

AIQI CUI

(814) 882-9127 | ac4788@columbia.edu | acui34.github.io

EDUCATION

Columbia University

Candidate for M.S. in Data Science

Coursework: Applied Deep Learning, Cloud Computing & Big Data, Personalization

New York, NY

Expected Dec 2021

Carnegie Mellon University (University Honor, GPA: 3.8/4.0)

B.S. in Information Systems, Secondary Major in Statistics & Machine Learning

Coursework: Machine Learning, Natural Language Processing, Data Mining, Advanced Data Analysis, Statistical Inference, Database, Data Visualization

Pittsburgh, PA

May 2020

SKILLS

Data Science Tools: SQL/SQLite/PostgreSQL, D3, Spark, Hadoop, MapReduce, Tableau, Excel Pivot Tables

Programming Tools: Python (NumPy, Scikit-Learn, Pandas, Keras/TensorFlow, NLTK, Matplotlib, ggplot), R, C

Software Development: React, Ruby on Rails, Swift, HTML, CSS, TypeScript, Git, A/B Testing, AWS

RELEVANT PROJECTS

Learning Pattern Mining and Visualization with Website Log (*Submitted for Initial Review*) Nov 2019 - Aug 2020

- Automated and managed extraction, cleaning, and reformation of server data with Linux commands and Python
- Analyzed 10-day time-series log data through pattern recognition and clustering to evaluate project main features
- Designed and implemented a dynamic system with D3.js to visualize learning pattern differences between novice and professional users
- Delivered under tight deadline on converting data insights into new features of an iOS app

NLP Auto Question Generation & Answering System

Feb 2020 - Present

- Built an NLP system to generate questions and answers automatically from Wiki articles as a team leader
- Applied techniques such as lemmatization, POS tagging, and named entity recognition to boost answering system accuracy from 50% baseline to 90%

House Prices Modeling and Prediction (*Kaggle Competition*)

May 2018

- Tested multiple supervised models such as Random Forest, Lasso Regression and Gradient Boosting Machine and performed hyper-parameter tuning to predict house prices in Ames, Iowa
- Placed in top 15% on worldwide leaderboard for final result

Cultural Difference Analysis on YouTube Trending Videos

April 2018

- Analyzed trending video data via Pandas and derived area-specific business development insights for youtubers
- Presented research results with innovative graphics (e.g. cloud map, heat map) with Tableau and Seaborn

WORK EXPERIENCE

Handshake

San Francisco, CA

Software Engineering Intern

May 2019 - Aug 2019

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

Microsoft

Guangzhou, CN

Software Engineering Intern (Quantitative)

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 3,000+ utterance data to support 95% business FAQs
- Analyzed weekly user responses via Excel Pivot Tables and Tableau to assess major functional changes