

Yanxi Li

Tel: 5714001652 Email: yanxi.li@gwu.edu
Address: 2803 Poag St, Alexandria, VA 22303

EDUCATION

George Washington University, Washington D.C., US 08/2023—06/2025

Master of Science in Computer Science (in progress)

- **Core Coursework:** Machine Learning, Neural Networks and Deep Learn, Design & Analysis of Algorithm

Chengdu University, Chengdu, China 09/2017—06/2021

Bachelor of Engineering in Computer Science and Technology received in June 2021

- **Core Coursework:** Data Structures, Cloud Computing, Database Principles, Computer Network, Computer Operating System Principles, Principles of Computer Composition
- **Academic Highlight:** University-level third-class scholarship

SKILLS

- **Programming Language:** Python; Java; JavaScript; C/C++
- **Data Processing & ML:** Pytorch; Sklearn; NumPy; Pandas; ElementTree; BS4; Hadoop; Spark
- **Database:** MongoDB; MySQL; SQLite; SQLServer
- **Operation System:** Ubuntu 18; Centos; Windows 10; Windows 11

WORK EXPERIENCE

Data Engineer; Chongqing Vip Technology Co., LTD 06/2022--06/2023

- Developed data collection processes for **resource integration systems**.
- Utilized **Python** and Python's **ElementTree** as a web crawler to gather data from various websites, encompassing academic journals, video, books, news, magazines, dissertations, reports and more.
- Implemented **multithreaded tasking** to enable concurrent crawling from different web pages.
- Deployed the crawler programs onto the distributed planning platform known as **AirFlow**.
- Devised an Extract, Transform, Load (ETL) procedure to cleanse and standardize data from various sources.
- Stored the refined data in **MongoDB** and generated sample data using **SQLite**.
- Facilitated data processing and transfer between **MongoDB** and **SQLite** as needed.
- Designed programs for calculating **similarity** between keywords and related data in company data pools.
- Based on **similarity degree**, initiated actions such as blocking or activating corresponding data in data pools.
- Leveraged Python libraries, including **Pandas** and **NumPy** to carry out data processing.

Database System Maintenance Intern; Sichuan Suojie Technology Co. LTD 09/2021--03/2022

- Designed and implemented **MySQL** databases to support efficient data querying for software systems.
- Developed **SQL** scripts to extract and export relevant data, and to remove invalid data from databases.
- Managed database backup procedures using **SQL**, securely storing backup files in designated directories.
- Created **Shell** scripts, scheduling regular intervals for critical log and data file backups.

Parking Lot Path Planning System University Project; Chengdu University 11/2020--04/2021

- Designed and implemented a parking lot management system with path planning functionality.
- Designed an interactive visual interface using Qt, implemented core logic in C++, and incorporated 3 user roles: super administrator, administrator, and car owner, with a server and multiple clients.
- In 10 parking scenarios, each with 300 randomly configured spaces, 5 simulated car owners simultaneously entered through 4 entrances. While maintaining constant speed control, owners traveling took 39.7% less time to find an available parking space when utilizing the system compared to when they didn't use it.

MeowNet: CatSound Classification University Project; George Washington University 02/2024--05/2024

- Created a dataset featuring various cat calls and their associated meanings, and then trained 3 neural networks to recognize cat language based on this dataset and compare their performance.
- **Paper Link:** <https://raphael.quarto.pub/meownetwebpage/>
- **GitHub Link:** https://github.com/Yanxi1024/Classification_Tasks