Haoxue Li

hl3664@nyu.edu 347-828-0470
https://github.com/Gilemore https://www.linkedin.com/in/haoxue-li

EDUCATION

New York University *Sept. 2019 - present*
Master of Science, Data Science, GPA: 3.9/4.0 *Expected grad date: May 2021*
Courses: Big Data, Machine Learning, Database, Natural Language Processing *NYC, NY*
Soochow University *Sept. 2015 - Jun. 2019*
Bachelor of Science, Mathematics and Applied Mathematics, GPA: 3.8/4.0 *Suzhou, China*
Won second prize at National Model APEC Competition in 2018, among 10,000+ participants
Participated Summer Leadership Program at Earnst & Young and received final offer

SKILLS

Programming: Java, Python, Spark, Hadoop, PHP, JavaScript, HTML, CSS
Focus Areas: Data Engineering, Machine Learning, Big Data

EXPERIENCE

Software Engineer Intern *May. 2020 - present*
MingJian Technology *NYC, New York*

- Migrated the whole website to Amazon Web Service (AWS) and maintained the database service
- Designed the main layout and upgrading the website using React, JavaScript, HTML, and CSS
- Programmed an automatically updated pricing script with external APIs to crawl real-time price of different products

Bio Statistical Research Analytic Intern *Dec. 2018 - Aug. 2019*
Duke Kunshan University, Global Health Research Center *Suzhou, China*

- Designed China's first study: Association between Blood Pressure and Death among Frail Older Adults, which focused on association between blood pressure and mortality among frail older adults
- Applied Cox model with penalized spline to detect the non-linear relationship between continuous blood pressure and mortality and then examined the linear association between three-level categorical blood pressure and mortality
- Conducted sensitive analysis and drew the counter-intuitive conclusion that higher blood pressure associated with lower death rate among frail older adults in China

Business Analysis Intern *Jul. 2018 - Oct. 2018*
Costa Crociere S.p.A *Shanghai, China*

- Made weekly reports on revenues and costs of cruiser ships and provided constructive suggestions based on the comparison with previous performance in routine sessions
- Conducted on-site inspection on the crew, scrutinized the outliers in reports and presented solutions
- Implemented price and order adjustments according to on-site inspection experience, resulting 10% increase in the revenue of Food & Beverage department for two months

PROJECTS

Web for Books Recommendation *May. 2019 - present*
Big Data course project *NYC, NY*

- Implemented a books recommendation system with 17m interactions using alternating least squares (ALS) method in Spark
- Built linear models, Gradient Boosting model, and KNN models to learn latent factor representation for items, achieving 10% accuracy improvement compared with the full model.
- Developed a website based on our books recommendation system to recommend books for users, which allows users to add their reading history and hobbits.

Neural Language Modeling with a Recurrent Neural Network (RNN) *Sept. 2019 - Oct. 2019*
Natural Language Processing project in class *NYC, NY*

- Trained a RNN model on Wikitext-2 dataset to predict whether a sentence makes sense in real life conversation, reaching the validation perplexity as 206.96
- Trained a Long Short Term Memory (LSTM) model, performed hyper-parameter tuning in both hidden size and dropout rate, and reduced its validation perplexity greatly by 17%
- Analyzed the projection embeddings, sampled 1,000 sentences from the best LSTM model, and concluded that most sampled sentences were similar with real-life sentences in validation dataset