QUOC-HUY NGUYEN
(714) 837-0205 | quochuyn1999@gmail.com | linkedin.com/in/quochuyn | quochuyn.github.io

EDUCATION

University of Michigan, Ann Arbor May 2022 - Aug 2023
Master of Applied Data Science 3.9 GPA

University of California, Irvine Sep 2017 - Jun 2021
Bachelor of Science in Mathematics - Data Science Specialization, Statistics Minor Summa Cum Laude, 3.9 GPA

EXPERIENCE

Upwork | California May 2022 - Present
Data Scientist (Freelancer)

- Evaluated and fine-tuned an overfitting ResNet-based TensorFlow convolutional neural network (CNN) aligning it with the specific requirements of a client's medical research within the healthcare domain.
- Developed Python scripts using GeoPandas for analytical insights and geographic visualizations, offering a comprehensive understanding of patterns and trends in malicious hacking attempts.
- Implemented SQL queries to extract valuable insights from customer billing transactions that were presented to a stakeholder.
- Delivered expert consulting on best Python programming practices, particularly on object-oriented methodologies.

Curacao | Los Angeles, California Sep 2021 - Apr 2022
Data Analyst

- Engineered a data pipeline for our mobile application product, enabling the extraction, transformation, and loading (ETL) of 16+ GB of Open Banking data from FinTech APIs to underwrite personal loans.
- Designed and developed a scikit-learn random forest classification model that surpassed a 70% prediction accuracy, discerning key thresholds of diverse credit risk levels.
- Leveraged customer data analytics to guide data-driven decision making, identifying high-volume user segments and optimizing physical space utilization.
- Optimized routine reports in SAS and Excel that resulted in saving about 10% of time spent.
- Acquired proficiency in SQL and SAS within 2 months, leveraging these tools to execute data analytic tasks.

ACADEMIC DATA SCIENCE PROJECTS

Youtube NLP Recommender June 2023 - Present

- Developed an end-to-end solution that integrated with the Youtube API and a website for real-time content viewing and filtering with Natural Language Processing (NLP) techniques.
- Configured an AWS RDS PostgreSQL database for user preferences that integrated with the Streamlit web app.

NHL Draft Predictions May 2023 - June 2023

- Created a machine learning model to predict NHL draft positions for prospective hockey players and cluster players together into groups based on the nearest neighbor algorithm.
- Harnessed the cutting-edge power of Language Models (LLMs) including Google's BERT and OpenAI's Chat-GPT to encode word embeddings as model features.

RESEARCH

Atwood Lab | Irvine, California Aug 2019 - Mar 2021
Undergraduate Researcher

- Implemented a Pytorch graph-based convolutional neural network (CNN) for semi-supervised classification model from a highly cited research article for identifying cancerous cells.
- Researched skin cancer single-cell data using methods for principal graphs and graph neural networks.

MathBioU & Math EXPLR | Irvine, California Jul 2019 - Aug 2019
Undergraduate Researcher

- Performed computational biology research at a 6-week summer research program.
- Collaborated as a co-author in a research report to present findings to a diverse >15 person audience.

SKILLS

Programming: Python (4 years), SQL (2 years), PostgreSQL, R, SAS, C++, Matlab, Bash, Shell, Unix/Linux Terminal
Libraries: Numpy, Pandas, Scikit-learn, PyTorch, TensorFlow, Hugging Face, PySpark, Statsmodels, NLTK, Matplotlib, Seaborn, Altair, Geopandas, Psycopg2, Streamlit
Datasets: Customer, Financial, B2B, Healthcare/Clinical, Geospatial, Causal Inference, JSON, Experiments, Web Scraping
Tools: Version Control (Git, Github, DVC), Jupyter Notebook, Business Intelligence (Power BI, Domo), Microsoft Excel, Cloud Infrastructure (AWS S3/EC2/RDS), DBeaver, RStudio, LaTeX