論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	A Gaussian Process-based Incremental Neural Network
著者(和文)	WANG Xiaoyu
Author(English)	Xiaoyu Wang
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11417号, 授与年月日:2020年3月26日, 学位の種別:課程博士, 審査員:井村 順一,藤田 政之,三平 満司,山北 昌毅,早川 朋久
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第11417号, Conferred date:2020/3/26, 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

系・コース： Department of, Graduate major in	Systems and Control Engineering	系 コース	申請学位（専攻分野）： Academic Degree Requested	博士 Doctor of （ 工学 ）
学生氏名： Student's Name	WANG Xiaoyu		指導教員（主）： Academic Supervisor(main)	井村 順一
			指導教員（副）： Academic Supervisor(sub)	

要旨（英文 300 語程度）
Thesis Summary（approx.300 English Words ）

A Gaussian Process-based Incremental Neural Network

Chapter 1: Introduction

Chapter 1 introduces the research background, previous related studies and their shortcomings. First, we present the disadvantages of some classical algorithms that explore the partition space (e.g., K-means, Gaussian mixture models): 1) the information of the underlying local distribution cannot be effectively conserved and 2) the component parameters are extremely sensitive to noise. Thus, we consider from the perspective of topology-preservation. However, the conventional self-organizing map has a fixed structure and cannot perform well in lifelong learning. And furthermore, we present the drawbacks of the recently proposed self-organizing incremental neural network algorithms. Finally, we explain the main contributions of this thesis, i.e., propose novel increment neural network algorithms for three online machine learning tasks (online density estimation, online clustering and online regression) and carefully discuss their properties and advantages.

Chapter 2: A Gaussian Process-based Incremental Neural Network for Online Density Estimation

Chapter 2 introduces a Gaussian process-based incremental neural network algorithm for online density estimation. First, when locating the two winning nodes for an input sample, instead of employing Euclidean distance and Mahalanobis distance, we present a novel similarity based on Gaussian Processes. Second, we derive the optimal bandwidth matrix using cross validation and EM algorithm. Third, the threshold region of one node is obtained by setting a threshold to the probability of linking to this node given by Gaussian process classification. When a new sample arrives, we check whether this sample falls into the intersection of the threshold regions of its winning nodes, if so, an edge is generated between its winning nodes; otherwise, this sample is maintained as a node. Fourth, the asymptotic mean squared error is analyzed for the first time that proves the convergence of self-organizing incremental neural network to the underlying

probability density function.

Chapter 3: A Gaussian Process-based Incremental Neural Network for Online Clustering

As an extension of Chapter 2, Chapter 3 introduces an online clustering algorithm. First, we present the quantitative relationship between the edge weight and the underlying probability density, it shows that the weight of an inappropriate edge with a segment in low density area is much larger than that of a normal edge and proves the rationality of detecting clusters by constructing the minimum spanning tree of the proposed Gaussian process-based incremental neural network. Second, we discuss some properties of the threshold regions that demonstrate some advantages of our method over previous algorithms.

Chapter 4: A Gaussian Process-based Incremental Neural Network for Online Regression

Chapter 4 proposes a Gaussian process-based incremental neural network algorithm for online regression. First, different from the unsupervised learning algorithms in Chapter 2 and Chapter 3, we derive the optimal bandwidth matrix for each node by minimizing the leave-one-out cross validation error in estimating the conditional expectation of dependent variable. Second, as the input of samples, we update not only the network parameters (e.g., edges and weight vectors), but also the approximate posterior over the regression function values at nodes to guarantee that the approximate posterior is close to the exact posterior given the previous input samples. Third, the threshold region of a node is obtained and related to the regression function values at this node and its neighbors. Specifically, we set a threshold for the variance given by the local Gaussian process around this node. Fourth, we discuss some of the properties of the optimal posterior and prove that our approach can be regarded as a novel class of statistical ensembles of networks that not only includes combinatorial effects, but also considers the similarity between the weight vectors of nodes.

For each machine learning task, various experiments are conducted to demonstrate the performance of our algorithms and shown in the corresponding sections of Chapter 2, Chapter 3, and Chapter 4, respectively.

Chapter 5: Conclusion

Finally, we conclude in Chapter 5 and some problems we are working on are stated (e.g., online classification).