类 dropout 的具有新型栈式结构的层次支持向量机

卢 剑 伟

(常州工业职业技术学院信息工程与技术学院 江苏 常州 213164)
(东南大学信息科学与工程学院 江苏 南京 211189)

摘 要 基于对抗学习提出一种类 dropout 的具有新型栈式结构的层次支持向量机(D-S-SVM)。随机抽取一定比例的样本攻击其标签类型使其成为对抗样本,利用支持向量机对包含对抗样本的训练集进行对抗学习生成对抗支持向量机(A-SVM)。通过栈式结构原理逐层级联一定数量的子分类器(即 A-SVM)构建 D-S-SVM。在该模型中计算子分类器输出误差对输入样本的一阶梯度信息,并结合 dropout 将部分一阶梯度信息嵌入到原输入样本特征中生成新样本作为下一个子分类器的输入。该模型不仅提供了一种新颖的层次结构级联方式,且实验结果表明它能够逐层提高数据分类精度且具有较强的泛化性能。

关键词 支持向量机 对抗样本 对抗学习 堆栈结构原理 dropout 分类算法

中图分类号 TP391 **文献标志码** A DOI: 10.3969/j.issn.1000-386x.2021.02.044

DROPOUT-LIKE HIERARCHICAL SUPPORT VECTOR MACHINE WITH NOVEL STACKED STRUCTURE

Lu Jianwei

(*School of Information Technology and Engineering, Changzhou Institute of Industry Technology, Changzhou 213164, Jiangsu, China*)
(*School of Information Science and Engineering, Southeast University, Nanjing 211189, Jiangsu, China*)

Abstract A dropout-like hierarchical support vector machine (SVM) with novel stacked structure (D-S-SVM) based on adversarial learning is proposed in this paper. The adversarial attacks were carried out on a certain percentage of training samples, which were randomly selected from the given training dataset, and the adversarial samples can be obtained. By training the support vector machine (SVM) on the training dataset including the adversarial samples, the resultant adversarial support vector machine (A-SVM) could be generated. Take the A-SVMs as sub-classifiers in a deep learning model, the D-S-SVM could be constructed by stacking a certain number of sub-classifiers based on the stacked generalization principle. In the D-S-SVM, the first-order gradient information corresponding to the output error of the sub-classifier in current layer with respect to the features of all inputs were calculated, and then they were integrated with the ideology of the dropout. The inputs were updated by embedding the resultant first-order gradient information such that the updated inputs could be taken as the inputs of the sub-classifier in next layer. The D-S-SVM provides a novel stacked way and extensive experimental results demonstrate that it can improve the classification precision in a layer-by-layer manner and has better generalization performance.

Keywords Support vector machine Adversarial samples Adversarial learning Stacked structure principle Dropout Classification algorithms

---

收稿日期: 2019-07-26。**江苏高校"青蓝工程"项目;常州工业职业技术学院新一代信息技术团队项目**(YB201813101005);**校企横向课题**(B018108)。**卢剑伟,副教授,主研领域:智能信息处理。**

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

0 引　言

对抗样本已被证明存在于大部分真实数据集中，且对抗样本的本质是数据集中容易被忽视的噪声或扰动样本[1-6]。一方面，对抗样本的本质决定了对抗样本会降低分类算法的分类性能; 另一方面，合理的对抗样本学习可有效地提高分类算法的分类性能，能够较好地应用于脑电信号识别[1]、手写体识别[2]、车辆识别[3]等。对抗样本学习的一般方式是将少量的噪声或扰动样本直接加入到数据集中[4]，或者对样本特征本身进行较小幅度的改动[1,5]，继而利用相应的分类算法进行学习以提高算法的分类性能。不同于对抗样本学习的一般方式，本文将尝试修改一定比例样本的标签类型使之成为对抗样本进行对抗学习以提高所提分类算法的分类性能。

神经网络[1,4]在人工智能、机器学习以及深度学习等领域展示着强大的样本学习能力，已被广泛应用于图像处理、人工智能信息处理、无人驾驶技术、自适应智能控制等诸多实际应用[7-12]。神经网络通过前向传播和反向传播算法不断优化隐藏层中的参数使之达到最佳权值来提高整体网络的学习性能，尤其是网络的泛化性。神经网络的层次结构模型使其能够更好地表示样本特征以及具备更强的函数模拟能力[11]。具有更多隐藏层数的神经网络能够更好地模拟神经网络模型与真实样本特征之间的关系，但却会带来较为严重的过拟合现象[13-14]。因此，文献[13]提出 dropout 技术来解决因需要神经网络增加隐藏层数大幅提升网络模型学习能力过程中产生的过拟合问题。在使用 dropout 技术的过程中，隐藏层中的节点在训练时将会以一定的概率被移出隐藏层，以这种方式来提高神经网络的泛化性。

受神经网络层级结构模型启发，本文选择一种成熟且性能较好的分类器——支持向量机(Support Vector Machine，SVM)[15-18]作为子分类器，且基于堆栈结构原理[19]和对抗样本学习[1-6]构建一种具有新型栈式结构的层次结构支持向量机(D-S-SVM)。通过对抗样本学习，每个子分类器输出误差对输入的一阶梯度信息将被嵌入到原输入特征中以此更新原输入。更新后的输入将被作为下一个子分类器的输入，通过堆栈结构原理逐层更新每个子分类器的输入，逐层提高每个子分类器的分类性能。特别地，在逐层更新原输入样本的过程中引入 dropout，即将每个子分类器输出误差对输入样本部分特征的一阶梯度信息嵌入到原输入特征中，以提高所提出的分类模型 D-S-SVM 的泛化性。

1 层次支持向量机

1.1 对抗学习

不同于对抗样本学习的一般方式，即直接将噪声或扰动样本添加到数据集中，或者对样本特征本身做小幅度修改。本文首先从训练集中随机选取一定比例的样本，修改选定样本的标签类型使之成为对抗样本，即用标签集中的其他标签代替所选定样本的真实标签。假设给定数据集 $\boldsymbol{X}=\{\boldsymbol{x}_i \mid \boldsymbol{x}_i \in \mathbf{R}^d, i=1,2,\cdots,N\}$，与之对应的真实标签集 $\boldsymbol{Y}=\{y_i \mid y_i \in \mathbf{R}, i=1,2,\cdots,N\}$。对于二分类问题，即 $y_i \in \{-1,1\}$，那么第 i 个样本 $\boldsymbol{x}_i$ 的标签 y_i 将从 1 变为 -1 或从 -1 变为 1; 对于多分类问题，即 $y_i \in \{1,2,\cdots,C\}$，那么第 i 个样本 $\boldsymbol{x}_i$ 的标签 y_i 将从当前标签类型变为 $\{1,2,\cdots,C\}-\{y_i\}$ 中的任意一个标签，记为 $\tilde{y}_i$。基于成熟且分类性能较好的 SVM[15-18]，本文将通过 SVM 训练包含对抗样本的数据集进行对抗样本学习，通过训练生成的 SVM 模型输出误差更新原样本特征。对抗样本学习具体过程描述如下:

SVM 的分类函数可表示为:

$$f(\boldsymbol{x};\boldsymbol{w},b)=\operatorname{sign}(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}+b) \tag{1}$$

式中: $\operatorname{sign}(\cdot)$ 为符号函数，当表达式大于 0 时 $\operatorname{sign}(\cdot)=1$，否则 $\operatorname{sign}(\cdot)=0$; $\boldsymbol{w}\in\mathbf{R}^d$ 及 $b\in\mathbf{R}$ 分别代表线性/非线性分类情况下的直线/超平面的法向量和截距，旨在寻找不同类数据的最大划分间隔。由式(1)可知，SVM 分类函数对样本特征本身敏感，意味着 SVM 输出对输入样本蕴含着丰富的梯度信息。对于样本 $\boldsymbol{x}_i$，假设训练包含对抗样本的数据集生成的 SVM 模型输出为 $\bar{y}_i$，输出误差记为 E，即:

$$E=(\bar{y}_i-y_i)^2 \tag{2}$$

输出误差对样本 $\boldsymbol{x}_i$ 的一阶梯度可计算为:

$$\frac{\partial E}{\partial \boldsymbol{x}_i}=2(\bar{y}_i-y_i) \tag{3}$$

进而可具体地计算输出误差 E 对于样本 $\boldsymbol{x}_i$ 中每个特征的一阶梯度组成一阶梯度向量 $\left(\frac{\partial E}{\partial x_{i1}},\frac{\partial E}{\partial x_{i2}},\cdots,\frac{\partial E}{\partial x_{id}}\right)$。

本文将基于对抗样本学习的一阶梯度信息嵌入到

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

原输入样本中以更新每个原输入样本，即：

$$x'_i = x_i + \gamma \frac{\partial E}{\partial x_i} \tag{4}$$

式中：参数 γ 代表特征学习率[20]，其值的大小会影响样本特征在特征空间的真实意义，较大的 γ 值会严重影响每个样本在特征空间中的真实位置关系，因此，必须根据真实实验效果结合交叉验证[21]的方法来确定 γ 值。当输出误差 E 对所有样本的一阶梯度计算完后便可得到基于对抗样本学习的 $N \times d$ 维一阶梯度信息矩阵 $\boldsymbol{G}$，即：

$$\boldsymbol{G}(i,j) = \frac{\partial E}{\partial x_{ij}} \tag{5}$$

式中：x_{ij} 代表数据集中第 i 个样本的第 j 个特征且有 $1 \leqslant j \leqslant d$。

同样，利用一阶梯度信息矩阵更新原数据集可得到更新后的数据集，即：

$$X' = X + \gamma \boldsymbol{G} \tag{6}$$

由 SVM 的分类函数(式(1))可知式(4)中的一阶梯度信息对样本特征本身敏感，这种敏感特性决定了本文利用对抗样本学习更新原样本特征进而提高所提分类模型的有效性。与基于训练原样本的 SVM 模型相比，基于训练更新后样本的 SVM 模型性能将会得到有效提高。

1.2 基于栈式结构的层次支持向量机

神经网络的层次结构特点使得神经网络能够较好地模拟网络模型与样本特征之间的关系。而且，随着网络中隐藏层数以及每一层中节点数的增加，神经网络的分类性能将会变得更好。类似地，本文将通过基于对抗样本学习的一阶梯度信息构造一种基于栈式结构的层次支持向量机，如图 1 所示。图 1 中：X 代表给定的真实数据集；Y 为与 X 相对应的真实标签集；$\overline{Y}_K$ 为层次支持向量机的最终输出标签集。

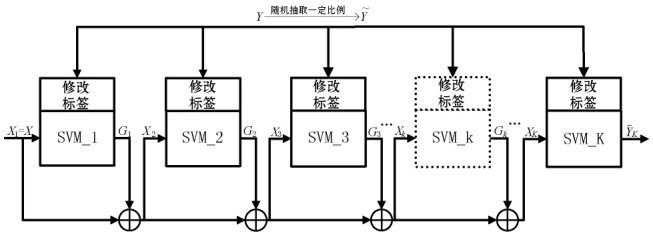


图 1 具有栈式结构的层次支持向量机

基于这一栈式结构[19]，层次支持向量机由 K 个基于对抗样本学习的 SVM 级联而成。对于第 1 层，首先从数据集 X 中随机抽取一定比例的样本，修改选定样本的标签类型，即将选定样本的标签类型修改为标签集中的其他标签类型，使选定的样本成为对抗样本。其次，利用 SVM 算法对包含对抗样本的数据集进行训练生成 SVM_1 模型，计算 SVM_1 的输出误差对所有输入样本 $X_1 = X$ 的一阶梯度信息矩阵 $\boldsymbol{G}_1 = X_1 + \gamma \frac{\partial E}{\partial X_1}$。最后，将 $\boldsymbol{G}_1$ 嵌入到数据集 X_1 中以更新 X_1 得到数据集 X_2，即 $X_2 = X_1 + \boldsymbol{G}_1$，且将 X_2 作为第二层中 SVM_2 的训练集。如此循环操作，直到层次支持向量机取得最佳分类效果或者达到最大层数值 K。

与神经网络模型层次结构相比，该层次支持向量机中的每一层 SVM_k 可类比于神经网络模型中的输入输出层/隐藏层，对每个原输入样本所有特征的一阶梯度信息可类比于神经网络模型输入输出层/隐藏层中的节点。特别地，本文提出的层次支持向量机前后两层由基于对抗样本学习的一阶梯度信息级联而成，这为有监督学习提供了一种新的实现方式。

2 D-S-SVM

第 1 节利用基于对抗样本学习的一阶梯度信息构建了一种具有新型栈式结构的层次支持向量机。然而，在利用一阶梯度信息更新每一层输入样本作为下一层输入样本的过程中涉及到 SVM_k 输出误差对每条输入样本中所有特征的一阶梯度。类似地，当增加神经网络的隐藏层数或增加每一隐藏层中的节点数时，更多的模型参数极易导致神经网络模型的过拟合现象。具体而言，在训练神经网络实施反向传播算法的过程中，随机选择每一隐藏层中的部分节点参数根据神经网络的输出误差进行修正，未被选择的节点参数将保持之前的状态。

根据 dropout 技术，本文在利用基于对抗样本学习的一阶梯度信息更新层次支持向量机中每一个SVM_k 输入过程中，随机选择每一个输入样本部分特征利用式(4)进行更新，其他未被选择的样本特征保持原有的特征值，以此逐层更新层次支持向量机的每一层输入并实现提高每一层的分类性能。图 2 展示了基于对抗样本学习的类 dropout 输入样本更新过程。其中，所提类 dropout 的具有新型栈式结构的层次支持向量机(D-S-SVM)具有 K 层，样本 $x_i = (x_{i,1}, x_{i,2}, \cdots, x_{i,12})$ 为一个 12 维的特征向量，将随机选择样本 x_i 的一半特

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

征(如图 2 中灰色圆形,实验中可根据实际分类效果进行确定) 利用式(4) 进行更新,其他样本特征保持不变(如图 2 中白色圆形) ,$\bar{y}_i$ 为 D-S-SVM 的最终输出。

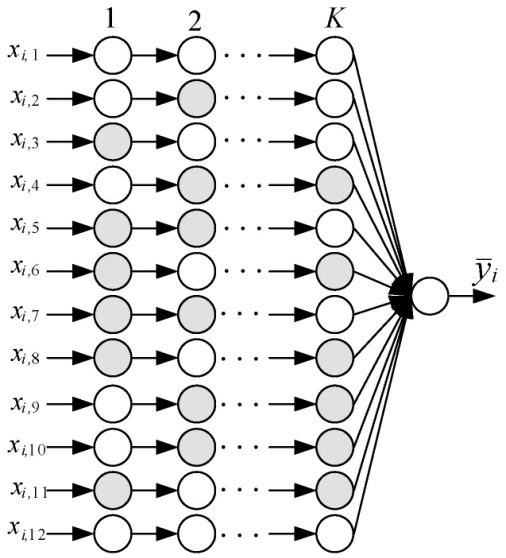


图 2　类 dropout 的层次支持向量机

对于神经网络,在每一次执行反向算法的过程中,神经网络模型隐藏层中的所有节点参数将被逐渐优化到最佳状态,神经网络模型将有效模拟模型与样本特征之间的关系。结合 dropout 技术,神经网络模型的过拟合问题也将得到有效解决。类似地,在运行 D-S-SVM的过程中,利用基于对抗样本学习的一阶梯度信息逐层更新每一层输入样本,使得不同类样本在特征空间中逐渐被隔开,因此,D-S-SVM 的分类性能逐渐被提高。同时,由于在逐层更新输入样本的过程中引入了 dropout 技术,D-S-SVM 的泛化性也将逐渐被提高。

3　实　验

由于所提 D-S-SVM 分类模型中的每一子分类器基于 SVM 实现,因此实验中将主要讨论线性(Linear) 和高斯核(Gaussian) 情况下的分类器分类性能。实验主要针对 SVM 和 D-S-SVM 的实际分类性能进行对比来验证利用基于对抗样本学习的一阶梯度信息构建具有新型栈式结构的层次分类模型的有效性。另外,将讨论 D-S-SVM 精简版(记为 D-S-SVM_0) 的分类性能,即在利用基于对抗样本学习的一阶梯度信息逐层更新输入样本的过程中不引入 dropout 技术,以此来突出 D-S-SVM 的泛化性。

3.1　实验数据集

表 1 详细列出了实验中选择的真实数据集,所选择的数据集均可从 UCI [22] 或 KEEL [23] 网站下载。

表 1　真实数据集描述

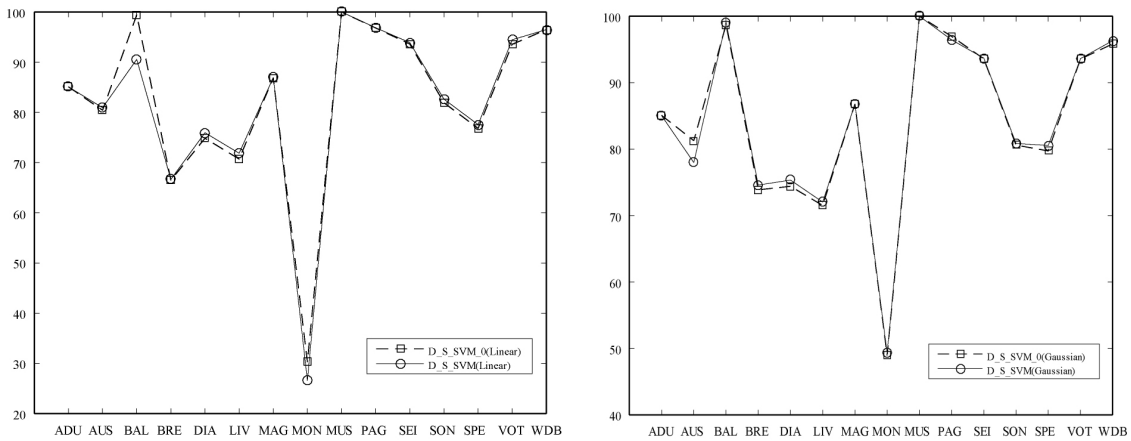
数据集	样本数	特征数
Adult(ADU)	48 841	14
Australian(AUS)	690	14
Balance(BAL)	625	4
Breast(BRE)	277	9
Diabetes(DIA)	768	8
Liver(LIV)	345	6
Magic(MAG)	19 020	10
Monks1(MON)	432	6
Mushroom(MUS)	8 124	21
Page_blocks(PAG)	5 473	10
Seismic_bumps(SEI)	2 584	18
Sonar(SON)	208	60
Spectfheart(SPE)	267	44
Vote(VOT)	435	16
Wdbc(WDB)	569	30

表 2 详细展示了各个分类算法在真实数据集上的测试精度,其中: 对于线性情况下的三种分类算法的正则化参数 c 的搜索范围为$\{10^{-5},10^{-4},\cdots,10^{4},10^{5}\}$; 对于高斯核情况下的三种分类算法的正则化参数 c 的搜索范围与线性情况下相同,高斯核宽度 σ 的搜索范围为$\{10^{-5},10^{-4},\cdots,10^{4},10^{5}\}$。对于每一个真实数据集,将随机选择 60% 的样本作为训练样本,其余作为测试样本,测试结果为运行 10 次后取得的平均分类性能。另外,D-S-SVM 共有 5 层,除了给出取得最佳分类性能情况下的参数 c 和 σ,还具体给出了 D-S-SVM 在第 k 层取得最佳分类性能。D-S-SVM 在执行 dropout 的过程中,将利用基于对抗样本学习的一阶梯度信息对每一个样本一半的特征进行更新。在执行 D-S-SVM_0 以及 D-S-SVM 过程中,将随机修改 5% 比例样本的标签使其成为对抗样本进行对抗学习,且根据文献[20],D-S-SVM_0 以及 D-S-SVM 中的样本特征学习率选择推荐值 $\gamma = 0.001$,两者的分类性能比较如图 3 所示。表 2 中已将对比算法最好的分类性能用粗体表示。

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

表 2 几种分类算法的详细分类性能

数据集	SVM(Linear) (c)	D-S-SVM_0(Linear) (c,k)	D-S-SVM(Linear) (c,k)	SVM(Gaussian) (c, σ)	D-S-SVM_0(Gaussian) (c,σ,k)	D-S-SVM(Gaussian) (c,σ,k)
ADU	85.08±0.156 6 (10^5)	85.10±0.087 1 (10^5,3)	85.09±0.146 5 (10^5,3)	85.06±0.251 2 (10^5,10^5)	85.07±0.179 6 (10^5,10^5,4)	85.04±0.213 9 (10^5,10^5,3)
AUS	79.57±1.328 3 (10^5)	80.43±2.701 6 (10^5,2)	**80.94**±1.693 2 (10^5,5)	78.19±1.344 0 (10^5,10^5)	81.27±1.440 6 (10^5,10^5,3)	77.97±3.115 9 (10^5,10^5,3)
BAL	98.56±0.823 7 (10^5)	99.44±0.196 0 (10^5,2)	**99.56**±1.488 1 (10^5,3)	98.56±1.148 2 (10^5,10^5)	98.72±1.547 1 (10^5,10^5,3)	**99.12**±0.640 0 (10^5,10^5,3)
BRE	65.41±2.022 5 (10^5)	66.49±5.077 1 (10^5,3)	**66.67**±1.507 5 (10^5,4)	72.79±3.599 1 (10^2,10^5)	73.87±2.933 1 (10^2,10^{-2},2)	**74.59**±3.036 5 (10,10^5,2)
DIA	73.36±1.097 9 (10^5)	74.85±1.587 8 (10^5,3)	**75.96**±0.521 2 (10^5,2)	75.54±1.517 0 (10^5,10^5)	74.40±2.205 4 (10^5,10^3,2)	**75.34**±1.910 8 (10^5,10^4,5)
LIV	70.00±1.751 2 (10^5)	70.72±3.323 9 (10^5,2)	**71.88**±2.726 8 (10^5,5)	69.49±2.347 0 (10^5,10^5)	71.52±3.476 0 (10^5,10^5,4)	**72.03**±2.429 4 (10^5,10^5,5)
MAG	86.48±0.162 0 (10^5)	86.87±0.429 4 (10^5,3)	**86.94**±0.421 9 (10^5,5)	65.07±0.363 0 (10^{-5},10^{-5})	86.77±0.202 0 (10^5,10^5,2)	**86.79**±0.378 0 (10^5,10^5,2)
MON	25.66±1.849 7 (10^5)	30.28±1.190 2 (10^5,4)	26.58±1.7147 (10^5,3)	47.63±1.372 8 (10^{-5},10^{-2})	48.95±0.969 0 (10^{-1},1,2)	**49.31**±1.735 1 (10^{-5},10^{-3},3)
MUS	100±0.000 0 (10^5)	100±0.000 0 (10^5, 1)	100±0.000 0 (10^5, 1)	100±0.000 0 (10^5,10^5)	100±0.000 0 (10^5,10^5,1)	100±0.000 0 (10^5,10^5,1)
PAG	96.66±0.142 7 (10^5)	96.82±0.416 3 (10^5,4)	**96.87**±0.251 2 (10^5,3)	96.61±0.331 8 (10^5,10^5)	96.97±0.301 0 (10^5,10^5,3)	96.44±0.333 8 (10^5,10^5,3)
SEI	93.25±0.331 7 (10^5)	93.62±0.266 6 (10^5,2)	**93.90**±0.453 6 (10^5,4)	93.53±0.307 2 (10^5,1)	93.59±0.237 1 (10,10^{-5},5)	**93.69**±0.591 1 (10^3,10^{-4},2)
SON	73.94±2.789 4 (10^5)	81.93±4.819 3 (10^5,3)	**82.65**±3.373 5 (10^5,4)	79.28±4.337 3 (10^5,10^5)	80.60±2.553 0 (10^5,10^5,2)	**80.84**±3.069 3 (10^5,10^5,5)
SPE	77.94±2.179 8 (10^5)	76.82±3.982 7 (10^5,5)	**77.57**±6.338 6 (10^5,5)	75.79±3.532 9 (10^4,10^5)	79.72±1.871 5 (10^{-5},10^{-5},3)	**80.56**±1.906 2 (10^{-5},10^{-5},4)
VOT	93.56±0.988 8 (10^5)	93.56±1.330 6 (10^5,4)	**94.48**±1.874 7 (10^5,3)	93.05±1.217 8 (10^5,10^5)	93.56±1.919 9 (10^5,10^5,2)	**93.62**±1.674 6 (10^5,10^5,5)
WDB	95.88±1.289 2 (10^5)	96.49±0.554 8 (10^5,2)	96.49±1.387 0 (10^5,2)	95.39±1.290 0 (10^5,10^5)	95.83±1.690 2 (10^5,10^5,2)	**96.32**±1.163 7 (10^5,10^5,4)



(a) Linear (b) Gaussian

图 3 D-S-SVM 与 D-S-SVM_0 的分类性能比较

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

根据表 2 可得:

(1) 将 D-S-SVM_0 与 SVM 相比,不论是线性还是高斯核情况下,D-S-SVM_0 在绝大部分真实数据集上的分类性能优于 SVM,在其他真实数据集上 D-S-SVM_0 至少能够保持相当的分类性能。这充分表明本文利用基于对抗样本学习的一阶梯度信息构建具有新型栈式结构分类模型的有效性,且该结构能够提高所提分类模型的分类性能。

(2) 结合表 2 和图 3,将 D-S-SVM 与其简易版本 D-S-SVM_0 相比,在绝大多数真实数据集上 D-S-SVM 的分类性能都优于 D-S-SVM_0,在小部分真实数据集上 D-S-SVM 至少能够保持相当的分类性能。这充分验证了本文模型在逐层更新输入样本的过程中引入 dropout 概念的有效性。

(3) 关于栈式结构,D-S-SVM 基本在 2 ~ 4 层内取得最佳分类性能,较少的层数一方面使 D-S-SVM 模型变得简单; 另一方面当超过一定层数时,由于利用基于对抗样本学习的一阶梯度信息逐层更新输入样本的缘故,原输入样本的特征空间将会被破坏,再增加层数只会降低 D-S-SVM 的分类性能。

3.2　逐层分类性能

表 3 详细列出了 D-S-SVM(Linear) 在真实数据集 AUS、PAG 与 WDB 以及 D-S-SVM(Gaussian) 在真实数据集 BAL、SEI 与 VOT 上每一层子分类器的训练和测试精度。其中,D-S-SVM 由 5 个子分类器级联而成,最佳的测试精度加粗显示。

表 3　D-S-SVM 逐层分类性能

D-S-SVM(Linear)

layer	AUS			layer	PAG			layer	WDB		
	c	训练	测试		c	训练	测试		c	训练	测试
1	10^5	98.84 ±0.723 0	78.99 ±3.360 0	1	10^5	97.64 ±0.242 5	96.81 ±0.127 3	1	10^5	100 ±0.000 0	94.91 ±1.767 5
2	10^5	98.55 ±0.305 5	80.65 ±1.897 9	2	10^5	97.64 ±0.147 4	96.71 ±0.362 0	2	10^5	100 ±0.000 0	**96.49** ±1.387 0
3	10^5	98.74 ±0.355 0	79.35 ±1.099 0	3	10^5	97.73 ±0.280 8	**96.87** ±0.251 2	3	10^5	100 ±0.000 0	95.96 ±1.452 0
4	10^5	98.89 ±0.246 3	77.54 ±0.916 6	4	10^5	96.76 ±0.255 2	96.35 ±0.050 0	4	10^5	100 ±0.000 0	95.61 ±2.019 4
5	10^5	98.41 ±0.497 4	**80.94** ±1.693 2	5	10^5	97.77 ±0.199 6	96.66 ±0.183 6	5	10^5	100 ±0.000 0	94.04 ±0.903 1

D-S-SVM(Gaussian)

layer	BAL			layer	SEI			layer	VOT		
	(c,σ)	训练	测试		(c,σ)	训练	测试		(c,σ)	训练	测试
1	$10^5,10^5$	100 ±0.000 0	98.72 ±1.055 3	1	$10^5,10^{-3}$	93.53 ±0.323 9	93.26 ±0.485 6	1	$10^5,10^5$	100 ±0.000 0	92.87 ±3.064 9
2	$10^5,10^5$	100 ±0.000 0	98.48 ±0.960 0	2	$10^3,10^{-4}$	93.24 ±0.394 3	**93.69** ±0.591 1	2	$10^5,10^5$	100 ±0.000 0	92.93 ±1.521 6
3	$10^5,10^5$	100 ±0.000 0	**99.12** ±0.640 0	3	10,10	93.58 ±0.271 0	93.18 ±0.406 3	3	$10^5,10^5$	100 ±0.000 0	92.64 ±2.656 1
4	$10^5,10^5$	100 ±0.000 0	98.20 ±1.255 4	4	$10^{-5},10^{-5}$	93.34 ±0.342 7	93.54 ±0.513 7	4	$10^5,10^5$	100 ±0.000 0	92.30 ±1.524 9
5	$10^5,10^5$	100 ±0.000 0	98.96 ±0.674 1	5	$1,10^{-1}$	93.40 ±0.371 8	93.45 ±0.557 3	5	$10^5,10^5$	100 ±0.000 0	**93.62** ±1.674 6

可以看出,当 D-S-SVM 未取得最佳分类性能前,利用基于对抗样本学习的一阶梯度信息逐层更新每一个子分类器的输入以及 dropout 技术的引入能够逐层提高 D-S-SVM 的分类性能。当 D-S-SVM 达到某一层取得最佳分类性能后,再增加 D-S-SVM 的层数会降低其分类性能,意味着 D-S-SVM 在该层输入样本特征空间已被优化到最佳状态。另外,D-S-SVM 在大部分情况下只需要 2 ~ 4 层的栈式结构便能取得最佳分类性能。

3.3　dropout 分析

表 4 展现了在执行 dropout 过程中不同比例的样本特征更新对 D-S-SVM 分类性能的影响。其中,比例选择 1/2、1/3、1/4,即利用基于对抗样本学习的一阶梯度信息对每一个样本 1/2、1/3、1/4 的特征进行更新。

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

表 4 dropout 分析

	D-S-SVM(Linear)			D-S-SVM(Gaussian)		
dropout 比例	1 /2	1 /3	1 /4	1 /2	1 /3	1 /4
AUS	80.94 ± 1.693 2	76.09 ± 2.582 4	77.39 ± 2.904 0	77.97 ± 3.115 9	77.14 ± 1.913 4	77.28 ± 2.239 6
BAL	90.56 ± 1.488 1	89.60 ± 1.910 0	89.36 ± 3.409 2	99.12 ± 0.640 0	89.48 ± 2.542 8	89.48 ± 2.071 1
DIA	75.96 ± 0.521 2	74.46 ± 1.375 8	73.55 ± 1.678 7	75.34 ± 1.910 8	74.66 ± 2.099 9	73.45 ± 1.687 8
LIV	71.88 ± 2.726 8	70.43 ± 3.285 7	70.29 ± 2.197 9	72.03 ± 2.429 4	69.63 ± 4.699 0	69.86 ± 4.011 1
PAG	96.87 ± 0.251 2	96.26 ± 0.404 9	96.24 ± 0.414 3	96.44 ± 0.333 8	96.17 ± 0.258 0	96.29 ± 0.266 7
SEI	93.90 ± 0.453 6	93.66 ± 0.408 5	93.64 ± 0.848 4	93.69 ± 0.591 1	93.70 ± 0.450 4	93.60 ± 0.564 7
VOT	94.48 ± 1.874 7	88.97 ± 1.831 9	87.59 ± 1.943 9	93.62 ± 1.674 6	87.82 ± 3.038 9	87.58 ± 2.611 0
WDB	96.49 ± 1.387 0	92.19 ± 2.548 4	91.67 ± 1.240 5	96.32 ± 1.163 7	91.71 ± 2.061 4	90.92 ± 1.455 3

可以看出，当在 D-S-SVM 模型中引入 dropout 概念逐层更新样本输入的过程中，随着样本特征更新比例的降低，D-S-SVM 的分类性能降低不明显，或至少保持相当的分类性能。这种情况可能是由于本文利用基于对抗样本学习逐层更新 D-S-SVM 输入样本并以此构造具备新型栈式结构的分类模型所导致。当在利用一阶梯度信息逐层更新输入样本的过程中，每一个子分类器的样本特征空间逐渐被隔离，当引入 dropout 概念对少量的样本特征进行更新时，D-S-SVM 也能够保持较好的分类性能。

4 结 语

针对真实数据集中存在对抗样本的事实，本文首先随机修改样本的真实标签类型，构造对抗样本进行对抗样本学习。在 SVM 的基础上，将经对抗样本学习的 SVM 模型输出误差对输入样本特征的一阶梯度信息嵌入到输入样本特征中以更新输入样本。其次，结合栈式结构原理，将经对抗样本学习的 SVM 模型作为子分类器构建层次支持向量机 D-S-SVM。特别地，利用基于对抗样本学习的一阶梯度信息更新输入样本为前后两个子分类器提供了一种新颖的栈式结构级联方式。将其与 SVM 相比，在真实数据集上的实验结果验证了该级联方式更加有效。最后，为提高 D-S-SVM 的泛化性，在逐层更新每一层子分类器输入的过程中引入神经网络中的 dropout 概念，即利用基于对抗样本学习的一阶梯度信息随机更新输入样本的部分特征，令输入样本的其他特征保持原有状态。在真实数据集上的分类结果有力地证明了 dropout 概念的引入确实能够有效增强 D-S-SVM 的泛化性。

今后将进一步研究如何从数据集中有效选择样本进行对抗样本学习以及如何有效选择样本的部分特征利用对抗样本学习更新样本特征。此外，如何将所提 D-S-SVM 分类算法推广到无监督学习也是今后研究内容之一。

参 考 文 献

[1] Ozdenizci O, Wang Y, Koike-Akino T, et al. Adversarial deep learning in EEG biometrics [J]. IEEE Signal Processing Letters, 2019, 26(5): 710 - 714.

[2] Chivukula A S, Liu W. Adversarial deep learning models with multiple adversaries [J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 31(6): 1066 - 1079.

[3] Lou Y H, Bai Y, Liu J, et al. Embedding adversarial learning for vehicle re-indentification [J]. IEEE Transactions on Image Processing, 2019, 28(8): 3794 - 3807.

[4] Zhang S C, Ji R R, Hu J, et al. Face sketch synthesis by multidomain adversarial learning [J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(5): 1419 - 1428.

[5] Yin Z Z, Wang F, Liu W, et al. Sparse feature attacks in adversarial learning [J]. IEEE Transactions on Knowledge and Data Engineering, 2018, 30(6): 1164 - 1177.

[6] Akhtar N, Mian A. Threat of adversarial attacks on deep learning in computer vision: A survey [J]. IEEE Access, 2018, 6: 14410 - 14430.

[7] Kobayashi M. Decomposition of rotor Hopfield neural networks using complex numbers [J]. IEEE Transactions on Neural networks and Learning Systems, 2018, 29(4): 1366 - 1370.

[8] Rakkiyappan R, Cao J D, Velmurugan G. Existence and uniform stability analysis of fractional-order complex-valued neural networks with time delays [J]. IEEE Transactions on Neural networks and Learning Systems, 2015, 26(1): 84 - 97.

[9] Huang W, Oh S K, Pedrycz W. Hybrid fuzzy wavelet neural networks architecture based on polynomial neural networks and fuzzy set/relation inference-based wavelet neurons [J].

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

IEEE Transactions on Neural networks and Learning Systems, 2018,29(8):3452 – 3462.

[10] Lü J C, Yi Z, Li Y X. Non-divergence of stochastic discrete time algorithms for PCA neural networks [J]. IEEE Transactions on Neural networks and Learning Systems, 2015, 26(2):394 – 399.

[11] Xiang W M, Tran H D, Johnson T T. Output reachable set estimation and verification for multilayer neural networks [J]. IEEE Transactions on Neural networks and Learning Systems, 2018,29(11):5777 – 5783.

[12] Zhang W, Li C D, Huang T W, et al. Synchronization of memristor-based coupling recurrent neural networks with time-varying delays and impulses [J]. IEEE Transactions on Neural networks and Learning Systems,2015,26(12): 3308 – 3313.

[13] Srivastava N, Hinton G E, Krizhevsky A, et al. Dropout: A simple way to prevent neural networks from overfitting [J]. The Journal of Machine Learning Research,2014,15(1): 1929 – 1958.

[14] Shao L, Wu D, Li X. Learning deep and wide: A spectral method for learning deep networks [J]. IEEE Transactions on Neural networks and Learning Systems,2014,25(12): 2303 – 2308.

[15] Cortes C, Vapnik V. Support-vector networks [J]. Machine Learning,1995,20(3):273 – 297.

[16] Xu Y T. Maximum margin of twin spheres support vector machine for imbalanced data classification [J]. IEEE Transactions on Cybernetics,2017,47(6):1540 – 1550.

[17] 王新艳,潘巍. 支持向量机算法应用于 2FSK 信号分类 [J]. 计算机应用与软件,2017,34(9):262 – 266.

[18] 刘铭,黄凡玲,傅彦铭,等. 改进的人工蜂群优化支持向量机算法在入侵检测中的应用 [J]. 计算机应用与软件,2017,34(1):230 – 235.

[19] Wolpert D H. Stacked generalization [J]. Neural Networks, 1992(5):241 – 259.

[20] Mosca A, Magoulas G D. Hardening against adversarial examples with the smooth gradient method [J]. Soft Computing,2018,22(10):3203 – 3213.

[21] Zhang Y P, Chung F L, Wang S T. Fast reduced set-based exemplar finding and cluster assignment [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2019,49(5):917 – 931.

[22] Dua D, Graff C. UCI machine learning repository [DS/OL]. CA: University of California. (2019) [2019-07-26]. http://archive.ics.uci.edu/ml.

[23] Alcala-Fdez J, Fernandez A, Luengo J, et al. KEEL data-mining software tool: Data set repository, integration of algorithms and experimental analysis framework [J]. Journal of Multiple-Valued Logic and Soft Computing,2011,17(2/3): 255 – 287.

(上接第 263 页)

参 考 文 献

[1] Samanta B, Al-Balushi K R, Al-Araimi S A. Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection [J]. Engineering Applications of Artificial Intelligence, 2003, 16(7/8): 657 – 665.

[2] Friedrichs F, Igel C. Evolutionary tuning of multiple SVM parameters [J]. Neurocomputing, 2005, 64:107 – 117.

[3] Ito K, Nakano R. Optimizing support vector regression hyperparameters based on cross-validation [C]//International Joint Conference on Neural Networks, 2003:2077 – 2082.

[4] 薛浩然, 张珂珩,李斌,等. 基于布谷鸟算法和支持向量机的变压器故障诊断 [J]. 电力系统保护与控制, 2015(8):8 – 13.

[5] Karaboga D, Akay B. Artificial bee colony(ABC) algorithm on training artificial neural networks [C]//2007 IEEE 15th Signal Processing and Communications Applications, 2007: 1 – 4.

[6] 刘铭,黄凡玲,傅彦铭,等. 改进的人工蜂群优化支持向量机算法在入侵检测中的应用 [J]. 计算机应用与软件, 2017,34(1):230 – 235,246.

[7] 安吉宇,杨瑜,刘志中. 基于人工蜂群优化的支持向量机模型在 Web 服务 QoS 预测中的应用 [J]. 计算机应用与软件,2016, 33(1):273 – 277,290.

[8] 邱正,钱玉良,张云,等. 基于人工蜂群算法优化支持向量机的燃气轮机故障诊断 [J]. 热能动力工程,2018,33(9): 39 – 43,57.

[9] 李航. 统计学习方法 [M]. 北京: 清华大学出版社,2012.

[10] 刘路,王太勇. 基于人工蜂群算法的支持向量机优化 [J]. 天津大学学报, 2011, 44(9):803 – 809.

[11] 喻金平,郑杰,梅宏标. 基于改进人工蜂群算法的 K 均值聚类算法 [J]. 计算机应用,2014,34(4):1065 – 1069, 1088.

[12] Aydin I, Karakose M, Akin E. A multi-objective artificial immune algorithm for parameter optimization in support vector machine [J]. Applied Soft Computing, 2011, 11(1): 120 – 129.

[13] 强光明. 船舶压载系统仿真软件研究与设计 [D]. 上海: 上海交通大学,2012.

[14] 聂立新. 基于优化支持向量机模型的发动机故障诊断 [D]. 沈阳: 东北大学,2015.

[15] 张金城. 船舶压载水系统故障诊断方法研究 [D]. 大连: 大连海事大学,2018.

[16] 聂立新,张天侠,张丽萍,等. 基于 DNPSO 的支持向量机的发动机故障诊断 [J]. 东北大学学报(自然科学版), 2012,33(4):571 – 575.

[17] 沈绍辉. 基于人工蜂群算法优化支持向量机的柴油机故障诊断研究 [D]. 太原: 中北大学,2016.

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net