Haoran Zhao

(445) 208 8781 | 3737 Lancaster Ave, Philadelphia, PA, 19104, USA
haoran.harry.zhao@outlook.com| haoranzhao.com

Motivated and passionate student with solid academic skills and extensive project experience

EDUCATION

Drexel University | Philadelphia, Pennsylvania September 2021- June 2023
Bachelor of Science in Data Science (cooperative education program with Lanzhou University)
GPA: 4.0/4.0
Relevant coursework: Advance Programming Techniques, Algorithms and Analysis, Deep Learning

Lanzhou University | Lanzhou, China September 2019- June 2023
Bachelor of Engineering in Computer Science and Technology (Data science direction)
GPA: 89.89/100
Relevant coursework: Machine Learning, Information Visualization, Data Structures

INSTERNSHIP EXPERIENCE

Vivid Seats Inc. | *Data Engineering Co-Op* Sep 2022 – Mar 2023
- Participated in the design, development, and enhancement of the company's data platform.
- Processed data requests submitted by different departments of the company.
- Learned to use SQL & Python programming languages as well as tools such as Snowflake, AWS, DataDog, and AirFlow DAG.
- Participated in developing innovative solutions that gave Vivid Seats Inc. a competitive advantage and technology improvements on the data engineering team.
- Helped to tackle challenges such as building out new data migration pipelines and optimizing warehouse and transaction queries.

Hangzhou Yunge Data Technology Co., Ltd | *Student intern* Aug 2019 – Sep 2019
- Assisted the project team members to carry out development and maintenance related works for the Qilu traffic big data management platform.
- Participated in project code testing and data maintenance related works.
- Participated in the development of an accident emergency rescue module and used the Dijkstra algorithm to find the nearest emergency car repair points on the highways.

RESEARCH AND PROJECT EXPERIENCE

Drexel University Metadata Research Center | *Research Assistant, supervised by Professor Jane Greenberg* Jan 2022 – June 2023
- Took charge of building the Pipeline, looking up & collecting the literature documents, and data processing.
- Took charge of constructing the neural learning model and named-entity recognition relevant tasks.
- Managed the acquired academic articles by using Python and processed the data into machine-readable form.
- Maintained the HIVE4MAT system and solved the database & front-end bugs, etc.

Material Literatures NER Based on BiLSTM-CNN-CRF | *Team Leader* April 2022
- Took charge of building the BILSTM-CNN-CRF model which was used for naming the entity recognition in materials science literature and making comparisons with the Bert-Base model.

Traffic Prediction for New York City Based on Graph Neural Network | *Team Leader* April 2022
- Completed the Data Science Project course assignment on traffic prediction in New York City by using graphic neural network and Uber movement data.

Movie Recommendation Based on Tweets | *Team Leader* April 2022
- Completed the Social Media Analysis course assignment by designing and building a movie recommendation & box office earnings prediction system.

Anime Content-Based Recommendation System | *Team Leader* April 2022
- Completed the Recommender System course assignment by designing and building a content-based anime recommendation system.

Google Open Source Blockly Teaching Cases Systems | *Team Leader, supervised by Prof. Changyan Di* Mar 2021 – Jan 2022
- Participate in the project design and group discussions.
- Determined the four curriculum data systems of the project, including epidemic transmission simulation, butterfly effect simulation, visual simulation of sorting in computers, and simulation of different game strategies.
- Learned and used mathematical modeling thinking to design the cases' database.
- Simulated the phenomena in life by combining mathematics and computer simulation methods.

Air Quality Prediction of Lanzhou Based on LSTM | *Team Leader, supervised by Prof. Binbin Yong* Mar 2021 – June 2021

- Took charge of organizing group discussions, determining the research topics, and assigning tasks to the team members in the early stage.
- Took charge of proposing and organizing the works related to the implementation of the project and controlling the project schedule in the middle stage.
- Adopted the CV method for repeated samplings and compared the prediction model with the traditional RNN model.
- Recorded the experimental data and completed the experimental report.

A New Type Box Drawing Design Based on Matplotlib | *Project Leader, supervised by Prof. Rui Zhou* Mar 2021 – June 2021

- Participated in the construction of the power boxplot module, drew new boxplot type methods (info boxplot, hist boxplot, and creative boxplot), improved the shortcomings of the existing methods, and built the module.
- Learned to use the Matplotlib package knowledge and skills.

Visual Analysis and Improvement of Post-secondary Education in Hong Kong Mar 2021 – June 2021

Team Member, supervised by Prof. Rui Zhou

- Participated in the study of the report on Post-secondary Education in Hong Kong published in China Daily, using graphic analysis to study the trends of youth participation in four post-secondary education programs in Hong Kong from 2000 to 2017.
- Worked with other team members to study the advantages and disadvantages of information visualization, its impacts on society, as well as the existing problems and solutions, etc.
- Learned to use Matplotlib to visualize and comprehensively copy the original information, to improve the theories and principles of information visualization through critical thinking.

Programmer Evaluation Model Based on Linux-stable Operational Data Mar 2020 – June 2020

Project Leader, supervised by Prof. Nicholas Mc Guire

- Participate in the establishment of the nonlinear regression model and used the nonlinear regression model to represent the individual ability, which was set as the weight value.
- Used people status in V4 to generate the weight value tables while using the existing data to validate and search for our assumptions.
- Utilized the time zones to judge the programmer level and code level and calculated the statistical values to validate our assumptions.
- Took charge of modeling the information that we collected.

PUBLICATIONS AND SOFTWARE DESIGN

- Greenberg, J., McClellan, S., Zhao, X., Kellner, E. J., Venator, D., **Zhao, H**., Shen, J., Hu, X., & An, Y. (2022). Materials Science Ontology Design with an Analytico-Synthetic Facet Analysis Framework. arXiv. http://arxiv.org/abs/2211.10407
- **Zhao, H**., Li, Z., & Xu, S. (2021). Computer dynamic model and time series prediction of air by LSTM recurrent neural network. Journal of Physics: Conference Series, 2033(1), 012085. https://doi.org/10.1088/1742-6596/2033/1/012085
- **Taodu Mountain Agricultural Information Visualization Display Platform** Sep 2019 – Nov 2020
  Utilized Python Flask framework to build the platform, utilized crawler to collect the data, and finally visualized the information by using the Google Chart.
- **Simple SCM Car Motion Control Software** April 2020 – Oct 2020
  Utilized Python to achieve car control.
- **Ginker College Students Attendance Check-in App** April 2019 – May 2019
  Programmed for the WeChat mini-program using WXML and PHP, etc.
- **MOWO Desktop Phase of the Moon Simulation Software** April 2019 – May 2019
  Utilized Python to achieve back-end operations and utilized C# to achieve front-end page display.

AWARDS AND HONORS

A. J. Drexel Scholarship awarded by Drexel University	May 2021 – May 2022
Successful Participant awarded by the 2021 Mathematical Contest in Modeling	Feb 2021
Third Prize Scholarship awarded by Lanzhou University	Sep 2020 – Sep 2021
Second Prize from the Mathematical Contest in Modeling awarded by Lanzhou University	Sep 2020
Second Prize Scholarship awarded by Lanzhou University	Sep 2019

COMPUTER SKILLS

- **Programming languages:** Python, C, SQL, HTML, R, JavaScript, Racket
- **Software tools:** AWS, Airflow, Snowflake, DataDog, Git, Postgres, PyCharm, WebStorm, Apache Hadoop, Xcode, Weka, DrRacket, Gephi, Jupyter Notebook
- **Frameworks:** Pandas, NumPy, Flask, TensorFlow, PyTorch, Scikit-learn, Beautiful Soup