

Kiet C. Lam

klam@hamilton.edu

(773) - 681- 2591

EDUCATION

Hamilton College, Clinton, NY

August 2022 - May 2026

Bachelor of Arts in Data Science

- GPA: 3.70
- **Relevant coursework:** Design Principles, Linear Algebra, Multivariable Calculus, Probability and Statistical Inference, Algorithms and Data Structures

SKILLS

Programming language: Python, R, SQL, C++

Machine learning frameworks: Tensorflow, Keras, PyTorch

Data science tools: Pandas, Numpy, tidyverse, Scikit-learn, Matplotlib, Tableau, Seaborn, ggplot2, Microsoft Excel, Google Sheets

Database: MySQL and BigQuery

EXPERIENCE

Research analyst, **Hamilton College Neuroscience Department**, Clinton, NY

May - June 2023

- Visualized participants' demographics using Python's Matplotlib
- Designed a correlational matrix among participants' responses to different questions from multiple psychological scales (MAIA, PVD, QCAE, TDD, and UCLA) using R's tidyverse ggplot2
- Customized an R script imported from college collaborator's Github repository to compile HRV data from 300+ participants
- Encoded participant's ambiguous responses into appropriate values using R's dplyr package
- Constructed an R script that compiles all responses from 4 different Qualtrics questionnaires
- Implemented additional R packages (lmerTest, lme4, and arsenal) to build multiple regression models, produce statistical summaries, as well as perform significant testings

PROJECTS

Income prediction based on demographics

January 2024 - March 2024

- Created a Logistic Regression model that predicts whether or not a person makes an income of over 50k
- Optimized the model using cross-validation, scaling, and regularization (L1 and L2)
- Visualized the data set's demographics (age, gender, occupation, etc.) using pie charts and bar charts
- Created a pipeline to preprocess data using Scikit-learn

Is your banana ripe?

February 2024 - March 2024

- Utilized TensorFlow's Keras API to build a 2D multi-layered Convolutional Neural Network
- Visualized the model's training and validation loss and accuracy
- Integrated an Early Stopping Callback mechanism to prevent overfitting and enhancing model's accuracy from 80% to 95%
- Applied a 5-fold cross-validation technique, yielding an accuracy of 95.37% and a loss of 2%
- Tuned hyperparameters (L2 regularization rate, learning rate, and dropout rate, epochs) using GridSearchCV
- Evaluated the model's final performance using key metrics: precision (94.4%), recall (100%), and accuracy (96.68%)
- Leveraged OpenCV to download and preprocess over 3,000 images sourced from Google Images