

# Xiaoxuan Qin

+1(412)-956-2064/ Email: xiq33@pitt.edu

---

## SKILLS

---

**Programming Language:** Python, R, SQL, JAVA, JavaScript, HTML/CSS, C, LaTeX

**Tech Stacks:** MySQL, NoSQL, Hadoop (HDFS, MapReduce), Hive

## EDUCATION

---

**Hubei University of Economics**

*Sept 2017-Jun 2021*

- Bachelor of Management in Information Management and Information System

GPA: 3.7/4.0

- Scholarship: Canglong Student Scholarship (Twice) (Top 10%); Specialized Scholarship (Top 5%)

**University of Pittsburgh**

*Sept 2022-till now*

- Master of Information Science

GPA: 3.75/4.0

## PRACTICAL EXPERIENCE

---

**Hangzhou Xiaomawang Education Technology Co. Ltd.**

*May 2021-May 2022*

*Lecturer of Python Programming*

- Instructed students in Python programming, starting with fundamental concepts, followed by data structures, then progressing through algorithm development, modular programming, front-end design, and data scraping, and culminating in AI programming.
- Guided nearly 90% of students to achieve the Level 3 Python Certificate endorsed by the China Institute of Electronics and received almost 100% positive feedback from students.

## SAP

*Aug 2020-Sept 2020*

*PTA of Intelligent Manufacture Project Group*

- Developed in-depth analysis of Dissona's Value Stream Mapping, which encompassed customer and market demand, production management, and supply chain system management, and identified critical process bottlenecks.
- Designed and implemented robust information systems (CRM, ERP, SCM, MES) tailored to optimize process and resource management, thereby effectively resolving bottlenecks identified through Dissona's Value Stream Mapping, culminating in a 100% successful transformation towards an intelligent workshop.
- Attended AI+CRM products seminar, discussing how to integrate AI with CRM to realize personality service to increase efficiency, and manage risks.

## SELECTED ACADEMIC EXPERIENCE

### Database System for E-Commerce Implemented by Flask

*Oct 2022-Dec 2022*

- A database system designed as an online perfume and body-care shopping system, which displays six categories of products with eight types of scents - aiming at providing customers with an online perfume and body-care shopping experience.
- Database Design: Designed and implemented a comprehensive e-commerce database comprised of 10 entities and 10 relationships between them.
- Customer Interface Development: Developed a comprehensive customer interface, that allow customers to register, log in, modify personal information, browse and search items by category or scent, add products to their shopping cart, and complete checkout process securely.
- Manager Interface Development: Designed a comprehensive management portal for site administrators, enabling them to view product details, update product information, manage product availability, and monitor sales data.

(Project available at: [https://github.com/Qin99113/22fall\\_infsci2710\\_project](https://github.com/Qin99113/22fall_infsci2710_project) )

### PPG Project based on Machine Learning (Personal)

*Mar 2023-Apr 2023*

- Developed machine learning models to predict the important property and classify the popularity of paint colors based on inputs from two color models: RGB and HSL.
- Performed regression analysis and classification using both non-Bayesian and Bayesian linear models to decipher the influence of color model inputs on a concealed paint property and to discern the popularity of paint colors based on color model inputs.
- Trained and tuned models with more complex methods such as Generalized Linear Model, Elastic Net, Neural Network, Random Forest, SVM, and Gradient-Boosted Tree via resampling.
- Tested model performance using RMSE, Accuracy(for classification), and ROC (for classification) to ascertain the most effective predictive model.
- Leveraged the selection of the best models to successfully identify key variables that had a strong impact on the important paint property and the popularity of paint colors.

### Enhancement of Canvas Search Engine

*Sep 2022-Dec 2022*

- Contributed to addressing the challenge of insufficient global search capabilities; i.e., through enhancing the ability of the search engine to access and retrieve data from the entire platform of a specific course.
- Perfected the Search Algorithm of Canvas to increase its breadth and precision, implementing both exact search algorithm and fuzzy search algorithm (BM25) and evaluated by MAP, which recorded [relatively] a high score of 0.837.

- Refined search results display mechanism to provide users with more relevant, accurate, and intuitive results; this helped to enhance user experience and boost user satisfaction nearly by 50%.

(Project available at: <https://youtu.be/FShazMnyVy8> )

### **Chaotic Data Analysis based on Machine Learning**

*Jan 2020-Mar 2020*

- Utilized KNN and linear regression to predict the passenger capacity of taxis in New York City at various times of the day. The research evaluated model performance using RMSE and R-squared, to ensure the selection of the most reliable predictive model.
- Built a multiple linear regression model for a bicycle sharing system and predicted the total number of bicycle rentals in a day according to the attributes of a day (season, month, holiday, etc.), and evaluated the initial model using RMSE and R-squared metrics. Tuned the model by adjusting its parameters and utilized cross-validation and regularization to avoid overfitting. Re-evaluated the model using the same metrics to assess the effectiveness of these adjustments.
- Implemented Gradient Boosting and Gradient Boosting Decision Tree, and made an efficiency comparison between my decision tree and that in the sklearn. (Project available at: <https://github.com/dhbloo/GBDT-project> )

### **Big Data Analysis of Taobao's Singles Day Sales**

*Jan 2019-Feb 2019*

- Constructed a Hadoop cluster on a Linux system, built a Hive data warehouse to store 7027945 pieces of data, and then built MySQL relational database to run real-time query and information visualization, and integrated them using Sqoop for efficient data transmission.
- Tackled three research gaps pertaining to the sales data which allowed for a detailed understanding of consumer behavior and trends, critical for informing targeted marketing strategies and improving customer engagement. The gaps tackled are:
  - Categorized by gender across different provinces the number of participants involved in purchasing.
  - Analyzed the attractiveness of the most popular brands (i.e., LM, SVM) among different age groups
  - Pinpointed the province with the highest number of indecisive shoppers. These shoppers are characterized by the behavior of adding items to their shopping cart but not proceeding to checkout (SVM).