Duc Anh Bui

bdanh96@gmail.com | +(31) 610825205
Medium | Personal Website | Github | LinkedIn

EDUCATION

UNIVERSITY OF NOTTINGHAM

2017-2018
Master in Computer Science with AI, 1st

2014-2017
Bachelor in Mathematics and Economics, 2.1

ABBEY COLLEGE

A Level
Mathematics (A), Further Mathematics (A*), Economics (B)

SKILLS

PROGRAMMING

Python • Java • C# • R
SQL • HTML • CSS • PHP
Javascript • Golang • Matlab

MACHINE LEARNING

Numpy • Pandas • Scikit-learn
SciPy • TensorFlow • Keras
Streamlit • Pytorch • Prophet

TOOLS

VSCode • GIT • PostgreSQL
Pytest • Prometheus
Flask • GCP • Kubernetes
Jupyter Notebook • FastAPI
Airflow • Helm charts
Pantsbuild • Jenkins • Docker

ACHIEVEMENTS

• International Orientation Scholarship
• Participant of JENESYS student exchange program
• Third place at district chess tournament

EXPERIENCE

HOUSING ANYWHERE | Senior Data Scientist

March 2021 – Present | Rotterdam

- Automated all targets forecasting, reducing time spent on setting targets by 80% per month.
- Built a centralised dashboard for displaying data and targets using Streamlit.
- Responsible for onboarding, training and mentoring new Data Scientists.

HOUSING ANYWHERE | Data Scientist

January 2020 – March 2021 | Rotterdam

- Development of a new version of the platform rental price prediction model. The new version allows the model to work on 4 new cities, able to perform range predictions and Mean Average Error drops by 30%.
- Port the platform room classification service written in Go to Python. Improving F1 score from 0.74 to 0.99.
- Built an image quality assessment service for the Platform leveraging Deep Learning. Service is deployed using Flask, Gunicorn, K8, GCP, and Helm chart. The new service is able to increase conversion by 5% per month.
- Integrated a CI/CD pipeline for a monorepo using Jenkins and Pantsbuild. Responsible for writing tests for existing machine learning applications.

HOUSING ANYWHERE | Data Scientist Intern

July 2019 – December 2019 | Rotterdam

- Validate assumptions brought forward by the product team and other stakeholders using product data analysis.
- Improving the platform online fraud prevention system by cleaning data and employing Natural Language Processing technique. The new version achieves a 20% increase in precision, ensuring no scams slip through.
- Assisting with building a CI pipeline for the fraud detection model using Docker, Kubernetes and Google Cloud Platform.
- Improving the platform message sanitizer via leading a text annotation project. Implementing a new version of the sanitizer in Go, achieving a 40% increase in recall and generate additional 4k in revenue per month.
- Improving the platform existing image recognition algorithm, leveraging Machine Learning to achieve a 15% increase in overall precision.

PROJECTS

MACHINE LEARNING DISSERTATION | Python

- A 20,000 words dissertation involves tackling the problem of predicting the outcome of a match in the popular multiplayer video game Dota 2. Various classification models were applied to build a recommender system.
- Obtained a higher prediction accuracy compare to previous studies by performing intensive data preprocessing and feature engineering. Achieved a grade of 77%.