Lihua (Neo) Pei

• Email: lihua.peidata@gmail.com • Phone: +1 847-848-5805 • Portfolio of Projects: https://lihuapeineo.github.io

OBJECT

Data Scientist jobs: I am looking for the Data Scienist jobs.
Legal Worker: I am the US Permanent Resident.

EDUCATION

The George Washington University <div align="right">WASHINGTON, DC</div>
Master Degree in Data Analytics (Computer Science Track) *Expected May. 2020*

Stony Brook University <div align="right">STONY BROOK, NY</div>
Bachelor Degree in Applied Mathematic & Statistics (AMS) *Dec 2017*

SKILLS

Computer Skills:
- **Coding**: Python, SQL, R, Java.
- **Front-end**: HTML/CSS, JavaScript.
- **Back-end**: MySQL, MongoDB, Linux.

Data Analysis Technologies:
- Skilled in the Machine Learning, Data Visualization, Nature Language Processing and Database Management.

EXPERIENCE

Data Scientist Intern - Database Developer <div align="right">WASHINGTON DC</div>
Fresh Air DC Project *May 2019 - Present*
- **Data Mining**: Improve data mining processes, resulting in a 90% decrease in time cost to preprocess data by using Python.
- **Co-location Analysis**: Conduct the Time Series Analysis and the Correlation Analysis to visualize the data comparison between uRAD monitors' data and the Department of Energy and Environment's data in order to evaluate uRAD monitors for future productions improvement.
- **Database**: Develop the MySQL database collecting uRAD monitors' data updating every 30 seconds at the George Washington University's cloud servers.

High-Performance Computing Intern - Research Assistant <div align="right">JINAN, CHINA</div>
National Supercomputer Center in Jinan *May. 2018– July. 2018*
- Developed the clusters' monitor system to inspect the CPU temperature and the spin speed of fans for avoiding the bias of experimental results by CPU downgrading frequency.
- Assisted to design clusters' experiments and collected experimental datasets.

Database Developer <div align="right">HONG KONG, CHINA</div>
Fire Pillar Studio *June. 2017 – May. 2018*
- **Deep-Learning**: Assisted to develop a deep-learning program MIB (make it breath) which can auto recognize the cartoon characters in pictures and make characters have the action like breathing.
- **Database**: Developed and maintaining the local MySQL databases for the project which stored over 20,000 original user updated pictures and processed data.

Optimal Topologies Searching Research Project <div align="right">STONY BROOK UNIVERSITY</div>
Research Assistant Guided by Prof. Yuefan Deng *Aug. 2017 - March. 2018*
- **Algorithm**: Designed and implemented an algorithm to generate and search the optimal topologies with 64 nodes and degree 3. Successfully found the optimal 64k3 topologies by applying the algorithm on High-performance Computer.
- **Database**: Built and maintained the local database to store over 10 million generated topologies for research purpose.
- **Data Analysis**: Collected the performance data of new topologies and visualized experimental results. The 64K3 optimal topology achieved 150% to 320% better efficiency than commonly used ring topologies.

Data Analytics Projects

Donald Trump's Twitters Analysis <div align="right">THE GEORGE WASHINGTON UNIVERSITY</div>
Topic Modeling *Sept. 2018 – Dec. 2018*
- Collected over thirty thousands twitters of realDonaldTrump by Twitter's API.
- Created an NLP model based on LSI and LDA methods to assigned each Trump's twitter with a matched topic.

Stocks Investment Recommendation System <div align="right">THE GEORGE WASHINGTON UNIVERSITY</div>
SVM (Support Vector Machine) *May. 2018 – Oct. 2018*
- Collected data of top 500 stocks with 18 significant features crawling from Yahoo Finance websites.
- Created a machine learning model based on SVM to predict the market trend and stocks' price. Achieved average 78% accuracy.
- Created an investment strategy recommendation model for the maximum profit.