

# Huimin (Hannah) Han

✉ huimin.han3@gmail.com

☎ 301-454-9485

🌐 hannahhan3.github.io

## Education

### Master of Science, Business Analytics (STEM Program)

GPA:3.8/4.0

*University of Maryland, College Park, United States*

08/18 – 12/19

Coursework: Big Data and Artificial Intelligence for Business; Data Mining and Predictive Analytics; Data processing and Analysis in Python; Database Management Systems; Data, Models and Decisions; Decision Analytics

### Bachelor of Management, Public Health Administration

GPA:3.56/4.0 (Rank:2/38)

*Capital Medical University, Beijing, China*

09/14 – 07/18

Merit-based scholarship

Coursework: Advanced Mathematics; Linear Algebra; Probability & Statistics; Statistics in Health Management and Statistics Software Package; Operational Research; Data Collection; Generality of Clinical Medicine; Health Service Administration; Computer System and System Software; Health Information Management

## Experience

### Research Assistant

College Park, MD, United States

*The Center for Health Information and Decision Systems (CHIDS) at UMD*

08/19 – present

Advisor: Gordon Gao

- Did research on time series data prediction, compared different forecasting methods including classical algorithms and machine learning based algorithms (e.g., ARIMA, SVMs, neural networks and combination of different algorithms)
- Obtained data from technology company: Inovalon
- Developed code and tools to predict pharmacy sales volume
- Deployed machine learning models into production to support staff scheduling of pharmacies

### Teaching Assistant (of Master's Students)

College Park, MD, United States

*R.H. Smith School of Business at UMD*

03/19 – present

- BUDT758P: Decision Analytics (Fall 2019) by Prof. S. Raghu Raghavan
- BUSO758M: Decision Modeling (Summer 2019) by Prof. S. Raghu Raghavan
- BUSI634: Operations Management (Fall 2019) by Prof. Zhilong Chen
- BUSM762: Decision Analytics (Spring 2019) by Prof. Sujin Kim

### Data Analyst Intern

Beijing, China

*Xuanwu Hospital*

11/17 – 05/18

- Utilized multiple sources of healthcare data such as financial and clinical data, to meet both internal and external reporting requirements and support new initiatives
- Identified opportunities for improving clinical processes and outcomes by using sophisticated and creative approaches of analyzing outcomes and utilizing data
- Learned the data structures, wrote queries, linked, and accessed data for manipulation and analysis

## Selected Projects

### Capstone Project Collaborated with real estate investment company: Roofstock

09/19 – 12/19

- Extracted and processed data from the database of Roofstock Real Estate Company using Python
- Analyzed and explored a large collection of historical market transactions data (using Pandas) and visualized data to uncover suspicious patterns in data (using Matplotlib)
- Utilized survival analysis technique and applied machine learning algorithms to develop a 'time to sell' model; a tool used to predict how long it will take a listing posted on the website to be sold

### Machine Learning Application in Hospital Readmission Rate

03/19 – 05/19

- Obtained data from John Hopkins University School of Medicine, preprocessed the datasets according to the principle of statistics in R, and performed variable selection and feature engineering
- Built different models (regression models, tree models along with ensemble methods, KNN models, etc.) to predict if a discharged patient will return within 30 days
- Evaluated models trained with imbalanced datasets using their F1 score

### Exploratory Data Analysis and Text Mining of Google Play Store Apps

03/19 – 05/19

- Preprocessed the Google PlayStore App dataset (using Pandas), and performed visual exploratory data analysis to find common characteristics between high-ratings and highly downloaded apps (using Seaborn and Matplotlib)
- Extracted features from the dataset's "reviews" column (text data)
- Selected specific features using the bag-of-words technique and used topic modeling to discover any natural topics in the reviews
- Provided specific suggestions for the app-making business of different industries according to results of analysis

### Restaurants Around UMD Database Application

09/18 – 12/18

- Designed a database for restaurants around University of Maryland which contained information about restaurants, customers, reviews and coupons, following an Entity Relationship approach and normalization rules
- Conducted logical and physical design of the database in SQL Server
- Built web application based on the database with R Shiny

### Skills

---

- **Machine Learning Algorithms:** Hands on experience of Regression (Linear, Logistic, Stepwise), Instancebased (KNN), Regularization (LASSO, Ridge), Decision Tree (CART), Clustering(K-Means), Dimensionality Reduction (PCA, LDA, QDA), Ensemble (Bagging, Boosting, Random Forest)
- **Programming languages & Visualization tools:** Python (pandas, numpy, sklearn, matplotlib, seaborn), R, SQL, Tableau
- **Machine Learning Tools:** Pandas, NumPy, SK-learn, Keras, Matplotlib, Seaborn