論 文 の 内 容 の 要 旨

論文題目　　Learning High-dimensional Models with the Minimum Description Length Principle
(記述長最小化原理による高次元モデル学習)

氏　　名　　宮口　航平

High-dimensional models that have hundreds of thousands of parameters such as deep neural networks and sparse models are effective in machine learning and data mining tasks. Controlling the complexity of such high-dimensional models is necessary for attaining appropriate inductive inference, e.g., preventing overfit and making it easier to interpret the results.

On the contrary, there are numerous different principles for measuring the complexity of models. The minimum description length (MDL) principle is an information-theoretic principle proposed by Rissanen (1978), of which one salient feature is offering a unified framework of inductive inference without imposing any assumptions on the distribution of data. According to the MDL principle, the complexity of models is quantified depending on the minimax-regret code length. However, for high-dimensional models, the computation of the exact code length is intractable, and no analytic approximation method has been implemented till date to resolve this issue.
This is problematic in terms of two basic tasks of inductive inference, namely model selection and prediction: (i) High-dimensional model selection is difficult since the code length of each candidate model is intractable. (ii) Even if it is numerically tractable, designing high-dimensional prediction algorithms is difficult as the code length cannot be analytically evaluated.

Considering this, in this thesis, we propose three approaches to the problems of high-dimensional inductive inference under the MDL principle. (i) We address the

problem of high-dimensional model selection over exponentially many candidates leveraging the continuous relaxation of the minimax code lengths. The proposed algorithm overcomes the computational difficulty minimizing code lengths without computing them but sampling the stochastic gradients. (ii) We study the minimax code length of smooth models to derive a new analytic approximation. We demonstrate its effectiveness through the problem of hyperparameter selection. (iii) We study a novel complexity measure, namely the envelope complexity, that provides a more general framework for the analytic approximation of the minimax code length.
Its power is demonstrated by deriving an adaptive minimax predictor over high-dimensional L1-balls and systematic upper bounds on predictive risks.

These three approaches provide the tools and foundations for the MDL principle to deal with high-dimensional modeling and prediction.