論文 / 著書情報
Article / Book Information

題目(和文)	変化検知のための統計的機械学習
Title(English)	Statistical Machine Learning Approaches to Change Detection
著者(和文)	柳松
Author(English)	Song Liu
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9555号, 授与年月日:2014年3月26日, 学位の種別:課程博士, 審査員:杉山　将,秋山　泰,篠田　浩一,村田　剛志,藤井　敦
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第9555号, Conferred date:2014/3/26, 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： Department of	計算工学	専攻	申請学位（専攻分野）： Academic Degree Requested	博士 Doctor of	（ 工学 ）
学生氏名： Student's Name	柳松		指導教員（主）： Academic Advisor(main)	杉山将	
			指導教員（副）： Academic Advisor(sub)		

要旨（英文 800 語程度）
Thesis Summary（approx.800 English Words ）

The development of modern technologies has offered us easier ways of accessing and modifying digital data. The analysis on such changing data challenges our traditional view on data analysis in the incoming big-data era. With noise, outlier and other uncertainties included, statistical machine learning offers a principled way of extract patterns and make sense out of data using statistical methods. It helps us build such models where the uncertainty is treated as probability. Comparing to deterministic models, statistical models are more compact, since the deterministic models require more rules to describe exceptional cases while the statistical models just assign them with lower probability using one single model. If statistical methods are employed for learning the pattern, we refer to such machine learning processes as Statistical Machine Learning.

The modern age has not only produced a large amount of data, but also offers easy ways of accessing and updating data. In fact, change is one of the most important properties of big data. For example, satellite images taken on the same spot may be different due to lighting or clouding conditions; the trend of Twitter topics may rapidly shift after breaking news is reported; popular queries sent to search engines may change on an everyday basis. However, the traditional machine learning that learns a static pattern is no longer helpful in such circumstances, since the pattern observed today may be altered tomorrow. In contrast, the dynamic view of machine learning allows us to incorporate changing patterns in traditional learning tasks.

In this thesis, we focus on one of the dynamic learning tasks: unsupervised change detection, and propose two novel approaches in distributional and structural change detection respectively. Guided by Vapnik's Principle, both algorithms avoid comparing two separately learned patterns by learning the change directly.

Our first contribution is on non-parametric distributional change detection. The objective of change-point detection is to discover abrupt property changes lying behind time-series data. It has wide applications including climate change detection, genetic time-series analysis, and intrusion detection. In this paper, we present a novel

statistical change-point detection algorithm based on non-parametric divergence estimation between time-series samples from two retrospective segments. We first apply a recently-proposed density-ratio estimation method called the unconstrained least-squares importance fitting (uLSIF) to obtain Pearson divergence as change-point score, then further improve the uLSIF-based change-point detection method by using the relative Pearson divergence as a divergence measure instead of original Pearson divergence. It is accurately and efficiently estimated by relative-uLSIF method. Through experiments on artificial and real-world datasets including human-activity sensing, speech, and Twitter messages, we demonstrate the usefulness of the proposed method.

Our second contribution is on structural change detection between two Markov networks. We propose a new method for detecting changes in Markov network structure between two sets of samples. Instead of naively fitting two Markov network models separately to the two data sets and figuring out their divergence, we directly learn the network structure change by estimating the ratio of Markov network models. This density-ratio formulation naturally allows us to introduce sparsity in the network structure change, which highly contributes to enhancing interpretability. Furthermore, computation of the normalization term, which is a critical bottleneck of the naive approach, can be remarkably mitigated. We also give the dual formulation of the optimization problem, which further reduces the computation cost for large-scale Markov networks. Through experiments, we demonstrate the usefulness of our method.

As it is shown in this thesis, the unsupervised change detection can successfully capture the changes of patterns in many real-world applications. However, there are still many future works need to be done. For example, the theoretical analysis may explain the advantages of the proposed methods on change-detection in a rigorous mathematical manner. Applying the change-detection on high-dimensional data may also help the proposed method fit into many big-data scenarios. We believe such areas will be very promising fields of machine learning in the coming years.