

Kim Khai Nguyen

Mobile: (+65) 9175 3973/ Email: dylannguyenkk@gmail.com / [LinkedIn](#) / [My Personal Portfolio](#)

Availability: Full-time opportunities starting May 2025

EDUCATION

Nanyang Technological University (NTU)

Aug 2021 – Jun 2025

- **Bachelor of Engineering (Electrical and Electronic Engineering)**
- **Honours (Distinction)** (Expected)
- Specialisation: **Data Analysis and Machine Learning** (Expected)
- Relevant Modules: (1) *From Computational Thinking to Programming* (2) *Data Structure and Algorithms* (3) *Introduction to Data science and Artificial Intelligence* (4) *Statistics & Analysis* (5) *Machine Learning Design & Applications*

VIRTUAL TRAINING AND LEARNING & DEVELOPMENT

- Coursera Relevant Modules: (1) *IBM Data Science* (2) *Mathematics for Machine Learning* (3) *Machine Learning Specialization* (4) *Deep Learning Specialization*.
- Udemy: (1) *Learn Python Programming Masterclass* (2) *End-to-End Machine Learning: From Idea to Implementation* (3) *MLOps Bootcamp: Mastering AI Operations for Success – AIOps*.
- EdX: *Probability: The Science of Uncertainty and Data*.

WORK EXPERIENCE

Corning Singapore Holdings Private Limited, Data Analyst Intern

June 2023 – Dec 2023

- **Utilized** Jupiter Notebook and Natural Language Processing to analyze global spending data, uncovering opportunities for bulk purchase discounts and achieving enhanced cost savings.
- **Optimized** and standardized **29** dashboards development, implementing strategic optimizations to boost dashboard refreshing efficiency by **60%** in Power BI.
- **Developed** a highly efficient Python script for **web scraping** across six websites, achieving a **90%** reduction in processing time.
- **Constructed** an automated solution using advanced VBA code, reducing supplier performance report generation time to just **3 seconds**.
- **Implemented** a Power Automate flow that streamlined stakeholder communication, boosting collaboration efficiency and reducing dashboard issue resolution time by **40%**.

Meiyume (Singapore) Pte. Limited, Data Management & Analytics Intern

May 2024 – Aug 2024

- **Automated** reporting processes and integrated forecasting features into the Global HR Dashboard, improving efficiency by **100%** and streamlining **AWS** data pipeline.
- **Designed** a Power BI dashboard to monitor DEI, training & development and compensation metrics, improving transparency and tracking of sustainability initiatives.
- **Conducted** topic modeling and **sentiment analysis** based on product listings and consumer reviews related to luxury fragrance packaging across three markets (US, UK, DE) using Python and OpenAI. This study was commissioned by COTY and completed within 2-week.
- **Identified** key packaging design elements for **Jo Malone's** home fragrance products tailored to the China/APAC market, based on consumer reviews and sentiment analysis using Python and Claude 3.5 Sonnet API.
- **Analyzed** leading UK brands in natural skincare for **The Body Shop** using fine-tuned Llama 3, **identifying** key ingredients and consumer sentiment that informed product development.

Industrial Scientific Corporation, Business Analyst Intern

Sep 2024 – Jan 2025

- **Analyze sales data** to develop actionable metrics and dashboards that drive strategic decision-making and improve **sales performance**.
- Identify and **address process inefficiencies**, implementing solutions that streamline operations and enhance overall efficiency.
- Prepare, review, and ensure the accuracy of **tendering commercial documentation**, aligning with company standards and client requirements.
- Assist in planning, executing, and optimizing **digital marketing campaigns** across multiple channels, utilizing **market research** to enhance targeting and engagement strategies.

ACADEMIC PROJECTS / PERSONAL PROJECTS / RESEARCH EXPERIENCE

Advanced Pneumonia Detection System

Feb 2024 – Mar 2024

An MLOps-powered solution for automating pneumonia detection

- **Led** the development of an end-to-end pneumonia detection system using a custom **pretrained VGG16 model**, achieving **95%** diagnostic accuracy.
- **Established MLflow tracking** with DagsHub and lightweight DVC, ensuring reproducibility, scalability, and accountability in model development.
- **Developed** a user-friendly Flask application to facilitate model training and image classification, enhancing accessibility and user experience.
- **Streamlined deployment** with **AWS CICD** via GitHub Actions, automating Docker image creation and EC2 provisioning for seamless model deployment.

LLM-Powered Knowledge Retrieval System

Jan 2024 – Jan 2024

A Chatbot to query information from your documents

- **Led** the development of a chat system using **LangChain RetrievalQA** and **OpenAI API**, streamlining access to critical information from diverse PDF documents.
- **Deployed** the application on **Streamlit**, delivering an intuitive interface for effortless document navigation.

Deep Learning based Automatic Modulation Recognition

Jan 2024 – April 2024

An advanced Transformer-based model for robust signal modulation classification

- **Generated** synthetic signal datasets using **MATLAB**, simulating various modulation formats and real-world channel impairments.
- Developed a **data preprocessing pipeline** to consolidate signal data from multiple CSV files into a single pickle file, ensuring seamless integration for model input.
- Implemented a **Transformer-based model** with convolutional, LSTM, and self-attention layers using **TensorFlow**, achieving **92%** accuracy in **modulation signal classification**.
- **Documented** all project aspects, including data generation, model architecture, training procedures, and evaluation metrics, ensuring reproducibility and transparency.

Face Mask Detection

Dec 2023 – Jan 2024

A web app for real-time face mask detection

- Utilized images, annotation files processing and **Faster-RCNN model** (Region-based Convolutional Neural Network) to clean and train Kaggle Dataset for detecting if people are wearing mask or not based on pictures.
- Achieved a precision rate of **94%**, signifying the model's proficiency in correctly identifying individuals wearing masks in a picture.
- Integrated model seamlessly with **Streamlit**, enabling users to upload images or engage in **real-time face mask detection**.

Prediction of Cardiovascular Disease (Team Leader of 4 members)

Aug 2022 – Dec 2022

A machine learning model to predict cardiovascular disease

- Applied **Python** with **4 Machine Learning Models** (i.e., Random Forest, Decision Tree, Logistic Regression, KNN) to clean and train Kaggle Dataset for predicting if patient suffered **cardiovascular disease** based on critical parameters.
- Implemented **hyperparameter tuning** using **GridSearchCV** to determine optimal parameters values for Machine Learning Models.
- Concluded that **Decision Tree** was the most ideal model with **73% prediction accuracy** and was significantly correlated with critical factors (e.g., Blood Pressure, Cholesterol).

LEADERSHIP / CO-CURRICULAR ACTIVITIES / COMMUNITY INVOLVEMENT / VOLUNTARISM

NTU Vietnamese Freshmen Orientation, Camp Group Leader

Jul 2022 – Aug 2022

- Led 40 team members with freshmen in 3 Day-Camp.
- Planned and hosted Orientation Group's Events (e.g., Bonding Session, Outing Day, Camp's Activities).
- Generated bonding activities, social games, and cheering celebrations.
- Advised and guided freshmen to adapt within NTU campus life and academic aspects readily.

SKILLS / HOBBIES

- **Programming Languages:** Python, SQL, Visual Basic Code (VBA), MATLAB.
- **Data Science & Machine Learning:** TensorFlow, PyTorch, Keras, FastAI, Scikit-Learn, OpenCV, Hugging Face, LangChain, Pandas, NumPy, BeautifulSoup, Selenium.
- **Big Data & Cloud Computing:** AWS (S3, Redshift), Azure, Google Cloud Platform, IBM Cloud, Spark, Hadoop.
- **DevOps & MLOps:** Git, GitHub, MLflow, Docker, DagsHub, Dask, Streamlit, Flask, Data Versioning Control (DVC), Prometheus, Jenkins, Heroku.
- **Database:** MySQL, PostgreSQL, SQLite, Oracle, Firebase, MongoDB.
- **Tools:** Power BI, Tableau, Power Automate, Excel, Power Point, Octoparse.
- **Soft Skills:** Team player, Self-motivated, Detail-oriented, Problem-solving, Initiative.
- **Hobbies:** Travelling, Play Tennis, Piano & Launchpad.