

Support Vector Machines Based on Incomplete Privileged Information

Zhimin Zheng

School of Mathematics and Physics, China University of Geosciences, Wuhan Hubei
Email: 1105764689@qq.com

Received: Feb. 29th, 2020; accepted: Mar. 13th, 2020; published: Mar. 20th, 2020

Abstract

The importance of privileged information is particularly prominent in many fields, especially in bio-medicine. It is because researchers have proven that privileged information can improve classification accuracy. And the traditional approach usually uses SVM+ to solve the classification problem that contains privileged information. Unfortunately, in many cases the training data set contains only partial privileged information due to the high cost of privileged information collection. In the face of data with incomplete privilege information, the biggest challenge is how to use this partial privilege information to obtain the maximum hidden information of the data. Fortunately, a new fusion classification algorithm based on incomplete privileged information for Support Vector Machine (ICSVM+) that we proposed, can solve this problem well. In this paper, two ideas are proposed for the fusion classification algorithm, namely, cross-correction (CC) and linear weighting (LW). We call these two methods (CC) ICSVM+ and (LW) ICSVM+. Another important work is that a weight updating method based on window theory proposed by us can adapt well to different data sets. Through experimental comparison with SVM and SVM+, our ICSVM+ methods were able to effectively deal with the lack of privilege information.

Keywords

Privileged Information, Fusion, Weight Updating

基于部分特权信息的支持向量机

郑志敏

中国地质大学, 数学与物理学院, 湖北 武汉
Email: 1105764689@qq.com

收稿日期: 2020年2月29日; 录用日期: 2020年3月13日; 发布日期: 2020年3月20日

摘要

特权信息的重要性在许多领域尤为突出，特别是在生物医学领域。许多研究表明，特权信息可以提高分类的准确性。传统方法是使用SVM+来解决包含特权信息的分类问题。但是在许多情况下，由于特权信息收集成本高昂，训练集仅包含部分特权信息。面对特权信息不完整的数据，最大的挑战是如何使用此部分特权信息来获取数据的隐藏信息。在文章中，我们提出的基于支持向量机(ICSVM+)部分特权信息的新融合分类算法，可以很好地解决这个问题。对此融合分类算法本文提出了两个思想，分别为交叉校正(CC)和线性加权(LW)去实现此融合算法。我们称这两种方法为CC-ICSVM+和LW-ICSVM+。另一项重要工作是，我们提出的基于窗口理论的权重更新方法能够很好地适应不同的数据集。通过与SVM和SVM+的实验比较，我们的ICSVM+方法能够有效地解决特权信息的缺失问题。

关键词

特权信息，融合，权值更新

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

Vapnik (2009)提出了基于 SVM [1] [2]的新学习范式，即学习使用特权信息(LUPI) [3]，被称为基于特权信息的支持向量机(SVM+)。训练阶段提供特权信息，但不适用于测试集。Vapnik 对特权信息做了大量的研究，并且被运用在各个领域中[4] [5] [6] [7] [8]。实验结果表明，学习使用正确的特权信息有助于提高分类性能[9]。在某些情况下，特权信息很难收集而且昂贵，特别是在生物学和医学。因此，在实际数据中，有许多情况下只有部分训练样本具有特权信息[10]。由于 SVM+不适合这种情况，我们必须构建一个适应具有部分特权信息的数据集的新模型。部分特权信息类似于通常已知的缺失数据集。对于缺少的数据集，我们现阶段有许多方法完成[11] [12]。使用其他信息和特权信息的关联来完成缺失的特权信息也是一个可行的方法。这种情况更适合 Pu J 等人研究的医学数据[13]。然而，它需要获得特征和标签之间的关系，以补充特权信息的缺失数据，这可能在生物医学中具有更好的实验性能。但在其他领域，特征和标签之间的关系在特权信息中不能很好的体现。因此，这样来补充缺失的数据集可能无法获得稳定的实验结果。

我们不想对原始数据集做大量的填充，而是想要一个好的分类结果。在此情况下，Vapnik 习惯于将其放在 SVM+的模型扩展部分中，用类似于 SVM+的思想来解决它。他使用包含一些特权信息的空间来训练校正函数。与 SVM+不同，用于训练校正功能的空间非常小。但缺少特权信息的样本还没有校正。该模型的想法与 SVM+的非常相似，理论上它是 SVM+模型的较弱版本。基于此理念，我们能否使用具有特权信息的样本空间和没有特权信息的样本空间来相互校正？

事实上，如果具有特权信息的训练集小于原始训练集，则特权信息的影响会非常小。我们可以完全丢弃特权信息不使用 SVM 训练，而使用 SVM+训练。在某些情况下，我们不想丢失任何信息。例如，当具有特权信息的训练集占用原始训练集的 30%时，SVM 和 SVM+都不能实现更好的分类结果。

本文提出了一种基于部分特权信息的名为 ICSVM+的融合分类算法。实现此算法需要三个步骤。图

1 中详细介绍了实施过程。

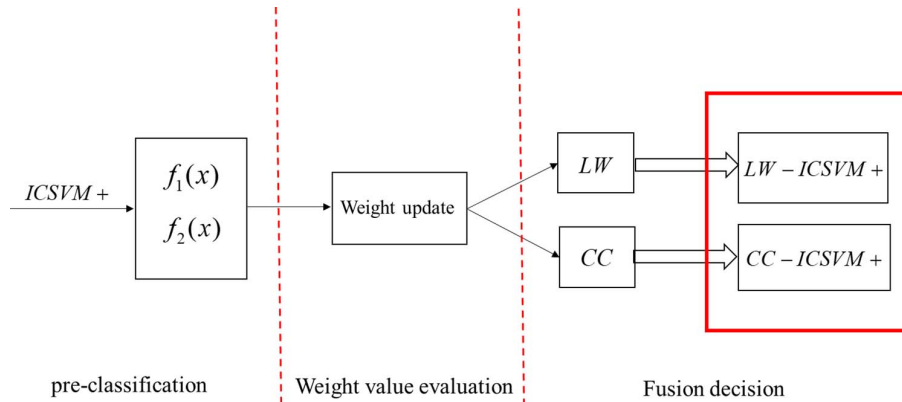


Figure 1. LW-ICSVM+ and CC-ICSVM+ implementation process. The red box is the two classification methods proposed in this paper. LW is a linear weighted fusion method and CC is a cross-correction fusion method

图 1. LW-ICSVM+和 CC-ICSVM+实施过程。红框是本文提出的两种分类方法，LW 是一种线性加权融合方法，CC 是一种交叉校正融合方法

2. 相关工作

特权信息在机器学习中得到了广泛的应用，这是 Vapnik 在 SVM+中首次提出的。后来的研究人员对特权信息在不同的模型上进行了应用[14] [15]。例如 adboost+就是将特权信息用在 adboost 算法上，以及用在决策树上也有很好的分类性能[16]。这些方法的原则是利用 X^* 特权信息产生的空间来校正原始空间 X 的信息。但是目前对于部分特权信息的数据还有待研究，因此我们将构建有特权信息样本和没有特权信息的样本之间的关系来解决这个问题。构建三个空间 X^* (特权信息)， X_1 (不含特权信息)， X_2 其中 $X_1 \cup X_2 = X$ ，利用权重融合去建立空间关系。事实上，权重融合[17] [18] [19]可以在很多方面有效地提高实验性能，如神经网络集成。

3. 算法介绍

3.1. 基于 SVM 的部分特权信息

SVM+实现了具有特权信息的数据分类问题，但没有解决部分特权信息的应用。本文提出的模型称为 ICSVM+可以更好地解决问题。当只有一部分样本有可用特权信息，其他 n 个样本没有特权信息，给定一组三维特征样本：

$$(x_1, y_1), \dots, (x_n, y_n), (x_{n+1}, x_{n+1}^*, y_{n+1}), \dots, (x_l, x_l^*, y_l), x \in X, x^* \in X^*, y \in \{-1, 1\}.$$

最小化间距由下式给出：

$$\begin{aligned} \min_{\omega, \omega^*, b, b^*} & \quad \frac{1}{2}[(w_1 \cdot w_1) + (w_2 \cdot w_2) + \gamma(w^* \cdot w^*)] + C \sum_{i=1}^n \xi_i + C^* \sum_{i=n+1}^l [(w^* \cdot z_i^*) + b^*] \\ \text{s.t.} & \quad y_i [(w_1 \cdot z_i) + b_1] \geq 1 - \xi_i, i = 1, \dots, n \\ & \quad y_i [(w_2 \cdot z_i) + b_2] \geq 1 - [(w^* \cdot z_i^*) + b^*], i = n + 1, \dots, l \\ & \quad \xi_i \geq 0, i = 1, \dots, n \\ & \quad (w^* \cdot z_i^*) + b^* \geq 0, i = n + 1, \dots, l. \end{aligned} \tag{1}$$

其中， w_1 ， w_2 ， w^* 由方程定义。

$$w_1 = \sum_{i=1}^n \alpha_i y_i z_i,$$

$$w_2 = \sum_{i=n+1}^l \alpha_i y_i z_i,$$

$$w^* = \frac{1}{\gamma} \sum_{i=n+1}^l (\alpha_i + \beta_i - C) z_i^*$$

其对偶问题表述如下:

$$\begin{aligned} \max_{\alpha, \beta} & \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ & - \frac{1}{2\gamma} \sum_{i, j=n+1}^l (\alpha_i + \beta_i - C)(\alpha_j + \beta_j - C) K^*(x_i^*, x_j^*), \\ \text{s.t.} & \sum_{i=n+1}^l (\alpha_i + \beta_i - C^*) = 0 \\ & \sum_{i=1}^n y_i \alpha_i = 0 \\ & \sum_{i=n+1}^l y_i \alpha_i = 0 \\ & \alpha_i \geq 0, \quad i = n+1, \dots, l \\ & 0 \leq \alpha_i \leq C, \quad i = 1, \dots, n. \end{aligned} \tag{2}$$

我们获得两个决策函数:

$$f_1(x) = (w_1, z) + b = \sum_{i=1}^n y_i \alpha_i K(x_i, x) + b_1, \tag{3}$$

$$f_2(x) = (w_2, z) + b = \sum_{i=n+1}^l y_i \alpha_i K(x_i, x) + b_2. \tag{4}$$

决策函数 $f_1(x)$ 是从没有特权信息的样本中获取的, 决策函数 $f_2(x)$ 是从特权信息样本中获取的。

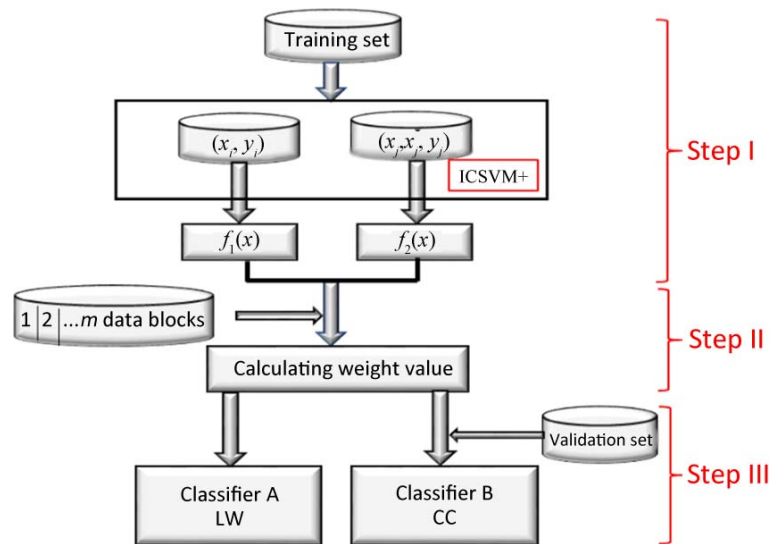


Figure 2. Process for dealing with the missing of privileged information
图 2. 处理特权信息缺失的流程

在最终获得决策算法之前，本文提出了一种基于滑动窗口的权重更新算法。它将有助于完成融合决策的两种部分特权信息的 SVM+算法，我们称之为 LW-ICSVM+和 CC-ICSVM+。图 2 显示了方法 LW-ICSVM+和 CC-ICSVM+方法的实现过程。

3.2. 权重评估

权重选择对于进行融合决策非常重要。在我们的实验中，我们发现基于滑动窗口的权重更新公式非常有效。

在训练集中，样本标签是已知的，我们将训练集划分 m 块，并表示为 $d_j (1 \leq j \leq m)$ 。对每个数据块使用上述模型，我们发现两组决策函数的值， f_1^j 以及 f_2^j 最佳权重可以通过以下公式获得：

$$q^j = \operatorname{arccmax}_{0 < q < 1} F [qf_1^j + (1 - q)f_2^j] \tag{5}$$

其中， $F[\cdot]$ 是这组决策值的准确率。

这些数据块中会有一些数据块，即使调整了参数，精度也不会很高。我们需要省略与这些数据块对应的参数。因此，我们将设计一个窗口 W ，并且窗口中的数据块数是固定的 ($W = \bigcup_{s=1}^u d_s$)。每次输入新的数据块，计算最佳参数及其相应的精度，记为 (q^j, A^j) ，其中 A^j 是精确率。比较原始窗口中权重的精度，并删除精度最低的相应权重(如果精度相同，则随机删除一个权重)，最后平均值 q 将作为最终权重。权重更新的步骤在算法 1 中给出。在图 3 中，我们显示了原始窗口与新窗口之间的关系：

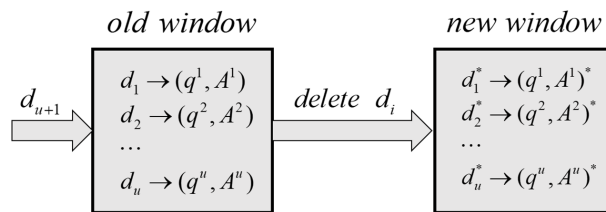


Figure 3. Weight value evaluated by a slide window
图 3. 滑动窗口计算的权重值

窗口大小为 u ，下一步我们的新窗口将成为旧窗口。

3.3. 融合方法

3.3.1. 线性加权融合法

线性加权融合获得的新决策函数为：

$$f(x) = qf_1(x) + (1 - q)f_2(x)$$

等价于：

$$f(x) = q \sum_{i=1}^n y_i \alpha_i K(x_i, x) + (1 - q) \sum_{i=n+1}^l y_i \alpha_i K(x_i, x) + (1 - q)b_2 + qb_1 \tag{6}$$

Algorithm 1

Input : $d = (d_1, \dots, d_m) (m \geq 2)$, Window capacity u

Output : q

$W \leftarrow \phi$

for $j = 1$ to u do

Continued

$$f_1^j, f_2^j \leftarrow \text{using ICSVM} +$$

$$q^j = \arg \max_{0 < q < 1} F[qf_1^j + (1-q)f_2^j] \quad \textcircled{1}$$

$$A^j = F[q^j f_1^j + (1-q^j) f_2^j] \quad \textcircled{2}$$

$$W \leftarrow W \cup (d_j, q^j, A^j)$$

end for

for $j = u + 1$ to m do

$$(d_j, q^j, A^j) \leftarrow \text{using } \textcircled{1} \text{ and } \textcircled{2}$$

$$A^{out} \leftarrow W \cup (d_j, q^j, A^j) \text{ with minimum } A^k$$

$$W \leftarrow W \cup (d_j, q^j, A^j) - (d_{out}, q^{out}, A^{out})$$

end for

$$q \leftarrow \text{got the average from } W$$

3.3.2. 交叉校正融合

在半监督学习中，有一类自训练算法[20]。该算法将测试集的置信部分返回到训练集进行再训练，直到所有测试集都分类。运用这个思想将测试集转换为训练集，提出了交叉校正融合方法。该方法可以得出一系列伪标签，删除不好的分类结果，保留较好的分类结果。与 Vapnik 的想法不同，我们的方法是双向相互修正，而不是单向修正。交叉校正融合方法有两个步骤：分配步骤和校正步骤。

标签预分配：

在 ICSVM+模型中，我们不仅得到两组决策值，而且得到两组标签 L_1 和 L_2 。以 L_2 为对象，我们可以获取 n 组伪标签 $L_2^1, L_2^2, \dots, L_2^n$ (算法 2 中显示了算法步骤)。

Algorithm 2

Input : q, L_1, L_2
 Output : ξ
 $\xi \leftarrow \emptyset$
 for $i = 1$ to n do
 $l = \text{selected different elements in } L_1 \text{ and } L_2$
 $l_1^i, \dots, l_{[q|l]}^i \leftarrow \text{randomly selected } [q, |l|] \text{ elements from } l$
 $[l_1^i, \dots, l_{[q|l]}^i] = -[l_1^i, \dots, l_{[q|l]}^i]$
 $L_2^i \leftarrow \text{return } [l_1^i, \dots, l_{[q|l]}^i] \text{ to } L_2$
 $\xi \leftarrow \xi \cup L_2^i$
 end for

标签校正

在监督学习过程中，通常数据集分为训练集和测试集，训练集用于生成分类器，测试集用于测试分类器的效果。在这里，我们添加一个验证集。从理论上讲，完整的数据集可以达到良好的分类效果，但当训练集有错误的数据(训练标签错误)时，测试集会得到一个坏的标签，从而降低分类精度。在这个问题上，我们提出了交叉校正融合方法(CC)，以选择 $(L_2^1, L_2^2, \dots, L_2^n)$ (上一步)中的最佳标签。

CC 通过以下步骤实现：

步骤 1：将 L_2^i 分配方法中的标签和测试集合并到新的训练集中，记为 D_{train}^i 。

步骤 2：从验证集中选择 k 个样本为新测试集，该测试集记为 D_{test} ，其标签记为 L_{test} 。

步骤 3: 由于分配步骤的测试集没有特权信息, 我们使用 SVM 来训练训练集 D_{train}^i 和测试集, D_{test} 并将预测标签标记为 L_{test}^* 。然后, 将步骤3中的标记与步骤2中的标记进行比较, 并将获得的准确率标记为 a_i 。

步骤 4: 重复上述步骤以获取对应于 n 个伪标签 $L_1^1, \dots, L_2^1, \dots, L_2^n$ 的 n 个准确率 $a_1, \dots, a_i, \dots, a_n$ 。选择准确率 $a_j = \max_{1 \leq i \leq n} a_i$, 其相应的标签 L_2^j 是最佳标签。

关于 CC 的实现步骤, 我们在算法 3 中给出了。在 CC 中, 我们没有在分配步骤中将测试集添加到具有已知标签的训练集中, 这与半监督学习的自训练算法不同, 因此很难选择最佳标签。同样, 我们试图在步骤 2 中放大 k 的值以放大精度之间的差异并选择最佳标签。因此, 在 CC 实现过程中, 数据集实际上是分批测试的, 输入数据块不应太小, 这样会降低 SVM 的精度, 也不应该太大, 太大会增加计算量。

Algorithm 3

```

Input :  $D_{test}^1, L_1, L_2$ 
Output :  $L_2^*$ 
 $\xi \leftarrow$  using the Algorithm 2
 $q \leftarrow$  using the Algorithm 1
 $a \leftarrow \phi$ 
for  $i = 1$  to  $n$  do
     $D_{train}^i \leftarrow$  add  $L_2^i$  to  $D_{test}^1$ 
     $D_{test} \leftarrow$  selected  $k$  samples from  $D - D_{test}^1$ 
     $a_i \leftarrow$  utilizing SVM got a accuracy from  $D_{train}^i \cup D_{test}$ 
     $a \leftarrow a \cup a_i$ 
end for
 $\tau = \arg \max_{1 \leq i \leq n} a_i$ 
 $L_2^* \leftarrow$  got the optional lable set
    
```

3.4. 数据块大小控制

在 CC-ICSVM+中, 我们需要选择最佳标签。预测标签的数量是分配步骤中的随机次数。因此, 为了节省计算量, 我们将尝试减少上述算法中的 n 。通过随机选择, 如果所有可能的组合都取到, 则一定可以得到最优标签。在某些情况下, 获得最优标签的成本非常高。实际上, 我们也可以选择次优标签。如果输入的测试数据块较大, 这样导致随机次数也会增加, 因此可以使用批量处理测试来节省计算量。从以下不等式可以看出, 对较大测试集的批量处理测试将节省计算量。

$$C_{|k_1|+|k_2|}^{n_1+n_2} > C_{|k_1|}^{n_1} + C_{|k_2|}^{n_2} \tag{7}$$

4. 实验设置

4.1. 实验准备

从 UCI 数据集中选出了八个数据集: Wine data set, seeds data set, Pima Indian Diabetes data set, Ionosphere data set, Abalone data set, statlog (heart) data set, Breast cancer Wisconsin data set, Liver Disorders data set。8 个数据集的样本数量不同, 为了统一化, 我们从 8 个数据集中提取了部分数据集且大小近似, 并将数据集的这一部分划分为训练集、测试集和验证集。训练集与测试集的比率为 7:3。此外, 将训练集分为两部分: 特权信息和非特权信息也就是普通特征, 然后分析特权信息样本比例的准确性影响。表 1 提供了八项实验数据的详细信息。

Table 1. Information about experimental data**表 1.** 有关实验数据的信息

数据集	原数据集大小	实验数据大小	特征个数	特权信息个数
Wine	178 * 14	116 * 5	4	1
seeds	210 * 8	200 * 7	5	2
pimaindiabetes	768 * 9	200 * 8	5	3
lonosphere	351 * 35	200 * 34	26	8
Abalone	4177 * 9	200 * 6	5	1
statlog (heart)	270 * 14	200 * 13	9	4
Breast cancer wisconsin	569 * 32	200 * 30	26	4
liverdisorders	345 * 7	200 * 6	5	1

在实验过程中，我们选择了 4 种方法：SVM、SVM+、LW-ICSVM+、CC-ICSVM+。在前两种方法中，数据无法充分利用。包含特权信息和没有包含特权信息的样本将分别在 SVM 和 SVM+ 中丢弃。值得注意的是，由于特权信息有助于分类，因此在真实数据中可能会丢失更多特权信息。尽管 SVM 训练样本比 SVM+ 训练样本多，但准确率不一定高于 SVM+。

4.2. 实验结果

在 LW-ICSVM+ 中，我们使用了线性加权融合方法，而在 CC-ICSVM+ 中，我们使用了交叉校正融合方法。在表 2 中，当特权信息争用占训练总量的 30% 时，我们分别使用四种方法获得八组数据的准确率。在实验过程中，30% 的特权信息在训练集中随机提取，因此训练集中的部分示例不包含特权信息。在这里，我们随机抽取了 3 次，最终准确率被平均。我们可以看到，对于八组数据集，LW-ICSVM+ 和 CC-ICSVM+ 方法均高于 SVM 和 SVM+，尽管 SVM 和 SVM+ 在某些情况下工作良好。对于 Wine data set, seeds data set, LW-ICSVM+ 方法优于 CC-ICSVM+。原因有两个：首先，因为 CC-ICSVM+ 对两个分类器进行相互更正。如果一个分类器的所有训练点的错误也是另一个分类器的训练点错误，则 CC-ICSVM+ 方法将不能得到准确的结果。其次，由于这两组数据集的样本数量几乎很少，而且新的测试集我们在校正步骤中产生的测试组也相对较小。因此，它导致校正步骤的精度差别很小，并且无法明确选择最佳标签。

Table 2. Experimental accuracy of eight groups of data**表 2.** 八组数据的实验准确率

数据集	SVM	SVM+	LW-ICSVM+CC-ICSVM+
Wine	86.11	91.67	97.22
seeds	90	87.5	93.75
pimaindiabetes	71.66	70	74.17
lonosphere	84.33	90	91.67
Abalone	90	81.67	91.67
heart	88.33	85	86.63
Breast	95	91.67	94.17
Liverdisorders	66.67	61.67	75

对于方法 CC-ICSVM+中的校正步骤, 我们知道每组预测标签对应的准确率, 对应的最高准确率的标签是最佳标签。表3显示了 Breast Cancer Wisconsin and Ionosphere datasets 中每个预测标签 L_2^i 的准确率。在分配步骤中, 我们取 $n = 5$, 获取 5 组伪标签, d_2, c_1 为最佳标签。

Table 3. The accuracy a_i of each prediction label L_2^i

表 3. 每个预测标签 L_2^i 的准确率 a_i

Breast Cancer Wisconsin		Lonosphere	
pseudo label	Accuracy (%)	pseudo label	Accuracy (%)
d_1	80	c_1	74.28
d_2	85	c_2	70.71
d_3	84.2	c_3	70.71
d_4	70	c_4	67.14
d_5	75.7	c_5	71.21

5. 结论

对于训练集只包含部分特权信息的情况, 以及几种研究提出的不同的方法, 但其性能和结果并不高。我们提出了 LW-ICSVM+和 CC-ICSVM+算法, 有效地解决了特权信息不完整的问题。在实现过程中, 我们采用了自适应权重更新公式的思想。实验表明, 与传统 SVM 和 SVM+相比, 该方法的准确率显著提高。在比较 LW-ICSVM+与 CC-ICSVM+时, 我们发现虽然后者的准确率高于前者, 但前者的复杂性更低。因此, 如果分类精度不需要很高, 可以选择 LW-ICSVM+。最后, 对于算法测试的复杂度上, 在 CC-ICSVM+中提出了对测试集的分区, 以控制算法的计算量。

在未来的工作中, 我们计划控制 CC-ICSVM+的计算量, 合理控制测试集数据块的大小, 对数据块的大小与分类准确率之间的关系进行实验研究, 从而获得能够保证在拥有较少的计算量的情况下产生较高的精度。

参考文献

- [1] Cortes, C. and Vapnik, V. (1995) Support-Vector Networks. *Machine Learning*, **20**, 273-297. <https://doi.org/10.1007/BF00994018>
- [2] Keerthi, S.S., Gilbert, et al. (2002) Convergence of a Generalized SMO Algorithm for SVM Classifier Design. *Machine Learning*, **46**, 351-360. <https://doi.org/10.1023/A:1012431217818>
- [3] Vapnik, V. and Vashist, A. (2009) A New Learning Paradigm: Learning Using Privileged Information. *Neural Networks*, **22**, 544-557. <https://doi.org/10.1016/j.neunet.2009.06.042>
- [4] Shiao, H.T. and Cherkassky, V. (2014) Learning Using Privileged Information (LUPI) for Modeling Survival Data. *International Joint Conference on Neural Networks*, Beijing, 6-11 July 2014, 1042-1049. <https://doi.org/10.1109/IJCNN.2014.6889517>
- [5] Yang, X., Wang, M. and Tao, D. (2018) Person Re-Identification with Metric Learning Using Privileged Information. *IEEE Transactions on Image Processing*, **27**, 791-805. <https://doi.org/10.1109/TIP.2017.2765836>
- [6] Motiian, S., Piccirilli, M., Adjeroh, D.A., et al. (2016) Information Bottleneck Learning Using Privileged Information for Visual Recognition. *Computer Vision & Pattern Recognition*, Las Vegas, 27-30 June 2016, 1496-1505. <https://doi.org/10.1109/CVPR.2016.166>
- [7] Lapin, M., Hein, M. and Schiele, B. (2014) Learning Using Privileged Information: SVM+ and Weighted SVM. *Neural Networks*, **53**, 95-108. <https://doi.org/10.1016/j.neunet.2014.02.002>
- [8] Feyereisl, J. and Aickelin, U. (2012) Privileged Information for Data Clustering. *Information Sciences*, **194**, 4-23.

-
- <https://doi.org/10.1016/j.ins.2011.04.025>
- [9] Pechyony, D. and Vapnik, V. (2010) On the Theory of Learning with Privileged Information. *Advances in Neural Information Processing Systems 23: 24th Annual Conference on Neural Information Processing Systems 2010*, Vancouver, 6-9 December 2010, 1894-1902.
- [10] Liu, D., Liang, D. and Wang, C. (2015) A Novel Three-Way Decision Model Based on Incomplete Information System. *Knowledge-Based Systems*, **91**, 32-45.
- [11] Emmanuel, J. and Candes, R.B. (2009) Exact Matrix Completion via Convex Optimization. *Foundations of Computational Mathematics*, **9**, 717-772. <https://doi.org/10.1007/s10208-009-9045-5>
- [12] Liu, J., Musialski, P., Wonka, P., et al. (2009) Tensor Completion for Estimating Missing Values in Visual Data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **35**, 208-220.
- [13] Pu, J., Wang, J., Zheng, Y., et al. (2017) Boosting Alzheimer Diagnosis Accuracy with the Help of Incomplete Privileged Information. *IEEE International Conference on Bioinformatics and Biomedicine*, Kansas City, 13-16 November 2017, 595-599. <https://doi.org/10.1109/BIBM.2017.8217718>
- [14] Fouad, S., Tino, P., Raychaudhury, S., et al. (2012) Learning Using Privileged Information in Prototype Based Models. *Proceedings of the 22nd International Conference on Artificial Neural Networks and Machine Learning*, Volume Part II, 322-329. https://doi.org/10.1007/978-3-642-33266-1_40
- [15] Wang, Z. and Qiang, J. (2015) Classifier Learning with Hidden Information. *Computer Vision & Pattern Recognition*, Boston, 7-12 June 2015, 4969-4977.
- [16] Chen, J., Liu, X. and Lyu, S. (2012) Boosting with Side Information. In: *Computer Vision ACCV 2012*, Springer, Berlin Heidelberg, 563-577.
- [17] Lin, H., Lin, Y., Yu, J., et al. (2014) Weighing Fusion Method for Truck Scales Based on Prior Knowledge and Neural Network Ensembles. *IEEE Transactions on Instrumentation & Measurement*, **63**, 250-259. <https://doi.org/10.1109/TIM.2013.2278577>
- [18] Ye, L.Y., et al. (2014) Multi-Sensor Weighted Data Fusion Method Using LMS Algorithm. *Computer Engineering & Applications*, **50**, 86-90.
- [19] Xu, X.L. and Tang, J.F. (2006) A New Sequential Weighed Fusion Method with Colored Noise and Time Delay. *2006 International Conference on Machine Learning and Cybernetics*, Dalian, 2006, 1879-1884. <https://doi.org/10.1109/ICMLC.2006.259055>
- [20] Mohammadpour, P., Sharifi, M. and Paikan, A. (2008) A Self-Training Algorithm for Load Balancing in Cluster Computing. *International Conference on Networked Computing & Advanced Information Management*, Gyeongju, 2-4 September 2008, 104-109. <https://doi.org/10.1109/NCM.2008.178>