Mai Ngoc Nong

1210 Vienna, Austria | nongngocmai22@gmail.com | +43 678 791 8857

linkedin.com/in/ngoc-mai-nong-351452199 | github.com/MiaHenstridge | miaworld.neocities.org

Personal Statement

Data Analyst and Data Scientist pursuing a Master's Degree in Social Data Science, with specialized experience in credit risk analytics for consumer finance and market research in the automotive industry. A proactive, efficiency-focused team player with strong technical expertise in data analytics, machine learning, and data processing. Skilled in Python, SQL, big data processing, and passionate about delivering innovative, data-driven solutions to tackle business problems. Actively seeking part-time positions or internships to further develop technical skills and industry knowledge, with a long-term goal of advancing to a Senior Data Scientist or Market Researcher role within the next five years.

Education

Central European University, MS in Social Data Science   Sept 2024 – Jun 2026

- **Specialization:** Economics
- **Coursework:** Advanced Topics in Applied Statistics, Scientific Python, Cloud Computing, Data Mining and Big Data Analytics, Data Management and Databases
- **Scholarships and Awards:** CEU Master's Scholarship, CultTech Scholar
- **Research Paper**: "Addressing Ethical Concerns of Algorithmic Credit Scoring: The Vietnam's Perspective", "Predatory Game Monetization: Going Beyond Loot Boxes and Gambling"

Tokyo International University, BS in Digital Business and Innovations   Sept 2019 – Mar 2023

- **GPA:** 3.9/4.0
- **Coursework:** Statistics, Mathematics for Business and Economics, Research Methods & Data Analysis in Business, Marketing Modeling, Python, R
- **Scholarships and Awards:** TIU 100% Tuition Reduction Scholarship, Monbukagakusho Honors Scholarship for Privately-Financed International Students
- **Research Paper**: "Data Analytics for Esports Development"

Experience

Associate ESG Analyst, onewoom GmbH – Vienna, Austria   Jan 2025 - Present

- Performed data collection, aggregation for greenhouse gas emission calculation and annual sustainability report
- Supported initiatives to foster inclusive and engaging working environment through in-depth analyses of employee collaboration and gender pay gap within the company
- Evaluated suppliers' alignment with company's sustainability strategy and identified areas for sustainability collaborative initiatives through quantitative and qualitative analysis of supplier sustainability

Analytics Engineer, MB Shinsei Finance Ltd. (Mcredit) – Hanoi, Vietnam   Apr 2023 – Aug 2024

- Managed monthly credit scorecard monitoring dashboards using Excel, Python, Spark SQL, and Apache Airflow, providing regular insights into the performance of scorecards and risk portfolio
- Developed and operated a feature store for scorecard development with Spark SQL, centralizing and automating the feature engineering process of over 1500 metrics on customers' payment behavior and debt collection, reducing the time spent on scorecard development and implementation by 60%
- Initiated and implemented daily data quality checkpoints, enabling timely detection and reaction to data quality issues, thereby ensuring reliable data for usage
- Developed and automated daily and monthly ETL jobs with Apache Airflow to populate Credit Risk Modeling Department's data mart for scorecard development and monitoring

Strategy and Business Development Intern, Mitsubishi FUSO Truck and Bus   Sept 2022 – Feb 2023
Corporation | Daimler Truck Asia – Kanagawa, Japan

- Supported the Strategy & Market Intelligence department by collecting, formatting, analyzing, and visualizing data, which enabled effective market trend research, competitor analysis, and the preparation of high-quality market intelligence reports and proposals
- Implemented and streamlined processes for data storage, querying, and weekly report with Tableau for the Daimler Truck Asia internal ideation project, improving efficiency and accessibility of project data
- Automated mass email communication with tailored content using VBA scripts

Projects

House Prices Prediction Kaggle Competition   Feb 2025
Personal Project

- Developed a house price prediction model using gradient boosting algorithms, applying advanced feature engineering and preprocessing techniques. Optimized performance through hyperparameter tuning, achieving a top 200 ranking on the leaderboard.
- Technologies Used: Scikit-learn, Optuna, OptBinning, LightGBM, XGBoost, CatBoost, seaborn, matplotlib

Deloitte & MachineHack Bank Loan Defaulter Prediction Hackathon   Oct 2024
Personal Project

- Conducted feature engineering and preprocessing to develop a robust classification model for predicting bank loan defaults. Employed hyperparameter tuning to optimize ensemble models, including Random Forest, LightGBM, XGBoost, and CatBoost, achieving a minimized log-loss of approximately 0.3.
- Technologies Used: Scikit-learn, Optuna, Feature-engine, OptBinning, LightGBM, XGBoost, CatBoost, seaborn, matplotlib

2024 US Presidential Debate Sentiment Analysis   Oct 2024
Personal Project

- Conducted sentiment analysis on the transcript of the 2024 US Presidential Debate on September 10 to reveal the underlying sentiment conveyed by Donald Trump and Kamala Harris on key topics, including: economy, abortion rights, foreign policies, and immigration.
- Technologies Used: nltk, TextBlob, NRCLex, SciPy, matplotlib

Skills

Core Competencies: Data Analytics, Big Data Processing, Machine Learning, Deep Learning, Data Governance, Time Series Analysis, Natural Language Processing, Sentiment Analysis, Network Analysis, Problem Solving

Programming Languages: Python, R, SQL, JavaScript

Frameworks & Libraries: Pandas, Numpy, Scikit-learn, TensorFlow 2.0, LightGBM, XGBoost, CatBoost, Optuna, PySpark, nltk, matplotlib, seaborn, graph-tool, Apache Airflow, Apache Iceberg, boto3

Software: MS Excel, Tableau

Databases & Cloud: MS SQL Server, PostgreSQL, Oracle, Neo4j, AWS (Athena, EC2, S3)

Certifications

- SQL Intermediate (HackerRank)
- Google's Professional Data Analytics Certificate (Coursera)
- Mathematics for Machine Learning Specialization (Coursera)
- Deep Learning Specialization (Coursera)
- Spark and Python for Big Data with PySpark (Udemy)
- TensorFlow 2.0: Artificial Intelligence and Deep Learning (Udemy)