New York City, NY
evanyo860126@gmail.com
(917)-940-2426

Chi-Yu Lin

linkedin.com/in/evanyo25
github.com/EvanYo25

Technical Skills

C++, Python, Numpy, Pandas, sklearn, Pytorch, Tensorflow, Selenium, Natural Language Processing, Machine Learning
Java, R, Matlab, HTML, CSS, JavaScript, SQL, PHP, Angular, Django, NodeJS, Json, Git, Haskell, Hadoop

---

Experience

Software Engineering Intern – eBay Inc. 2021 - 2021

Search & Discovery Page for eBay Stores, ranking stores by sales, participation, and traffic.

- Utilized **SparkSQL** to gather over **1.2 million rows of data per day** from more than **10 tables**.
- Preprocessed the data into N-day window, and created tables in **Hadoop File System** for storing.
- Set up **automation** for data loading, used **Elasticsearch** to store the finalized data, and created **APIs** for them.
- Developed various filters and ranking options on the frond-end using **Angular** framework with **NodeJS**.
- Created a tool to analyze stores internally, and APIs for choosing top stores to create a new entrance.

Research Assistant - National Taiwan University 2017 - 2018

NameSex on PyPI, Package to predict gender tendency from a Chinese given name (in Traditional Chinese).

- Trained a word embedding using Word2Vec algo in gensim on Yahoo News with **87 million Chinese characters**.
- Trained predictive models such as **K-NN and logistic regression**.
- Performed model tuning and validation by **k-fold** and reached an **accuracy of 89.57%** on cross validation set.

Side Projects

Improving L2 Regularization Based on AutoL2 2021 - 2021

Modification on AutoL2, Referenced paper: On the training dynamics of deep networks with L2 regularization.

- Utilized **Pytorch** to build **Convolutional Neural Network** to reproduce and run further experiments on CIFAR-10.
- Multiplied a fixed parameter with MINLOSS & MINERROR resp. as a new threshold deciding whether to update λ.
- Added a feature to increase λ in order to jump out of local minimum and possibly reach a better minimum.
- Experiment achieved a **better accuracy of 6%** when compared with both AutoL2 and optimal λ we found.

Unsupervised Classification for Super Basketball League Players 2019 - 2019

Alternative categorization for basketball players rather than traditional categories of guard, forward, and center.

- Collected players' game performance data from the official site of SBL using Python Selenium.
- Performed dimensionality reduction using **Principal Component Analysis , t-SNE, and LDA**.
- Applied **K-means** clustering and **hierarchical agglomerative clustering** for strategic team formation.

Foxconn Tech. Stock Price Prediction Using Online News 2018 - 2018

- Searched for positive and negative keywords from the news using **N-gram** (2-gram ~ 6-gram).
- Used the keywords selected through **TF-IDF as features** of news to fit into various **regression algorithms**.
- Reached an **accuracy of 82%** predicting a rise or fall in stock price on cross validation set using **k-fold**.

Basketball Statistics System 2017 - 2017

- Utilized **Django** to develop a log in and statistical system for more than **10 teams** at National Taiwan University.
- Designed and constructed the **database** storing teams, players and games data **using SQLite**.

Education

Cornell University - M.Eng., Computer Science New York City, NY 2022
New York University - M.S., Computer Science - GPA 3.9, Transferred to Cornell New York City, NY 2021
National Taiwan University - B.B.A., Information Management - Presidential Academic Award Taipei, Taiwan 2019

- Class Representative, Graduate Representative, Captain of NTUIM's Basketball Team, Softball Varsity Team