**EDUCATION**

**Columbia University New York, NY**

**M.S. in Data Science** Aug 2021 – Dec 2022

Coursework:Algorithms, Data Visualizations, Applied Machine Learning, Applied Deep Learning

**University of California, Berkeley Berkeley, CA**

**B.A. in Data Science**, Business Analytics concentration, GPA: 3.93/4.0, Phi Beta Kappa society Aug 2019 – May 2021

Coursework: Data Structures, Artificial Intelligence, Data Science Techniques, Probability, Time Series, Causal Inference

**East Los Angeles College Monterey Park, CA**

**A.A. in Mathematics & A.S. in Economics**, GPA: 4.0/4.0, Summa Cum Laude Aug 2017 – June 2019

Coursework:Linear Algebra, Ordinary Differential Equations, Calculus, C++, Java

**EXPERIENCE**

**iQIYI, Inc. Beijing, China**

*Ads Algorithm Backend Intern* May 2021 – Aug 2021

* Developed a testing framework for the ads allocation emulator in assisting automatic validation process.
* Created a time series module for inventory prediction, achieving a cross-validation accuracy of over 80%.
* Developed High Water Mark (HWM) algorithm from scratch for contract allocation confliction testing.
* Conducted research on real-time bidding and reported to manager possible algorithms based on more than 20 articles.
* Updated and shared codes with the team using git.

**Goodly Labs, Inc. Berkeley, CA**

*Validation Team Lead* Aug 2020 – May 2021

* Lead a team of 3 undergraduate students supporting the validation application of the Public Editor system.
* Developed architecture based on Krippendorff’s alpha, providing insights into the reliability of the program to clients.
* Report to the CEO and the mentor of Goodly Labs regularly to ensure maximum efficiency of the team’s progress.
* Link:<https://github.com/Hegelim/PE-Validation-Fall-20>

**PROJECTS**

**The solveminmax Python package, open source** May 2021 – Present

* Implemented a Python module that solves a sum of min and max equations using regular expressions, numpy, sympy, matplotlib, etc.
* Created unit testing using pytest to validate the module extensively with more than 30 testing cases.
* Link: <https://github.com/Hegelim/solveminmax>

**Hand-written digits classification, UC Berkeley** June 2020 – Aug 2020

* Classified MNIST handwritten digits using a fine-tuned two-layer neural network, reaching a cross-validation accuracy of 97.7%.

**Gitlet, UC Berkeley** Apr 2020 – May 2020

* Built a version-control system in Java from scratch that mimics the features of Git using data structures, OOP, graph-traversal, file-reading/writing, serialization, and error-handling.
* Extended features as in git remote and earned full extra credits.

**Spam/ham email classification, UC Berkeley** Apr 2020 – May 2020

* Developed a logistic regression model in Jupyter notebook using one-hot-encoding and automated feature selection based on word occurrences, improving classification accuracy from 87% to 99%.
* Participated in the course competition and ranked No.3 among more than 1000 submissions.

**SKILLS AND TECHNOLOGIES**

Programming:Python (main), Jupyter (main), SQL, R, Java, HTML5/CSS/JavaScript, C++

Python Packages:pandas, numpy, scipy, scikit-learn, sympy, matplotlib, seaborn, re

Other: Markdown, reStructuredText, LaTex, Git, Excel