TRAN THI BICH NGOC

(+84) 987 654 321 | tran.thi.bich.ngoc@gmail.com | linkedin.com/in/tran-thi-bich-ngoc | kaggle.com/bichngoc | github.com/bichngoc

PROFESSIONAL SUMMARY

A dedicated Data Analyst with over 4 years of experience in extracting, analyzing, and visualizing data to deliver actionable business insights that drive strategic decision-making. Proficient in Python, R, SQL, and BI tools such as Tableau and Power BI. Passionate about leveraging machine learning techniques and data storytelling to optimize business outcomes. Seeking to join a forward-thinking organization where I can apply my analytical expertise and data-driven approach to create meaningful impact.

PROFESSIONAL EXPERIENCE

Data Analyst | Shopee Vietnam Co., Ltd. | July 2021 – Present

- Analyzed large-scale customer behavior datasets (over 10 million records) to identify purchasing trends, enabling optimization of marketing, operations, and product strategies.
- Designed and deployed over **15 interactive dashboards** using **Tableau** and **Power BI**, providing real-time insights into business performance and reducing decision-making time by **30%**.
- Wrote complex **SQL** queries on **PostgreSQL**, **Google BigQuery**, and **AWS Redshift** to extract, aggregate, and normalize data from diverse sources, ensuring accuracy and consistency.
- Developed revenue forecasting models using **Python (Pandas, Scikit-learn)** and **R**, applying regression and Random Forest algorithms, improving forecast accuracy by **15%**.
- Collaborated with data engineering teams to optimize **ETL pipelines**, enhancing data quality and reducing processing time by **25%** through automated workflows.
- Conducted **A/B testing** to evaluate advertising campaign performance, increasing conversion rates by **12%**.
- Presented data-driven insights to senior leadership using storytelling techniques, ensuring clear and impactful communication of findings.

Junior Data Analyst | Tiki Corporation | June 2019 – June 2021

- Analyzed sales and logistics data to identify bottlenecks in the supply chain, resulting in a **10% reduction** in operational costs.
- Built recurring analytical reports using **Excel** and **Google Data Studio**, supporting product management teams in optimizing product portfolios.
- Utilized **Python (Pandas, Matplotlib)** for descriptive and inferential statistical analysis, providing insights into customer segmentation.
- Assisted in developing simple data pipelines using **SQL** and **Airflow**, streamlining internal data processing workflows.
- Performed cohort analysis to assess customer retention rates, contributing to user growth strategies.

Data Analyst Intern | Techcombank | June 2018 – December 2018

- Supported data cleaning and preprocessing for customer transaction datasets (over 500,000 records) for credit risk analysis projects.
- Created basic performance reports for financial products using **Excel (Pivot Tables, VLOOKUP)** and **Google Sheets**.
- Applied basic statistical methods (linear regression, variance analysis) to support risk analysis projects.
- Participated in internal training on **SQL** and **Tableau**, enhancing skills in querying and data visualization.

SKILLS

- **Programming Languages**: Python (Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, Plotly), R (ggplot2, dplyr, tidyr)
- **Databases & Querying**: SQL (MySQL, PostgreSQL, MS SQL Server, Google BigQuery, AWS Redshift), NoSQL (MongoDB, Cassandra)
- **BI & Visualization Tools**: Tableau, Power BI, Google Data Studio, Excel (Advanced, VBA, Power Query), Looker
- **Statistics & Machine Learning**: Descriptive Statistics, Inferential Statistics, Regression (Linear, Logistic), Classification (SVM, Decision Trees, Random Forest), Clustering (K-Means, DBSCAN), Time Series Analysis, Natural Language Processing (NLP)
- **Tools & Platforms**: Jupyter Notebook, RStudio, Git, Docker, Airflow, AWS (S3, Redshift, Athena), Google Cloud Platform (BigQuery, Dataflow), Azure (Data Factory)
- **Other**: A/B Testing, Cohort Analysis, Data Storytelling, ETL Processes, Data Governance, Data Quality Assurance

EDUCATION

Master of Science in Data Science | University of Science, Ho Chi Minh City | September 2021 – June 2023

- **GPA**: 8.8/10
- **Key Courses**: Advanced Machine Learning, Data Mining, Natural Language Processing, Time Series Analysis, Big Data Analytics, Artificial Intelligence.
- **Master's Thesis**: "Predicting Online Shopping Trends Using Machine Learning" – Developed a predictive model for customer behavior using Random Forest and LSTM, integrated with real-time API data, achieving a score of 9.2/10.

Bachelor of Science in Statistics | National Economics University | September 2015 – June 2019

- **GPA**: 8.3/10
- **Key Courses**: Probability and Statistics, Linear Algebra, Calculus, Econometrics, Multivariate Data Analysis, Statistical Programming with R.
- **Capstone Project**: "Analyzing ROI of Digital Advertising Campaigns" – Utilized multivariate regression and cohort analysis to evaluate campaign effectiveness, scoring 8.8/10.

CERTIFICATIONS

- Google Data Analytics Professional Certificate | Coursera | May 2022
- Microsoft Certified: Power BI Data Analyst Associate | Microsoft | November 2023
- AWS Certified Data Analytics – Specialty | Amazon Web Services | March 2024
- Professional Certificate in Machine Learning | IBM | August 2022

PERSONAL PROJECTS

- **Retail Customer Data Analysis** (Kaggle: [link-to-kaggle])
  - Analyzed a retail dataset with over 1 million records using **Python (Pandas, Seaborn)** and **Tableau** to identify customer segments and propose revenue growth strategies.
  - Applied **K-Means** and **DBSCAN** clustering models, achieving **85% accuracy** in customer segmentation.
  - Published an interactive report on Kaggle, receiving over 500 views and 50 upvotes.
- **House Price Prediction in Ho Chi Minh City** (GitHub: [link-to-repo])
  - Built linear regression, Gradient Boosting, and Random Forest models using **Python (Scikit-learn)** to predict house prices based on real estate data.
  - Preprocessed over 10,000 records using **Pandas** for handling missing values and feature normalization.
  - Achieved an R^2 score of 0.82 and reduced RMSE by 10% compared to baseline models.
- **Sentiment Analysis of Customer Feedback on Social Media** (GitHub: [link-to-repo])
  - Collected data from Twitter and forums using **Python (Tweepy, BeautifulSoup)**, applying **NLP** with **NLTK** and **BERT** for sentiment analysis.
  - Developed a **Power BI** dashboard to visualize real-time sentiment trends.
  - Proposed customer experience improvements based on analysis, implemented in an internal pilot project.
- **Customer Churn Prediction for E-commerce** (Kaggle: [link-to-kaggle])
  - Developed a classification model using **Python (Scikit-learn, XGBoost)** to predict customer churn based on transactional and demographic data.
  - Achieved **88% accuracy** and improved model performance using hyperparameter tuning.
  - Visualized results with **Seaborn** and presented findings in a comprehensive report.

EXTRACURRICULAR ACTIVITIES

- **Speaker at Vietnam Data Summit 2023** | October 2023
  - Delivered a talk on "Data Storytelling: From Numbers to Business Decisions" to over 200 data professionals.
- **Member of Vietnam Data Science Community** | January 2020 – Present
  - Participated in hackathons and workshops, contributing to open-source data analysis projects.
- **Mentor at Online Data Analytics Bootcamp** | June 2022 – December 2022
  - Guided 15 learners in mastering **Python** and **Tableau** for data analysis and visualization.
- **Volunteer Data Analyst at Local NGO** | March 2020 – August 2020
  - Analyzed donation and outreach data to optimize fundraising campaigns, increasing donations by **20%**.

LANGUAGES

- Vietnamese: Native

- English: Fluent (IELTS 7.5, 2022)

INTERESTS

- Data visualization competitions, open-source data projects, marathon running, and blogging about data science trends.