

基于无监督学习算法训练径向基概率神经网络*

赵温波

(中国科学技术大学 自动化系 合肥 230031)
(中国科学院合肥智能机械研究所 合肥 230031)

黄德双

(中国科学院合肥智能机械研究所 合肥 230031)

王树坤

(炮兵学院 合肥 230031)

摘 要 本文研究了径向基概率神经网络(Radial Basis Probabilistic Neural Networks, RBPNN)的一种新的无监督学习算法,该算法整合了径向基神经网络的结构原理与动态聚类算法的特点,使得在对训练样本的聚类分析并正确划分其类别属性的同时,自动完成径向基概率神经网络的训练过程.本算法在对 IRIS 和双螺旋分类问题的应用中,取得了较好的分类效果,而且在推广能力方面,由本文算法训练的 RBPNN 要明显好于有监督训练的径向基函数神经网络(RBFNN).

关键词 径向基概率神经网络,双螺旋分类问题,无监督学习算法,Parzen 窗函数

中图法分类号 TP18

1 引 言

径向基概率神经网络(RBPNN^[1])是由径向基函数神经网络(RBFNN^[2])与概率神经网络(PNN^[3])发展而来的,因此,RBPNN 吸收了径向基函数神经网络与经典的概率密度估计的优点,如快速的收敛性能、在网络的输出端形成了输入样本的后验概率估计等等.与传统的前馈神经网络(FNN)相比,RBPNN 在模式分类、时间序列预测等方面有较为显著的优势^[4,5]. RBPNN 在网络结构上可以分为四层(如图 1),其中,

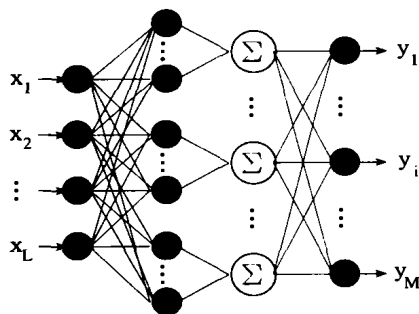


图 1 径向基概率神经网络的拓扑结构
第一层和最后一层分别为输入层和输出层,中间两

层为隐层,第一隐层是以 Parzen 窗函数^[6](或称为核函数)为激活函数,由隐中心矢量构成,第二隐层根据第一隐层隐中心矢量的类别属性对其输出进行求和.一般来说,RBPNN 是一种前馈神经网络,因为 RBPNN 的分类与模式识别原理是建立在对已知类别的训练样本集之上.例如,当训练样本集 Φ 中第 k 个模式类别样本为 RBPNN 的输入时,在输出端应该以向量 $[0, \dots, \underset{k}{1}, \dots, 0]^T$ 为期望信号,使得 RBPNN 的第二隐层与输出层的连接权重沿着使期望输出与实际输出误差最小的方向改变. RBPNN 的训练方法主要有递归正交最小二乘算法(Recursive Orthogonal Least Square Algorithms, ROLSA)^[5]和递推最小二乘方法(RLSA)^[7].当然,传统的 BP 训练算法也同样适用于 RBPNN,但与前两种方法相比,BP 算法在训练速度以及收敛精度方面没有优势.尽管 ROLSA 与 RLSA 都有较快的收敛速度与收敛精度,但两种方法也有区别,主要在于对被求解问题的训练样本集容量的要求不同. RLSA 需要合适的初始条件,比较适合大规模训练样本集.而 ROLSA 可以利用成熟的矩阵算法(正交分解法),与 RLSA 相

* 国家自然科学基金资助项目(No. 60173050)

收稿日期:2003-02-26;修回日期:2003-06-09

比训练速度更快而且收敛精度更高,因此我们首选ROLSA用于本文的RBPNN的无监督学习算法。

对于输入向量 x_j ,RBPNN输出层第 i 个神经元的输出值 $y_i^a(x_j)$ (或 y_{ij}^a)可以表示为

$$y_i^a(x_j) = \sum_{k=1}^M \omega_{ik} h_k(x_j), \quad (1)$$

$$h_k(x_j) = \sum_{i=1}^{n_k} \phi_i(\|x_j - c_{ki}\|_2), \quad (2)$$

其中, $h_k(x_j)$ (或 h_{kj})是对应于输入样本 x_j 的第二隐层第 k 个神经元的输出, ω_{ik} 是第二隐层第 k 个神经元和输出层第 i 个神经元的连接权重, $\phi(\cdot)$ 是Parzen窗函数(核函数), c_{ki} 表示第一隐层第 k 个类别的第 i 个隐中心矢量, n_k 是第一隐层第 k 个类别的隐中心矢量, $\|\cdot\|_2$ 指欧拉范数, M 表示输出层的神经元数目。

对于 L 个训练样本,(1)式用向量—矩阵形式表示为

$$[y_{ij}^a]_{L \times M} = [h_{kj}]_{L \times M} \times [\omega_{ik}]_{M \times M}, \quad (3)$$

$$\text{或 } Y^a = HW, \quad (4)$$

其中 Y^a 、 H 与 W 分别为对应 L 个训练样本RBPNN的实际输出矩阵、第二隐层输出矩阵以及第二隐层与输出层的连接权重矩阵。

根据文献[5],输出层与第二隐层间的连接权重矩阵 W 由下式求得,

$$W = R^{-1} \hat{Y}, \quad (5)$$

其中 R 、 \hat{Y} 分别由以下两式求得,

$$H = Q \begin{bmatrix} R \\ \cdots \\ 0 \end{bmatrix}, \quad (6)$$

$$Q^T Y = \begin{bmatrix} \hat{Y} \\ \tilde{Y} \end{bmatrix}. \quad (7)$$

(6)式为第二隐层的输出 H 的正交分解过程, Q 为 H 的正交矩阵, R 为上三角方阵, Y 为期望信号矩阵, \hat{Y} 和 \tilde{Y} 分别为正交化后的期望信号矩阵和误差驻留矩阵。

对于RBPNN的有监督学习算法来说,当训练RBPNN时,必须首先明确训练样本集 Φ 中每一个样本的类别属性,只有这样,RBPNN才能被用于分类或模式识别。但是,在解决实际问题的过程中,有时候我们并不知道训练集中每个样本的类别属性,也就无法训练RBPNN。因此能否在不明确训练样本确切类别属性的情况下,利用RBPNN的映射机理对样本集进行训练,即如何实现RBPNN的无监督训练,

这是本文讨论的重点。

实现无监督训练RBPNN的首要任务是如何实现训练样本的类别划分,实现样本类别划分是有很多方法可以利用^[8],从实现机理上来看,这些方法可以区分为基于概率密度估计的直接方法和基于样本间相似度的间接方法。前一种方法是基于大样本集,从神经网络训练的计算复杂度及实际问题的样本容量方面来考虑,这种方法显然不能满足要求。间接方法中的典型代表就是动态聚类方法,这种方法无需考虑训练样本的容量限制,尤其在小样本集的情况下也不影响训练效果。单纯研究动态聚类方法不是本文的重点。本文的目标在于如何针对RBPNN的映射特点,把RBPNN的训练与动态聚类过程结合起来,使得在对训练样本动态聚类的同时完成RBPNN的训练过程。

2 无监督学习算法的构成与原理

本文的无监督学习算法由以下三部分组成:(1)典型样本选取算法;(2)类别初始划分算法;(3)类别调整算法。

2.1 典型样本选取算法

典型样本,即在训练样本集中能够体现同一类别样本共同特征的代表样本。典型样本的选取也有很多方法可以利用^[8],如根据先验知识选取、密度法选取、重心法选取等。本文采用密度法。该方法主要依据样本密度来实现的。所谓样本密度(P_i),就是以该样本(S_i)为超球心(对于二维、三维训练样本,分别为圆心和球心),与其它样本的距离落入以某一正数 R 为半径的超球形(圆形或球形)邻域内的样本数即是该样本的密度。样本密度是依据样本距离来度量的。一般来说,样本距离有多种度量方法,如海明距离、曼哈坦距离、明考斯距离、欧拉距离等等。本文采用欧拉距离,也就是样本矢量间的2范数,即 $D_{ij} = \|S_i - S_j\|_2$ (D_{ij} 为样本 S_i 与样本 S_j 的欧拉距离)。所以,样本密度就可以表示为 $P_i = \sum_{i,j} \text{sgn}(D_{ij} \leq R)$ ($\text{sgn}(\cdot)$ 为符号函数)。一般来说,一个类别中的典型样本有很多,如果把所有的典型样本都能选到,可能会增加运算量。因此本文算法对于训练集中的每一个模式类别只选择1个典型样本。

为讨论的方便,我们假定训练样本集 Φ 中有 L 个样本且归属于 M 个类别($L \geq M$),算法选取的每个类别的典型样本分别为 I_1, I_2, \dots, I_M 。选取典型样本的过程如下:

(1) 设定邻域半径 R 和已知类别数 M , 初始化样本集类别指示器 I^d 为空, 计算训练集中所有的样本距离 $D_{ij}(i, j = 1, 2, \dots, L)$.

(2) 计算训练集样本密度 $P_i(i = 1, 2, \dots, L)$.

(3) 在训练集样本密度中寻找密度值最大的样本为第一类别的典型样本 $I_1(I_1 = S_{i_1}, i_1 = \arg \max_i P_i)$, 并将类别指示器 I^d 中相应索引的内容置为 1, 即 $I^d(i_1) = 1$, 然后将该样本密度置为最小(一般为 0), 即 $P_{i_1} = 0$.

(4) 继续寻找样本密度中最大的样本 $S_{i_k}(i_k = \arg \max_i P_i)$, 并计算与其它典型样本 I_1, I_2, \dots, I_{k-1} 的距离 $D_1^R, D_2^R, \dots, D_{k-1}^R$, 如果所有距离都小于等于超球形(圆形或球形)邻域半径 R , 则不把它作为已知类别的典型样本, 同时也将该样本的密度置为最小, 否则, 把它作为新类别的典型样本 $I_k(I_k = S_{i_k}, i_k = \arg \max_i P_i)$, 同时将该样本密度置为最小 $P_{i_k} = 0$, 并将类别指示器中相应索引的内容置为新类别数, 即 $I^d(i_k) = k$, 转向下一步.

(5) 判断已选典型样本数目是否达到已知类别数 M , 若达到, 则结束程序, 否则, 将转向(4).

需要说明的是, 在上述算法中, 样本类别指示器 I^d 用来标示训练集中所有样本的类别归属, 它是一个向量, 维数即是所有样本数 L . 这就要求将所有样本在训练之前要按照固定的顺序存放, 该指示器也是下文初始分类算法和类别调整算法所共同使用的参量. 超球形(圆形或球形)邻域是用来区分训练集中类别差异的标准. 如果在当前样本中, 样本密度最大的个体与已知典型样本的所有距离大于这一邻域, 即认为该样本属于一个新类别. 表达式 $i_k = \arg \max_i P_i$, 表示取样本密度中的最大值的索引, 即所有样本密度 $P_i(i = 1, 2, \dots, L)$, 最大值的样本索引为 i_k .

2.2 类别初始划分算法

根据训练样本集中所有类别的典型样本, 利用训练样本与典型样本的核相似度来初始划分所有的训练样本的类别. 所谓核相似度, 就是按照 RBPNN 第一隐层的映射机理, 求取某一训练样本与每个类别中已知成员核函数值的和. 实际上, 样本的核相似度是一个向量, 它有 M 个分量, 即 $C_i = [c_{i1}, c_{i2}, \dots, c_{iM}]$. 在核相似度向量中, 取分量最大值的索引为该样本的类别属性, 同时, 将该样本增加为最大值索引所指定类别的成员. 具体的计算过程如下:

(1) 在训练样本集中选择一个样本 S_i , 计算该

样本的核相似度 C_i .

(2) 取该样本的核相似度 C_i 分量最大值的索引为该样本的类别属性, 即

$$S_i \in j' = \arg \max_{j=1}^M c_{ij}, \text{ 其中 } j' \in [1, M],$$

并将类别指示器 I^d 中相应索引的内容置为该样本的类别序数, 即 $I^d(i) = j'$.

(3) 如果样本集中的每一个样本都有类别归属, 则停止计算, 存贮类别指示器 I^d , 否则, 转向(1), 继续计算.

在上述算法, 每一样本的核相似度向量 C_i 又可以表示为 $C_i = [f(S_i, K_1), f(S_i, K_2), \dots, f(S_i, K_M)]$, 其中 K_j 为第 j 个类别的核, 它是由第 j 个类别中的已知成员和核函数控制参数组成, 即 $K_j = \{ \langle I'_m, \sigma_l \rangle \mid m = 1, 2, \dots, M_j, l = 1, 2, \dots, n \}$, 其中, I'_m 为第 j 个类别中的已知成员, σ_l 为核函数控制参数, M_j 是第 j 个类别中的样本数目, n 为样本的空间维数. 根据 RBPNN 的工作机理, 即可得到该样本的核相似度向量 C_i 的第 i 个分量的具体表达式

$$f(S_i, K_j) = \sum_{m=1}^{M_j} \exp\left(-\sum_{l=1}^n \left(\frac{S_{il} - I'_{il}}{\sigma_l}\right)^2\right). \quad (8)$$

式(8)中核函数控制参数 σ_l 是未知量, 从几何意义上来看, 该参数是由训练集中各样本的空间分布特性所决定, 该参数的取值大小直接影响了初始样本的分类结果. 根据文献[9], 该参数的启发性取值有如下的表达式:

$$\sigma_l = \frac{D_l^*}{\sqrt{L}}, l = 1, 2, \dots, n, \quad (9)$$

其中, D_l^* 为训练样本集中各个样本间第 l 个分量的最大欧拉距离, L 为样本总数, n 为样本空间的维数.

2.3 类别调整算法

类别调整算法的目的是将类别初始划分算法中错误划分类别的样本进行重新划分, 使其能够接近该样本的真实类别. 这种调整是基于训练集能量来实现的. 所谓训练集能量就是训练集中的每个样本与其它样本(除本身以外)最为接近类别中的所有成员的核相似度值的总和, 它是衡量调整过程是否收敛的标准. 只有当类别调整以后, 训练集能量趋于稳定时, 才认为这种调整是成功的, 否则, 需要重新进行类别调整, 直到满足能量稳定为止, 具体的调整过程如下:

(1) 初始化停止误差标准 ϵ , 调整迭代索引 P , 初始化相似度矩阵 $C_{L \times M} = [C_1, C_2, \dots, C_L]^T$ 为零矩阵, 初始化期望信号矩阵 $Y_{L \times M}$ 为 $L \times M$ 零矩阵. 根

据公式(9),计算核函数控制参数 $\sigma_l(l = 1, 2, \dots, n)$.

(2) 选取训练集中的一样本 S_i , 计算该样本的核相似向量 C_i .

(3) 在类别初始划分算法得到的类别指示器 I^d 中, 如果样本 S_i 的类别属性为 k , 除去其在核相似向量 C_i 中的贡献(即 $c'_{ik} = c_{ik} - 1$), 而其它的核相似分量保持不变(即 $c'_{i1} = c_{i1}, \dots, c'_{i(k-1)} = c_{i(k-1)}, c'_{i(k+1)} = c_{i(k+1)}, \dots, c'_{iM} = c_{iM}$).

(4) 在样本 S_i 新的核相似度 $c'_{i1}, \dots, c'_{ik}, \dots, c'_{iN}$ 中取分量最大值的索引为其新类别属性, 即 $S_i \in j'' = \arg \max_{j=1}^M c_{ij}$, 且 $j'' \in [1, M]$, 重新调整核相似度值, 即除第 j'' 个核相似度分量值加 1 以外, 其余核相似度分量保持不变, 新的核相似向量 C''_i 的各分量为 $c''_{i1} = c'_{i1}, \dots, c''_{ij''} = c'_{ij''} + 1, \dots, c''_{iM} = c'_{iM}$, 并将核相似向量 C''_i 设置为相似矩阵 C 的第 i 个成员, 并存储.

(5) 判断样本 S_i 新类别属性与初始分类后类别指示器 I^d 中的类别属性是否相同, 如果不相同, 则将样本 S_i 从初始分类中转移到新的类别中, 并将类别指示器作相应的调整, 否则, 不做类别成员的移动, 直接转到下一步.

(6) 判断是否将所有训练样本的类别属性都进行了类别的重新调整, 如果没有则转向(2), 否则, 转向下一步.

(7) 计算本次的训练集能量, 即 $E_n^p = \sum_{i=1}^L \max_{j=1}^M (c''_{ij})$, 并比较与上次计算(或初始分类的)的训练集能量 E_n^{p-1} 的差值是否满足给定的停止计算误差条件 $\epsilon \geq |E_n^p - E_n^{p-1}|$, 如果满足, 则转向下一步, 否则, 转向(2), 重新对训练样本集所有样本的类别属性进行评估.

(8) 令 $H = C$, 根据公式(5) ~ (7), 计算 RBPNN 的连接权重.

本文算法只对训练样本集进行类别划分以及把这些标记类别的样本都作为 RBPNN 的隐中心矢量直接训练该网络, 而没有对隐中心矢量进行优化, 因为对 RBPNN 隐中心矢量的优化(或 RBPNN 结构优化)不是本文研究的重点, 相关内容可以参考文献[4]和文献[5].

3 模拟结果与讨论

3.1 IRIS 分类问题

我们把本文算法应用于对 IRIS 的三种花型的

识别, 即 setosa、versicolor 和 virginica, 根据文献[8]知, setosa 是一种具有 38 个染色体的“二倍体”, virginica 是具有 70 个染色体的“四倍体”, 而第三种花型 versicolor 是一个“六倍体”, 而且, 文献[10]也推断第三种花型可能是前两种花型多倍体杂交的产物. 文中的识别工作从花型的四个参数入手, 即花萼的长度与宽度、花瓣的长度与宽度. 具体的训练数据见文献[10], 共有 150 个样本, 其中每种花型各有 50 个样本. 本文将 150 个样本平均分成两部分, 即在每种花型中各抽取 25 个样本(共 75 样本), 组成训练样本集和测试样本集.

超球形邻域半径为 0.78. 典型样本选取算法在 75 个训练样本中选取的三个类别的典型样本分别为 $I_1 = [5.1, 3.5, 1.4, 0.2]^T$, $I_2 = [6.3, 2.7, 4.9, 1.8]^T$, $I_3 = [5.9, 3.0, 4.2, 1.5]^T$.

类别初始划分算法的识别率为 85.33%. 运用类别调整算法, 正确识别率为 97.33%, 即只有 3 个样本没有被正确识别. 类别调整过程中的训练集能量变化及正确识别率如图 2 及图 3.

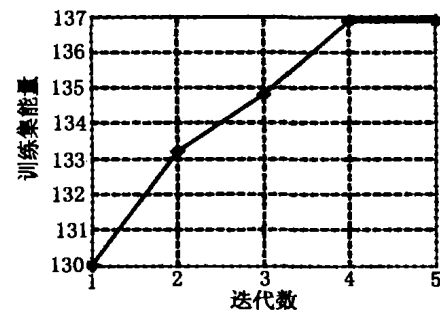


图2 类别调整算法用于分类 IRIS 问题时的训练集能量变化曲线

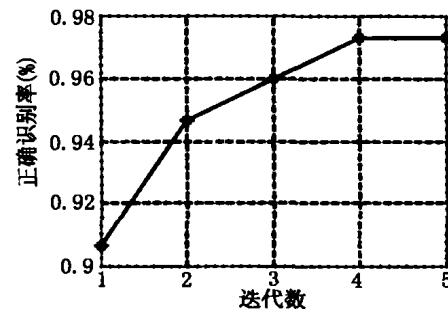


图3 类别调整算法用于分类 IRIS 问题时的正确识别率的变化情况

算法得到的收敛权重 W 如下:

$$W = \begin{bmatrix} 1.0000 & 0 & 0 \\ 0 & 1.0273 & -0.0273 \\ 0 & -0.0319 & 1.0320 \end{bmatrix}$$

为了进一步检测本文算法的有效性,我们将原始 75 个测试样本混合 20 组零均值不同方差的高斯白噪声组成测试集,以测试算法得到的 RBPNN 的推广能力,同时,我们用这 20 组测试集来测试有监督训练的 RBFNN 的推广能力,并与 RBPNN 形成比较,具体的结果见图 4.从图中我们可以看出,尽管对于无噪声测试样本的正确率 RBFNN 优于 RBPNN,但在噪声条件下,RBPNN 明显优于 RBFNN.

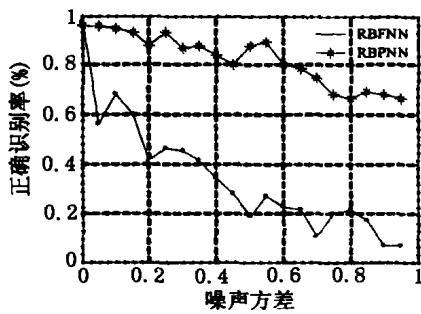


图 4 无监督训练的 RBPNN 与有监督训练的 RBFNN 在分类 IRIS 问题时的网络推广能力比较

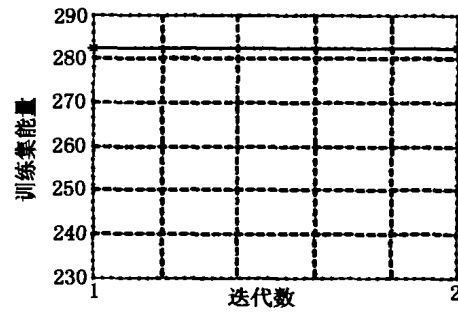


图 5 类别调整算法用于分类双螺旋问题时的训练能量变化曲线

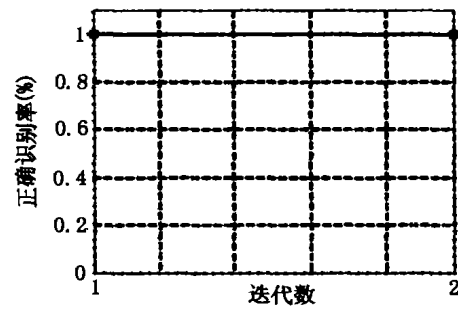


图 6 类别调整算法用于分类双螺旋问题时的正确识别率的变化情况

3.2 双螺旋分类问题

我们来讨论使用本文无监督算法训练的 RBPNN 来分类双螺旋问题.假定双旋中的每一旋的长度都选取 98,训练样本的采样步长为 1,则共有 196 个样本点参与训练,即每旋有 98 个采样点组成训练样本集(样本数据的产生见文献[4]).

圆形邻域半径为 0.40.典型样本选取算法在 196 个训练样本中选取的二个类别的典型样本分别为

$$I_1 = [0.4548, 0.5213]^T,$$

$$I_2 = [0.4024, 0.1200]^T.$$

类别初始划分算法的正确识别率为 99.49%,即只有一个样本被错误识别.运用类别调整算法得到的正确识别率为 100%,即所有 196 个训练样本全部正确识别.调整过程中的训练集能量变化及正确识别率如图 5 及图 6,可以看出,类别调整算法只经过一次迭代就完成对所有训练样本的正确识别.由本算法得到的双螺旋分界面如图 7 所示,该分界面平滑、均匀.由此可见,本文算法训练的 RBPNN 对双螺旋分类问题有较好的分类性能.类别调整算法所得到的收敛权重 W 如下:

$$W = \begin{bmatrix} 1.0003 & -0.0003 \\ -0.0003 & 1.0003 \end{bmatrix}$$

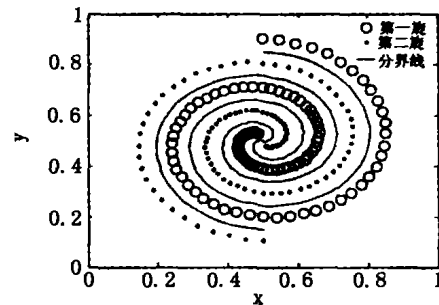


图 7 无监督训练的 RBPNN 分类双螺旋问题得到的分类界面

我们用测试样本集来检测无监督算法训练的 RBPNN 对测试样本的正确识别能力.所使用的测试样本集是这样构造的,双旋的长度不变(98),而采样步长由 1 缩减到 0.1,这样每旋有 971 个采样点,共有 1942 个采样点组成原始的测试样本集.然后把 1942 个原始测试样本增加 20 组 0 均值不同方差(0,0.005, 0.01, ..., 0.09)的高斯白噪声,即共有 20 组每组 1942 个样本点参与测试.同时,我们也使用这 20 组测试集来测试有标识类别训练的 RBFNN.具体的测试结果如图 8 所示,可以看出,本文无监督学习算法训练的 RBPNN,尤其是在有噪声扰动的条件下,训练后的 RBPNN 的推广能力要明显好于 RBFNN.

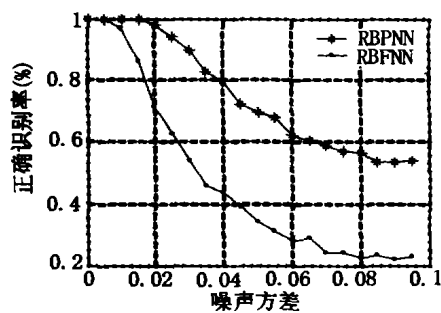


图8 无监督训练的 RBPNN 与有监督训练的 RBFNN 分类双螺旋问题时的网络推广能力比较

4 结束语

本文把基于 Parzen 窗函数的聚类算法与径向基概率神经网络的映射机理结合起来,提出了使用新的无监督算法来训练径向基概率神经网络,该算法的特点是在不明确训练样本类别属性的情况下,能够实现对样本类别的划分并同时完成 RBPNN 的训练过程.从文中使用本文算法训练的 RBPNNs 对 IRIS 和双螺旋分类问题的分类效果来看,本文的无监督算法对 RBPNN 的训练是可行的、有效的.

参 考 文 献

- [1] Huang D S. Radial Basis Probabilistic Neural Networks: Model and Application. International Journal of Pattern Recognition and Artificial Intelligence, 1999, 13(7): 1083 - 1101
- [2] Lowe D. Adaptive Radial Basis Function Nonlinearities and the Problem of Generalization. In: Proc of 1st International Conference on Artificial Neural Networks. London, 1989, 171 - 175
- [3] Specht D F. Probabilistic Neural Networks. Neural Networks, 1990, 3: 109 - 118
- [4] Zhao Wenbo, Huang D S. The Structure Optimization of Radial Basis Probabilistic Neural Networks Based on Genetic Algorithms. In: Proc of International Joint Conference on Neural Networks. Honolulu, Hawaii, 2002, 1086 - 1091
- [5] Zhao Wenbo, Huang D S. Application of Recursive Orthogonal Least Squares Algorithm to the Structure Optimization of Radial Basis Probabilistic Neural Networks. In: Proc of International Conference on Signal Processing. Beijing, China, 2002, 1211 - 1214
- [6] Parzen E. On Estimation of a Probability Density Function an Mode. Annals of Mathematical Statistics, 1962, 33: 1065 - 1076
- [7] 黄德双. 神经网络模式识别系统理论. 北京: 电子工业出版社, 1996, 119 - 137
- [8] 边肇祺, 张学工, 等. 模式识别. 第二版. 北京: 清华大学出版社, 1999, 230 - 248
- [9] Haykin S. Adaptive Filter Theory. 3rd edition. Upper Sadle River, NJ: Prentice-Hall, 1996
- [10] Fisher R A. The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics, 1936, 7: 179 - 188

TRAINING THE RADIAL BASIS PROBABILISTIC NEURAL NETWORK BASED ON A UNSUPERVISED LEARNING ALGORITHM

Zhao Wenbo

(Department of Automation, University of Science and Technology of China, Hefei 230031)

(Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei 230031)

Huang Deshuang

(Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei 230031)

Wang Shukun

(Artillery Academy of People Liberation Army, Hefei 230031)

ABSTRACT

This paper investigates a novel unsupervised learning algorithm of the radial basis probabilistic neural network (RBPNN). This novel algorithm, which integrates the structure principle of the RBPNN and the characteristics of dynamically clustering algorithms, not only carries out clustering analysis and correctly labeling the classes properties for all training samples, but also at the same time automatically completes training RBPNN. In the application of IRIS classification problem and two-spiral discrimination problem, the good recognition results are achieved in the training stage, furthermore in the generalization ability the RBPNNs learned by the proposed algorithm are better than the results by the radial basis function neural networks (RBFNNs) trained by the samples with the teaching signals.

Key Words Radial Basis Probabilistic Neural Networks, Two-Spiral Discrimination Problem, Unsupervised Learning Algorithm, Parzen Windows Function