

# Hoang Ngoc Thuy Thuong

📍 Go Vap District, Ho Chi Minh City, Vietnam ✉️ [thuythuong2501@gmail.com](mailto:thuythuong2501@gmail.com) 📞 (+84) 707 250 103

🌐 [berylhoang2501.github.io](https://berylhoang2501.github.io) in [LinkedIn Profile](#) 🌐 [github.com/berylhoang2501](https://github.com/berylhoang2501)

## Professional Summary

I am a final-year student skilled in Python, Machine Learning, and Big Data analytics, with hands-on experience in developing predictive models and analytical solutions to drive business insights. I am proficient in extracting actionable intelligence from complex datasets and eager to contribute as a Data Scientist by delivering data-driven strategies that solve real-world challenges.

## Education

**Hoa Sen University, Ho Chi Minh City, Vietnam**

2021 – Present

Dual Degrees: B.A. in International Business (E-commerce) and B.A. in English Language Studies (Business English)

- GPA: 3.69/4.0

## Professional Experience

**Data Analyst Intern**

Bien Hoa, Vietnam

OPDA Co., Ltd.

Jul 2024 – Oct 2024

- Built a customer churn prediction model using LightGBM, achieving an 85% F1-score, enabling retention strategies that supported business growth for the Smart Business Platform.
- Solved the Traveling Salesman Problem using Simulated Annealing, optimizing routes across 64 Vietnamese cities, reducing travel distance to 5016.4 km.
- Documented data visualization libraries (Matplotlib, Seaborn) to streamline training for future interns.
- Tools: Python, Pandas, Scikit-learn, PyCaret, Plotly, Google Colab

## Technical Skills

- **Python and Libraries:** NumPy, Pandas, Scikit-learn, PyCaret, PySpark
- **Machine Learning:** Regression, Classification, Clustering, PCA, ALS
- **Data Analysis:** EDA, Feature Engineering, Data Cleaning, NLP (underthesea, Gensim)
- **Data Visualization:** Matplotlib, Seaborn, Plotly, WordCloud, PowerBI
- **Databases and Big Data:** MySQL, SQL Server, PostgreSQL, FugueSQL, PySpark (RDDs, DataFrames, MLlib)
- **Data Collection:** Web Scraping (BeautifulSoup, Selenium)
- **Version Control & Model Deployment:** Git, GitHub, Streamlit

## Certifications

**Data Science and Machine Learning Certificate** [🔗](#)

Dec 2023 – Dec 2024

VNUHCM - University of Science, Ho Chi Minh City (Scores: 9–10/10)

- Capstone Projects (Nov 2024): Sentiment Analysis, Hasaki.vn (93.35% accuracy, Random Forest), providing insights for product improvements using PySpark; Recommender System, Hasaki.vn (achieved 78% top-5 accuracy using Collaborative Filtering with KNNBaseline and ALS, correctly recommending at least one preferred product), enhancing user experience through personalized product recommendations. Developed on Google Colab with Surprise and PySpark.
- Tools: Python, Pandas, Scikit-learn, PySpark, underthesea, Gensim, Matplotlib, Seaborn

## Leadership and Extracurricular Activities

**Head of Finance, E-commerce Club**

Jan 2024 – Present

Hoa Sen University

- Managed budgeting using Excel and Python to optimize resources, demonstrating strong problem-solving skills.
- Analyzed event metrics, boosting attendance by 20% with data-driven strategies, showcasing teamwork and adaptability.

## Languages

- **Vietnamese:** Native
- **English:** Full Professional Proficiency ([TOEIC 905/990](#) [🔗](#), May 2025)
- **Korean:** Elementary Proficiency