介绍了迁移学习的重要性,并且举了三个例子说明迁移学习当前是被迫切需要的。在最后一段介绍了文章的总体结构。首先是一些总体概念和一些 一大一接下来在文章中会用到的定义。然后是迁移学习的历史。接着是给出了迁移学习的定义,再把迁移学习划分成了三个不同的情况。再后来是给出一 INTRODUCTION 些关于negative transfer的当下的调查。最后举一些迁移学习成功应用的案例以及一些工具。最后是展望 半监督学习解决了标签数据太少而不能建立一个很好地分类器的问题(通过使用大量的未标记样本和已经标记的样本),祝贺吴已经解决如何应对 嘈杂标签分类问题。杨考虑到了花费感情的学习问题当额外的测试被用到未来的样本上去的时候。并且,他们大部分都假设标记和未标记的样本有 着相同的分布。但是迁移学习与他们不同,他允许训练样本和测试样本可以在分布区域、任务、分布不同的情况下进行 迁移学习的灵感来自人们可以机智的利用以往学过的知识到新的领域中去,或者对于之前的问题他们想出了更好的解决方案。迁移学习最重要的动 力来自于在nips大会上讨论的关于终身机器学习的思想,这种思想来源于要保持和利用过去学过的知识。 A Brief History of Transfer Learning 与迁移学习很类似的学习技术是多任务学习,这种技术致力于同时学习多个任务,即使这些任务不相同。多任务学习最典型的一个方法是找到可以 有益于多个个体的共同(隐含)特征 ipto给了迁移学习一个官方定义,即把在以往样本学到的知识和技术应用到新的任务中去。与多任务学习不同的是迁移学习更加关注目标任务,而 不是同时学习元领域和任务领域的任务。对于迁移学习来说原领域和目标领域的任务不再均衡。 介绍了domin和task的概念。domin包含了两个部分一个是特征空间X,一个是边缘分布P(X)。为了简易我们这里只考虑一个原领域和一个目 Notations and Definitions 定义1:迁移学习是通过利用从Ds与Ts中学到的知识来优化目标领域中的目标预测函数ft(i)。且目标领域和原领域不重合或者原领域的任务和目 标领域不同。 迁移学习三大问题: what to tranfer ,how to transfer,when to transfer 传统的机器学习是在领域和任务都相同的情况下工作。归纳迁移学习是在同一个领域解决不同任务的学习,无监督学习是在都不同的情况下迁移 传感器迁移是在任务相同的情况下进行。 源领域有很多的标记样本可以使用,这种情况的归纳迁移学习类似于多任务学习,不同的是迁移学习只注重通过从源领域学到的知识去在目标领域 中取得很好地成果。但是多任务学习尝试同时学习源领域和任务领域的任务 inductive transfer learning 源领域没有可用的标记样本 这种情况类似于self-taught学习 源领域和目标领域的特征空间不同 transductive transfer learning A Categorization of Transfer Learning Techniques 源领域和目标领域的特征空间相同但是输入数据的边缘分布不同 与文本分类、样本选择偏差和协变量位移中的知识迁移的领域适应有关 我们在三种子环境下对迁移学习进行了分类(这个分类是基于源和目标领域和任务的不同情况来划分) 致力于解决目标领域的无监督学习问题,比如聚类,降维,密度估计。这种情况下,源领域和目标领域都没有可以使用的标签样本 1基于实例的迁移学习 寻找一个好的特征表示来减少源领域和目标领域的分类和回归模型错误 在以上三种子环境下的迁移学习的方法可以根据"迁移什么"来划分为四种情况 假设源任务和目标任务共享一些参数或者是模型超参数的先验分布,因此通过共享参数和先验分布知识可以在任务中间进行很好地传递 假设源领域和目标领域的数据关系是相似的。数据的关联学习技术主导者这个领域 尽管源领域中的数据不能直接被使用但是仍然可以喝目标领域的少量标签样本一起使用。d提出了一个激增算法,叫triadaboost这是adaboost的 一个扩展。triadaboost假设源领域和目标领域使用完全相同的特征集和标签,但是这两个领域数据的分布不同。此外由于这些数据的分布不同, 就会导致有些源领域的样本数据源可能有益于目标样本但是有一些可能会有害。他反复的迭代源领域数据的去减少这些坏的样本的影响并且鼓励好 Transferring Knowledge of Instances 的样本去给目标领域的任务增益。并且tri使用和ada相同的方法去更新那些在目标领域错误分类的样本,但使用不同的方式去更新源领域中被错误 分类的样本。jiang提出了基于源领域和目标领域条件概率不同来移除那些在源领域错误领导的训练样本.liao提出了可以在源领域样本的帮助下去 标记一些目标领域没有标签的样本。。die将源领域数据汇集成了一个SVM(space vector model我猜的)框架以改善分类表现 3纳迁移学习背景下的监督特征构造特征的思想类似于多任务学习中构造特征表示时使用的思想。它的基本方法是去学习那些有关联的任务中共享 移学习的背景下,共同的特征可以通过解决一个优化问题来得知。在这个等式中,S,T代表源领域和目标领域的任务。A是一个参数矩阵,U是一个 supervised feature construction d*d的正交矩阵主要是用来把原来高纬度的数据转换为低维度表示。(r,p)的范数A被定义为。这个优化问题1是估计模型的源领域和目标领域 Transferring Knowledge of Feature Representations ,可以使用监督学习去构建一个特征表示。这种方式类似于多任务学习领域的共同特征学习。如果原领域的标记样本不可用,那么可以用无监督读 的基于矩阵的频谱正则化框架。nn提出了一个凸优化算法,用于从一组相关的预测任务中学习原优先级和特征权重。元优先级可以在不同的任务 学习去构造一个特征表示。 中传递。xx提出用svm为多任务学习选择特征。y为归纳迁移设计了一个基于核的方法,这种方法旨在为目标数据找到一个合适的核。 A Survey on Transfer Learning xx提出了应用稀疏编码去为迁移学习去学习更高级的特征,稀疏编码是一种无监督的特征构造方法。这个方法包含两个步骤。第一步通过解决优化 unsupervised feature construction 问题在源领域中学习更高级的偏差向量。第二步骤如下: INDUCTIVE TRANSFER LEARNING • 归纳迁移学习背景下的大多数参数传递的方法都认为关联任务的个体模型共享一些参数或者超参数的先验分布。本节描述的大部分方法,包括 正则化框架,分层贝叶斯框架都是为了在多任务学习环境下工作。但是可以轻易的调整他们以进行迁移学习。正如上文所说,多任务学习试图 同时并且完美学习原领域和目标领域的任务,但是迁移学习仅仅旨在通过利用源领域的数据去改善目标领域的工作效果。因此,在多任务学习 中,原领域和目标领域的loss函数的权重是相同的。迁移学习中不同领域的权重是不同的。直观的来讲,我们可能会分配更大的权重给目标领域 的loss函数以保证我们可以在目标领域中获得更好的表现。【】基于高斯过程提出了一个高效的算法,用以掌控多任务学习案例。{}尝试从多个 Transferring Knowledge of Parameters 任务的高斯过程中学习参数,主要通过共享一些相同的GP先验分布。【】也探究了GP背景下的多任务学习。作者提出在任务上使用自由形式的 协方差矩阵去为任务间的关系建模,而使用GP先验概率去诱导人物之间的内在联系。【】提出了将分层次贝叶斯框架和GP共同作用于多任务学 习。除了迁移GP模型的先验分布,一些研究者还提出了在正则化框架中转换SVM的参数。【】在多任务学习中借鉴了HB的思想在SVM上。他 提出的思想认为SVM里面每个任务可以被分成两个term'的的参数,一个是任务中term的共同参数,另一个是一个特定任务的参数。很多学 者进一步推进了参数学习的方法。【】提出了一个本地加权模型的集成学习框架以结合用于迁移学习多种模型,其中权重是根据模型在任务领 域中队每个测试模型的预测能力来动态分配的。 与其他三个情况不同,相关知识迁移的方法解决的是相关领域的迁移学习问题,其中这些数据并不满足独立同分布,而且可以用丰富的关系来表示 比如书网络数据和社会网络数据。这种方法并不像传统的假设认为从每个领域取出的数据是独立同分布的。他尝试从源领域到目标领域去迁移数 Transferring Relational Knowledge 据之间的关系。在这种背景下,统计的相关学习技术被提出使用来解决这些问题。【提出了一种TAMAR算法,这种算法MLNs在不同的领域间迁 移相关的数据】 转导性迁移学习环境中的实例迁移方法是由重要抽样所激发的。为了了解基于重要性抽样的方法在这种情况下是如何有帮助的,我们首先回顾了经 Transferring the Knowledge of Instances $_{\neg}$ 验风险最小化(ERM)[60]问题。一般来说,我们可能会想学习最优参数 θ *通过最小化模型的预期的风险, In [65], Xue et al. proposed a cross-domain text classification algorithm that extended the traditional probabilistic latent (PLSA) algorithm to integrate labeled ndunlabeled data from different but related domains, into a 直推是迁移学习是由【】首次提出,在直推式迁移学习中原领域和目标领域的任务是一致的,尽管他们的领域可能不同。在这些条件之上,他们进 一步要求在训练时,目标域中所有未标记的数据都是可用的,但我们相信这个条件可以放宽;相反,在我们对转换学习设置的定义中,我们只要求在 RANSDUCTIVE TRANSFER LEARNING unified probabilistic model. 训练时看到部分未标记的目标数据,以获得目标数据的边际概率。 Transfering knowledge of Feature representations 跨导转移学习设置的大多数特征表示转移方法都是在无监督的学习框架下进行的。 Blitzer等人[38] 提出了一种结构对应学习(SCL)算法,该算 法进行了扩展[37],以利用来自目标域的未标记数据来提取一些可以减少差异的真实特征。SCL的第一步是在每个领域的无标记数据上定义一个枢 纽特征集(枢纽特征集合的个数用m来记录),然后SCL将这些枢纽特征集合从这些数据中移出,并把每个枢纽特征看做一个新的标签向量。m个 分类问题就被构建出来了。通过假设每个问题可以通过线性分类器解决,如下所示;SCL可以学到一个参数为w的矩阵。在第三步,用SVD对矩阵 W做分解。θ是一个矩阵它的行是w最左侧单值向量。(降维)最后标准的。增强特征向量包含所有的附加了新的共享特征向量的原是特征向量Xi 。如[38]所述,如果枢轴特征设计合理,则 学习映射θ编码来自不同域的特征之间的对应关系。【】【】【】在[64]中,Ling等人。 提出了用于跨域转移学习问题的频谱分类框架,其中引入 了目标函数以寻求域内监管与域外内在结构之间的一致性。 在[65]中, 薛等。 提出了一种跨域文本分类算法, 该算法扩展了传统的概率潜在语义 分析(PLSA)算法,可将来自不同但相关域的标记和未标记数据集成到统一的概率模型中。 新模型称为主题桥接PLSA或TPLSA。 戴等。 [26]研究了一种新的聚类问题案例,称为自学聚类。 自学式聚类是无监督迁移学习的一个实例,其目的是借助源域中的大量未标记数据将 根据无监督转移学习设置的定义,在训练中在源域和目标域中未观察到标记数据。到目前为止,关于这种情况的研究很少。最近,提出了自学聚 目标域中的一小部分未标记数据聚类。 STC尝试学习跨域的公共特征空间,这有助于在目标域中进行聚类。 STC的目标函数如下所示。同样,[27] UNSUPERVISED TRANSFER LEARNING) 类(STC)[26]和转移判别分析(TDA)[27]算法,分别用于转移聚类和转移降维问题。 提出了一种转移判别分析(TDA)算法来解决转移维数减少的问题。 TDA首先应用聚类方法为目标未标记数据生成伪类标签。 然后,它将降维方 法应用于目标数据和带标签的源数据以降低维数。 这两个步骤反复运行,以找到目标数据的最佳子空间。 在[68]中,哈桑·马哈茂德(Hassan Mahmud)和雷(Ray)使用Kolmogorov复杂度分析了转移学习的情况,并证明了一些理论上的界限。 特别 TRANSFER BOUNDS AND NEGATIVE 是,作者使用条件Kolmogorov复杂度来度量任务之间的相关性,并在贝叶斯框架下的顺序转移学习任务中转移"正确"的信息量。最近,Eaton 等人[69]提出了一种新的基于图的知识转移方法,该方法通过将学习到的源模型集嵌入一个以可转移性为度量的图中来建模源任务之间的关系。将 TRANSFER 问题映射到图中,然后学习图上的一个函数,该函数自动确定要转移到新学习任务的参数。 垃圾过滤和泛华2WiFi实例。到目前为止,已经发表了几个数据集用于迁移学习研究。我们分别用文本挖掘数据集、邮件垃圾邮件过滤数据集、Wi Fi随时间段定位数据集和情绪分类数据集来表示。 迁移学习三大背景分类:归纳迁移学习,传感迁移学习,无监督迁移学习;基于what to transfer 又将迁移学习的方法分为四种:实力迁移;特

一 征表示迁移;参数迁移;相关知识迁移;未来的问题:如何解决negative transfer;我们希望在不同特征空间迁移知识,而不仅仅是在泛华连个

CONCLUSIONS

相同领域;引用到大型的数据迁移学习