Anthony Nguyen

Vancouver, BC • (236) 234-7843 • hoangan.np@gmail.com • LinkedIn • E-portfolio

PROFILE

- Current data analytics student and research assistant in Chemistry Lab of Langara College.
- Former R&D engineer in semiconductors from Taiwan with 5+ years of experience in data analysis. Module engineer in charge of thin-film and bonding process.
- Track records in statistics proficiency, RDBMS management, programming in R and Python, machine learning, predive model development, dashboard visualization.

EDUCATION

Data Analytics, Post-degree Diploma • Langara College • BC   Oct 2023 – Dec 2024

- Dean's Honor Roll recipient for outstanding students. GPA 3.69
- Statistics, data mining, EDA, data manipulation. Programming languages (R, Python). Machine learning, deep learning (scikit-learn, Tensorflow). Data visualization (Tableau, PowerBI). API on LLMs, Generative AI (BERT, GPT, langchain). Database management (postgreSQL)
- Supervised, unsupervised algorithms and regression. Hyperparameter tuning on ML models

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

- Distinguished International Student Scholarship recipient. GPA 4.05
- 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
Skills: Python, machine learning, deep learning, supervised classification

- Replicated study from Nature Communication by Vogelstein et al. (2021).
- Multiple dimension reduction methods comparison (LOL, RP, PCA, LDA). Evaluated 3D CNN performance vs machine learning 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)
Skills: Python, MS Excel, HPLC-UV, GC-MS, transformation, software integration, statistics

- 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

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

ML classification pipeline for fatality prediction on ICU patients (Langara College)

Github: https://github.com/realanguyen96/eICU_database

Skills: SQL, snowflake, Matplotlib, Seaborn, feature selection, dimension reduction, cross-validation

- Replicated study from J Per Med by Na Pattalung et al. (2021) using eICU database
- Database uploaded to cloud-based tool (snowflake) and extracted by SQL query
- EDA, data manipulation, imputation by ML. Final classifier model has accuracy up to 60%

PROFESSIONAL EXPERIENCE

Research Assistant • Applied Research Center • Langara College • Canada   Oct 2023 - Present
- Successfully optimized and statistically validated new methodology for purification and quantification of organic compounds in hops with 30% higher accuracy, 2x shorter run time
- Developed automated data processing pipeline for output data from GC-MS, HPLC-UV, Q-TOF.
- WHMIS and Lab Safety certificate for safety protocols, chemical handling, record maintenance.

Research & Development Engineer • Xintec Company • 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 & Development Engineer • Everwide Company • 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

Data Analyst Intern • Olam Vietnam Ltd. • 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