To Duc Anh

toducanh2001@gmail.com
My Phone: 0983691898
My LinkedIn: To Duc Anh

Hyprnx
</> Python, PyTorch, Docker, SQL, AWS

See light theme

Introduction

Management Associate at Techcombank Vietnam. Graduated from National Economics University with distinction degree, major in Data Science. Detail-oriented, Mathematics, Artificial Intelligence, and Data Science enthusiast. Excellent at self-study, and communication skills, capable of finishing works with superb quality.

Experiences

- **Techcombank Vietnam - Data and Analytics Division** — Hanoi, Vietnam
  Management Trainee Associate - Key Talent — *January 2023 - Now*
  Advanced Analytics and Innovation — *August 2023 - Now*
  - Analyze and build Machine Learning/Deep Learning models to analyze and predict customer's transactions purpose.
  - Analyze business needs and build AWS ad-group for business users to enable them building and migrating Machine Learning pipelines to on-cloud enironment

  Business Intelligence — *June 2023 - July 2023*
  - Play a deciding role in developing a strategic initiative project for Techcombank as Technology Advisor for Accenture Outsourcing Solution.
  - Take part in building a Power Workbench, a service that enable business users to build and analyze isolated data which dedicated built for them.

  Data Engineering and Delivery — *January 2023 - May 2023*
  - Analyze data updating pattern - CDC and SCD types.
  - Delivered 30 data ETLs that run daily.
  - Familiar with Techcombank's tech stack and on cloud environments.

- **KiotViet (Former Citigo Software) - AI team, Data Department** — Hanoi, Vietnam
  Intern Data Scientist — *August 2022 - December 2022*
  - Take part in constructing text classification machine learning model to predict product's category utilizing product name.
  - Design and implemented a package that be able to extract brand name from product name, base on fuzzy search, regex and Trie Data Structure. Published to Gitlab for internal use.
  - Upgraded existing Address Suggestion module to improve suggestion accuracy based on existing module and Elasticsearch.

  Technologies/Tools: Sklearn, Regex, Pytorch and Keras

- **DSLAB NEU** — Hanoi, Vietnam
  Research Assistant — *November 2021 - Present*
  - Responsible for building a face attendance system on nVidia Jetson Nano. The model recorded a 95% accuracy on testset. Ready to bedeployed.

- **Vietsearch Foundation** — Hanoi, Vietnam
  Intern & Collaborator — *August 2020 - March 2022*
  - Designed, implemented crawlers to crawl data from LinkedIn, Google, Wikipedia, etc. with BeautifulSoup and requests library in Python.
  - Parallelized existing crawlers, lower three times running time.
  - Developed query data APIs to serve customer's demand based 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** Hanoi, Vietnam
  Major: Data Science in Economics and Business *September 2019 – May 2023*
  - GPA: 3.63/4.0
  - Third prize in student scientific research contest
  - Forth consecutive school scholarship for outstanding results
  - Achieved A+ in Data Structure and Algorithm, Machine Learning and Data Preparation and Visualization course

Following sections items are clickable on the PDF version.

Courses

- **Advanced Architecting on AWS** Online
  August, 2023

  Key Topics Covered: Advanced cloud architecture design, scalability, fault tolerance, security, cost optimization, AWS services (like ECR, ECS, DynamoDB, RDS, Step Functions,..), infrastructure as code (IaC) utilizing AWS CloudFormation and Terraform/Terragrunt for managing infrastructures, advanced networking concepts (like VPC, Direct Connect), and more are some of the main topics covered.

  Achievements: Successfully designing and implementing highly available and scalable AWS architectures on lab-based environment. Creating affordable solutions for infrastructure optimization, developing expertise in utilizing different AWS services to satisfy business needs. Managing Machine Learning Infrastructure using Terragrunt as IaC solution.

- **AWS Security Governance at Scale** Online
  August, 2023

  Key Topics Covered: AWS security automation and orchestration using AWS Security Hub and AWS Systems Manager; AWS Identity and Access Management (IAM), AWS Key Management Service (KMS); AWS CloudTrail for auditing and compliance; AWS Config for resource inventory and configuration management, etc.

  Achievements: AWS security automation and orchestration using AWS Security Hub and AWS Systems Manager,understanding AWS Identity and Access Management (IAM), securing AWS resources using AWS Key Management Service (KMS), implementing AWS CloudTrail for auditing and compliance, utilizing AWS Config for resource inventory and configuration management. Combined with AWS Cost Management, IaC solution and Attribute Based Access Control (ABAC) to enable business department billing.

- **Architecting on AWS** Online
  April, 2023

  Key Topics Covered: AWS architecture best practices, creating robust and scalable applications on AWS, comprehending AWS services (such as EC2, S3, and RDS), security considerations, cost-saving techniques, fault-tolerant and highly available systems, etc.

  Achievements: Successfully implemented services that run on cloud environments. Understand and be able to develop simple robust and scalable on-cloud services for data processing.

- **AWS Cloud Practitioner Essentials** Online
  April, 2023

  Key Topics Covered: The basics of AWS cloud computing; AWS pricing and billing; AWS services and their use cases; AWS security and compliance; how to utilize the AWS Management Console, etc.

  Achievements: Mastered the fundamentals of AWS cloud computing, comprehended the main AWS services and how they are used, became familiar with AWS security best practices, became adept at utilizing the AWS Management Console, etc.

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*

Personal Projects

- Used car prediction: 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.
- Car Specification: An open-source project that was created with the primary purpose of saving time for the cars-related research community. The project offers the community the dataset that includes all car models and their variations that were mass produced from 1985 to early 2022. The project also comprised a Scrapy-based crawler that have clear instruction to re-crawl the data again if needed.
- Text Classification: Directed a team of four to design and implement a classification module to classify product names into four distinct categories. The module was later deployed on Streamlit Cloud for demonstration. Models that are based on transformer architecture (sBERT and phoBERT) were used to calculate text embedding before passing to custom-made two layers Neural-Net for the classification task. The ONNX version of the models was used to accelerate the inference time.
- Personal Website: A website were made using HTML and TailwindCSS framework, with the main purpose of making a portfolio/introduction page for myself.

Personal researches and publications

- **Used Cars Prices Prediction** Published on National Scientific Conference
  Hanoi, April, 2022

  Archived Third Prize in NEU Scientific Research Competition

  ISBN: 978-604-358-602-2

  -Investigated and extracted used cars information that are selling on Vietnamese E-commerce site.

  -Constructed a Machine Learning model to help Vietnamese choosing the right used cars for their demands.

  -Conducted a dataset consists of all cars specifications (with variants) while doing Data Prepossessing.