SONGPENG ZU

Department of Statistics, Harvard University

(+1) 8572962544 ⋄ songpengzu@g.harvard.edu ⋄ https://beyondpie.github.io

EDUCATION

PhD on Control Science and Engineer	Tsinghua University	*2011.09 - 2017.01*
Bachelor on Biology	Tsinghua University	*2007.09 - 2011.07*

RESEARCH EXPERIENCE

Post-Doctoral Fellow Department of Statistics, Harvard University *2019, 09 - Now*
Advisor: Jun Liu, PhD

- Algorithm design and data analysis for single-cell sequencing and tumor immunology.

PhD Department of Automation, Tsinghua University *2011.07 - 2017.01*
Advisor: Shao Li, PhD
Studied the compound-protein interactions from the machine learning perspective.

- Predicted latent chemogenomic features from drug-target interactions by the global-optimized probabilistic model under EM framework.
- Inferred the binding affinities of compound-protein interactions by the hierarchical Bayesian model under multi-task learning view.
- Analyzed the molecule-based evaluation of clinical drug responses in cancer by ensemble method to solve the high-dimensional small and imbalanced data.

Visiting Scholar Department of Statistics, Harvard University *2014.03 - 2014.09*
Advisor: Jun Liu, PhD

- Applied the Bayesian Nonparametric algorithm via sliced inverse model to detect the non-linear relationships on cis-eQTLs.

Undergraduate research training School of Life Science, Tsinghua University *2009 - 2010*
Advisor: Li Yu, PhD

- Molecular mechanisms about cell autophagy.

FUNDING

- Harvard Data Science Initiative Postdoctoral Fellow Research Fund
  Principle Investigator: Songpeng Zu $6,933 *2020.03 - 2021.02*
  Title: Learning Peptide-specific T Cell Receptors in Human Cancers by Deep Neural Network and Structural Modeling

PUBLICATIONS

- **Zu S.***, **Li S.*** (equal contribution) *et al.*, *REPlY: representation learning of B cell receptor sequences.* Under preparation, 2021
- **Zu S.***, **Sahu A.*** (equal contribution) *et al.*, *MSSC: multiple-sample differential expression analysis of single-cell RNA sequencing data. Under preparation*, 2021
- Hu Z.*, **Zu S.*** (equal contribution), and Liu J.S. *SIMPLEs: single-cell RNA sequencing imputation and cell clustering methods by modeling gene module variation.* Accepted by **NAR Genomics and Bioinformatics**, 2020
- Ding Q., Hou S., **Zu S.**, *et al., VISAR: an interactive tool for dissecting chemical feature learned by deep neural network QSAR models.* **Bioinformatics**,36 (11), pp3610-3612., 2020.
- Xu X., **Zu S.**, *et al., Modeling Attention Flow on Graphs.* **NeurIPS 2018** Relational Representation Learning **Workshop**.

- Xu X., Chen L., **Zu S.**, *et al., Hulu video recommendation: from relevance to reasoning.* **RecSys'18** Proceedings of the 12th **ACM Conference** on Recommender Systems, pages 482-482, 2018
- Xu X., **Zu S.**, *et al., Backprop-Q: Generalized Backpropagation for Stochastic Computation Graphs.* **NeurIPS 2018** Deep Reinforcement Learning **Workshop**.
- Ding Z., **Zu S.**, and Gu J. *Evaluating the molecule-based prediction of clinical drug response in cancer.* **Bioinformatics**, 32(19), 2891-2895, 2016.
- **Zu S.**, Chen T., Li S. *Global optimization-based inference of chemogenomic features from drug-target interactions.* **Bioinformatics**, 31 (15), 2523 - 2529, 2015.

SOFTWARE

- **SIMPLEs**: an R package for single cell RNASeq data imputation based on cell similarities and gene correlations.
- **GIFT**: A C++ package to infer the chemogenomic information based on drug-protein interactions.

INDUSTRIAL EXPERIENCE

Algorithm Expert (P7) City-Brain Lab, DAMO Academy, Alibaba, China *2018.12 - 2019.08*

- Studied the deep Poisson-Gamma probabilistic model for traffic prediction.
- Developed one graphical neural network based online pipeline for transportation system with C++ language.

Researcher (IC3) Recommendation Team, HULU Beijing, China *2017.01 - 2018.11*

- Applied the recurrent neural networks for online real-time recommendation
- Lead the development of the multimodal algorithm for personalized recommendation by integrating user sequential watch behavior and shows' metadata.
- Studied the stochastic computational graph optimization with reinforcement learning strategies like Q-learning.

TEACHING EXPERIENCE

- *Guest Speaker* of Stat 221 at Depeartment of Statistics, Harvard *2020.10. 2020.11*
  - Introduction to variational auto-encoder with pytorch.
  - Introduction to STAN, a probabilistic programming language.
- *Teaching Assistant* of Probabilistic Graphical Models for graduate students *2013.09 - 2014.01*
- *Teaching Assistant* of Systems Biology for undergraduate students *2014.09 - 2015.01*

OTHER EXPERIENCE

- Undergraduate Affair Counselor (for scholarship and financial aid assessment) *2011.08 - 2013.01*
- Completed the Full Marathon *2009.10, 2010.10, 2011.10*
- The volunteer of 2008 Beijing Olympic Games *2008.08*

AWARDS

- National Scholarship for Graduate Students *2015*
- Tsinghua Scholarship for Overseas Graduate Studies *2014*
- Tsinghua Excellent Undergraduate Affair Counselor *2013*
- Tsinghua Zhongying Tang Scholarship *2008, 2009, 2010*