Yichen Li

Ann Arbor, MI 48105
patronusd366_shc@indeedemail.com
+1 734 881 8107

• Currently studying at the University of Michigan, Ann Arbor, majoring in Data Science and Machine Learning.
• Interested and experienced in Data Analysis and Engineering. I am good at programming using Python, Java, SQL, and C++, distributed system frameworks as Hadoop, Hive, and Spark. I have done projects and researches using Python, Matlab, and SQL; During internships, I have utilized distributed systems and Java to make my contributions, and I am currently learning C++ and improving my SQL and Java skills.

Willing to relocate: Anywhere

Work Experience

Machine Learning Software Engineer Intern

Intel Corporation - Shanghai
March 2021 to August 2021

Duty: Manage server cluster; Program for the end-to-end machine learning pipeline project in Python; Optimizing distributed system and big data frame based on benchmark tools
• Build server cluster with several nodes based on CentOS 7. Configure distributed framework environment.
• Based on TPCx-BB and TPCx-DS, optimize big data frameworks like Hive, Spark, yarn. Participate in the full-stack optimization of some machine learning and big data projects oriented to customer needs.
• Program for machine learning pipeline project implementing Numpy, Pandas, TensorFlow, and Modin, making the performance of the ETL part 10 times better.

Software Engineer Intern

GrapeCity Inc. - Xi'an, Shaanxi
August 2020 to December 2020

Duty: Test and help develop Document for Excel, one of GrapeCity’s flagship software products
• Test more than 20 features, including support Json I/O with options, cellpadding and labels, and PDF export, etc. And some features interact with other products like SpreadJS and Document for Word
• Help develop cellpadding feature and Calculation Engine, program some APIs for basic package utilized in the releasing version, improve their accuracy and regularity
• Together with the team, we released a version V3.2.4 and a main version V4.0.

Undergraduate Research Assistant

Fuel Cell Research Center of Huazhong University of Science and Technology - Wuhan, Hubei
March 2019 to June 2020

Statement: A patent based on this method is being reviewed; Provincial college student innovation and entrepreneurship project

Objectives: Combining two main fields of SOFC (Solid Oxide Fuel Cell) health prediction and fault detection, data-based and model-based, to reach a better timely manner and better accuracy than any single method.
• Build a SOFC mathematical model with knowledge of physics and chemistry.
• The PSO (Particle Swarm Optimization algorithm) is applied to optimize coefficients. And transfer the model to state space.
• Granger causality test is performed on each variable in the actual system operating data to obtain the causality relationship of each variable in the actual situation, resulting in improving the model performance by around 8%.
• Explored, analyzed, and deduced hidden relationships among system data with the help of pattern recognition and big data prediction obtained from historical data

Education

Minor in Data Science
University of Michigan Ann Arbor - Ann Arbor, MI
September 2021 to Present

Master's degree in Data Science and Machine Learning
University of Michigan-Ann Arbor - Ann Arbor, MI
January 2021 to Present

Bachelor's degree in Artificial Intelligence and Automation
Huazhong University of Science and Technology - Wuhan, Hubei
September 2016 to June 2020

Summer Program in Information Science and Engineering
University of Illinois at Champaign-Urbana - Champaign, IL
July 2017 to August 2017

Skills

- Python
- SQL
- Java
- C#
- C++
- Jira
- Git
- Hadoop
- Apache Hive
- Spark
- Machine learning
- Big data
- Data management

- MySQL
- AWS
- Data Science
- Linux

Languages

- English - Fluent

Links

https://www.linkedin.com/in/yichen-li-7178041b0

Additional Information

I would need visa sponsors, CPT, and in the future, H1B to work legally in the United States.