

# DI (HEDY) HE

dh3171@nyu.edu • linkedin.com/in/hedi • (+1) 718-300-2709

## EDUCATION

### New York University

Sep 20 – Exp: May 22

*Master of Science in Data Science. GPA 3.95 / 4*

- Courses: Data Science, Machine Learning, Natural Language Processing with Representation Learning, Deep Learning, Big Data, Probability and Statistics, Optimization and Linear Algebra

### Xiamen University

Sep 13 – Jun 17

*Bachelor of Management in Accounting. GPA 3.80 / 4. Certified Public Accountant in China.*

## WORKING EXPERIENCES

### Fidelity Investments

Jun 21 – Aug 21

*Data Science Summer Intern, Advanced Technology in Investment Management* Boston, MA

- Researched on **change point detection** models to identify changes of equity business, based on equity's risk exposures on market risk factors that were measured by **correlation** and **mutual information**. Based on related period, trained **XGBoost** models to use market factors to predict equity prices. Automated the whole pipeline on **AWS** via **PySpark** for 10000+ assets using 30 years' daily data.
- Imputed quarterly data to daily data using gradient-based **optimization** on top of **PyTorch**.

### Charles & Keith Group

Jul 19 – Jun 20

*Data analyst, Retail Data Science* Shanghai, China

- Predicted product-level sales for new season using **random forests** and interpreted it with **SHAP**. Built a website using **Dash** and deployed on Heroku to help users interact with model outcomes.
- Recommended new store locations for expansion team using **k-means** and **DBSCAN Clustering** with parsed transportation and competitor stores' geolocation data getting from Gaode map **API**.
- Designed and maintained tables in **Hive** and built **Tableau** dashboards for stores and products.
- To help set daily sales targets, first forecasted monthly sales via **exponential smoothing** and **Facebook prophet**. Then allocated them daily based on holiday and week seasonality measured by **regression**.

### Deloitte Consulting LLP

Jul 17 – May 19

*Business Analyst, Enterprise Application* Shanghai, China

- As a consultant, prepared requirements documents and function specifications documents for client's financial **information system**. Designed prototypes using Axure RP and tested final applications.
- Increased client's operation efficiency by building management **dashboards** for 200+ subsidiaries.

## ACADEMIC PROJECTS

### Self/Semi-Supervised Learning (SSL) Competition

Apr 21 – May 21

- Competed on finding the best SSL algorithm for a 512,000-images classification task with only 5% labelled from 800 classes. Researched and trained **contrastive learning** models (SimCLR, SimSiam, BarlowTwins), and **pseudo labeling** models (FixMatch, CoMatch) via **PyTorch** on Greene Cluster. Achieved 51% accuracy on the testing set and got 2<sup>nd</sup> place.

### Music Recommendation System

Mar 21 – Apr 21

- Learnt latent representation for users and items using **alternating least squares** (ALS) model via **PySpark** on the Million Song Dataset. Used **PCA** to reduce dimension and visualize music genres.

### Aspect-Based Sentiment Analysis (ABSA)

Oct 20 – Dec 20

- Predicted aspect-level sentiment polarities on a restaurant reviews dataset (MAMS) that contains multiple aspects with multiple sentiment polarities. Implemented **LSTM** model and fine-tuned **BERT** based model. Achieved 84% accuracy on MAMS and performed error analysis.

## SKILLS

- Data Science: Visualization, Feature Engineering, Regression, Classification, Clustering, Statistical Modeling, A/B Testing, Time Series, Text Analysis, NLP, Deep Learning, Data Structure, Big Data
- Skills: Python (Pandas, Matplotlib, scikit-learn, Torch, Dash), Spark, AWS, SQL, Hive, Tableau