Liguang Zhu
Data Scientist

Liguang (Leon) Zhu
leon.l.zhu@gmail.com
+61 433 082 233

Highlights

- Strong skills in machine learning and solution prototyping with Python and Weka.
- Extensive experience in cross-disciplinary research, analytics and predictive modeling.
- Experienced in Hadoop and its ecosystem
- Experienced in high performance computing, grid computing and cloud service such as AWS.
- Experienced in SQL, NoSQL, and data warehousing.

Experience

Jul 2017 - Present, Customer Experience, BigData Telstra
Data Scientist

Specialise in improving customer experience with Telstra call centres.

- Played a key role in design, implementation and deployment of predictive model pipelines in Apache Spark to improve Telstra call centre agents routing priority. Improved agent resolution, sales and retention rates by 2-4 percentage point.
- Improved current customer intent prediction performance by 0.2 AUC.

Mar 2010 - June 2017, Clatyon School of IT, Monash University
Researcher & PhD

Specialise in data science disciplines including machine learning, probabilistic modeling, predictive modeling, and graph theory.

- Played a key role in researching and protyping of a novel feature engineering algorithm: Subsumption Resolution.
- Independently researched various predictive models on highly imbalanced biological data and implemented a tool for enzyme specificity prediction.
- Leading the design and development of a novel protein design pipeline by study high dimensional biological data.
- Designed one of the most stable protein molecules: FN3con.
- Designed a pharmaceutically promising molecule that challenges the current understanding of structure-folding relationship of proteins.

Nov 2016 - Mar 2017, South East Water
Data Scientist Intern

Responsible for exploring statistical and machine learning solutions to risk prediction and budget modeling of the gravity sewerage system.

- Utilised customer retention analysis for system failure modeling.
- Sourced multiple datasets of CCTV inspections, historical renewal programs, and system failure records from data warehouses.
- Designed and implemented a budget optimisation software to maximise ROI.

Sep 2012 - Present, Monash University
Teaching Associate

Responsible for delivering, facilitating, and managing tutorials, discussion, and practical laboratories.

- Advanced algorithms and data structures
- Business intelligence and data warehousing
- Introduction to data science
- Introduction to computer science
- Programming foundations in Python

Sep 2010 - Dec 2010, School of Biological Science, Monash University
Research Assistant

Responsible for the development and deployment of a portable database system on customised data collection devices for ecology research.

Sep 2010 - Dec 2012, Fleet Software & Services Pty Ltd
Freelancer Data Science Consultant

Involved in the design and development of a value prediction system based on historical fleet exchange data.

- Managed data cleaning and predictive modeling.
- Minimised the average prediction error to $600.

Education

2012 - 2017, Monash University
Doctorate of Philosophy Candidature

2008 - 2010, Monash University
Master of Computer Science (Minor Thesis in Data Science)

2003 - 2007, Beijing Jiaotong University, China
Bachelor of Engineering in Computer Science with Honours

References

Provide upon request.

Publications

- B. Porebski, S. Keleher, J. Hollins, A. Nickson, **L. Zhu**, et. al. (2016) *Smoothing a rugged protein folding landscape by sequence-based redesign.* Scientific Report, 6, Art. no. 33958.

- B.T. Porebski, A.A. Nickson, D.E. Hoke, M.R. Hunter, **L. Zhu**, S. McGowan, G.I. Webb, and A.M. Buckle. (2015). *Structural and dynamic properties that govern the stability of an engineered fibronectin type III domain.* Protein Engineering, Design and Selection. 28(3): 67-78. Oxford University Press.

- **L. Zhu**, B.T. Porebski, A.M. Buckle, G.I. Webb. (2013). *A probabilistic approach to In Silico protein design.* QMB E3: Enzyme Engineering and Evolution. Queenstown, New Zealand.

- B.T. Porebski, **L. Zhu**, D.E. Hoke, W. Dai, S. Keleher, N.A. Borg, S.P. Bottomley, G.I.Webb, A.M. Buckle. (2013). *A structural, biophysical and computational investigation of two sequence-based protein engineering methods.* Poster session presented at: The 38th Lorne Conference on Protein Structure and Function. Lorne, Victoria.

- F. Zheng, G.I. Webb, P. Suraweera, and **L. Zhu**. (2012). *Subsumption Resolution: An Efficient and Effective Technique for Semi-Naive Bayesian Learning.* Machine Learning 87(1): 93-125. Springer Netherlands.