

# Anthony Nguyen

Vancouver, BC • (236) 234-7843 • [hoangan.np@gmail.com](mailto:hoangan.np@gmail.com) • [LinkedIn](#) • [E-portfolio](#)

## PROFILE

- Data analyst and research assistant in Chemistry Lab of Langara College.
- R&D engineer in semiconductors from Taiwan with 5+ years of experience in data analysis and chemical engineering.
- Track records in statistics by R, RDBMS management, data pipeline, ETL, programming in Python, machine learning, dashboard visualization, AWS cloud service (Bedrock, E3), DevOps

## EDUCATION

**Data Analytics, Post-degree Diploma • Langara College • Vancouver, BC** Sep 2023 – Dec 2024

- Dean's Honor Roll recipient for outstanding academic achievement x 2
- Scholarship: Langara Student Union (LSU) Volunteer Leadership Award
- Statistics, data mining, EDA, R, Python, machine learning, deep learning (scikit-learn, Tensorflow). Data visualization (Tableau, PowerBI). API on LLM, Generative AI (BERT, GPT, langchain). Relational database management (postgresql), ETL (Apache Airflow)

**Master of Science, Biotechnology • Cheng Kung University • Tainan, Taiwan** Feb 2018 – Feb 2020

- Distinguished International Student Scholarship recipient
- Structural biology and molecular assays for drug development. Cloning, qPCR, gel electrophoresis, cryo-EM, ITC, microarray, X-ray crystallography

## PROJECTS

**ASD positive classification by 3D brain MRI using ML and deep learning** (Langara College)

**Github:** [https://github.com/realanguyen96/3D\\_brain\\_MRI](https://github.com/realanguyen96/3D_brain_MRI)

**Skills:** Python, machine learning, deep learning, supervised classification

- Replicated study from Nature Communication by [Vogelstein et al. \(2021\)](#).
- Multiple dimension reduction methods comparison. Evaluated 3D CNN performance vs ML classifiers (RandomForest, LDA).
- Building predictive ML model for ASD positive cases that had accuracy up to 60%.

**Automation program for chemical data processing** (Applied Research Center - ARC) – NDA included

**Skills:** Python, MS Excel, HPLC-UV, GC-MS, transformation, DevOps, statistics, SDLC, Github

- Writing, debugging python-based applications to process, extracting target information automatically from analytical instruments' results in workbook format (MS Excel)
- Using feature selection algorithms to identify significant parameters for different methods comparison

**Image annotation by CNN and NLP modeling to predict tweets sentiment by Gensim** (Langara College)**Github:** [https://github.com/realanguyen96/image\\_annotation](https://github.com/realanguyen96/image_annotation)**Skills:** convolutional neural network (CNN), natural language processing (NLP), word cloud, embedding

- Tensorflow CNN model was trained by CIFAR-100 dataset for input image recognition
- Text-to-embedding by Gensim library, classifying tweets at various sentiment levels from 1 to 5 with 50-74% accuracy

**Machine learning model to predict fatality of intensive care patients** (Langara College)**Github:** [https://github.com/realanguyen96/elCU\\_database](https://github.com/realanguyen96/elCU_database)**Skills:** cloud-based RDBMS, snowflake, SQL, feature selection, dimension reduction, algorithm

- Replicated study from J Per Meds by [Na Pattalung T et al. \(2021\)](#)
- Moved database to cloud, queried target data in SQL, EDA, transformed and identified significant explanatory variables. Final model's accuracy was up to 60%

## PROFESSIONAL EXPERIENCE

**Research Assistant • [Applied Research Center](#) • Vancouver, BC**

Oct 2023 - Present

- Successfully optimized and statistically validated new SOP for Mitacs and NSERC funded project in developing new methodology for quantification of target compounds in cannabis and hops using GC-MS, HPLC-UV, Q-TOF
- WHMIS and Lab Safety certificate for safety protocols, chemical handling, record maintenance.

**Research and Development Engineer • [Xintec Technology](#) • Taoyuan, Taiwan**

Nov 2022 – Jul 2023

- Module R&D in-charged of developing new recipes for electroplating (thin-film) and printing (bonding) processing on 8" and 12" Si wafer, mainly for CMOS optical sensor devices.
- Monitored pilot runs and supervised quality control to achieve refinery strategies using warehouse and production control systems (MES, ERP, SPC)

**Research and Development Engineer • [Everwide Company](#) • Douliu, Taiwan**

Mar 2020 - Oct 2022

- Extracted relevant data for experimental design and identified data trends to develop new epoxy resin products customized specifically for client's requirements.
- Researched literature and pattern review (OneNote)

**Data Analyst Intern • [Olam Vietnam Ltd.](#) • HCMC, Vietnam**

Jul 2015 - Feb 2016

- Descriptive analysis, EDA and hypothesis testing using statistical approaches (ANOVA, two-way t-test, etc.) in MS Excel and JMP to develop export strategy for Olam Ltd.

## MICRO-CREDENTIALS

- Complete Guide to Power BI for Data Analysts by Microsoft Press (LinkedIn Learning) Sep 2024
- PostgreSQL Essential Training (LinkedIn Learning) Sep 2024
- Python Data Analysis (NASBA) Jul 2024
- MS 365 Essential Training (Word, Excel, PowerPoint) Sep 2023