

Beijing Forest Studio
北京理工大学信息系统及安全对抗实验中心



多示例多标记学习(Multi- Instance Multi-Label Learning)

白崇有 硕士研究生

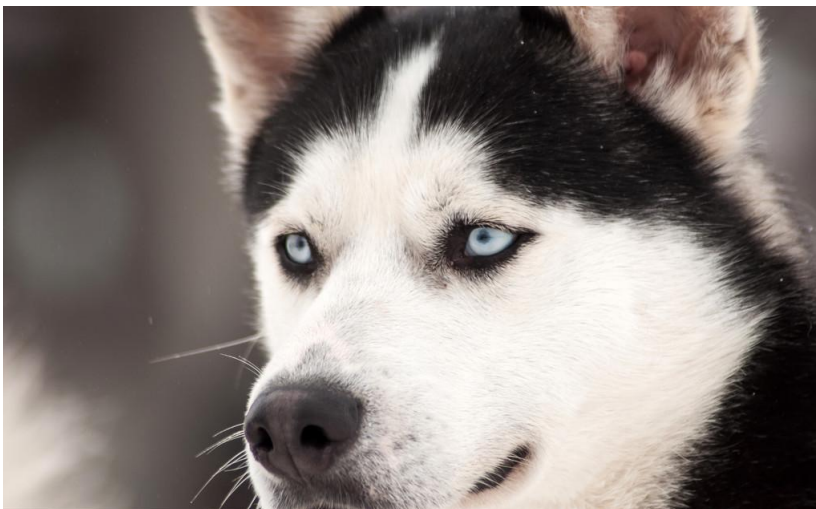
2018年03月11日

- 背景简介
- 基本概念
- 算法原理
- 优劣分析
- 应用总结
- 参考文献

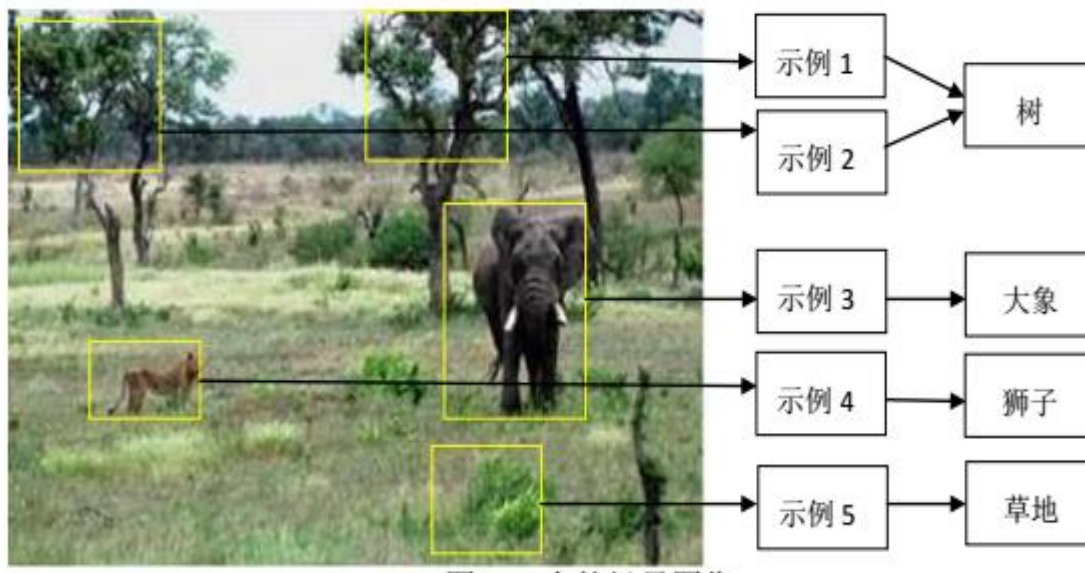


背景简介

- 在传统监督学习框架下，认为真实世界的对象与其描述及概念标记之间都是一一对应的关系，一般认为，这样的学习问题是没有歧义性的。
- 一个对象可以用**一个示例**来描述；
- 一个对象对应着**一个标记**。



- 然而，歧义性对象在现实世界是广泛存在的。
- 一个对象可以用**多个示例**来描述；
- 一个对象对应着**多个标记**。





基本概念

- 多示例学习

给定数据集 $\{(X_1, y_1), (X_2, y_2), \dots, (X_m, y_m)\}$, 目标是学得:

$$f: 2^{\mathcal{X}} \rightarrow \mathcal{Y}$$

其中, $X_i \subseteq \mathcal{X}$ 为一组示例 $\{x_{i1}, x_{i2}, \dots, x_{i, n_i}\}, x_{ij} \in \mathcal{X}$

$(j=1, 2, \dots, n_i), y_i \in \mathcal{Y}$ 为 X_i 的类别标记, n_i 为 X_i 中所含示例的个数。

- 多标记学习

给定数据集 $\{ (x_1, Y_1), (x_2, Y_2), \dots, (x_m, Y_m) \}$, 目标是学得:

$$f: \mathcal{X} \rightarrow 2^{\mathcal{Y}}$$

其中, $x_i \in \mathcal{X}$ 为一个示例, $Y_i \subseteq \mathcal{Y}$ 为 x_i 的一组合适类别标记

$\{y_{i1}, y_{i2}, \dots, y_{i, l_i}\}, y_{ik} \in \mathcal{Y} (k=1, 2, \dots, l_i)$, l_i 为 Y_i 中所含类别标记的个数。

- 多示例多标记学习

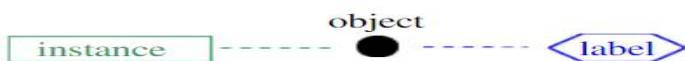
给定数据集 $\{(X_1, Y_1), (X_2, Y_2), \dots, (X_m, Y_m)\}$, 目标是学得:

$$f: 2^X \rightarrow 2^Y$$

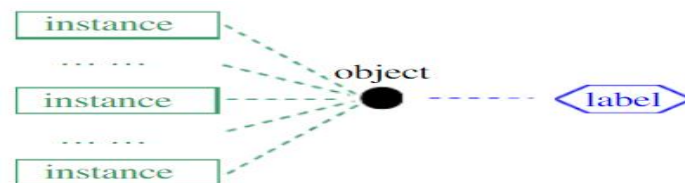
其中, $X_i \subseteq X$ 为一组示例 $\{x_{i1}, x_{i2}, \dots, x_{i, n_i}\}, x_{ij} \in X$
($j=1, 2, \dots, n_i$), 而 $Y_i \subseteq Y$ 为 X_i 的一组合适类别标记 $\{y_{i1}, y_{i2}, \dots, y_{i, l_i}\}, y_{ik} \in Y$
($k=1, 2, \dots, l_i$), n_i 为 X_i 中所含示例的个数, l_i 为 Y_i 中所含类别标记的个数。

• 四种机器学习框架的直观比较

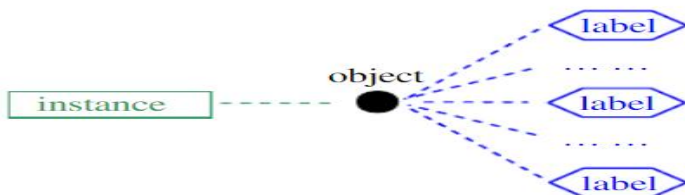
- 多示例学习从输入空间，即内容表示上来考察对象的歧义性；
- 多标记学习则是从输出空间，也就是概念标记上来考察对象的歧义性；
- 多示例多标记学习从输入空间和输出空间两个方面同时考察对象的歧义性。



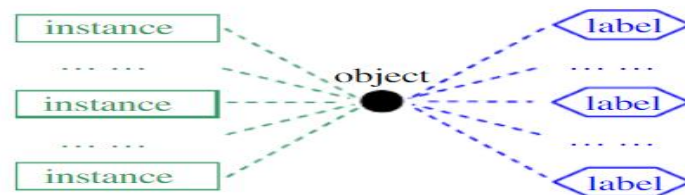
(a) 传统监督学习（单示例、单标记）



(b) 多示例学习（多示例、单标记）



(c) 多标记学习（单示例、多标记）

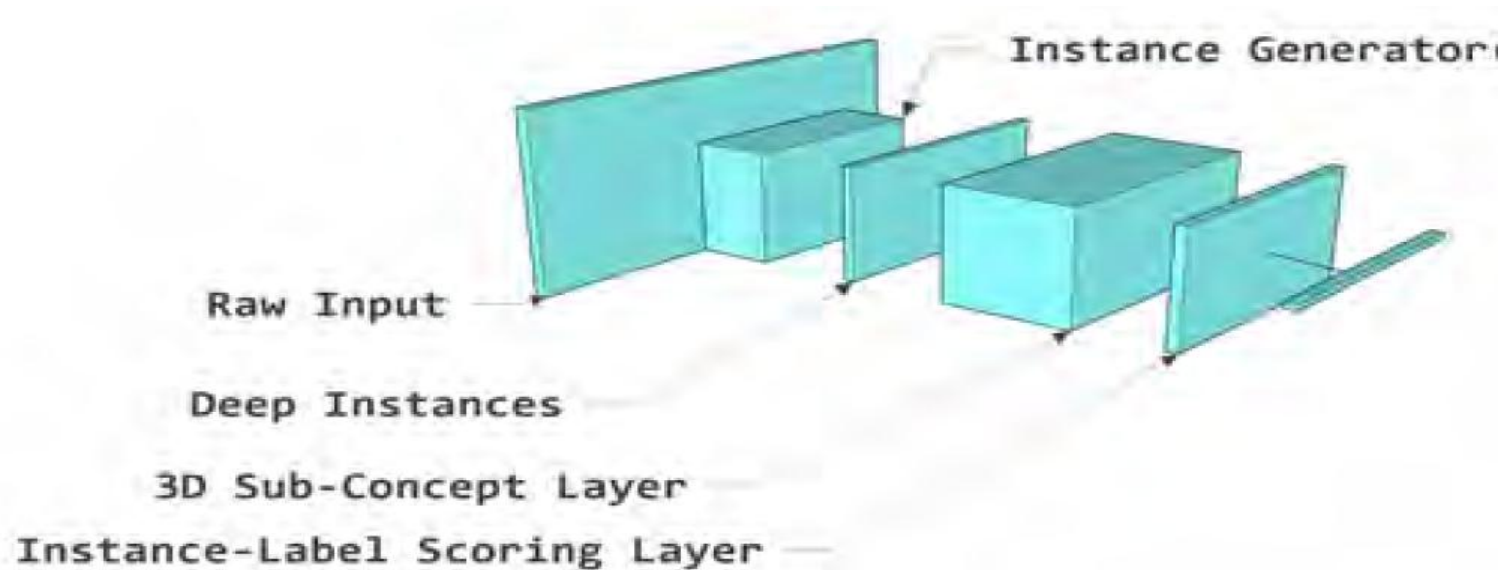


(d) 多示例多标记学习



算法原理

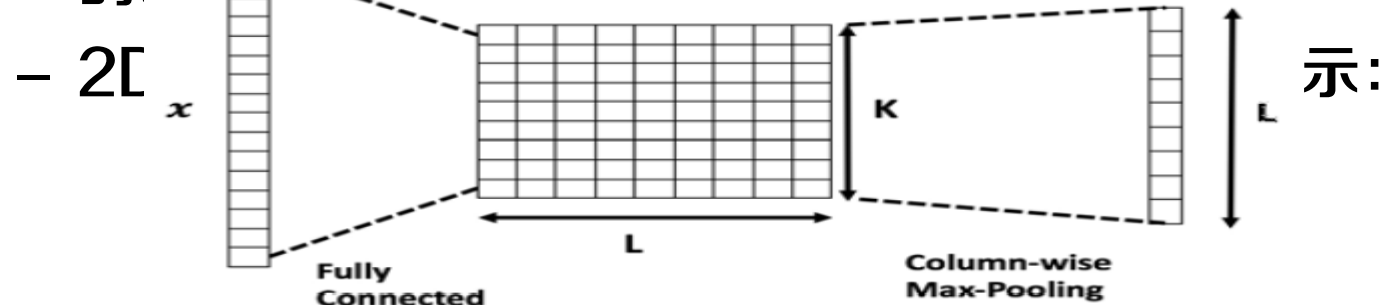
- Deep MIML Network
 - Deep MIML Network直观结构如下图所示：



- Deep MIML Network的功能
 - 1.能够自动从原始输入中产生示例集合；
 - 2.能够学习到每个示例和每个标签的对应关系；
 - 3.能够学习出对象最终所属的类别。

- **第一步：将原始数据表示成示例集合**
 - 将原始数据输入到一个示例生成装置（instance generator）中，该装置依赖特定领域。
 - 以图像任务为例，去掉最后一层全连接层的深度卷积神经网络结构（CNN）就可以作为一个示例生成装置。

- 第二步：示例集合作为3D Sub-Concept层的输入
- 2D Sub-Concept Layer for Single Instance
 - 是一种新的2维神经网络层，把它称作Sub-Concept Layer；
 - 可以为一个示例和每一个标签的Sub-Concept部分匹配分数。
 - 为了得到标签级别的预测，需要在该层进行一次列池化操作，得到一个 $L \times 1$ 的分数层，每一项代表示例 x 和相应输出标签

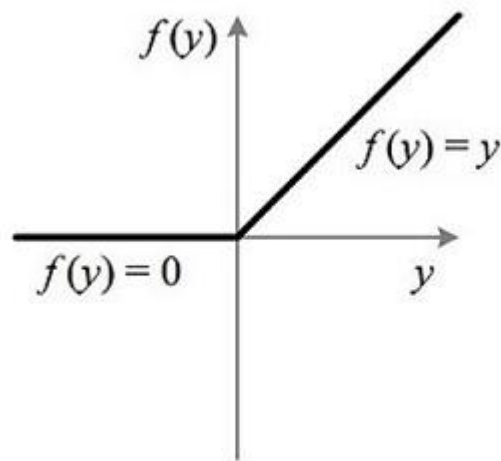


其中，在该层中的第 (i,j) 个节点表示该示例和第 j 个标签的第 i 个 sub-concept 之间的匹配分数，该节点的激活形式为：

$$c_{i,j} = f(w_{i,j}x + b_{i,j})$$

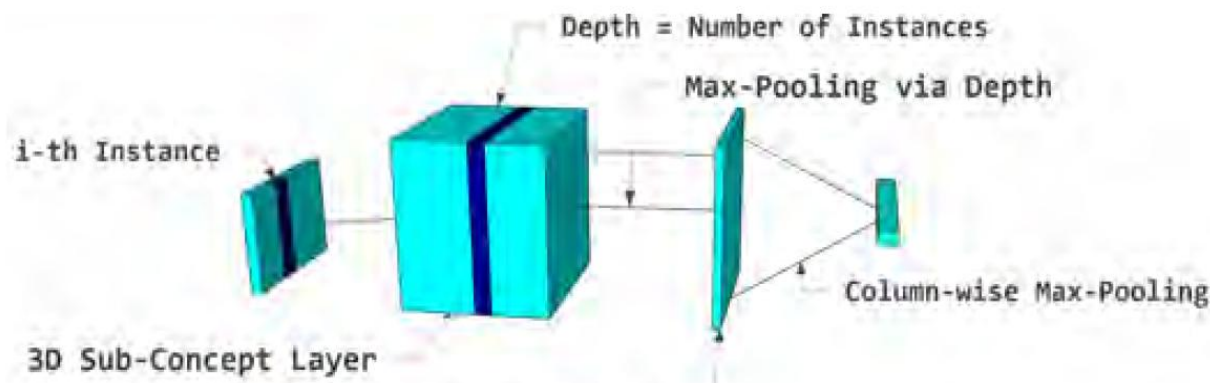
$w_{i,j}$ 是权值向量， $f(.)$ 是一种 ReLU 激活函数，其形式如下：

$$f(y) = \max$$



- 与其他常见的网络结构比较
 - 不像一般的卷积特征图（有局部连接滤波器），2D sub-concept层和输入示例向量是全连接的；激励可以表示为每个标签的sub-concept和示例的匹配得分；这里每个结点的权重是不一样的，而卷积层共享权重；
 - 2D sub-concept 层以一种可解释的方式排列，也就是说，该层的每一列都是一个标签的得分向量；

- 3D sub-concept layer for Multiple Instances
 - 该层的输入是示例集合，假定每个示例集合含有相同数量的示例；
 - 该层由多个2D sub-concept layer堆积而成，深度等于输入示例的数量；
 - 需要进行两次池化操作。
 - 3D sub-concept layer直观表示如下图所示：



- **第三步：进行两次池化操作**
 - **第一次池化操作：**在该层进行垂直池化操作，结果变成一个大小为 $L \times M$ 的2维层，其中，在位置 (i, j) 的节点表示第 i 个示例和第 j 个标签的匹配得分。我们把这层称作Instance-Label Scoring Layer。
 - **第二次池化操作：**在Instance-Label Scoring Layer进行一次水平的池化操作，结果产生一个大小为 $L \times 1$ 的一维层，其中第 j 项表示该对象和第 j 个标签的匹配分数。

- **第四步：全连接层**
 - 全连接层的大小等于标签的数量。
 - 损失函数选择的是cross-entropy;
 - 训练的过程中采用的是随机梯度下降算法（SGD）。



优劣分析

- 优点：
 - 与其他MIML算法相比较，大多数先前的MIML算法都是假设这些示例是预先给定的，或者由一些手动设计的示例生成器生成的，而**本方法是由模型本身完成示例的生成**。
 - 此外，本方法的sub-concept层可以发现并且保留输入与输出标签之间的潜在关系。
- 缺点：
 - 与比较复杂的CNN-RNN模型相比，本方法表现出了次优性能。这是因为本方法使用的示例生成器非常简单，而像停车计时器这样的小对象无法有效地编码到示例中。



应用总结

- 自然场景分类问题
- 文档分类问题
- 基于内容的图像检索
- 生物信息学等



参考文献

- [1] Feng J, Zhou Z H. Deep MIML Network[C]//AAAI. 2017: 1884–1890.
- [2] 张敏灵. 多示例学习与多标记学习的研究[D]. 南京大学, 2007.
- [3]<https://wenku.baidu.com/view/74d80328915f804d2b16c11e.html> MIML: 多示例多标记学习.
- [4]<http://blog.csdn.net/w5688414/article/details/78990301>.



道可道，非常道。名可
名，非常名。无名天地
之始。有名万物之母。
故常无欲以观其妙。常
有欲以观其徼。此两者
同出而异名，同谓之玄。
玄之又玄，众妙之门。

谢谢！

