the best features based on clients' input configuration. In addition to meeting many different types of models (classification, regression), input data quality control, and data processing, this engine must also ensure that overfitting is automatically resolved.

Coderschool **Hanoi, Vietnam**
Data Science Mentor *(Part-time)* *Jun. 2022 - Current*
Help 20+ students learn Data Science courses in a better way based on my experience and best practices. The data science courses last 6 months, organized in the form of 1 mentor supporting 1 mentee to build skills related to Python, SQL, pandas, data visualization and practical data analysis projects.

Vin3S (a subsidiary of Vingroup) **Hanoi, Vietnam**
Data Scientist *Apr. 2021 - May 2022*
My work focused on building data warehouses and solving data science problems for VinHomes (a real estate company).

- *ETL data.* Ingested and cleansed data from different sources of VinHomes then built corresponding data warehouses.
- *Data Analysis.* Built analytical reports and proposed detailed actions to improve data quality.

- *Identity Matching.* Developed identity matching rules and aligned with business users to consolidate each dimensional dataset, especially customer profile data.
- *Lead Scoring.* Built a Machine Learning model with over 95% precision to automatically classify potential customers (leads) into groups based on online behavior on the website, responses to the agent as well as transactions and interactions with Vinhomes. This model helps increase monthly hot lead conversion rate by 20%.

Viettel Digital Services (a subsidiary of Viettel Group) **Hanoi, Vietnam**
Data Scientist *Mar. 2020 - Mar. 2021*

My work focused on building Machine Learning models to improve the operation of products on ViettelPay (an e-wallet product).

- *Lead Generation*. Built a Machine Learning model that follows a look-alike scenario to find potential users for new products on ViettelPay. The model delivered 2 times higher average conversion rate by messaging than traditional methods for products from a variety of industries.
- *Credit Scoring.* Built a credit scoring model from raw telecommunications data, achieving Gini efficiency greater than 0.7. The model was created by features from customer profiles, telecommunications consumer behaviors and user locations.

Developer Circle Program (*Part-time*) **Hanoi, Vietnam**
Data Science Instructor *Sep. 2020 - Nov. 2020*

Taught a basic data science course that introduces the process of performing a data science problem as well as modeling algorithms (logistic regression, support vector machine, tree-based modeling, ensemble models).

3S Intersoft **Hanoi, Vietnam**
Machine Learning Engineer *Apr. 2019 - Mar. 2020*

- *Sentiment Analysis.* Built and developed a multi-labels classification model on self-crawled newspaper data. This problem had an imbalanced dataset with 20k samples and 8 output labels. The goal is to filter out important, strongly nuanced information and the final model achieved 94% accuracy and a macro F1 score of 0.83.
- *Computer Vision.* Surveyed AI solutions which are best fit with fore container number (both vertical and horizontal texts). Tested solutions included:
  – Identify the text area (object detection): YOLOv3 / Tensorflow model zoo (RCNN, SSD…)
  – Text Segmentation: Tesseract / OpenCV
  – Text Recognition: CNN/RNN + CTC loss (+ attention scheme)

Viettel Aerospace Institute (a subsidiary of Viettel Group) **Hanoi, Vietnam**
R&D Engineer *Jul. 2017 - Mar. 2019*

- Research, preliminary design and simulation of satellite communication systems.
- Researched and built up theoretical basis about Proportional Navigation (PN) missile guidance law; simulated missile-target engagement on MATLAB.

SKILLS

Technical skills:

- *Programming:* Python/Pyspark, SQL.
- *Machine Learning libraries:* Scikit-learn, Numpy, Pandas, Tensorflow/Keras.
- *Tools:* Google Cloud Platform, data visualization tools (Tableau / Power BI / Google Data Studio / Matplotlib / Seaborn).

Foreign language proficiency: English - PTE 74 *Jun. 2023*

HONORS AND AWARDS

1st Prize, Kalapa Credit Scoring Challenge for Students. *Oct. 2020*
2nd Prize, Vietnam Mathematics Olympiad for University Students *2015*
1st Prize, Vietnam Mathematics Olympiad for University Students *2013*

CERTIFICATIONS

Google Cloud - Professional Machine Learning Certification *Sep. 2022*
Google Cloud - Professional Data Engineer Certification *Dec. 2022*

Last updated: Jun. 2024