

To Duc Anh

toducanh2001@gmail.com

My LinkedIn: To Duc Anh

 HyprnX

</> Python, C/C++, Docker, R, SQL, HTML

See Dark theme

INTRODUCTION

I am a third year Data Science student at National Economics University (NEU) who is learning and gaining knowledge and experience in my field. With the passion for data science, economics and programming, i believe that i am capable of finishing assigned task at the finest level.

EXPERIENCE

• DSLAB NEU

Member

Hanoi, Vietnam

November 2021 - Present

- DSLAB is a laboratory belong to the Faculty of Mathematics Economics (MFE) , National Economics University (NEU), specialized in the fields of Data Science and Artificial Intelligent (DS AI)).

• IBM's Qiskit Localization Team

Translator/Proofreader

(remote) Hanoi, Vietnam

March 2021 - Present

- Translate IBM's Qiskit documentation from English to Vietnamese for the community.

Technologies: Python, Qiskit, Crowdin.

Theory: Programming, Quantum Physics, Quantum Programming.

• Vietsearch.org

Intern

Hanoi, Vietnam

August 2020 - Present

- Designed, implemented crawlers to crawl data from LinkedIn, Google, Wikipedia, etc. with BeautifulSoup and requests library in Python.
- Optimized existing crawlers, lower 3 times running time.
- Developed query data APIs to serve customer's demand base on Elasticsearch search engine.
- Designed unit and intergration tests for APIs.

Technologies: Docker, Python, Swagger, Flask, Beautifulsoup, Selenium, Unittest, etc.

Database: Elasticsearch, MongoDB.

Theory: Crawling data, cleaning data, systematize data, Query, API, Testing.

EDUCATION

• National Economics University

Major: Data Science in Economics and Business

Hanoi, Vietnam

September 2019 – May 2023

- GPA: 3.45/4.0
- The second semester honored student of the school year 2021
- Achieved A+ (10.0) in Data Structure and Algorithm course
- Achieved A+ grade in Machine Learning 1 course
- Achieved A+ grade in Data Preparation and Visualization

• Vietnam - Germany High school

Hanoi, Vietnam

September 2016 – May 2019

- GPA: 8.6/10
- Third Prize in Biology Contest (2017, 2018)
- Third Prize in Geographic Contest (2018)

CERTIFICATIONS

Certificate	Received Date
(Name and score(s) are given, if possible below)	
- IELTS 7.0	Issued April 2019
- The Internet and Computing Core Certification: 2520	Issued December 2019
- BLOCKCHAIN MATHEMATICS AND COMPUTING	Issued July 2021
- Qiskit Localization Contributor	Issued May 2021

PROJECTS

- Used car prediction project: A personal project participated in a Kaggle Competition, to get the lowest Root Mean Squared Error(RMSE) when predicting the used car's price. NaNs in Numeric columns were replaced with mean. One-hot Encoding and Target Encoding were also used to encode categorical data. There were ten regression models used to train the data with the help of GridsearchCV. The best model was LGMBRegressor, which archived the RMSE of 119k, rank fourth in the competition.
- Exploratory Data Analysis for Hotel dataset: Work with a group of three as the leader for the midterm project of the Data Preparation and Visualization course to create an EDA for the Hotel dataset. The group used Python libraries such as NumPy for calculation, Pandas for data manipulation, Matplotlib, and Plotly for plotting data and creating interactive charts. The final result was 9 on a 10 scale.
- Database Management System final project: Directed a team of four to design and implement a database to manage a telecommunication company using SQL as a final project for Database Management System (DBMS). The database allows the company to create plans to serve prepaid and postpaid customers, manage cash flow, and log errors. The design also allows users to create feedback and rating. The final result received a nine on a ten scale..
- Personal Website: I want to make an introduction website for myself, so I taught myself HTML and CSS to build one and then used Github Pages to publish my website.

COURSES

- Data Structure and Algorithm (DSA)** NEU, Hanoi, Vietnam
OOP: Understand about four concepts of OOP
Data Structures & Theories: Understand about Time Complexity, Space Complexity, Recursion and the concepts and usages of basic data structures such as Array, Queue, Stack, Tree, Heap.
Sorting Algorithms: Experience using Binary Search, Insertion sort, Heap sort, Bubble sort. Basic understanding of Radix sort
- Database Management System (DBMS)** NEU, Hanoi, Vietnam
Designing database: Understand Business Rules to design tables and database.
Query: Experienced how to get data from one or multiple table using SELECT, JOIN.
View, Procedure and Trigger: Know how to create a view to quickly see data. Understand the usages of Procedure and Trigger, applied those usages in the final project.
- Data Preparation and Visualization** NEU, Hanoi, Vietnam
Preparing data: Experienced using Python libraries such as Numpy and Pandas to calculate, manipulate and clean data. Implemented other encodings to encode categorical data
Visualizing data: Used Matplotlib and Plotly to visualize interactive charts
- Machine Learning 1 (Base on Stanford CS229)** DSLAB, NEU, Hanoi, Vietnam
Maths: Understand how the model is created, experience proving the algorithms.
Models: Familiar with different types of regression models such as Linear, Ridge, Lasso, ElasticNet Regression, as well as classification models like SVM, K-mean, Decision Tree. Experienced Ensemble methods such as Decision Tree.