

Yena Lee

4441 Purves St, Long Island City, NY, 11101 | (917)-999-8489 | yl4315@columbia.edu

EDUCATION

MS in Computer Information Sciences, HARRISBURG UNIVERSITY *Aug 2022 – May 2024 (Expected)*

- Relevant coursework: Data Structures & Algorithms, Computing Systems Architecture

MA in Statistics, COLUMBIA UNIVERSITY *Sep 2019 – May 2021*

- Relevant coursework: Bayesian Statistics, Machine Learning, Neural Networks, Financial Modelling
- Extracurricular: Vice President of Columbia Statistics Club

BS in Applied Math & Statistics, STONY BROOK UNIVERSITY *Aug 2014 – May 2018*

- Relevant coursework: Probability, Linear Algebra, Data Analysis, Game Theory, Financial Economics
- Second major: Economics
- Honors: Dean's list, Cum Laude
- Extracurricular: President of Korean Traditional Drumming Club (DDKY)

SKILLS

Data Analysis Python (NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn, Keras, TensorFlow) | R (dplyr, tidyr, ggplot2) | SAS | Tableau | D3 | Excel

Data Engineering SQL (MySQL, SQLite) | PySpark | GCP

Web Analytics Google Analytics

PROJECT

Predicting The Term Deposit Subscription [Python] *Oct 2022*

- Developed classification models using **Logistic**, **LDA**, **SVM**, **Decision Tree**, **KNN**, and **Random Forest** to identify customers who are likely to respond to the campaign and predict response rate.
- Cleaned and normalized bank campaign data of 40k observations and 17 variables, and analyzed missing values. Identified the imbalanced nature of the dataset with a heavy skew toward negative responses and applied Upsampling method to resolve the issue.
- Classified the target customers with 89% test accuracy and 62% recall, and pinpointed age, average yearly balance, and a number of campaign touchpoints as influential predictors for responding positively to the marketing campaign.

Image Colorization with Neural Network [Python] *Apr 2021 – May 2021*

- Designed Convolutional Neural Network algorithm that automates the colorization of grayscale images into RGB using **TensorFlow2.0** and **Keras**, and developed it using features extracted from the pre-trained **Inception-ResNet-v2** model.
- Preprocessed 35k images with 1k categories and trained the model with 0.025 MSE loss.
- Predicted accurate colors for high-level images, such as the sky and forest, but not for specific objects, such as butterflies.

Head and Shoulders Pattern Recognition and Prediction [R] *Mar 2021 – May 2021*

- Developed Head and Shoulders Pattern Algorithm to recognize and predict future patterns of stock markets using Kernel regression with **Random Forest** model.
- Predicted Head and Shoulders Patterns with 78% recall and 66% precision accuracy.

RELEVANT EXPERIENCE

Teaching Assistant, STONY BROOK UNIVERSITY *Jan 2018 – May 2018*

- Assisted Applied Linear Algebra course with solution development, assignment grading, and exam proctoring. Resolved questions from over 200 students by holding office hours twice a week.