

基于数据场的概率神经网络算法

李春芳^{1,4}, 刘连忠^{2,3}, 陆震¹

(1. 北京航空航天大学自动化科学与电气工程学院, 北京 100191; 2. 北京航空航天大学计算机学院, 北京 100191;
3. 北京航空航天大学网络技术北京市重点实验室, 北京 100191; 4. 河北体育学院网络中心, 河北石家庄 050041)

摘要: 提出基于数据场高斯势约简概率神经网络结构, 基本思路: 引入数据场估计训练集各类概率密度, 选择局部极大密度估计样本构造网络; 对初始网络迭代训练, 依次扩展各类具有最大密度估计值的误分样本至模式层并调整权重参数, 直至满足指定精度. 采用增量密度计算, 保证快速迭代和高概率收敛. 基于重采样技术进一步提升泛化精度. 实验表明, 提出的算法解释精练、拟合优度适中、计算高效.

关键词: 概率神经网络; 数据场; Parzen窗; 重采样

中图分类号: TP18 **文献标识码:** A **文章编号:** 0372-2112 (2011) 08-1739-07

Probabilistic Neural Network Based on Data Field

LI Chun-fang^{1,4}, LIU Lian-zhong^{2,3}, LU Zhen¹

(1. School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China;
2. School of Computer Science and Engineering, Beihang University, Beijing 100191, China;
3. Key Laboratory of Beijing Network Technology, Beihang University, Beijing 100191, China;
4. Network Center, Hebei Institute of Physical Education, Shijiazhuang, Hebei 050041, China)

Abstract: This paper proposed to decrease the structure of probabilistic neural network based on Gaussian potential of data field. Core idea is following: Introduce data field to estimate probabilistic density of training set of each class and select their maximum to construct the network; iteratively train the initial network by appending the maximum density sample unrecognized of each class to pattern layer and modify the weight of samples until satisfying desired accuracy. Incremental computing density ensures faster iteration and higher possible convergence. And introduce resampling technique to boost the generalization accuracy. Experiments show that the proposed algorithms have concise explanation, moderate fitness and effective calculation.

Key words: probabilistic neural network; data field; Parzen window; resampling

1 引言

1990年Specht基于贝叶斯决策和Parzen窗(核密度估计)^[1]提出的概率神经网络^[2](Probabilistic Neural Network, PNN)是一个四层、前馈、一次训练的高效分类器. 许多研究表明^[3-7], PNN有较强的泛化能力, 分类精度超过广泛使用的BP(Back Propagation)算法^[3,4]. 由于采用全部训练集构造网络, 不做任何的去冗余操作, PNN识别速度慢, 其空间和时间复杂度都为 $O(nm)$ (n 是样本数, m 是特征维数). 对此, 已有文献研究了使用各种聚类技术, 包括LVQ(Learning Vector Quantization)^[6]、K-means、Fuzzy C-Means^[7]算法约简网络规模. 聚类技术将每类样本划分成若干个类别, 用类中心作为PNN的模式层, 降低神经元数量, 以损失少量的精度为代价提升

计算速度, 是一种先聚类后分类的两阶段算法.

从数据场思想^[8,9]出发, 本文提出一种基于高斯势的优化算法 Gaussian-PNN (G-PNN), 通过高斯势发现类内样本密集区域, 迭代反馈纠错增量密度计算构造 G-PNN. 为进一步提升泛化精度, 基于重采样技术, 提出了3分类器集成的G3-PNN算法.

2 非参数估计与概率神经网络

2.1 样本选择

迄今为止, 解决大样本处理的方法包括随机抽样、聚类^[6-8]、遗传算法、近邻压缩法^[10,11]、动态选择法等. 动态选择按照某种规则在训练过程中根据模型的精度需要, 决定将哪些样本添加到训练集中, 能适应数据集的差异, 适用于规模适中的样本库, 本文采用了动态选

择法.

直接挑选部分样本代表全部从而剔除冗余的典型方法是 Hart 提出的压缩近邻算法 (Condensed Nearest Neighbor, CNN)^[10], 以及郝红卫改进的加权 WCNN 算法^[11].

CNN 优先选中编号靠前的样本. 为此, WCNN 采用迭代并投票评估样本的代表性, 一定程度弥补了 CNN 的缺陷. 为此, 我们提出一种可行的随机选择方法, 从每类中随机选择若干个样本, 构造分类模型并计算对训练集的识别率, 再从每类错分集合中随机选择部分样本添加到分类模型中, 迭代直到一定精度. 如果每次迭代添加全部错分样本称为选择概率 $\varphi = 1$ 的完全纠错法. 下文以 PNN 为分类器的随机选择法称为 R-PNN (Random-PNN).

针对 PNN 样本约简, 我们在文献 [12] 中提出中心近邻 C-PNN (Center-PNN) 算法, 用每类中心近邻作为 PNN 的模式层, 计算识别率, 迭代添加每类错分集合的中心近邻到模式层, 直到预定精度. 中心近邻可以看作样本概率密度极大值的一种估计, 但在数据分布交差非线性情况, 偏差较大.

2.2 非参数估计

参数估计要求已知密度函数形式, 常见函数很难拟合实际数据复杂的密度分布. 非参数统计对数据分布的具体形式不做细致假定, 尽量从数据本身获得其结构关系, 逐渐建立研究对象的数学模型和统计模型. 非参数估计主要分为多维直方图和 Parzen 窗法. 前者划分区间数随维数指数增长, 不适合大规模数据分析; 后者基于任何概率分布可以由正态分布叠加逼近, 平滑拟合密度函数, 对规则、非规则、单峰、多峰分布都适用. 基于 Parzen 窗, 已提出了若干分类和聚类算法, 如 PNN 分类算法^[1], 核密度估计聚类算法^[8], 基于数据场的聚类算法^[9], 以及 Mean shift 聚类算法.

Parzen 窗法的窗函数是一种径向基函数, 最常选用高斯函数. 图 1 是用 200 个样本拟合的标准正态分布密度, 窗宽分别取 $\sigma = 0.3, 0.1, 0.05$, 直观显示 σ 合适时 ($\sigma = 0.3$) 拟合误差很小.

2.3 概率神经网络

PNN 基于 Bayes 和 Parzen 窗理论直接用已知类别样本的先验知识预测新样本类别, 渐进达到 Bayes 最优决策边界.

假设分类问题具有 c 个类别: C_1, C_2, \dots, C_c , 其分类规则由 m 维的特征向量 $\mathbf{X} = (X_1, X_2, \dots, X_m)$ 所决定, 在 m 维样本空间中, 各分类的概率密度函数为特征向量的函数: $f_1(\mathbf{X}), f_2(\mathbf{X}), \dots, f_c(\mathbf{X})$.

判别样本为第 i 类的 Bayes 分类决策公式:

$$h_i c_i f_i(\mathbf{X}) > h_j c_j f_j(\mathbf{X}), \quad \text{对所有的 } j \neq i \quad (1)$$

f_i 为第 i 类概率密度函数, c_i 表示为第 i 类, 但被判判的风险函数值 (取决于应用背景, 一般可以认为误判风险函数相同), h_i 是第 i 类的先验概率.

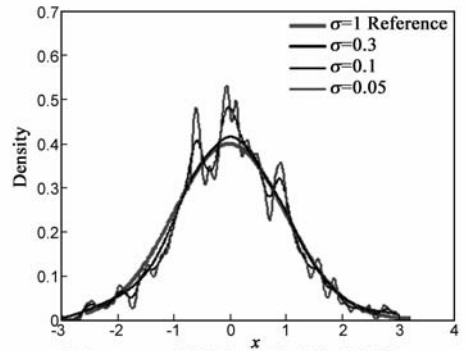


图1 Parzen窗拟合的正态分布密度

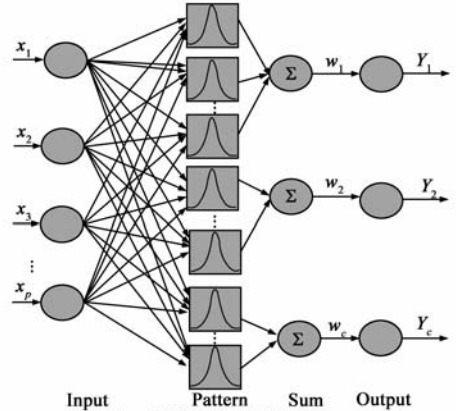


图2 概率神经网络结构

PNN 的四层结构如图 2 所示^[1], 包含 Input, Pattern (Hidden), Sum, 和 Output 层, 直接从训练样本载入数据, 无迭代过程, 除收集和存储样本外无学习代价. 采用公式 (2) 计算新样本点的概率密度函数 f_i .

$$f_i(\mathbf{X}) = \frac{1}{(2\pi)^{m/2} \sigma^m} \left(\frac{1}{n_i} \sum_{j=1}^{n_i} e^{-\frac{-(\mathbf{X} - \mathbf{Y}_{ij})^T (\mathbf{X} - \mathbf{Y}_{ij})}{2\sigma^2}} \right) \quad (2)$$

$f_i(\mathbf{X})$: 分类 i 在 \mathbf{X} 点的概率密度函数值

m : 训练样本维数

σ : 窗宽

n_i : 在 i 分类中的训练样本个数

\mathbf{X} : 测试分类样本

\mathbf{Y}_{ij} : 在分类 i 中第 j 个训练样本

3 基于数据场的 PNN 算法: G-PNN

3.1 定义

① 设 T 表示训练集, S 表示训练集的一个非空真子集, 给定一个代表阈值 e , 如果 S 具有如下特征, 则称 S 为 T 在 e 下的代表集 (Representative Set, RS): 用 S 作为训练集, 用 PNN 识别 T 中的每个样本, 其识别率不低于 e .

②当 $e = 1$ 时,定义 S 为 T 的完全代表集(Complete Representative Set, CRS).

③一个代表集的压缩率定义为 $Y, |S|$ 为代表集样本数, $|T|$ 为训练集样本数

$$Y = |S| / |T| \quad (3)$$

完全代表集具有下列性质:

(1)CRS 不唯一;(2)代表阈值 e 与训练集的认识率 R 满足, $e \leq R$.

3.2 数据场

场的概念最早是 1837 年由英国物理学家法拉第提出的,用于描述物质粒子间的非接触相互作用.根据物理学中稳定有源场的势函数性质,李德毅院士提出数据场的概念.研究发现短程场的势函数能很好地描述数据对象间的相互作用,并证明了具有短程场特点的拟核力场势函数与概率密度函数至多相差一个归一化常数^[9],即拟核力场势函数可以用做总体分布的估计,最常用的拟核力场是高斯场.

数据场方法按照数据分布的本来面貌抽象模型,对数据解释自然简单,与 Parzen 窗法有异曲同工之妙,已成功用于聚类^[8]和网络社区发现^[15].空间一点的叠加高斯势值用式(4)计算^[8,15]

$$G(\mathbf{X}) = \sum_{i=1}^n m_i e^{-\frac{\|\mathbf{X} - \mathbf{X}_i\|^2}{\sigma^2}} \quad (4)$$

比较式(2)和(4)容易发现,PNN 也是基于数据场的分类算法,使用等势线可以形象模拟决策边界,计算势场分布容易发现数据分布的高密度区域.

3.3 G-PNN 算法

G-PNN 采用数据场全局的“预览”技术发现最有代表性的样本.其启发式规则为:在每类数据密集的样本点高斯势较大,最大势值样本点代表本类样本最密集的区域,选用该点作为 PNN 的模式层有很好的局部代表性.G-PNN 算法如图 3 所示,通过一次计算类内训练样本的高斯势,不断将错分集合中最大势值样本添加到模式层,迭代直到一定精度.

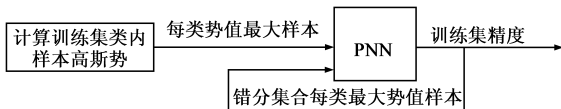
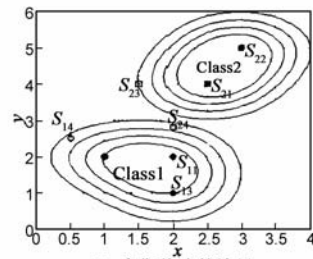


图3 G-PNN的训练流程

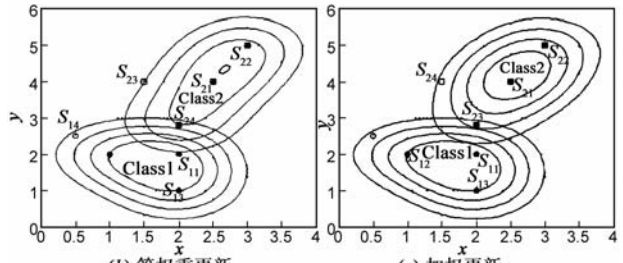
G-PNN 采用加权策略拟合全部样本的密度分布,直观想法是一个样本的代表性同它能提升的拟合精度相关,提升精度越高,应分配的权值越大.

等势线描述的决策边界和加权作用如图 4 所示.图 4(a)是某次迭代中样本 S_{24} (应属于 Class2)被错分,在该点 Class1 产生的势值大于 Class2.图 4(b)是不加权将 S_{24} 添加到模式层,决策边界发生了较大变化,可能导致新的分类错误,对保留较高的泛化精度不利.图 4(c)是

采用加权策略的决策边界,减小对原决策的影响,其目的是纠错尽量不引起新的错误,保留较大的分类间隔.



(a) 高斯势决策边界



(b) 等权重更新

(c) 加权更新

图4 样本更新前后等势线分布

算法 1 G-PNN 算法

输入:训练集 T ,代表阈值 e ,调整指数 k ,窗宽 σ

输出:G-PNN 网络

算法步骤:

①用式(2)计算 T 中样本在各类内的高斯势,其势值为类内所有样本在该点的势值叠加,权重 $m_i = 1$;

②将每类中最高势值的样本作为 PNN 的模式层,构造初始网络;

③用构造的 PNN,计算对 T 中所有样本的识别率 R_i , if $R_i \geq e$ then 结束, else 转④;

④选择每类中被错分集合中具有最高势值的样本添加到模式层,每轮迭代新添加样本的权重 $w_{new} = 1 - R_i^k$ (k 为调整指数),原有样本的权重 $w_{i+1} = R_i^k w_i$,令 $r_i = R_i^k$.保证权重衰减,使新样本权重始终小于已有样本权重,采用如下局部调整

If $r_{i+1} \leq 1/(2 - r_i)$

Then $r_{i+1} = 1/(2 - r_i) + \epsilon$ (ϵ 是小正数, $0 < \epsilon < 1$)

按公式(6)更新 PNN,转步骤③

$$f_j^{(i+1)}(\mathbf{X}) = R_i^k f_j^{(i)}(\mathbf{X}) + (1 - R_i^k) f_j^{(new)}(\mathbf{X}) \quad (6)$$

分析上述算法,步骤①根据给定 σ 值计算每类各样本的高斯势,其时间复杂度为 $O(n^2/(2c))$, n 为总样本数, c 为分类数,假设每类的样本数均为 n/c .由于计算每类中各样本高斯势值的相对大小,用公式(2)和(4)计算结果只差一个常数,鉴于一致性,算法采用公式(2)计算高斯势.

步骤②第一次迭代每类有一个样本,权重为 1.

步骤③计算对 T 中所有样本的识别率.

步骤④中新加样本的密度贡献权值较小为 $1 - R_i^k$,已有样本的权值乘以衰减系数 R_i^k .用识别率作为权重的变量,反映了新样本的代表能力.只计算新样本对密

代精度 R 调整权重 k^k 的作用如图 6 所示. 参数 k 取值在第一次迭代中初步确定, 如果数据集各类分布比较集中, 第一次迭代识别率 R_1 较高, 期望增加新样本的权重 $1 - R_1^k$, 取值 $k \geq 1$; 在数据集各类分布非线性交叉, 第一次迭代识别率较低, 如识别率 $R < 0.5$ 时, 期望减小新加入样本的作用, 选用 $k \leq 1$. 由于 G-PNN 采用迭代反馈纠错, 实验结果表明初步确定 k 值后, 具体值对迭代速度和泛化精度影响不大.

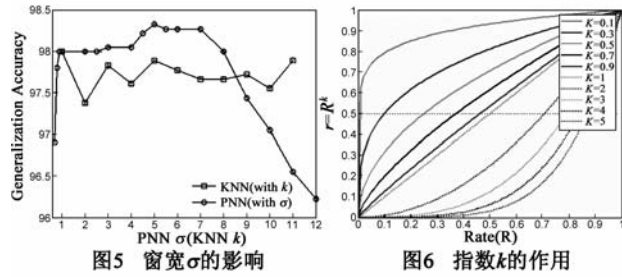


图5 窗宽 σ 的影响

图6 指数 k 的作用

5 实验

在 AMD Athlon 64 位 CPU, 1GHz 主频, 1.25G 内存台式机环境, Matlab 7.0 软件下实现了 G-PNN 算法, 以下实验结果训练集拟合精度大于 0.999.

5.1 Iris

Iris 包含 3 类, 4 个属性, 每类 50 个样本. 图 7 是四种算法生成的决策边界, 用全部特征计算欧式距离, 用 sepal length 和 petal length 描述数据场. 图 7(a) 所示为 Iris 全部样本作为场源用式(2)计算的高斯势场分布. 图 7(b) 是 FCM 算法每类划分成 3 类, 聚类中心做场源的数据场分布和 4 个错分样本, 训练集精度 $R = 0.9733$. 图 7(c) 是 C-PNN 的 12 个样本的数据场分布, $R = 1$. 图 7(d) 是 G-PNN 产生的 12 个样本的数据场, $R = 1$. 极小化势熵法计算得 $\sigma = 2.9$, 此时, G-PNN 和 C-PNN 都取得

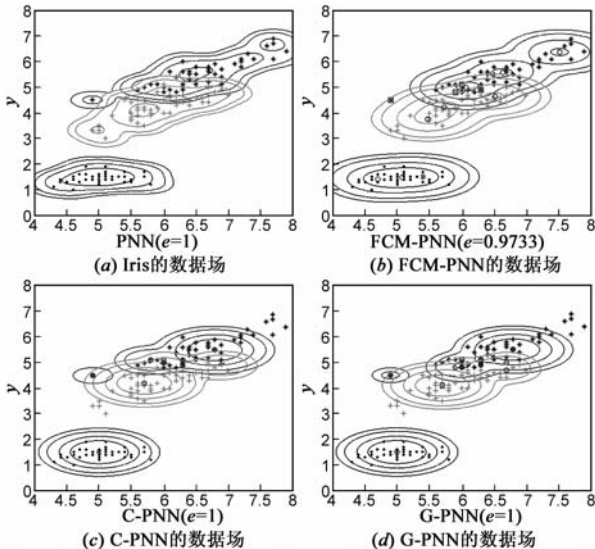


图7 数据场描述的决策边界比较(○:隐层样本, □:错分样本)

最好的压缩率 0.08.

FCM-PNN 的聚类并不考虑分类需要, 在极易区分的第一类使用样本有冗余. C-PNN 和 G-PNN 采用反馈纠错, 直接考虑分类问题选择样本.

5.2 Optdigits

Optdigits 是 UCI 库中一个手写数字数据集, 共 10 类, 样本为 8×8 维 0-16 级灰度图像, 训练样本 3823 个, 测试样本 1797 个. 部分训练样本如图 8 所示.



图8 Optdigits 的样本图像

G-PNN 的收敛性如图 9 所示. 由于 G-PNN 选择局部密度极大值样本, 代表性较好, 比 C-PNN 收敛快. R-PNN 选择的样本量较多, 收敛速度快, 其中完全纠错的收敛最快, 但网络样本仍有较大冗余.

从图 10 中可以看出, 在 k 取定值, σ 对泛化精度有较大影响, 存在一个 σ 最优值区间, 而 σ 取定值, k 对泛化精度影响很小.

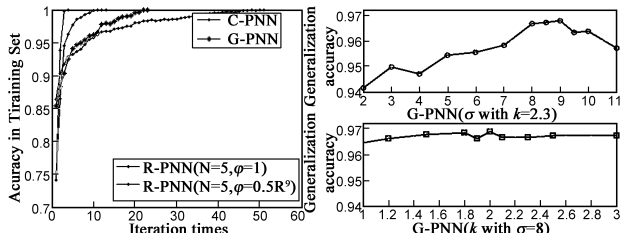


图9 G-PNN 的收敛性

图10 参数的影响

表 2 列出 G-PNN 与相关算法的比较, 识别时间是对全部测试集. PNN 和 KNN 训练时间为 0, 识别时间较长, PNN 的泛化精度最高. FCM-PNN 在聚类参数最优情况下, 只用原训练集 0.026 倍的样本, 泛化精度比 PNN 降低 1% 左右. G-PNN 的去冗余效果非常明显, 无需聚类, 网络规模是原来的 0.074 倍, 泛化精度比 C-PNN 高, 与 FCM-PNN 相当. G-PNN 第一次迭代的识别率达到 0.8 左右, 选择指数 k 增大公式(6)新样本权重, 取 $k = 2.3$. G3-PNN 算法 20 次平均精度达到 97.3%. 需要指出的是基于统计学习理论直接考虑最大分类间隔的 Support Vector Machine (SVM) 算法的泛化精度为 96.88% [17], 足见 G-PNN 和 G3-PNN 算法的实用性.

表 2 Optdigits 的实验结果

Algorithm	Parameters	Training Time(sec.)	Recognizing Time(sec.)	Y	Accuracy (%)
PNN	$\sigma = 5$	0	40	1	98.27
KNN	$K = 1$	0	70	1	98.0
FCM-PNN	$\sigma = 8, s = 10$	40	3	0.026	97.11
C-PNN	$\sigma = 7, k = 2.4$	66	14	0.052	96.7
G-PNN	$\sigma = 8, k = 2.3$	60	10	0.074	96.99
G3-PNN	$\sigma = 8, k = 2.3$	190	35	0.22	97.3 ± 0.1

5.3 USPS

USPS 是 Buffalo 邮政信封的手写数字数据, 16×16 维图像, 训练样本 7291 个, 测试样本 2007 个, 图 11 是部分训练样本。



图 11 USPS 数据集样本图像

极小化势熵优化 $\sigma = 3.2$, 在 $\sigma = 1.2$ 时 PNN 取得最大泛化精度 0.9467. G-PNN 第一次迭代识别率在 0.3, 取 $k \leq 1$, 选择 $k = 0.5$. 各算法表现的训练时间、识别时间以及泛化精度顺序基本与 Optdigits 相同, 详见表 3. G-PNN 用原网络规模的 0.11 倍, 识别率达到 0.92, G3-PNN 的平均精度超过 FCM-PNN. 而 WCNN 算法在 0.768 倍压缩率下 BP 算法的识别率为 0.916^[11].

表 3 USPS 的实验结果

Algorithm	Parameters	Training Time (sec.)	Recognizing Time(sec.)	Y	Accuracy (%)
PNN	$\sigma = 1.2$	0	320	1	94.67
KNN	$K = 1$	0	230	1	94.37
FCM-PNN	$\sigma = 1.2, s = 20$	370	12	0.027	93.02
C-PNN	$\sigma = 0.5, k = 0.5$	1090	250	0.094	90.01
G-PNN	$\sigma = 0.5, k = 0.5$	1250	240	0.114	92.03
G3-PNN	$\sigma = 0.5, k = 0.5$	3250	670	0.344	93.3 ± 0.08

5.4 Letters

Letters 共 26 类, 为 A-Z 大写字母, 每个样本提取 16 个特征, 其值为 0-15 之间的整数, 共 20000 个样本, 前 16000 为训练集, 后 4000 为测试集。

在 20000 样本中随机选取 500 个, 用极小化势熵法计算 $\sigma = 2.3$. PNN 在 $\sigma = 0.8$ 取得最高精度 0.9613. 实验结果见表 4. 而 SVM 算法的泛化精度为 0.9423^[17].

表 4 Letters 的实验结果

Algorithm	Parameters	Training Time (sec.)	Recognizing Time(sec.)	Y	Accuracy (%)
PNN	$\sigma = 0.8$	0	600	1	96.13
KNN	$K = 1$	0	110	1	95.65
FCM-PNN	$\sigma = 1.5, s = 100$	330	80	0.153	94.73
C-PNN	$\sigma = 1, k = 0.1$	1100	90	0.141	93.03
G-PNN	$\sigma = 1, k = 0.1$	1150	50	0.115	94.05
G3-PNN	$\sigma = 1, k = 0.1$	3400	160	0.336	94.9 ± 0.11

综上所述, 属于先聚类后分类的集成算法 FCM-PNN 泛化精度略微高于单分类器 G-PNN. 而 G3-PNN 的泛化精度高于 FCM-PNN 和 G-PNN.

6 结束语

与聚类优化相比, G-PNN 考虑了数据本身的结构特点, 采用与 Parzen 窗本质上相同的高斯数据场发现数据密集区, 样本代表性强, 通过权重衰减, 增量迭代并自

动归一化保证算法高概率收敛和较高的泛化精度. G3-PNN 通过数据建模降低模型建模的复杂性, 能有效提升 G-PNN 的泛化精度. G-PNN 迭代构造, 在大样本集情况, 采取其他技术减少迭代次数或减少扫描原训练集次数有望提升训练速度。

参考文献

- [1] E Parzen. On estimation of a probability density function and mode [J]. Annals of Mathematical Statistics, 1962, 33(3): 1065 - 1076.
- [2] D F Specht. Probabilistic neural networks [J]. Neural Networks, 1990, 2(1): 109 - 118.
- [3] D F Specht. Enhancements to probabilistic neural networks [A]. In Proc IEEE Int Joint Conf on Neural Networks [C]. Baltimore, MD, 1992. 761 - 768.
- [4] D F Specht, P D Shapiro. Generalization accuracy of probabilistic neural networks compared with back-propagation networks [A]. Proc of the Int'l Joint Conf on Neural Networks [C]. Seattle, 1991. 887 - 892.
- [5] D J Montana. A weighted probabilistic neural network [A]. In Proc of the Advances in Neural Information Processing Systems 4 [C]. San Francisco: Morgan Kaufmann, 1992. 1110 - 1117.
- [6] P Burrascano. Learning vector quantization for the probabilistic neural network [J]. IEEE Trans on Neural Networks, 1991, 2(4): 458 - 461.
- [7] 蔡曲林, 刘普寅. 一种新的概率神经网络有监督学习算法 [J]. 模糊系统与数学, 2006, 20(6): 83 - 87.
Cai Qu-lin, Liu Pu-yin. A new supervised learning algorithm for probabilistic neural network [J]. Fuzzy Systems and Mathematics, 2006, 20(6): 83 - 87. (in Chinese)
- [8] 涂文燕, 李德毅, 王建民. 一种基于数据场的层次聚类方法 [J]. 电子学报, 2006, 34(2): 258 - 262.
Gan Wen-yan, Li De-yi, Wang Jian-min. An hierarchical clustering method based on data fields [J]. Acta Electronica Sinica, 2006, 34(2): 258 - 262. (in Chinese)
- [9] 李德毅, 杜 ■. 不确定性人工智能 [M]. 北京: 国防工业出版社, 2005. 207 - 217.
- [10] P E Hart. The condensed nearest neighbor rule [J]. IEEE Trans on Information Theory, 1968, 14(3): 515 - 516.
- [11] 郝红卫, 蒋蓉蓉. 基于最近邻规则的神经网络训练样本选择方法 [J]. 自动化学报, 2007, 33(12): 1247 - 2151.
Hao Hong-wei, Jiang Rong-rong. Training sample selection method for neural networks based on nearest neighbor rule [J]. ACTA Automatica Sinica, 2007, 33(12): 1247 - 1251. (in Chinese)
- [12] Qian Li-pu, Li Chun-fang, Zhang Yan-ping. Optimization of probabilistic neural networks based on center neighbor and KLT [A]. Proc of the Int'l Conf on Machine Learning and

Cybernetics [C]. Liverpool: World Academic Press, 2010. 1447 - 1452.

- [13] 毕华, 梁洪力, 王珏. 重采样方法与机器学习[J]. 计算机学报, 2009, 32(5): 862 - 877.
Bi Hua, Liang Hong-li, Wang Jue. Resampling methods and machine learning [J]. Chinese Journal of Computers, 2009, 32(5): 862 - 877. (in Chinese)
- [14] Zhou ZH, Li M. Tri-training: Exploiting unlabeled data using three classifiers [J]. IEEE Trans on Knowledge and Data Engineering, 2005, 17(11): 1529 - 1541.
- [15] 涂文燕, 赫南, 李德毅, 王建民. 一种基于拓扑势的网络社区发现方法[J]. 软件学报, 2009, 20(8): 2241 - 2254.
Gan Wen-yan, He Nan, Li De-yi, Wang Jian-min. Community discovery method in networks based on topological potential [J]. Journal of Software, 2009, 20(8): 2241 - 2254. (in Chinese)
- [16] 范九伦, 吴成茂. FCM 算法中隶属度的新解释及其应用[J]. 电子学报, 2004, 32(1): 157 - 159.
Fan Jiu-lun, Wu Cheng-mao. The new explanation of membership degree in FCM and its applications [J]. Acta Electronica Sinica, 2004, 34(2): 258 - 262. (in Chinese)
- [17] I W Tsang, A Kocsor, T J Kwok. Simpler core vector machines with enclosing balls [A]. Proc of the 24th int'l Conf. on Machine Learning [C]. New York: ACM Press, 2007. 911 - 918.

作者简介



李春芳 女, 1974 年生于河北, 北京航空航天大学博士研究生, 副教授, 主要研究方向为机器学习、信息安全审计。

E-mail: LCF1974@126.com



刘连忠 男, 1958 年生于河北, 北京航空航天大学副教授, 硕士生导师, 主要研究方向为信息安全、电子政务。



陆震 男, 1942 年生于北京, 北京航空航天大学教授, 博士生导师, 主要研究领域为机器人、人工智能。