

# Đào Duy Đại

Hà Nội, Việt Nam

0911825666 | [aiaoduy@gmail.com](mailto:aiaoduy@gmail.com) | [hankniel.github.io](https://hankniel.github.io) | [linkedin.com/in/daoduydai](https://linkedin.com/in/daoduydai)

## EDUCATION

### Foreign Trade University - Class of 2020

Hà Nội, Việt Nam

*Bachelor of Business in Japanese-Style International Business (JIB)*

Expected Jul. 2024

- **Current GPA:** 3.59 / 4.0;
- **Relevant Coursework:** Financial Analysis, Econometrics, Business Plan V-BIZ, Business Communication.

### Achievements

- **Score 29/30** (A department) in the National High School Graduation Exam 2020 with 10/10 in Math;
- **Valedictorian** of A department in Tuyen Quang province (2020);
- **Top 30** RMIT Business Analytics Champion Competition 2023.

## SKILLS

**Technical Skills:** Data Manipulation, Data Cleaning, and Data Wrangling in Python; Query (complex join, subqueries, unions..) and Design Database in SQL; Data Processing in PySpark; Linux Command at User Level

**Visualization Skills:** Design Dashboard, Visualize and Model Data in Power BI; Python (Matplotlib, Seaborn)

**Machine Learning:** Supervised Learning (XGBoost, Regression); Unsupervised Learning (Clustering); Recommendation Engines (Collaborative Filtering)

**Soft Skills:** Problem-solving Skills, Systematically Analyzing Skills, Organizational Skills, Logical Thinking

**Languages:** Conversational Proficiency in English, Basic Communication in Japanese (Equivalent N3)

## FEATURED PROJECTS (more details at <https://hankniel.github.io/>)

### Customer Segmentation Analysis and Churn Prediction | *Python*

- Clean, transform, visualize, and analyze Pizza Hut's fictitious customer data in a competitive market context.
- Provide insights and recommendations for each Cluster based on KMeans - RFM Segmentation.
- Perform Customer Lifetime Value and Churn Probability predictions using Machine Learning (XGBRegressor) and BG/NBD - Gamma-Gamma Model.

### Fraud Detection | *Python*

- Leverage Python to clean, transform, visualize, and analyze credit card transaction data of users.
- Build a Machine Learning model to detect fraudulent transactions using the XGBoost algorithm - XGBClassifier Model. The Model can detect 91.6% of fraudulent transactions.

### Anime Recommendation System | *PySpark, Python*

- Build two anime recommendation models with Collaborative Filtering (ALS Algorithm) and Content-based Recommendation (TF-IDF) using data from MyAnimeList Database 2020.

### Business Growth Strategy Analysis | *Power BI*

- Analyze data of a fictional company that needs to grow and expand its business using several metrics and time-series analysis; Visualize metrics with insightful charts.
- Provide insights and recommendations for the company's growth plan regarding details on the market, time of penetration, target customer profile, etc.

## WORK EXPERIENCE

### Personal Tutor | *Mathematics*

Apr. 2022 - Jun. 2023

- Tutor for four students including two individual students and one group of two students.
- Design lectures and learning programs to suit each student's learning ability and goals (national high school exam, school exam, etc.) and closely follow the learning roadmap.
- All four students achieved their set goals within a period of six months - one year and all passed into their first desired university.

## CERTIFICATIONS

- DataCamp Data Analyst Professional
- DataCamp Data Engineer Associate
- Hackerrank SQL Advanced

## ACTIVITIES AND INTERESTS

**Activities:** FTU Guitar Club - Organizing Department.

**Interests:** Interested in Math and Finance; Data and Programming.