Quantum Annealing based Feature Selection in Machine Learning

Daniel Pranjić[1], Bharadwaj Chowdary Mummaneni[1], Horst Stühler[2], and Christian Tutschku[1]

[1] *Fraunhofer IAO, Nobelstraße 12, 70569 Stuttgart, Germany*
[2] *Zeppelin GmbH, Graf-Zeppelin-Platz 1, 85766 Garching, Germany*

May 23, 2024

Abstract

Feature selection is a pivotal step in enhancing the accuracy and efficiency of machine learning models. This study explores the use of quantum annealing [1] to address the challenge of selecting features that maximize mutual information (MI) [2] and conditional mutual information (CMI). Traditional methods for calculating the optimal feature set for maximum MI are computationally prohibitive for large datasets, even with approximation techniques. We used mutual information quadratic unconstrained binary optimization (MIQUBO) formulation to efficiently solve on a D-Wave quantum computer.

Our research demonstrates the effectiveness of the MIQUBO approach in identifying the optimal feature combinations that maximize MI. To validate its real-world applicability, we applied MIQUBO to the task of forecasting the price of used excavators. The results show a significant improvement in the predictive performance of machine learning models, achieved by focusing on a smaller subset of features with high MI concentration. This study highlights the potential of quantum annealing in solving complex feature selection problems, paving the way for more accurate and efficient machine learning models in various applications.

References

[1] Catherine C. McGeoch. Adiabatic quantum computation and quantum annealing. 2014.

[2] Hemanth Venkateswara, Prasanth Lade, Binbin Lin, Jieping Ye, and Sethuraman Panchanathan. Efficient approximate solutions to mutual information based global feature selection. In *2015 IEEE International Conference on Data Mining*, pages 1009–1014. IEEE, 2015.