

Evan (Yuheng) Yang

Email: yy2597@nyu.edu | Mobile: +1 (201) 748-9638 | LinkedIn: www.linkedin.com/in/evan-yang-a8825a1a7

EDUCATION

New York University

New York City, NY

M.S. in Data Science | **Cumulative GPA: 4.0**

September 2022 – May 2024 (Expected)

- Coursework: NLP with Representation Learning, Optimization, ...

New York University Shanghai

Shanghai, China

B.S. in Data Science; B.S. in Business and Finance | **Cumulative GPA: 3.927**

September 2018 – May 2022

- Awards: **Summa Cum Laude (Top 5% ranking)**, University Honors Scholar, NYU Shanghai Global Quintessence Scholarship
- Coursework: Probability and Statistics, Databases, Machine Learning, Econometrics, Investments, Derivatives, Corporate Finance, ...

WORK EXPERIENCE

Ziroom | *the largest unicorn in IoT & housing management in China*

Beijing, China

Data Scientist Intern, AI and Business Intelligence Dept.

June-August 2021

- **Data modeling with machine learning:** Worked on an inventory turnover diagnosis project, developed a well-performing classification model based on **Random Forest** and **Neural Networks** to predict whether the inventory unit can be leased with a test ROC-AUC of 0.8. The model was integrated into operations applications and used by retailing teams with **over 1500 workers**.
- **Data analysis:** Analyzed and evaluated the factors affecting inventory turnover. Built a factor improvement advising model, and reported the analysis results to retailing teams for refinement and adjustments to inventory units. The prototype with the classification model helped improve overall inventory turnover performance by **20%**.
- **Data management and visualization:** Performed data retrieval and analytics using **Hive**, assisted a data warehouse architect with schema modeling. Completed visualization tasks using **Tableau** and assisted building 2 dashboards using **Youdata BI**.

Essence Securities | *Top 10 investment bank in China*

Shanghai, China

Security Analyst Intern, Research Dept.

June-August 2020

- **Industry and company research:** Participated in the research of the company POPMART, composed industry analysis reports (partially published).
- **Research report composing and data analysis:** Completed qualitative research on supply chains and logistics systems for Yonghui Superstores (SH: 601933) under advising. Assisted data processing and analytics for valuation modeling.

PROJECTS & RESEARCH

Kaggle *Super Soaker* Performance Prediction Challenge

August 2022

- **Result:** Ranked **top 7%** (131/1888) solo.
- **ML pipeline implementation:** Obtained insights among different features, performed feature engineering, and applied an ensemble model using Logistic Regression, Random Forest, and Neural Networks.

Financial Distress Prediction based on a Time-weighted Ensemble Model and Its Application

January-May 2022

- **Features Construction:** Preprocessed over 50 financial gauges of 490 companies across different years, applied feature engineering using PCA, and performed sentiment analysis to MD&A from the financial statements of these companies to construct features pool.
- **Model design:** To predict default risks of these companies, created a **time-weighted ensemble classifier** that dynamically assigns different importance to different base classifiers (logistic regression, Random Forest, etc.) with optimal feature choices across different years. The results show the proposed algorithm could predict the potential default three years after at a test ROC-AUC of 0.8693.

MBTI Personality Categorization Prediction with Social Media Speech Posts

November-December 2021

- **Unstructured data processing:** Applied text preprocessing techniques using NLTK to unstructured forum posts by **over 8600** individuals for the four-dimensional classification tasks.
- **Model training and evaluation:** Trained and fine-tuned classifiers based on Random Forest and LSTM with the preprocessed data to complete the personality prediction. Achieved **test accuracy of 80-90%** in each of the four theoretical dimensions.

SKILLS

Programming and Toolkits: Python, Scikit-Learn, PyTorch, Numpy, Pandas, Stata, SQL, Basic AWS Architecture, JavaScript, D3

Language: Mandarin (Native), English (Fluent, TOEFL 112), Japanese (Beginner)