

ZHENG MA

Email: zheng.ma@duke.edu | Tel: 919-606-1611 | LinkedIn: <https://www.linkedin.com/in/zanecode>

Github: <https://github.com/zanemarkson/> | Website: <http://zanecode.io>

Objective

Data science enthusiast and PhD (expected to graduate 2017) in computational chemistry seeking for data scientist position to leverage my knowledge of machine learning/data mining and my experiences of data analysis using big data technologies.

Skill Set

C, Bash, Python, Scala, JavaScript/ES 2015, Java, MATLAB/Octave, Markdown;
Machine Learning/Deep Learning, Hadoop/YARN, Apache Spark, TensorFlow, SLURM, MongoDB, MPI;
Node.js, AngularJS, Express.js, socket.io, React/JSX, Django, Twitter Bootstrap, HTML5, CSS3, SASS/SCSS;
Linux, Git, LaTeX, Github Electron, Quantum Chemistry Simulation;

Education

PhD in Computational Chemistry, March 2017 (expected), DUKE UNIVERSITY – DURHAM, NC

Bachelor of Science in Chemical Physics, July 2010, UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA (USTC)

Projects & Experiences

Data Analysis & Management

- Managing cloud-based SQL database for e-commerce company ApolloBox (www.theapollobox.com) inventory and order system;
- Building a product recommendation system for ApolloBox using Apache Spark Streaming and MLlib;
- Set up and managed a sharded NoSQL database to store, query and backup TB scale of data from research;
- Performed data analysis/mining using distributed computing platform such as SLURM and Hadoop/YARN;
- Facilitated data analysis by integrating different simulation steps into an automated workflow;

Front/Back-End Web Development at ApolloBox

- Creating responsive and interactive front-end UI as well as administrative tools for marketing team;
- Implementing a back-end inventory managing system for automatic order/inventory processing with Node.js and Express.js;
- Building next generation of ApolloBox website using python Django, React and PostgreSQL.

Scientific Computing & Software Engineering

- Developed NEqMD-ToolKit (<http://bit.ly/2c8uX47>), a C/Fortran-based comprehensive Unix/Linux command-line tool set, to facilitate scientific research and data analysis in computational chemistry;
- Prepared lib-NEqMD, a C style library/API to assist further NEqMD-ToolKit development;
- Implemented an MPI wrapper to enable parallel computing features for NEqMD-ToolKit;

Miscellaneous

- Developed an interactive drawing board (<http://bit.ly/2ctPS0g>) web/desktop application using MEAN stack, Twitter Bootstrap and Github Electron;
- Developed a remote control plugin for reveal.js presentation framework and provided API;