Shuo Wang, Ph.D.

shuo@shuo-wang.com :: xxx-xxx-xxxx

Professional Summary

I solve business questions and deliver value by combining data science principles, software engineering, and computing technology to ask/answer the right questions, implement effective and scalable solutions, and execute on priorities.

Technical Skillset

Machine Learning:	Ensemble learning, bootstrap aggregation, weka classification, neural networks, k-means, random forest, GLM, GAM
Data Science:	Jupyterhub, Jupyter, Zeppelin
Software Engineering:	R, Python, Agile methodologies, RESTful APIs, git/svn, bash, Matlab, C, Fortran, Spark, DistributedR
Data Visualization:	ggplot, PowerBI, Tableau, Avizo, Amira, Drshti, Meshlab, CTAn/CTVol
Databases:	Vertica, SQL, Oracle, MySQL
Cloud Computing:	Amazon Web Services (AWS), Microsoft's Azure ML, HPC
Operating Systems:	Windows, Linux, OS X

Experience

Carlson Wagonlit Travel *Jan 2018-Present*

Principal Data Scientist

- Developing metrics and ML models centered around Fair Market Share expectations for airlines in a big data environment.
- Investigating and implementing the technology stack for data science production deployments

Trane Intelligent Services *Oct 2014-Dec 2017*

Data Alchemist/Software Engineer

- Led the development of data science and machine learning understanding both in terms of general knowledge and the applicability to our domain. This included:
    - Understanding fundamental principles of data science such as feature engineering, data splitting, cross validation, bootstrap aggregation, and ensemble methods.
    - Mentoring two undergraduate interns applying various ML models, e.g. random forests, neural networks, SVMs, in Azure ML in terms of scoring and performance. Lead to the development of domain specific error metrics that better met SME requirements vs standard metrics (RMSE, MAE, MAPE).
    - Developed a weather based random forest model for building level energy consumptions at 15 minute predictions with 10% error and confidence bands. This model was applicable to many types of building – office space, theaters, hospitals, hotels, etc.
- Designed, tested, and implemented R-based analytic backend for commercial HVAC fault detection algorithms designed by SMEs. Data flow leveraged existing REST-based APIs implemented by Software Engineering team for a highly scalable cloud-based solution.
- Leveraged R's Shiny package to develop interactive, UI-based framework for testing and validating new analytics. Using this approach reduced development time from weeks/months to days. Approved prototype code could be added directly to the Engine for rapid deployment.
- Developed a robust understanding of end-to-end data flow from equipment to XML to Database to UI using available methods including SQL queries, parsing XML feeds stored on S3, manual

execution of API calls on AWS EC2 via curl, and sniffing UI api calls using Chrome DevTools. Understanding this was critical to understanding the context of the data.
- Active participant in the Minneapolis Analytics and Big Data Communities. Presentation at IoTFuse received a user feedback rating of 9.2/10. 3 presentations at Twin Cities R User's Group Meetups. Technical Judge at MinneMUDAC analytics competition.

Centre for Grain Food Innovation/University of Western Australia
PhD Candidate / Data Scientist / Research Scientist *2010-2014*
Visiting Scientist *2009-2010*

- Developed new value added opportunities for Western Australian wheat via dough processing control through research demonstrating impact of pore structure on quality.
- Built digital twins of breads.
- Obtained AUD$120,000 in funding for Ph.D. research in co-written grant proposal.
- Supervised undergraduate student projects on automated image segmentation and analysis.

Dr. David A. Yuen's Research Group
Minnesota Supercomputing Institute, Minneapolis, MN, USA *2004-2009*
Undergraduate Research Assistant

- Predicted impact of various geophysical phenomena and answered research questions using Supercomputers pertinent to visiting professors, post-doctoral researchers, and PhD Candidates, including:
    - Impact of tsunami wave propagation originating in South China Sea on Chinese and Japanese seaboards. Numerical modeling and visualization of >5GB datasets.
    - Different perturbation conditions in 3-D spherical mantle convection simulation
    - Seismological phenomena at the core mantle boundaries highlighting anomalies possibly leading to the formation of the Appalachian Mountains.
    - Validating the feasibility of PS3's Cell BE engine for Supercomputing (2007) and new CUDA language for nVidia GPU processers as low-cost supercomputers (2009).
    - Submission of geophysical phenomena codes written in Fortran to different supercomputing clusters around the world.

Professional Certifications

Certified Scrum Product Owner (CSPO), Scrum Alliance® (2014)

Education

Doctor of Philosophy (Food Science), University of Western Australia, Perth, Australia (2014)
Bachelor of Arts (Computer Science), University of Minnesota, Minneapolis, USA (2009)

Awards and Honors

- June 2012, Featured at the Institute of Agriculture's Postgraduate Showcase
- July 2010, Grains Research and Development Corporation, Ph.D. Scholarship
- July 2010, Scholarship for International Research Fees, University of Western Australia
- May 2007, Best Poster Paper Award, awarded at ICCS 2007, Beijing China.
- May 2004, President's Distinguished Student Scholarship, University of Minnesota.