

Michael Pham

pham_michael@outlook.com | (503) 744 - 1839 | github.com/mpham8

Education

Swarthmore College

Swarthmore, PA

Pursuing Bachelor's of Arts in Mathematics/Statistics and Economics, 3.75 GPA

Class of 2025

- Relevant coursework: Linear Algebra, Statistical Methods II, Econometrics II, Data Structures & Algorithms
- Online lectures: CS229 Machine Learning, CS230 Deep Learning
- Cross Country/Track & Field Student-Athlete, Economics Teaching Assistant, Economics Department Coordinator of Community-Building Events, Effective Altruism Club

Work Experience

UCSF Microbiome Sequencing Lab

Remote

Part-Time Data Analyst | R, Python (Tensorflow/Keras, SciKit)

November 2022 - Present

- Working with colleagues under guidance of Dr. Renuka Nayak to analyze high-dimensional sparse dataset to find how gut microbiome in SKG mouse colony alter immune phenotype and arthritis development
- Reading bioinformatics/ML research literature, using data generation techniques such as Poisson RV sampling and machine learning models such as Random Forests, CNNs to predict host phenotype from microbiome data

Karner Blue Capital

Remote

Quantitative Research Intern | Python (Pandas, SciKit)

May 2022 - August 2022

- Read mathematical financial literature to implement models such as Fama-French and CLA Semivariance into firm's Portfolio Optimizer, outperforming previous optimizer by 19 ppt on average over 6-year backtests
- Built machine learning models to predict stock over/underperformance relative to Portfolio Optimizer predictions
- Interpreted results and presented summary reports of findings to Vice President of KBC regularly

Technical Skills

Programming Languages: C++, JavaScript (Express.js, Node.js), Solidity, HTML/CSS

Data/ML Technologies: Python, Hugging Face, SQL, R, Stata, Jupyter, Git, Docker, AWS (EC2, S3)

Projects

Low-Resource Bam-Fr Neural Machine Translation | Python (Tensorflow, Pandas)

February 2023

- Worked with partner on using Transformer architecture to machine translate neglected low-resource language
- Achieved BLEU score of 14.81 on Bambara to French translation with limited data and computing power
- Placed 5th in Orange Silicon Valley Bambara-French machine translation competition

SKG Machine Learning Phenotype Prediction | Python (Tensorflow, SciKit-Learn, Pandas)

January 2023

- Built Model that takes in sparse microbiome data, uses GANs to generate artificial data, phylogenetic tree data augmentation encoding, ResNet architecture to make phenotype prediction
- Achieved 0.74, 0.75, 0.31 F1 score on predicting SKG/WT, Gender, Antibiotic treatment phenotypes/traits

Generating Pictures of Numbers | Python (Tensorflow, Pandas)

December 2022

- Created Generative Adversarial Network that generates new digits from MNIST number dataset

Neural Network using No Tensorflow | Python (Numpy)

November 2022

- Built 3 Layer ANN from scratch without ML libraries to better understand math behind neural networks
- Tuned hyperparameters on MNIST number dataset and identified numbers with 0.91 F1 Score