Jian Li | Ph.D. student in Statistical Machine Learning

Institute of Information Engineering, CAS – Beijing, China

(+86)13020058233 • lijian9026@iie.ac.cn • lijian.ac.cn
Google Scholar • github.com/superlj666

Education

Institute of Information Engineering, Chinese Academy of Sciences — **Beijing, China**
Ph.D. student, advisor: Prof. Weiping Wang and Associate Prof. Yong Liu — *Sep. 2015–Present*

Northeastern University — **Shenyang, China**
Undergraduate, Software College — *Sep. 2011–Jun. 2015*

Research Interests

My research interests mainly lie in **efficient large scale machine learning with theoretical guarantee**, but also include kernel methods, semi-supervised learning (SSL) and interpretability of neural networks. Indeed, my works focus on **generalization** analysis of those areas and building effective and scalable **optimization** tools for them, to channel theory and algorithms into applications. Current works:

- **Algorithms**: Design of efficient algorithms for semi-supervised settings, by making using of random projections, stochastic gradient methods and distributed learning.
- **Theory**: Statistical learning of large scale algorithms to apply to semi-supervised settings via data-dependent measures, including local Rademacher complexity and integral operator.
- **Interpretability of Neural Networks**: Understand neural networks in spectral kernels way, based on random Fourier features and Rademacher complexity.

Publications (👉 Google Scholar Profile)

- Multi-Class Learning using Unlabeled Samples: Theory and Algorithm.
  Jian Li, Yong Liu, Rong Yin, Weiping Wang.
  In Proceedings of the 28th International Joint Conference on Artificial Intelligence (**IJCAI 2019**).
- Approximate Manifold Regularization: Scalable Algorithm and Generalization Analysis.
  Jian Li, Yong Liu, Rong Yin, Weiping Wang.
  In Proceedings of the 28th International Joint Conference on Artificial Intelligence (**IJCAI 2019**).
- Multi-Class Learning: From Theory to Algorithm.
  Jian Li, Yong Liu, Rong Yin, Hua Zhang, Lizhong Ding, Weiping Wang.
  Advances in Neural Information Processing Systems 31 (**NeurIPS 2018**).
- Efficient Kernel Selection via Spectral Analysis.
  Jian Li, Yong Liu, Hailun Lin, Yinliang Yue, Weiping Wang.
  In Proceedings of the 26th International Joint Conference on Artificial Intelligence (**IJCAI 2017**).

Preprints

- Automated Spectral Kernel Learning. (Submission in AAAI)
  Jian Li, Yong Liu, Weiping Wang.
  arXiv preprint arXiv:1909.04894, 2019.
- Learning Vector-valued Functions with Local Rademacher Complexity. (Submission in TPAMI)
  Jian Li, Yong Liu, Weiping Wang.
  arXiv preprint arXiv:1909.04883, 2019.
- Distributed Learning with Random Features.
  Jian Li, Yong Liu, Weiping Wang.
  arXiv preprint arXiv:1906.03155, 2019.
- Efficient Cross-Validation for Semi-Supervised Learning.
  Yong Liu, **Jian Li**, Guangjun Wu, Lizhong Ding, Weiping Wang.
  arXiv preprint arXiv:1902.04768, 2019.
- Max-Diversity Distributed Learning: Theory and Algorithms. (Submission in AAAI)
  Yong Liu, **Jian Li**, Weiping Wang.
  arXiv preprint arXiv:1812.07738, 2018.

Engineering Experience

Large-scale Data Analysis and Process Platform **Institute of Information Engineering, CAS**
Data mining algorithms engineer *Apr. 2017 - Dec. 2017*

- Design and implement general machine learning algorithms.
- Apply general algorithms to specific large-scale tasks.
- Improve efficiency and scalability of both algorithms and applications.

Mass Email Analysis Subsystem **Institute of Information Engineering, CAS**
Big data algorithms engineer *Jul. 2016 - Mar. 2017*

- Build and maintain parallel computing engine (Spark cluster).
- Design and implement large scale graph mining algorithms (including community discovery, Spread back, behavior analysis and key nodes finding) on Spark.
- Apply graph mining algorithms to practical applications, e.g. analysis of community, spread and key notes.

Sword and Magic (Mobile game) **UEGame Cooperation**
U3D Game Developer *Dec. 2014 - Sep. 2015*

- Implement the state machine to control behavior of game role.
- Implement dynamic 2D and 3D effects UI for passing a stage.
- Accelerate efficiency of packaging and communication.

Expertise

- **Generalization Theory**: Rademacher complexity, Integral operator and approximate kernel theory.
- **Large Scale Optimization Algorithms**: random projections, gradient methods and distributed learning.
- **Programming Languages**: Python, Matlab, C/C++, Java.
- **Development Environments**: Pytorch, Tensorflow, Spark, sklearn, Pandas, Linux.
- **Languages**: English, Fluent. Chinese, Mother Tongue.

Honors and Awards

- **The UCAS Joint PhD Training Program (almost 100/50,000, USD $22,800)**
  University of Chinese Academy of Sciences (UCAS). 2019.
- **CAS Presidential Scholarship (Top 1%, RMB ¥5,000)**
  Chinese Academy of Sciences (CAS). 2019.
- **National Scholarship for Doctoral students (Top 2%, RMB ¥30,000)**
  Ministry of Education of P.R. China. 2018.
- **IIE Presidential Scholarship (Top 10%, RMB ¥2,000)**
  Institute of Information Engineering, CAS. 2018.
- **Merit Student**, University of Chinese Academy of Sciences (UCAS). 2018.
- **Merit Student**, University of Chinese Academy of Sciences (UCAS). 2019.
- **Laboratory Excellent Student Scholarship**, Institute of Information Engineering, CAS. 2017.
- **Laboratory Excellent Student Scholarship**, Institute of Information Engineering, CAS. 2018.