

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Discriminative Methods with Imperfect Supervision in Machine Learning
著者(和文)	NiuGang
Author(English)	Gang Niu
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第9330号, 授与年月日:2013年9月25日, 学位の種別:課程博士, 審査員:杉山 将,佐藤 泰介,秋山 泰,篠田 浩一,瀬々 潤,工藤 一 浩
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第9330号, Conferred date:2013/9/25, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

(博士課程)
Doctoral Program

論文要旨

THESIS SUMMARY

専攻 :	計算工学	専攻	申請学位 (専攻分野) :	博士 (工学)
Department of			Academic Degree Requested	Doctor of
学籍番号 :			指導教員 (主) :	杉山 将 (准教授)
Student ID Number			Academic Advisor(main)	
学生氏名 :	Gang Niu		指導教員 (副) :	
Student's Name			Academic Advisor(sub)	

要旨 (英文 800 語程度)

Thesis Summary (approx.800 English Words)

We have entered the age of information for more than two decades, and now we are entering the age of big data. Traditional supervised learning, which requires all labels for training a model, costs too many resources and thus becomes inapplicable nowadays. In vivid contrast, imperfectly supervised learning, which does not have such a requirement, is much less expensive and then more practical for real-world applications. To this end, this doctoral thesis is devoted to developing discriminative methods with imperfect supervision in machine learning.

We have found six major problem settings that can be viewed as learning with imperfect supervision. This thesis focuses on three of them. Three learning models each with one or two algorithms are proposed for certain learning problems, namely, clustering, metric learning, and semi-supervised classification.

Firstly, we investigate the clustering problem and propose *maximum volume clustering* (MVC). State-of-the-art clustering methods have many advantages, for example, the cluster shape can be very flexible. However, all of existing methods lack two important theoretical guarantees: Finite sample stability which analyzes when different local optima induce the same data partition, and clustering error bound which theoretically bound the clustering error from above. MVC employs the soft response vector as the hypothesis rather than the centroid or hyperplane, and is approximately solved by sophisticated optimization methods. Consequently, MVC is theoretically guaranteed with the finite sample stability and clustering error bound. Experiments demonstrate MVC is promising.

Secondly, we investigate the metric learning problem and propose *semi-supervised metric learning paradigm with hyper-sparsity* (SERAPH). State-of-the-art semi-supervised metric learning methods are all based on manifold regularization and manifold embedding. These methods can successfully extract the similarity information of unlabeled data, but the dissimilarity information is simply ignored. Nonetheless, most unlabeled data should be dissimilar for rich enough input data domains. To this end, SERAPH learns a metric by learning a probability parameterized by that metric, while it employs entropy regularization so that it can also extract the dissimilarity information of unlabeled data. Experiments demonstrate SERAPH is promising.

Thirdly, we investigate the semi-supervised classification problem and propose *squared-loss mutual information regularization* (SMIR). Information maximization methods for semi-supervised classification, as the state-of-the-art methods, can directly deal with the multi-class out-of-sample classification problem. However, all of existing information maximization methods are non-convex, and thus they have no access to the globally optimal solution. SMIR replaces the ordinary mutual information with the squared-loss mutual information, and the optimization involved in SMIR is strongly convex under mild conditions and then has the analytic expression of the unique globally optimal solution. Experiments demonstrate SMIR is promising.

Given the encouraging experimental results of the proposed methods, we finally conclude that discriminative methods with imperfect supervision in machine learning are successful and worth a further study in the future.

備考 : 論文要旨は、和文 2000 字と英文 300 語を 1 部ずつ提出するか、もしくは英文 800 語を 2 部提出してください。

Note : Thesis Summary should be submitted in either a copy of 2000 Japanese Characters and 300 Words (English) or 2 copies of 800 Words (English).