University of Technology Sydney

Faculty of Engineering and Information Technology

The Hong Kong Polytechnic University

Department of Computing

Learning Sparse Graphical Models for Data Restoration and Multi-Label Classification

Qiang Li

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

April 2017

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____(Signed)

Qiang Li (Name of student)

To my loving parents.

Abstract

Sparse probabilistic graphical models play an important role in structured prediction when the dependency structure is unknown. By inducing sparsity over edge parameters, a typical sparse graphical model can combine structure learning and parameter estimation under a unified optimization framework. In this thesis, we propose three specific sparse graphical models accompanied by their applications in data restoration and multi-label classification respectively.

For the data restoration task, we propose random mixed field (RMF) model to explore mixed-attribute correlations among data. The RMF model is capable of handling mixed-attribute data denoising and imputation simultaneously. Meanwhile, RMF employs a structured mean-field variational approach to decouple continuous-discrete interactions to achieve approximate inference. The effectiveness of this model is evaluated on both synthetic and real-world data.

For the multi-label classification task, we propose correlated logistic model (CorrLog) and conditional graphical lasso (CGL), to learn conditional label correlations. (1) The CorrLog model characterizes pairwise label correlations via scalar parameters, thus effects in an explicit (or direct) fashion. More specifically, CorrLog extends conventional logistic regression by jointly modelling label correlations. In addition, elastic-net regularization is employed to induce sparsity over the scalar parameters that define label correlations. CorrLog can be efficiently learned by regularized maximum pseudo likelihood estimation which

enjoys a satisfying generalization bound. Besides, message passing algorithm is applied to solve the multi-label prediction problem. (2) The CGL model further leverages features in modelling pairwise label correlations in terms of parametric functions of the input features, which effects in an implicit (or indirect) fashion. In general, CGL provides a unified Bayesian framework for structure and parameter learning conditioned on input features. We formulate the multi-label prediction as CGL inference problem, which is solved by a mean field variational approach. Meanwhile, CGL learning is efficient after applying the maximum a posterior (MAP) methodology and solved by a proximal gradient procedure. The effectiveness of CorrLog and CGL are evaluated on several benchmark multi-label classification datasets.

Acknowledgements

I feel so grateful to have Prof. Dacheng Tao and Prof. Jane You as my supervisors in UTS and PolyU respectively. I have learned a lot from their immense and extensive knowledge, meticulous attitude to academic research, and genial personalities. Besides, I am grateful to have Dr. Wei Bian and Dr. Richard Yi Da Xu as my de-facto co-supervisors in UTS. I have also learned a lot from their insightful guidance and elaborate suggestions during the second and third years of my PhD candidature. Actually, without those fruitful discussions, I can hardly achieve the research outcomes which constitute the main content of this thesis.

I would like to thank my friends in UTS, including Maoying Qiao, Mingming Gong, Ruxin Wang, Zhibin Hong, Tongliang Liu, Changxing Ding, Meng Fang, Zhe Xu, Shaoli Huang, Hao Xiong, Xiyu Yu, Guoliang Kang, Liu Liu, Chen Gong, Huan Fu, Baosheng Yu, Zhe Chen, Zijing Chen, Yali Du, Xun Yang, Jiankang Deng, Jing Yang, Dr. Jun Li, Dr. Tianyi Zhou, Dr. Yong Luo, Chang Xu, A/Prof. Shengzheng Wang, A/Prof. Xianhua Zeng, Prof. Bo Du, A/Prof. Shigang Liu, A/Prof. Wankou Yang, A/Prof. Xianye Ben, A/Prof. Tao Lei, Liping Xie, Jun Li, Long Lan, Chunyang Liu, Bozhong Liu, Sujuan Hou, Wuxia Yan, Haishuai Wang, Qin Zhang, Zhaofeng Su, Zhiguo Long, Lianyang Ma, Nannan Wang, Fei Gao, Weilong Hou, Mingjin Zhang, Chao Ma and many others. I learned a lot from the discussions and interactions with them in both academic research and

daily life. In particular, I really enjoyed the relaxed atmosphere of the afternoon meet-ups with Dr. Wei Bian and Maoying Qiao. Thanks for their understanding and tolerance to my tedious complaints on research, and thanks for their help in my preparation of research papers. Another particular thanks goes to Mingming Gong and Ruxin Wang, thanks for their encouragements and help on my PhD application, comments and suggestions on my research, and considerate assistance in daily life.

I would also like to thank my friends in PolyU, including A/Prof. Lefei Zhang, Ruohan Zhao, Siwei Hu, Yanxin Hu, A/Prof. Risheng Liu, Dr. Xianbiao Qi, Hui Li, Runjie Tan, Wengen Li, Ruosong Yang, Wei Lu, Yumeng Guo, Xiao Shen, Sitong Mao, Jiaxin Chen, Minglei Li, Edison Chan, Qiang Zhang, Liang Zhang, Quanyu Dai, Zimu Zheng, Yu Lei, Lei Han, Lei Xue, Zhijian He, Xingye Lu and many others. Those occasional discussions and talks with them made the fourth year of my PhD study a nice journey.

Finally, my special thanks goes to my family, including my parents, my elder brother, my grandmother and also in memory of my grandfather who left us three years ago. Thanks for their endless love, encouragements, support and blessing that helped me make it through the hard times of my study and life.

Table of Contents

List of Figures

List of Tables