**姓名:刘鑫杰 班级:04011806**

***The second week***

**《Learning To Detect Unseen Object Classes by Between-Class Attribute Transfer》阅读报告**

**前言：**

本paper为ZSL的开山之作，主要介绍了一个问题：检测未在训练集标签中出现的类标签的实例检测，即ZSL，同时介绍了两种算法DAP以及IAP。上一周对ZSL有了一个大概的全局了解，这周开始研读相关论文。并尝试复现其中的算法，以加深自己对算法的理解。同时也学习MATLAB相关的知识。

**目录：**

**1.定义**

**2.DAP和IAP算法**

**3.总结**

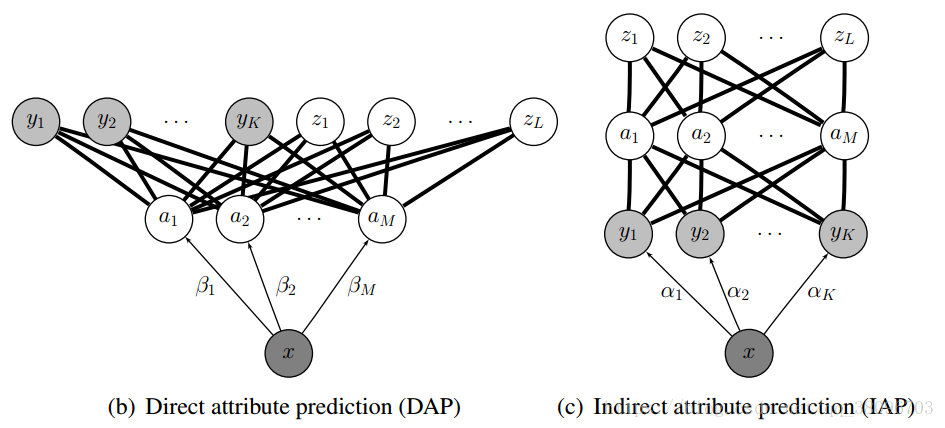
**4.额外篇(补充数学知识)**

**1.定义**

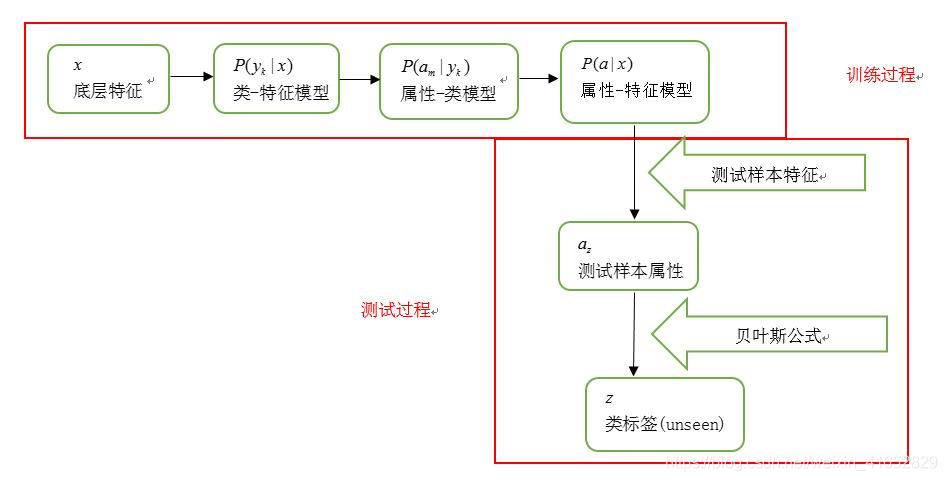
x为实例集合，y为可见的标签集合（训练集），z为不可见的标签集合（测试集），y与z的交为空。a为属性向量，M维，每个维度代表一个类的一种属性且为{0，1}取值，对于每个标签，皆对应一个M维向量a作为其的属性向量

**2.DAP和IAP算法**

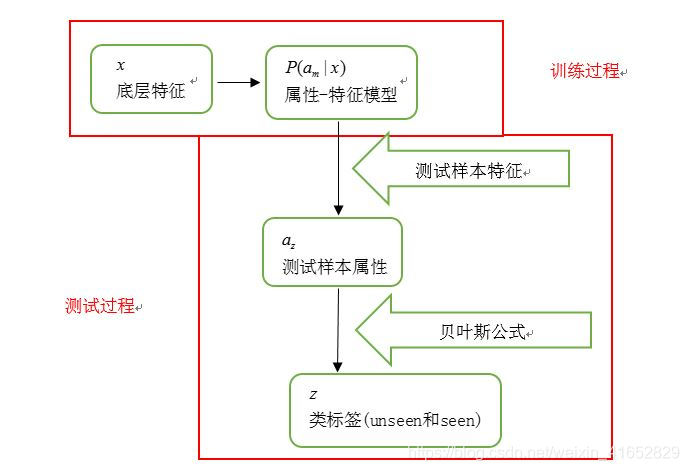
传统的多分类器无法胜任ZSL（因为其仅仅建立可见标签集与实例的联系），故介绍两种新算法来解决，DAP和IAP都属于基于类属性的分类，通过加入人类先验高级描述，达到零样本学习的目的，而有别于传统的分类方法。DAP与IAP的区别在于训练类与测试类之间的关系。在DAP中，基于属性层中的属性结果，对所有类一视同仁。相反，通过间接学习IAP中的属性，可以从作为中间特征层的训练类中归纳出属性的值。



*IAP训练和测试过程:*



*DAP的训练和测试过程:*



2.1 DAP (directed attribute prediction)

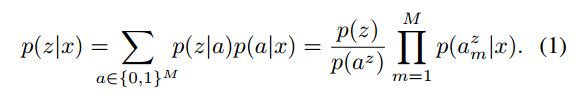
（1）添加了一个属性层在图像与类标签中用于解耦, DAP可以理解为一个三层模型：第一层是原始输入层，例如一张电子图片（可以用像素的方式进行描述）；第二层是p维特征空间，每一维代表一个特征（例如是否有尾巴、是否有毛等等）；第三层是输出层，输出模型对输出样本的类别判断。在第一层和第二层中间，训练p个分类器，用于对一张图片判断是否符合p维特征空间各个维度所对应的特征；在第二层和第三层间，有一个语料知识库，用于保存p维特征空间和输出y的对应关系。

简单来讲，就是对输入的每一个属性训练一个分类器(每个分类器判别该样本是否有该属性,有为1,无为0)，然后将训练得出的模型用于属性的预测，测试时，对测试样本的属性进行预测，再从属性向量空间里面找到和测试样本最接近的类别。

（2）自己通俗的理解：以AWA2数据集为例, x表示训练集的图片，经过数据预处理之后，每张图片变成了1x85的矩阵A，每列位置上的数是连续的，通过每个属性分类器将其变为01，每列位置上对应0（无）或者1（有）表示属性的有无，即对应于上图的a1~am,然后属性层和标签层之间有语料知识库🡪这是一个50x85的矩阵B，每一行代表一类动物，每一列代表一种属性，每个位置的值表示属性的有无或者对评估的重要性。通过在训练集上训练可以得到一个模型，这个模型可以通过输入的矩阵A对训练样本的类别进行预测（基于遍历矩阵B进行对比选择近邻），从而对未知类别Z进行预测（同样基于矩阵B）。

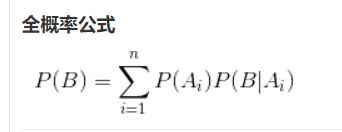
（3）形象举例：假设我们已经训练好了一个DAP模型，第一层和第二层间的分类器可以判断 是否黑眼圈、是否喜欢吃竹子 之类的特征，然后在语料知识库里面包含一个映射：黑眼圈 喜欢吃竹子--> 熊猫，那么即使我们的模型在训练时没有见过熊猫的图片，在遇到熊猫的图片时，我们可以直接通过对图片的特征进行分析，然后结合知识语料库判断出这张图片是熊猫。假设即使语料知识库里面不包含 黑眼圈 喜欢吃竹子--> 熊猫 的映射，我们也可以通过计算熊猫图片的特征与其他训练样本的特征的汉明距离度量，得到熊猫和什么动物比较类似的信息。整个DAP的运作思想就是类似于上述过程。

（4）数学推导



通常的最大后验估计（MAP）操作流程为利用训练数据，得到似然函数，再结合参数先验(先验概率是指根据以往经验和分析得到的概率, 如全概率公式，它往往作为"由因求果"问题中的"因"出现的概率)，得到后验概率，并求argmax得到参数的估计值，此处将测试实例x输入后的标签作为待估计的参数，而利用训练数据,学习得到实例空间与属性空间的映射p(a|x), 对于测试实例x,即可利用MAP的思想,找出概率最大的类为输出的估计类.

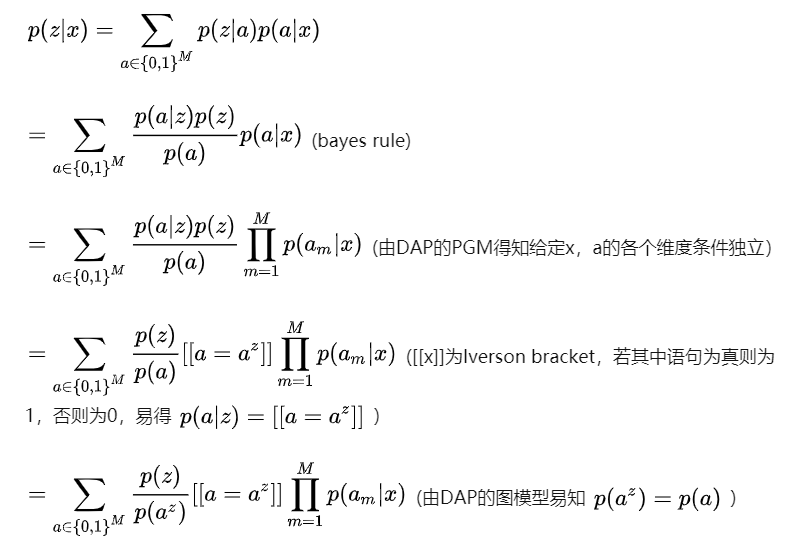
全概率公式:

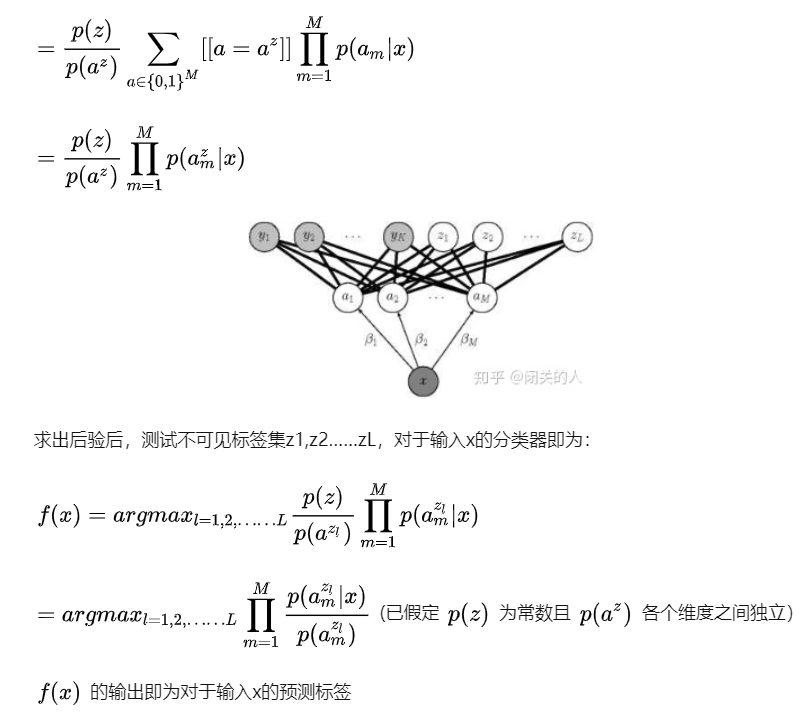
中p(A)即先验概率

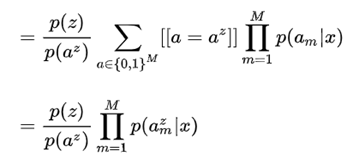
操作过程:要推断z, DAP采用MAP(最大后验概率), 通过输入训练数据, 可以学习到

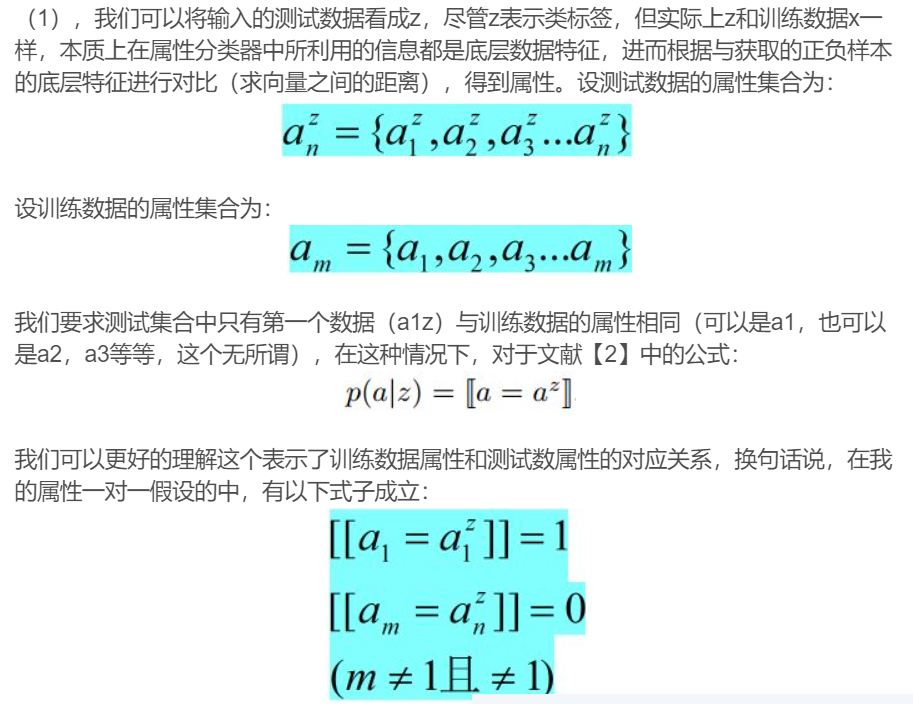
P(am|x) m=1,2,3,…M

进行推断的过程,先求出z的后验:

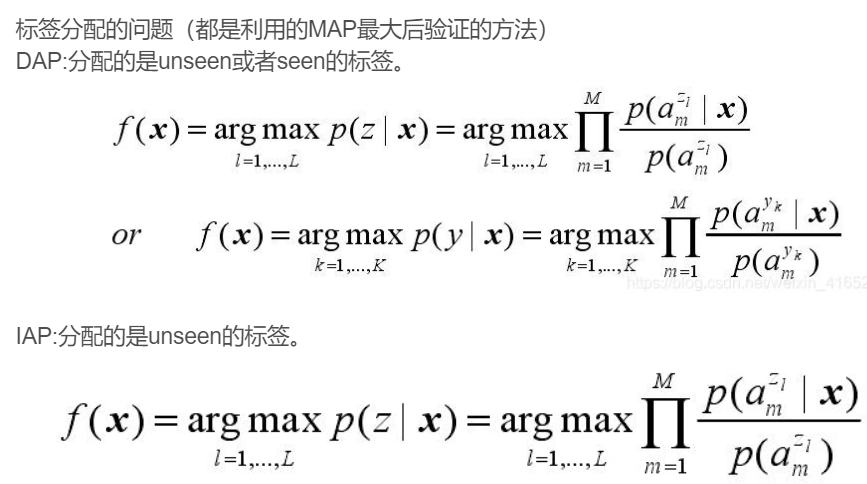




推导过程中对理解



即对[[a=az]]进行求和,其值为1+0+0…+0=1,而am=amz,所以上述推导正解.



（5）DAP的缺点：

1.无法克服邻域漂移问题；

2.训练很多学习器也是一件费时的事；

3.无法利用额外的属性信息；

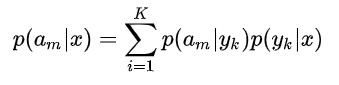
4.无法利用新的样本逐步改善分类器的功能；

5.算法引入了中间层，核心在于尽可能得判定好每幅图像所对应的特征，而不是直接去预测出类别；因此DAP模型在判定属性时可能会做得很好，但是在预测类别时却不一定；

2.2 IAP (indirect attribute prediction)

这一方法使用了两层标签,属性层作为中间层，但是属性在两层标签之间形成了一个连接层，一层用于训练时已知的类，另一层用于训练时未知的类。IAP的训练阶段是普通的多类分类。在测试时，对所有训练类的预测都会引导属性层的标记，从而可以推断出测试类上的标记。

IAP中a的后验为:

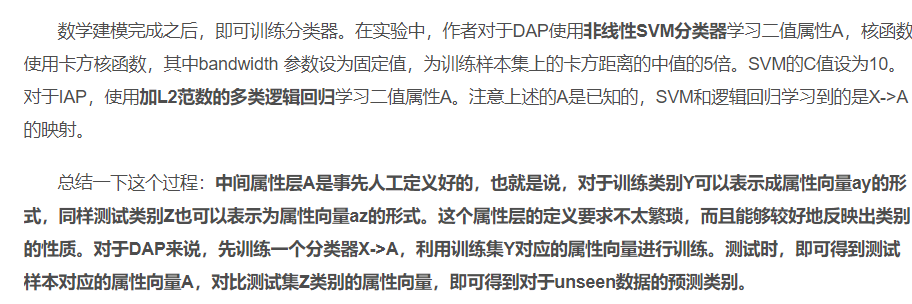


得到这个后验后,再求出z的后验,即可如同DAP中一样应用MAP即可

**3.总结**

参考这个博客,后面总结

https://blog.csdn.net/u011070272/article/details/73250102



**4.额外篇(补充数学知识)**

**前言**

对概率看法不同的两大派别频率学派(世界是确定的,直接为事件本身建模,即事件在多次重复试验中趋于一个稳定的值p,则p就是概率)与贝叶斯派(世界是不确定的,因获取的信息不同而异.假设对世界先有一个预先的估计，然后通过获取的信息来不断调整之前的预估计。 他们不试图对事件本身进行建模，而是从旁观者的角度来说。因此对于同一个事件，不同的人掌握的先验不同的话，那么他们所认为的事件状态也会不同。), 看待世界的视角不同,导致对产生数据的模型参数的理解不同.

4.1极大似然估计（Maximum Likelihood Estimate）

模型参数是个定值, 通过类似解方程组的方式从数据中求得该未知数. 🡪条件:在数据量很大的情况下效果很好.

4.2最大后验概率估计（Maximum A Posteriori estimation）

模型参数源自某种潜在的分布, 希望从数据中推知该分布。对于数据的观测方式不同或者假设不同，那么推知的该参数也会因此而存在差异。这就是贝叶斯派视角下用来估计参数的常用方法-最大后验概率估计（MAP），这种方法在先验假设比较靠谱的情况下效果显著，随着数据量的增加，先验假设对于模型参数的主导作用会逐渐削弱，相反真实的数据样例会大大占据有利地位。极端情况下，比如把先验假设去掉，或者假设先验满足均匀分布的话，那她和极大似然估计就如出一辙了。

4.3区别: MAP与MLE最大区别是MAP中加入了模型参数本身的概率分布，或者说。MLE中认为模型参数本身的概率的是均匀的，即该概率为一个固定值。

**《Zero-Shot Learning with Semantic Output Codes》阅读报告**

**前言：**

我们考虑零样本学习问题，目标是学习一个分类f:X→Y，它可以预测从训练集中忽略的Y的新值。为了实现这一点，我们定义了语义输出代码分类器（SOC）的概念，它利用Y的语义属性知识库将其外推到新的类中。

论文给出了这类分类器的形式并研究了在PAC框架下的理论性质，显示在一定条件下分类器可以准确预测新类。

本peper为带有语义输出码的零样本学习,在基于对前几周对ZSL的整体有了一个大概了解后,在细看相应解决算法.其提出的一个广义问题是：给定很多类的语义编码，我们能否建立以一个分类器来识别出训练集中没有类。

**目录：**

**1.语义特征空间（Semantic Feature Space）**

**2.语义知识库（Semantic Knowledge Base）**

**3.语义输出码分类器（Semantic Output Code Classifier）**

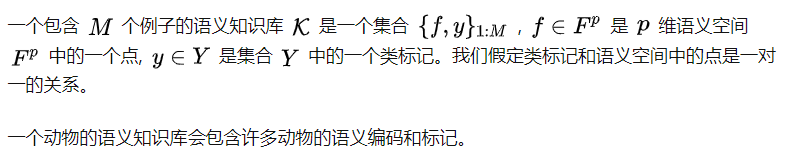
**4.理论分析**

**5.补充PAC学习框架**

**1.语义特征空间（Semantic Feature Space）**

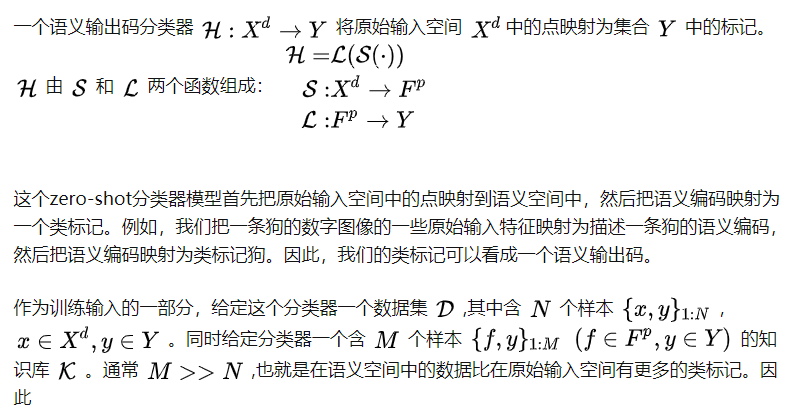
一个 p维的语义特征空间是一个度量空间，每一维编码一个语义性质的值。例如考虑一个描述动物高水平性质的语义空间。在这个例子中，我们考虑一个小空间 p=5 .每一维是一个是否特征：有没有毛皮？有没有尾巴？可以在水下呼吸吗？是否肉食？动作缓慢吗？在这个语义空间中典型的狗的概念可以表示为点 {1,1,0,1,0} .

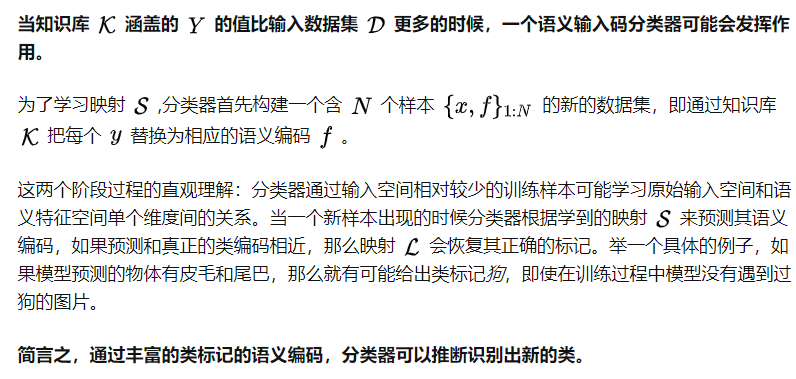
**2.语义知识库（Semantic Knowledge Base）**



如({1，1，0，1，0}，狗)就是知识库中的一个样本，一个知识库由多个这样的样本组成，如（{0，1，1，0，0}，鱼）也是一个样本。

**3.语义输出码分类器（Semantic Output Code Classifier）**





**4.理论分析**

在什么条件下，语义输出代码分类器将从训练集中省略的类中识别示例？

思想：如果分类器的第一阶段S（·）能够很好地预测语义属性，那么第二阶段L（·）将有很好的机会从新类中恢复正确的实例标签。

我们将考虑语义输出代码分类器的一个实例。我们假设语义特征是二元标签，第一阶段S（·）是PAC可学习线性分类器的集合（每个特征一个分类器），第二阶段L（·）是使用汉明距离度量的1-最近邻分类器。

**5.补充PAC（probably approximately correct，很可能接近正确的）学习框架**

5.1前言

PAC学习框架是机器学习的基础，主要回答了这几个问题：什么问题是可以高效学习的？ 什么问题本质上就难以学习？ 需要多少实例才能完成学习？ 是否存在一个通用的学习模型？

5.2 PAC原理

说一个问题是PAC可学习的，需要定义m个sample组成S空间，其中每个sample服从D分布，并且互相独立；如果存在一个算法A，在m（sample个数）有限的情况下，找到假设h；使得对于任意两个数x，y，概率P(h对S中sample预测错误次数大于x) < y；xy对应C:\Users\DELL\AppData\Roaming\Tencent\QQ\Temp\TempPic\24}]]ZP4Y]3W35S1~$NNS)X.hc 中两个奇怪的符号！注意上面说的是小于，截图中说的是相反事件的大于。其实是一回事。那么该问题是PAC可学习的。

PS：对于这篇论文还是没有理解他的两个映射函数如何得到的，以及数学部分不太清楚明了。后面再补充。