Privileged Machine Learning for Prediction

Yangyang Shu

Supervisor: Dr. Guandong Xu

Advanced Analytics Institute
Faculty of Engineering and Information Technology
University of Technology Sydney

A thesis submitted for the degree of PhD
October 2021

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Yangyang Shu declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. In addition, this document has not been submitted for qualifications at any other academic institution. This research is supported by the Australian Government Research Training Program.

Production Note:
Signature removed prior to publication.

11 October 2021

Acknowledgments

I would like to express my sincere gratitude towards everyone who made this thesis possible.

First and foremost, I would like to show my deepest gratitude to my supervisor, Prof. Guandong Xu, a respectable, responsible and resourceful scholar, who has provided me with valuable guidance in every stage of the writing of this thesis. I have learned a lot from him about what research is, and the patience he has shown towards me is genuinely exceptional. Without his enlightening instruction, impressive kindness and patience, I could not have completed my thesis. His keen and vigorous academic observation enlightens me not only in this thesis but also in my future study. I could not have imagined a better supervisor.

I would especially like to thank my Cooperating teachers Dr. Yulei Sui and Dr. Hongyu Zhang for helping me to grow as a researcher. I truly admire their passion, drive, and ability to express complicated things in the most straightforward way possible. They have continuously challenged me and has been a great source of motivation and advice. I sincerely appreciate all the efforts they have taken to shape me as a person.

I am grateful to my co-supervisor Dr. Qian Li and Dr. Shaowu Liu for many insightful ideas they provided when I was stuck. I am also grateful to all the members of the UTS Data Science and Machine Intelligence Lab for all the good times we had both during and outside research. I express my gratitude towards friends I made.

I am thankful to UTS for providing financial and technical support for my research.

I would also like to thank Dr. Baosheng Yu who has helped me to develop the fundamental and essential academic competence. My sincere appreciation also goes to Dr. Chang Xu, the University of Sydney, who helps my research with great cooperation.

I can not thank my girlfriend Yuying, enough for her sacrifice, support, advice, and live throughout this journey. She is the reason I smile despite when chips are down.

Finally, and most importantly, I would like to thank my parents. I could not have done this without their guidance, encouragement, selfless love and support.

Abstract

Machine learning for prediction suffers from asymmetric distribution, such as posterior information, future information and hidden information. With some additional information only available in training, how to learn a machine learning model with them remains a key challenge. Despite recent advances in important domains such as vision and medicine, the standard learning under privileged information paradigm does not offer a satisfactory solution for learning variational privileged information. In this thesis, I will introduce how to learn under variational privileged information by leveraging asymmetric distribution and machine learning algorithms, specifically via 1) semi-supervised learning, 2) adversarial learning, 3) multi-task learning, 4) slack variable learning in support vector regression. We evaluate the proposed method in three applications: photo aesthetic assessment enhanced by high-level aesthetic attributes hidden in photos; music emotion recognition from songs with the help of implicit information about music elements and musical styles judged by composers; and multiple object recognition from images with the help of implicit information about the object's importance conveyed by the list of manually annotated image tags. Experiment results demonstrate that the proposed methods are superior to the classic learning paradigm when solving practical problems. In summary, in this thesis, we propose privileged machine learning for prediction. The detailed contents are follow:

1. propose a unified framework to systematically address the aforementioned three forms of privileged information. The proposed V-SVR+ method integrates continuous, ordinal, and binary PI into the learning process of support vector regression (SVR) via three losses. For continuous privileged information, we define a linear correcting (slack) function in the privileged information space to estimate slack variables in the standard SVR method using privileged information. For the ordinal relations of privileged information, we first rank the privileged information and then, regard this ordinal privileged information as auxiliary information used in the learning process of the SVR model. For the binary or Boolean privileged information, we infer a probabilistic dependency between the privileged information and labels from the summarized privileged information knowledge. Then, we transfer the privileged information knowledge to constraints and form a constrained optimization problem.

2. propose a novel approach of photo aesthetic assessment under the help of

aesthetic attributes. The aesthetic attributes are used as privileged information (PI), which is often available during training phase but unavailable in prediction phase due to the high collection expense. The proposed framework consists of a deep multi-task network as generator and a fully connected network as discriminator. Deep multi-task network learns the aesthetic attributes and score simultaneously to capture their dependencies and extract better feature representations. Specifically, we use ranking constraint in the label space, similarity constraint and prior probabilities loss in the privileged information space to make the output of multi-task network converge to that of ground truth. Adversarial loss is used to identify and distinguish the predicted privileged information of deep multi-task network from the ground truth PI distribution. Experimental results on two benchmark databases demonstrate the superiority of the proposed method to state-of-the-art.

3. propose a novel privileged learning based framework that fully explores the musical domain knowledge thus to enhance the emotion recognition. Particularly, to best of our knowledge, we are the first to construct a systematic taxonomy of the dimension and style-related music information by applying a domain knowledge of musicology and psychology. Then, we customize Restricted Boltzmann Machine (RBM) to generate the informative feature representation that captures the intrinsic dependency between musical elements and music styles. Finally, we formulate the generated feature representation as privileged information (PI), and develop a PI-based support vector regression (i.e., SVR+) for music emotion recognition task. Extensive experimental results on two benchmark databases demonstrate the superior performance on emotion recognition compared against the state-of-the-art baselines.

4. propose a novel regression algorithm via ordinal Privileged Information, which takes into consideration of ordinal form of privileged information. By integrating ordinal constraint into the learning process, the privileged information and the dependencies in images features can enhance the object recognition. In optimization, we adopt maximum margin regression model. Alternating Direction Method of Multipliers (ADMM) is developed to optimize this proposed model. We evaluate the proposed method on multiple object recognition from images with the help of implicit information about the object's importance conveyed by the list of manually annotated image tags. Experimental results demonstrate that the proposed method can effectively take the advantage of ordinal privileged information.

Publications

The majority of the thesis has been published in peer-reviewed conference and journal proceedings.

Journal Publications

- **Yangyang Shu**, Qian Li and Guandong Xu. "Privileged Multi-Task Learning for Attribute-Aware Aesthetic Assessment." *Pattern Recognition*, 2021. *Under review.*
- **Yangyang Shu**, Qian Li and Guandong Xu. "Semi-Supervised Adversarial Learning for Attribute-Aware Photo Aesthetic Assessment." *IEEE Transactions on Multimedia (TMM)*, 2021. In press, 10.1109/TMM.2021.3117709
- **Yangyang Shu**, Qian Li, Chang Xu, Shaowu Liu and Guandong Xu. "V-SVR+: Support Vector Regression with Variational Privileged Information." *IEEE Transactions on Multimedia (TMM)*, 2021. In press, 10.1109/TMM.2021.3060955
- **Yangyang Shu** and Guandong Xu. "Learning with Privileged Information for Music Emotion Recognition." *IEEE Transactions on Affective Computing (TAC)*, 2021. *Under review.*
- **Yangyang Shu**, Qian Li, Shaowu Liu and Guandong Xu. "Learning with privileged information for photo aesthetic assessment." *Neurocomputing*, 2020, 404, pp.304-316.
- Shangfei Wang, Can Wang, Tanfang Chen, Yaxin Wang, **Yangyang Shu** and Qiang Ji. "Video affective content analysis by exploring domain knowledge." *IEEE Transactions on Affective Computing (TAC)*, 2019. 10.1109/TAFFC.2019.2912377

Conference Publications

- **Yangyang Shu**, Yulei Sui, Hongyu Zhang and Guandong Xu. "Perf-AL: Performance Prediction for Configurable Software through Adversarial Learning" *Empirical Software Engineering and Measurement (ESEM)*, Bari, Italy, Ictorber 5-7, 2020, pp. 1-11.

- **Yangyang Shu** and Guandong Xu. "Emotion Recognition from Music Enhanced by Domain Knowledge." *Pacific Rim International Conference on Artificial Intelligence (PRICAI)*, Yanuca Island, Fiji, AUgust 26-30, 2019, pp. 121-134.
- **Yangyang Shu** and Shangfei Wang. "Emotion recognition through integrating EEG and peripheral signals." *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, USA, March 5-9, 2017, pp. 2871-2875.

Contents

List of Figures

List of Tables