Introduction to Semi-Supervised Learning

[Introduction to Semi-Supervised Learning_下载链接1_](#)

著者:Xiaojin Zhu

出版者:Morgan and Claypool Publishers

出版时间:2009-6-29

装帧:Paperback

isbn:9781598295474

Semi-supervised learning is a learning paradigm concerned with the study of how computers and natural systems such as humans learn in the presence of both labeled and unlabeled data. Traditionally, learning has been studied either in the unsupervised paradigm (e.g., clustering, outlier detection) where all the data are unlabeled, or in the supervised paradigm (e.g., classification, regression) where all the data are labeled. The goal of semi-supervised learning is to understand how combining labeled and unlabeled data may change the learning behavior, and design algorithms that take advantage of such a combination. Semi-supervised learning is of great interest in machine learning and data mining because it can use readily available unlabeled data to improve supervised learning tasks when the labeled data are scarce or expensive. Semi-supervised learning also shows potential as a quantitative tool to understand human category learning, where most of the input is self-evidently unlabeled. In this introductory book, we present some popular semi-supervised learning models, including self-training, mixture models, co-training and multiview learning, graph-based methods, and semi-supervised support vector machines. For each model, we discuss its basic mathematical formulation. The success of semi-supervised learning depends critically on some underlying assumptions. We emphasize the assumptions made by each model and give counterexamples when appropriate to

demonstrate the limitations of the different models. In addition, we discuss semi-supervised learning for cognitive psychology. Finally, we give a computational learning theoretic perspective on semi-supervised learning, and we conclude the book with a brief discussion of open questions in the field. Table of Contents: Introduction to Statistical Machine Learning / Overview of Semi-Supervised Learning / Mixture Models and EM / Co-Training / Graph-Based Semi-Supervised Learning / Semi-Supervised Support Vector Machines / Human Semi-Supervised Learning / Theory and Outlook

作者介绍:

目录:

[Introduction to Semi-Supervised Learning_下载链接1_](#)

标签

机器学习

半监督学习

数据分析

算法

数据挖掘

计算机

CS

模式识别

评论

内容简洁、描述清晰、排版一流，一本入门的好书。

一句话semi-supervised learning就是基于各种assumption把unlabeled examples整合进regularization里。现在Jerry又开始鼓捣homology，祝一路走好。

对我来说核心问题是即使读完了也不知道应该用在哪里……望天

可以当故事书看

完备记录了跨越整个decade的东西，但这个时代几乎已经过去了

很有条理很好懂

说实话，介绍计算机算法的书很难评论，尤其是对于身处算法领域外的人而言，但是作为应用实践者，在茫茫多的算法书中指摘出自己的心仪之作仍不失为一种浪（强）漫（迫）感（症）。倘若你有机会了解一下机器学习的基础信息，会发现算法实现主要分为监督、无监督和强化三种学习范式，而近年来多位专业大牛则纷纷强调后两者。相比之下，半监督学习有点悲摧，虽然顶着"人类学习机制的最大可能性"这类帽子，可最为缺少关爱的样子，也许是由于其实现难度往往取决于监督或无监督的进展（也就是在这两者基础上改成半监督）。在为数不多的半监督学习相关书籍中，这本书的质量可算是上乘，全彩图，一共才130页，每一个具体算法配一个正面例子，加上许多的负面例子，将"算法表现取决于分析者对数据信息本质作出的假设与算法本身的匹配程度"的道理说了个明白。

除了概念什么都没讲，不如去看wikipedia

绝对的入门好书

[Introduction to Semi-Supervised Learning_下载链接1_](#)

书评