Hao Li

857-891-9461 | Boston, MA | li.hao6@husky.neu.edu | 🔗leot5.com 🔗 🔗

EDUCATION

Northeastern University, Candidate for M.S. in Data Analytics Engineering, Boston, MA   Expected Dec. 2020
- Relevant Courses: Big-Data System Engineering Using Scala, Machine Learning, Computer Networking

Hangzhou Normal University, M.S. in Computer Science, Hangzhou, China   June 2018
- Relevant Courses: Database, Operating Systems, Artificial Intelligence and Pattern Recognition

Xidian University, B.S. in Mathematics and Applied Mathematics, Xi'an, China   July 2015

SKILLS

- Programming Languages: Python, SQL, R, Java, Scala, C, JavaScript, HTML, CSS, MATLAB, bash
- Frameworks/Tools: Tensorflow, Spark, Hadoop MapReduce, MySQL, Nginx, Redis, Scrapy, Linux, Git

PROFESSIONAL EXPERIENCE

Lenovo   Beijing, China
Software Engineer Intern   May - August 2019
- Provided data for visualization modules by designing 29 Views and 51 Tables and executing using SQL
- Connected the client and the database with JDBC and solidified the database operation statement in the program
- Implemented Java Scheduled Task to update the data in the database to update the chart display of the front-end visual display board based on Spring Task Scheduler

Alibaba Research Center for Complexity Sciences   Hangzhou, China
Research Intern   May - Dec. 2017
- Established Python-based distributed web crawler engine by combining Scrapy with BeautifulSoup and Requests in Python and collected information of more than three million users
- Designed MySQL database to manage 2GB data, generated six relational tables under 3NF
- Produced a true diversity measurement by Python to better capture diversity of users' musical tastes
- Discovered factors (levels of economic development and so on) that greatly impact users' music tastes, paper accepted by International Journal of Modern Physics C as first author

Orimuse   Xi'an, China
Systems Engineer Intern   March - May 2014
- Developed a T-shirt customizing mobile application which has been thrived for 5 years
- Designed the database schema and split data into three shards by using horizontal partitioning
- Built MySQL Master/Slave replication, copied data on slaves and achieved Read/Write splitting
- Utilized Nginx as a load balancer to split system horizontally to run multi-backend

PROJECTS EXPERIENCE

Doge Invasion Game   March - May 2019
- Developed a 2D games in which the aim is to shoot down a fleet of aliens as they drop down the screen in levels that increase in speed and difficulty by using completely Object-oriented approach and Pygame

Movie Recommendation System   Feb. - March 2019
- Predicted ratings for movies with Tensorflow Content-based filtering and Spark Collaborative filtering, then recommended movies to users

Adult Income Prediction   Nov. - Dec. 2018
- Implemented Classification tree, Boosting tree, Random forests and Neural network by Python to achieve our goal in classifying whether an individual's annual salary is more than $50,000
- Estimated imbalanced data with synthetic sampling technique, SMOTE, by Python and improved accuracy of all models to 86% and higher

Raw Sockets *HTTPGET* Application   Oct. - Nov. 2018
- Created a Socket program in Python that takes a URL on the command line and downloads the associated file and saves it to the current directory
- Rebuilt the system's TCP/IP stack in socket, implemented all features of IP/TCP packets