

QUOC-HUY NGUYEN

(714)837-0205 • quochuyn1999@gmail.com • linkedin.com/in/quochuyn • github.com/quochuyn

EDUCATION

Master of Applied Data Science

May 2022 - Aug 2023

University of Michigan

B.S., Mathematics - Data Science Specialization; Statistics Minor

Sep 2017 - Jun 2021

University of California, Irvine

Summa Cum Laude, 3.94 GPA

Activities: Phi Beta Kappa

TECHNICAL SKILLS

Programming: Python, R, SAS, SQL, C++, Matlab, LaTeX

Data Analysis: NumPy, Pandas, Matplotlib, Seaborn, Vega-Altair

Machine Learning: Scikit-learn, PyTorch, TensorFlow

Software Tools: Postgres, Git, AWS, Bash, Power BI, Domo, Excel, Jupyter, RStudio, Amazon S3/EC2/RDS

PROFESSIONAL EXPERIENCE

Data Analyst

Sep 2021 - Apr 2022

Curacao

Los Angeles, CA

- Engineered a data pipeline to extract, transform, and load 16+ GB worth of Open Banking data from FinTech APIs.
- Designed and developed a scikit-learn random forest classification model that predicts loan payment with more than 70% accuracy.
- Performed business data analysis to provide insights and recommendations across the Customer Lifecycle.
- 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.

PROJECTS

Youtube Filtering Capstone - *Medium Article*

June 2023 - Present

- Implemented a Youtube video filtering system to identify and flag potentially inappropriate content using Natural Language Processing (NLP) techniques.
- Developed an end-to-end solution that integrated with the Youtube API and a Streamlit web application for real-time content analysis and filtering.

NHL Draft Predictions - *Deepnote Article*

May 2023 - June 2023

- Harnessed the cutting-edge power of Language Models (LLMs) including Google's BERT and OpenAI's Chat-GPT to encode word embeddings.
- Leveraged these models to predict NHL draft positions for prospective players and cluster players together into meaningful groups based on shared characteristics.

RESEARCH EXPERIENCE

Undergraduate Researcher

Aug 2019 - Mar 2021

Atwood Lab

Irvine, CA

- Researched methods for principal graphs and graph neural networks to solve problems in single-cell data science such as trajectory inference.
- Programmed a graph-based semi-supervised classification model using Pytorch.
- Assembled a list of parameters for reclaiming spatial information of high dimensional single-cell RNA-sequencing data lost during data collection.

Undergraduate Researcher

Jul 2019 - Aug 2019

MathBioU & Math EXPLR

Irvine, CA

- Performed computational biology scientific research at a 6-week summer research program.
- Collaborated as a co-author in a research report to present findings to a diverse audience.
- Mentored a high school student on conducting research.
- Analyzed single-cell RNA-sequencing data, using methods including quality control, dimension reduction, clustering, and cell type identification.