XIANGHAO CHEN

(517)-215-5089 · xianghc@amazon.com

EDUCATION

University of Michigan, Ann Arbor   *February 2021*
M.S. in Data Science (with Graduate Certificate in Computational Discovery & Engineering)
Coursework: Parallel Computing, Applied GPU Programming, Database Management Systems,
Web Systems, Machine Learning, Natural Language Processing, Data Structures, Algorithms
Overall GPA: 3.98

Tsinghua University   *July 2018*
B.Eng. in Environmental Engineering

TECHNICAL SKILLS

Computer Languages	C++, C#, Python, SQL, Bash, Javascript(React)
Libraries	CUDA, MPI, Keras, PyTorch, NumPy, Matplotlib, OpenCV

EXPERIENCE

Amazon.com, Inc.   March 2021 - Present
Software Development Engineer   *Seattle, WA*

· Helped build the data service for distribution nodes with **DynamoDB** and **ECS/Fargate**.

Futurewei Technologies, Inc.   May - September 2020
Database Kernel, Intern   *Bellevue, WA (Remote)*

· Enabled optimization for **multi-threaded** execution of distributed plan and introduced more flexibility
  to physic plan generation by implementing property enforcement for the Cascades **query optimizer**.
· Improved the performance of **cardinality estimation** to match that of PostgreSQL by rewriting column
  statistics collection and adding analyzer for filter logic.
· Contributed to the **open source C#** database prototype project by resolving TPC-H execution issue.

University of Michigan, Dept. of Bioinformatics   April 2019 - March 2020
Research Associate I   *Ann Arbor, MI*

· Built a prediction model by integrating different **machine learning** methods on EC2 instances and
  created new feature sets to increase performance. The model can predict sepsis onset probability with
  AUC of over 0.95. The work won DII National Data Science Challenge Task 1.
· Achieved denoising of microarray data through building convolutional neural network with **PyTorch**.

Ann Arbor Algorithms, Inc.   October 2018 - April 2019
Software Engineer   *Ann Arbor, MI*

· Implemented a 2D U-Net architecture **deep learning** model using **Keras** with Tensorflow backend to
  extract lesions in biomedical images and achieved performance of 0.91 dice score.
· Designed a 3D model for segmentation and classification of artery in CT scan, and developed geometric
  algorithm to refine coordinates of the centerline due to the challenge of identifying small tubular volume.
· Wrapped **C++** module into **Python** script for more efficient data preprocessing.
· Improved user interface of **Django**-based annotation system by adding data visualization features.

AWARDS

First Place of DII National Data Science Challenge	*October 2019*
First Prize in 35th Tsinghua "Challenge Cup" Technology Research Competition	*April 2017*
Second Prize in National Olympiad in Informatics in Provinces (NOIP)	*December 2013*