

Yichen Liu

ycliuleo0601@gmail.com · [Github](#) · [LinkedIn](#)

EDUCATION

Georgetown University

Master of Science: Data Science

Washington, District of Columbia

Aug 2022 - Present

Virginia Tech

Bachelor of Science: Computer Engineering

Focus: Control, Robotics, Autonomy and Networking and Cybersecurity

Blacksburg, Virginia

Aug 2018 - May 2022

SKILLS

Languages - C++, Python, R, SQL, BASH, Matlab, C, Verilog, VHDL, MIPS Assembly

Tools - Azure, AWS, Cmake, MongoDB, Quarto, Git, Linux, VMware ESXi, Arduino, GCP, Tableau

Frameworks - Pytorch, Keras, Scikit learn, OpenCV, NLTK, NumPy, Pandas, Tensorflow, plotly, altair

Selected Coursework - **Data Structure and Algorithm**, **Apply Software Design**, Fundamentals Digital Systems, Embedded Systems, **Digital Image Processing**, Telecommunication Networks, Artificial Intelligence and Engineering Applications, Computer and Networking Security Fundation, Network Application Design, **Probabilistic Modeling and Statistical Computing**, **Natural Language Processing**, **Statistical Learning**, **Time Series**, **Advanced Data Visualization**, **Big Data and Cloud Computing**

EXPERIENCE

Hipond Technology

Backend Developer - Part Time

June 2023 - Present

Remote

- Design and implement the back-end of a software application utilizing Python, Flask framework, and MongoDB database.
- Create and integrate modules for posting articles, conducting article searches, user management, and user login functionality.
- Foster collaboration with front-end developers to analyze and comprehend functional requirements.
- Develop a robust database architecture and optimize data accessibility to enhance overall system performance.

Georgetown University

Teaching Assistant - Computational Linguistic and

Advance Python

Aug 2023-Present

Washington, District of Columbia

PROJECTS

[A Time-Series Analysis of World Tourism Industry](#)

January 2023 - Present

- **Data collected** from Our World in Data website, including International tourist arrivals by region, GDP from tourism as a share of total GDP, etc.
- Perform exportary data analysis and visualization using **ARMA/ARIMA/SARIMA with Garch model, and deep learning such as CNN, GRU, and LSTM**.
- Final goal is to **forecast** the developmentt tourism industry. Find which region will be the popular region for tourism in the next decade, and what is the trend of global tourism industry.
- Visualize the data using plotly, ggplot, and Tableau.

[A Survery on Different Text Summarization Models](#)

August 2022 - December 2022

- Using different text summarization models, including TextRank, GPT-2, LSTM, and attention mechanism. Evaluating model performance using ROUGE and F1 scores. Training and fine-tuning the models on the CNN/DailyMail dataset.
- TextRank achieved a RougeL score of 0.29, GPT2 achieved a RougeL score of 0.16, LSTM+attention achieved a RougeL score of 0.09. Due to the calculation method of RougeL scores, generative models generally have lower scores compared to extractive models.

[Study About Development of Esports and Soccer](#)

August 2022 - December 2022

[Markets in America](#)

- **Collected social media data** using **Twitter and Reddit APIs**. Collected data from specific professional data website, such as, www.transfermarkt.co.uk, leagueoflegends.fandom.com, and general data from websites such as www.ourworldindata.org
- **Python and R were used to clean the data and dataframes were be created** for easier access.
- Machine learning strategies were applied to explor the data, such as **Naive Bayes, Decision Tree, Support Vector Machine, Clustering, etc.**
- Plots, wordclouds would be used to **visualize** the data.

A lisp compiler using C++

January 2022 - May 2022

- Coded a lisp compiler using **C++**. Input lisp formatted code, the program will parse the code and run the code, support simple math operation, pointer, for loop, if statement, while loop.
- A GUI is created using **Qt**. GUI allows the user to enter math expression and the program will return simple line diagram of it.

NASA Vertical Solar Array Technology (VSAT)

August 2021- May 2022

- Developed a hardware design, **control algorithm**, and working prototype for a Mast Assembly Controller capable of deploying and actively controlling the VSAT assembly.
- Leveraged Matlab **Simulink and Python** to develop the state flow model and program for the controlling algorithm.
- Converted the control signal from the MCU to control brushless motors using a digital-analog converter (DAC), as well as send feedback data from the motor back to the MCU using an analog-digital converter (ADC).
- Used Speedgoat system and NASA provided model to **simulate and optimize** the system.

Cuisine Recommendation System

October 2021 - December 2021

- Used two Raspberry Pis. One was a client machine, which was used to collect data from the peripheral, and hosted the web GUI to interact with the users. The other one was a server machine, which was used to collect and store the data online. The two Raspberry Pis would communicate via **RESTful** interactions.
- Used a GPS module to locate the current location, and an API from Yelp to search for certain cuisines.
- Developed a **web GUI** using Flask and HTML for the user to interact with the system.
- Used a **noSQL** database to record all the search histories as cache so that the system could recommend cuisines based on previous searches.

Smart Home System

August 2020 - December 2020

- Used different **sensors**, such as humidity sensors and thermometers to monitor the room condition so that the system could automatically control peripherals, such as air-conditioners and dehumidifiers.
- Used Arduino as the platform to **collect the data** from the sensors and host all the other code.
- Developed an **Android application** to help the user to switch between manual and automated control mode. A Google Assistant SDK was used to enable voice control on the Android application.
- A **noSQL** database was created to store log files, which recorded important data and timestamps.