

Thai Bao An

 baoan102 |  thaibaoan |  baoan102.github.io |  baoan.mis@gmail.com |  +84.36.302.6387

SUMMARY

I am a third-year Data Science student at Ho Chi Minh University of Banking. My passion is exploring and analyzing data to derive meaningful insights. My coursework and projects have equipped me with practical experience in using tools like Python, SQL Server and Power BI. I am eager to apply my analytical skills and knowledge in a Analyst Intern role to develop my expertise in data analysis.

EDUCATION

2021 - present Data Science at **Ho Chi Minh University of Banking** (GPA: 3.2/4.0)
Core modules: Data Analysis with SQL, Data Visualization with Google Data Studios,
Python Programming, Classical Machine Learning, Object-oriented Programming.

PROJECTS

Customer Onboarding Project

[Link to Project](#)

- Design banking data warehouse to optimize SQL query performance.
- Build ETL pipeline using SSIS to combine data from multiple sources and ingest it into the data warehouse.
- Design data quality metrics: accuracy, consistency, and timeliness.
- Build the report including metrics such as the rate of each step, cost per acquisition (CPA), and return on investment (ROI) for each source.
- Monitor onboarding efficiency, identify areas for improvement, and ensure a smooth and seamless customer onboarding experience.
- Using libraries Numpy, Pandas, Scikit-learn, Matplotlib, Seaborn,... to build and analyze models.
- Comparison and choosing the best model by confusion matrix, F1 score and area under ROC curve.
- Use the ensemble method to predict the class label and optimize the threshold. The model has Accuracy = 73% and F1 = 73%.

Future Retailer Project

[Link to Project](#)

- Utilize expertise in data modeling and transform data by using SQL.
- Analyze the reasons for customer order cancellations and provide suitable solutions.

ACTIVITY

Research - User Identification

[Link to Project](#)

- Collect user data from social media platform.
- Text preprocessing using the NLTK library.
- Use word embedding algorithms and keyword extraction for text processing.
- Use metrics such as cosine similarity and edit distance to identify whether accounts on different platforms belong to the same person.

SKILLS

Soft skills: Problem Solving, Analytical Thinking and Critical Thinking
Tools: Excel, Power BI, Looker Studio, Python and SQL.