Huy Ho

huyho@berkeley.edu | github.com/huy-ho | linkedin.com/in/huydhho

EDUCATION

University of California, Berkeley | Berkeley, CA — Expected Graduation May 2022
B.A. - Data Science and Statistics — *Cumulative GPA: 3.0*

San Diego Miramar College | San Diego, CA — *Aug 2018 - May 2020*
IGETC Transfer Certificate - Statistics — *Cumulative GPA: 3.87*

Courseworks: Principles and Techniques of Data Science, Data Structures, Algorithms, Multi-variable Calculus, Statistics, Probability, Macro/Micro Economics, Advanced Businesses Analytics

SKILLS & TECHNICAL TOOLS

Languages: Python, Java, SQL/SQLite, HTML, CSS, JavaScript
Technologies: Numpy, Pandas, Tensorflow, Keras, Scikit-Learn, Matplotlib, Seaborn, Geopandas
Others: Jupyter Notebook, Git, MS Office, G Suite, Adobe Suite

EXPERIENCES

Data Engineer Intern | Nozomi Networks — *June 2021- Present*

- Researched and experimented with supervised and unsupervised machine learning techniques for time-series data.
- Used Keras and Tensorflow to implement neural networks models and an ensemble method for anomaly detection.
- The forecasting model was put into production and introduced to the company's pipeline.

Curriculum Development Team | DataGood @ Berkeley — *Jan 2021- Present*

- Created and hosted workshops for club technical meetings
- Teach fundamental CS/DS skills and teachniques such as Python, Git, SQL, and relevant libraries.

Finance Intern | UC Berkeley Vietnamese Student Association — *Aug 2020 - Dec 2020*

- Coordinate meetings between the finance department.
- Created info-graphics to promote the club and upcoming events.
- Along with another intern, hosted a fundraiser that raised over $2000.

Academic Tutor| San Diego Miramar College — *Sep 2019 - May 2020*

- Helped students with mathematics ranging from Algebra to Differential Equations
- Facilitated study groups and review sessions for students

PROJECTS

Stock Market Forecasting Analysis using Machine Learning Techniques| *Python, Pandas, Keras, Tensorflow*

- Used Yahoo Finance time-series data to investigate major stocks (AMD, NVIDIA, and INTEL)
- Applied machine learning techniques for forecasting, clustering, and analysis.
- Implemented LSTM, CNN, XGBoost, ARIMA, etc models to test different approaches.

Washington DC and Chicago Crime Analysis | *Python, Pandas*

- Used Washington DC and Chicago crime data to explore trends across factors such as race, socioeconomic status.
- Analyzed data using Pandas, regression, and geospatial analysis.
- Final Deliverable: Analysis of crimes in Chicago and DC to explore trends across factors such as neighborhoods and time. GeoPandas maps to show crime trends with regards to poverty

Spam/Ham Email Classifier | *Python, Pandas, scikit-learn*

- Created a data pipeline to process the data and built a regression model to predict whether an email was spam or ham with a 85% training accuracy on the test set.