Kotaro Tanahashi
Recruit Communications Co., Ltd.
Feature Selection by Quantum Annealing Processor -- An Extension of the QBoost --

Feature selection is to extract important features in the data among many features, which is an NP-hard problem. Since feature selection is widely used in real-world data analysis, development of an algorithm to overcome the difficulty in feature selection is an important issue. Quantum annealing (QA) is expected to be an efficient method to obtain the best solution for combinatorial optimization problems [1,2]. Thus, we proposed a new quantum annealing algorithm for feature selection (QAFS). To use quantum annealing processor, besides, we developed a method to map QAFS to quadratic unconstrained binary optimization (QUBO). Our method is regarded as an extension of the QBoost [3]. QAFS finds the optimal combination of the weak classifiers by QA to minimize the error over the training data with the limited number of features. Since QAFS is composed of non-linear weak classifiers, it can capture non-linear interactions between features while conventional linear methods cannot. We performed QAFS to an actual online advertising problem and public data sets by quantum annealing processor, D-Wave 2X. We confirmed that QAFS outperforms the conventional methods such as L1 regularized logistic regression and feature selection by random forest. This work was done in collaboration with Shinichi Takayanagi (Recruit Communications Co., Ltd.) and Shu Tanaka (Waseda Institute for Advanced Study, Waseda University and JST, PRESTO).

[1] T. Kadowaki and H. Nishimori, Physical Review E, 58, 5355 (1998).
[2] T. Kadowaki, Ph. D Thesis (Tokyo Institute of Technology, 1999). arXiv:quant-ph/0205020
[3] H. Neven, V. S. Denchev, G. Rose, W. G. Macready, JMLR: Workshop and Conference Proceedings 25, 333 (2012).