

AIQI (CHELSEA) CUI

(814) 882-9127 | ac4788@columbia.edu | [acui34.github.io](https://github.com/acui34)

EDUCATION

Columbia University, New York, NY

Expected Dec 2021

Master of Science in Data Science

Carnegie Mellon University (GPA: 3.8/4.0), Pittsburgh, PA

May 2020

Bachelor of Science in Information Systems, Secondary Major in Statistics & Machine Learning

Coursework: Machine Learning (Graduate Level), Natural Language Processing (Graduate Level), Data Mining, Advanced Data Analysis, Database, Data Visualization

SKILLS

Software Development: Python, Ruby on Rails, JavaScript, React, Swift, HTML, CSS, Ajax, TypeScript, Git, Linux, C

Data Science: R, SQL, Scikit-Learn, NLTK, Hadoop, MapReduce, Pig, Tableau, Excel Pivot Table

WORKING EXPERIENCE

Handshake | Software Engineering Intern

San Francisco, CA

May 2019 – Aug 2019

- Used Elasticsearch to enhance job role search performance and increased conversion rate by 262% in 7 days
- Initiated company-wide web accessibility discussion on the platform user interface leading to 10+ major changes
- Conducted A/B tests on search relevance and emails with React and Ruby on Rails to improve application rate

Microsoft | Software Engineering Intern

Guangzhou, China

May 2018 – Aug 2018

- Created a filter with Python (Scikit-Learn) to raise chatbot's chitchat response accuracy from 70% to 96%
- Built and organized a bilingual knowledge base with 3000+ utterance data to support 95% business FAQs
- Analyzed weekly user responses via Excel Pivot Table and Tableau to assess the efficiency of each project sprint and major functional changes

SELECTED PROJECTS

NLP Auto Question Generation & Answering System (Language: Python)

Feb 2020 - Present

- Led the design of an NLP system which generates questions and answers automatically from Wiki articles
- Applied NLP techniques such as lemmatization, POS tagging, and named entity recognition to boost answering system accuracy to 90%

Learning Pattern Data Mining with Website User Log (Language: Python, JavaScript)

Nov 2018 – May 2019

- Utilized Linux commands and Python to access, clean and reformat raw server data
- Analyzed user log data through pattern recognition and clustering to evaluate project main features
- Designed and implemented a system with d3.js to visualize learning pattern differences between novice and professional users

Predicting House Prices with Machine Learning (Language: Python)

May 2018

- Developed a machine learning model with Python to predict house prices in Ames, Iowa
- Applied statistic analytical skills for data cleansing and feature engineering
- Placed in top 15% on the worldwide leaderboard for the final result

E-commerce Web Scraper (Language: Python)

March 2018

- Created a web scraper with Python (BeautifulSoup) to collect all Samsung phone sales information on a website
- Self-learned and implemented web driver with Selenium to trigger and collect JavaScript hidden information