HAO LIU

25 River Drive South, APT 2804, Jersey City, NJ 07310 | (206) 245-5556

hliu135@fordham.edu | www.linkedin.com/in/zluxyoe

EDUCATION

FORDHAM UNIVERSITY   *Aug. 2017 - Dec. 2018   New York, US*
- M.S. in Applied Statistics and Decision Making, Gabelli School of Business
- Selected Courses: Machine Learning, Big Data Programming, Statistical Methods and Computation
- Selected Awards: Dean's Scholarship

ESCP EUROPE   *Sept. 2014 - Jun. 2017   Paris, France*
- M.S. in Management, Marketing option
- Selected Courses: Algorithms and Programming, Customer Relationship Management, Marketing Research

XIAMEN UNIVERSITY   *Sept. 2010 - Jun. 2014   Xiamen, China*
- B.S. in Civil Engineering, School of Architecture and Civil Engineering; Minor in Advertisement
- Selected Awards: Excellent Academic Performance in 2010, 2011, 2012 (Top 10% among all students)

SELECTED EXPERIENCES

Data Analyst Intern | GfK France   *Sept. 2016 - Jan. 2017   Paris, France*
- Client *DPD Group*: Validated data and prepared KPI dashboard for brand awareness and satisfaction survey for 23 countries; visualized survey results based on indexes such as Net Promoter Score.
- Client *Sodexo*: Analyzed customer satisfaction level via text analysis; delivered full analysis reports.
- Client *L'oreal*: Checked questionnaire content and performed logical test for L'oreal panel strategy study in China.

Data Analyst Intern | Millward Brown   *Feb. 2016 - Jul. 2016   Paris, France*
- Used Millward Brown internal methodology in ad tracking (postview analysis) and brand performance analysis to generate metrics with higher readability.
- Extracted and processed data with SPSS, implemented Excel VBA to automate and simplify data validation pipeline.

Data Analyst Intern | Yidian Inc (Internet Startup) *May. 2015 - Jul. 2015   Beijing, China*
- Implemented script to monitor and analyze ads traffic with Baidu Analytics, e.g. total page view, unique user count.
- Launched an online campaign on the social media Wechat based on the user patterns found in traffic analytics.

SELECTED PROJECTS

Predict Readmission Probability for Diabetes Patients   *Spring 2018*
- Cleaned and processed raw data using scikit-learn package in Python
- Designed a Fully-Connected Neural Network to learn the relationship between medical features and readmission possibility using Tensorflow
- Achieved over 95% test accuracy and visualized the results using Matplotlib in Python

Analyzing StackOverflow Questions Using Spark   *Spring 2018*
- Built a Hadoop cluster on Google Cloud for analyzing dataset sizing around 3GB
- Analyzed the dynamic trend of programming languages and editors on StackOverflow over years using SparkSQL
- Implemented Word2Vec model in Spark MLlib to predict possible tags based on the question title

Web Analytics Course Project: Tweet Analysis   *Fall 2017*
- Crawled 10K+ tweets related to "thanksgiving" via twitter API using tweepy package in Python
- Tokenized tweet text into bag-of-words, then fitted with LDA topic model, clustered tweets with similar content to find out the general topics frequently discussed with "thanksgiving" and visualized results with Tableau

SKILLS

SQL, R, Python, Spark, Scala, Hive, Pig, Java, SPSS, Excel (VBA, lookup, pivot table), LaTeX, HTML