

FPNN:优化的过程神经网络模型及其应用

王蕾¹, 平静¹, 宋国杰²

WANG Lei¹, PING Jing¹, SONG Guo-jie²

1. 安阳师范学院 计算机与信息工程学院, 河南 安阳 455002

2. 北京大学 信息科学技术学院, 北京 100871

1. School of Computer and Information Engineering, Anyang Normal University, Anyang, Henan 455002, China

2. School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China

E-mail: wanglei9346@163.com

WANG Lei, PING Jing, SONG Guo-jie. FPNN: Optimized process neural network model and its application. *Computer Engineering and Applications*, 2009, 45(24): 136-139.

Abstract: Through a transformation under certain condition, this paper transforms a PNN model into an identical FNN model, named FPNN model. The equivalence also has been presented in detail. It achieves higher time efficiency without any decrease of representation power or forecast correctness.

Key words: Fourier Neural Network (FNN); Process Neural Network (PNN); Fourier Process Neural Network (FPNN); classification; model transformation

摘要: 利用既定条件下过程神经元与傅里叶神经元的等价性, 提出一种优化的过程神经网络模型 FPNN, 并对等价性进行了证明。FPNN 网络在保持了过程神经网络模型表达能力和预测准确率的同时, 继承了 FNN 的优点, 大大提高了模型的效率。

关键词: 傅里叶神经网络; 过程神经网络; 傅里叶过程神经网络; 分类; 模型转换

DOI: 10.3778/j.issn.1002-8331.2009.24.040 文章编号: 1002-8331(2009)24-0136-04 文献标识码: A 中图分类号: TP301.6

在传统人工神经网络模型中, 系统的输入和输出都是与时间无关的常量, 即网络的输入是一些几何点式的输入。但是, 在很多情况下系统的输入是一些随时间变化的值, 这些值会持续一定时间, 形成一个输入值的序列, 即一个输入过程或是一个依赖于时间的序列。这个序列中往往包含了一些隐藏的相对有用的信息, 描述了序列的变化规律和趋势。如果采用几何点式的输入, 则会忽略这些有用的信息, 从而降低模型的预测表达能力。

目前可以处理时间序列输入的神经网络有过程神经网络模型^[1-3]和傅里叶神经网络模型^[4-6]。两种模型都是传统人工神经网络在时间域上的扩展, 但相比之下, 两种模型各自有自己的优点和缺点: 过程神经网络模型通用性和表达能力更强, 可以准确地发现序列特征, 但模型比较复杂, 时间效率有待提高。傅里叶神经网络模型简单且效率更高, 但其表达能力弱于过程神经网络。面对实际应用中海量的具有序列特征的输入数据, 对目前已有的神经网络模型而言是一个很大的挑战。

将过程神经网络和傅里叶神经网络的优点结合起来, 从过程神经网络出发, 证明了既定条件下过程神经元与傅里叶神经元的等价性, 构造出一种优化的过程神经网络模型 FPNN。在保持了模型表达能力和预测准确率的同时, 时间效率得到很大提高。值得说明的是, 所提出的模型得到了真实的应用, 取得了

很好的实用效果。在移动领域真实大规模数据集上也进行了实验, 证明了该模型的有效性。

1 基本概念

1.1 过程神经网络

与传统神经网络类似, 过程神经网络由若干过程神经元按一定的拓扑结构组成, 可分为前向型和反馈型网络。一个过程神经网络模型可以表示为:

$$y = f \left[\int_0^T A(t) K(t) dt - \theta \right] \quad (1)$$

其中, θ 为神经元的阈值, f 为神经元的激励函数 (可取线性函数、Sigmoid 函数、Gauss 型函数等)。函数 $K(t)$ 为时间聚合基函数。

在上式中, $A(t) = \sum_{i=1}^n w_i(t) x_i(t)$, $w_i(t) = \sum_{k=1}^L w_i^{(k)} b_k(t)$ 。其中, $x_i(t)$, $0 < i \leq n$, 为过程神经元输入函数向量; $w_i(t)$, $0 < i \leq n$, 为相应的权函数; $b_k(t)$ 是基函数, 为有限基或者可数基, 可以是正交或者非正交的。

1.2 傅里叶神经网络

文献[5]给出了基于傅里叶函数的傅里叶神经网络, 可表示为:

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.60703066)。

作者简介: 王蕾(1975-), 女, 讲师, 研究方向: 数据挖掘; 平静(1975-), 男, 讲师, 研究方向: 数据仓库; 宋国杰(1975-), 男, 博士, 副教授, 研究方向: 智能数据分析。

收稿日期: 2009-02-05 修回日期: 2009-04-02

$$Y=f\left(\sum_{i=1}^L w_i(s_i \cdot x(t))\right)$$

其中, f 为激励函数, 一般可以用 Sigmoid 函数; w_i 为权值 ($1 \leq i \leq L, L=2n+1$); s_i 为傅里叶基函数; $x(t)$ 为长度为 T 的时间序列输入; n 为傅里叶展开的级数; Y 为经过激励的输出。

傅里叶基函数可表示为:

$$s_1 = \frac{1}{T}, \frac{2\cos \omega_i t}{T}, s_{2i+1} = \frac{2\sin \omega_i t}{T}$$

一个长度为 T 的时间序列 $x(t)$ 可以被表示为:

$$x(t) = \frac{\alpha_0}{2} + \sum_{i=1}^n (\alpha_i \cos(\omega_i t) + \beta_i \sin(\omega_i t))$$

用离散傅里叶变换值 c 表示 $x(t)$ 的傅里叶变换:

$$c = (c_1, c_2, \dots, c_{2n+1}) = (\alpha_0, \alpha_1, \beta_1, \dots, \alpha_n, \beta_n)$$

傅里叶基函数是一组高维空间的正交基, 满足: (1) 单位性: 当 $i=j$ 时 $s_i s_j = 1$; (2) 正交性: 当 $i \neq j$ 时, $s_i s_j = 0$ 。通过单位性

和正交性可知 $x(t) = \sum_{i=1}^n c_i s_i$ 。

由正交性可知, 对于任意不同的 i 和 $j, s_i s_j = 0$ 意味着 $c_i c_j$ 互不相关。化简后可得:

$$Y = f\left(\sum_{i=1}^L c_i w_i\right) \quad (2)$$

c 是和神经网络计算无关的常量, 可以通过预处理来避免在神经网络训练过程中反复计算。

2 FPNN 神经网络

首先给出一种过程神经网络结构, 然后给出过程神经元在既定条件下与傅里叶神经元的转换和等价性证明, 最后提出优化的神经网络模型 FPNN。

2.1 过程神经网络模型的拓扑结构

假设有一个四层的前馈过程神经网络模型(记做 PNN), 包含一个输入层、一个输出层和两个隐含层。

(1) 输入层

输入层对应两种属性: 具有过程式的连续输入和传统标号属性, 分别对应时间序列输入和标号属性输入。前者对应 n 个输入节点, 每个节点对应一个输入序列。后者有 $(A_1 + \dots + A_m)$ 个输入节点, 第 i 个属性对应 A_i 输入。

(2) 第一隐层

第一隐层结点对应 n 个过程神经元和 m 个传统神经元两类, 前者分别对应 n 个连续时间序列输入; 后者处理标号属性, 每个神经元对应(聚合)一个属性。

(3) 第二隐层

第二隐层神经元是普通神经元, 第二隐层与第一隐层中的神经元之间是全连接的。

(4) 输出层

输出层只有一个传统神经元结点, 与第二隐层中的所有结点全连接。

网络结构如图 1 所示, 学习过程与 BP 神经网络类似, 此处不再赘述。

2.2 过程神经元与傅里叶神经元在既定条件下的等价性

2.2.1 神经元转换

在公式(1)中, $b(t)$ 为基函数, 所以 $x(t)$ 也可以用这组基函

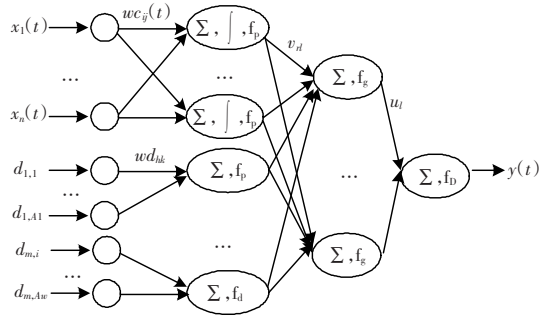


图1 PNN网络的拓扑结构

数展开, 得到 $x_i(t) = \sum_{j=1}^L c_j^{(i)} b_j(t)$ 。带入公式(1)可得

$$Y = f\left[\int_0^T \left(\sum_{i=1}^n \left(\sum_{k=1}^{k=L, j=L} c_j^{(i)} w_i^{(k)} b_k(t) b_j(t)\right)\right) K(t) dt - \theta\right]$$

采用傅里叶基函数 s_i 作为过程神经网络空间聚合的基函数时, 可得

$$Y = f\left[\int_0^T \left(\sum_{i=1}^n \left(\sum_{k=1}^{k=L, j=L} c_j^{(i)} w_i^{(k)} s_k(t) s_j(t)\right)\right) K(t) dt - \theta\right]$$

由于傅里叶基函数是正交基, 根据单位性和正交性, 可得

$$Y = f\left[\left(\int_0^T K(t) dt\right) \left(\sum_{i=1}^n \left(\sum_{k=1}^L c_k^{(i)} w_i^{(k)}\right)\right) - \theta\right]$$

现在, 时间聚合函数变成和输入无关的常量, 因此可以在不影响模型表达能力的情况下去掉时间聚合函数, 简化后模型

为 $Y = f\left[\left(\sum_{i=1, k=1}^{i=n, k=L} c_k^{(i)} w_i^{(k)}\right) - \theta\right]$ 。当只有一个输入时间序列时, 进一

步简化为 $Y = f\left[\left(\sum_{i=1}^L c_i w_i\right) - \theta\right]$ 。

θ 可以省略, 省略后的模型和原模型等价:

$$Y = f\left[\sum_{i=1}^L c_i w_i\right] \quad (3)$$

式(1)和式(3)是等价的, 因此有定理 1。

定理 1 以傅里叶基函数为空间聚合基函数的过程神经元模型与傅里叶神经元模型等价。

2.2.2 等价性证明

分别从计算能力和学习等价性方面给出定理 1 更加形式化的证明。

2.2.2.1 计算能力等价证明

对于计算能力的等价性, 有引理 1 和引理 2。

引理 1 对于一个以傅里叶基函数为空间聚合函数的过程神经元模型 P , 都存在一个傅里叶神经元 F , 满足 $Y^P(X) = Y^F(X)$ 。

引理 2 任何一个傅里叶神经元模型 F , 都存在一个以傅里叶基函数为空间聚合函数基函数的过程神经元模型 P , 满足 $Y^P(X) = Y^F(X)$ 。

引理 1 的证明可由 2.2.1 节的推导得证。由于推导过程的可逆性, 引理 2 也得证。

由引理 1 和引理 2 知, 傅里叶神经元和以傅里叶基函数为空间聚合基函数的过程神经元计算能力等价。

2.2.2.2 学习能力等价证明

引理 3 对任意一个以傅里叶基函数为空间聚合基函数的过程神经元 P 和任意训练迭代次数 $N(N \geq 0)$, 都存在一个傅

里叶神经元 F 和训练迭代次数 $M(M \geq 0)$, 使得对任意的训练集 $Z=(X, Y)^k$, 其中对于每个样本 X, F 都可以在学习能力上模拟 P , 即 $Y_M^F(Z)=Y_N^P(Z)$ 。

证明 按数学归纳法思路证明。

基始: 根据引理 1, 对任意的 P 和任意输入 X , 可以构造傅里叶神经元 F , 使得 $Y^F(X)=Y^P(X)$, 即经过 0 次训练后两者计算能力等价。

归纳: 下面证明如果经过 k 次训练后计算能力等价, 那么经过 $k+1$ 次训练后, 两者计算能力依然等价。

对于任意一个训练样本 $Z=(X, Y)$, 设 P 经过在 Z 上的一次训练后的模型是 P' , F 经过在 Z 上的一次训练后的模型是 F' , 那么如果 P 和 F 计算能力相同, 则需证明对任意的输入 X' , $Y^{F'}(X')=Y^{P'}(X')$ 成立。

在学习过程中, 对于过程神经元 P , 新的权值为:

$$\tilde{w}^P = w^P + \alpha c_i \Delta_{Y_i^P}(X)$$

所以神经元 P 新的输出为:

$$Y_i^P = f_P \left[Q \left(\sum_{s=1}^n \sum_{k=1}^L c_k^{(s)} w_k^{P(s)} \right) \right] = f_A \left[Q \left(\sum_{s=1}^n \sum_{k=1}^L c_k^{(s)} (w_k^P + \alpha c_i \Delta_{Y_i^P}(X)) \right) \right]$$

而对傅里叶神经元 F , 新的权值为:

$$\tilde{w}^F = w^F + \alpha c_i \Delta_{Y_i^F}(X) = Q w^P + \alpha c_i \Delta_{Y_i^P}(X)$$

所以神经元 F 新的输出为:

$$Y_i^F = f_A \left[\sum_{j=1}^n \sum_{k=1}^L c_k^{(j)} w_k^F \right] = f_A \left[\sum_{j=1}^n \sum_{k=1}^L c_k^{(j)} (Q w_k^P + \alpha c_i \Delta_{Y_i^P}(X)) \right] = Y_i^P$$

通过以上数学归纳, 对样本 Z 学习后, 神经元 F 将完全可以模拟神经元 P 的学习能力, 所以 F 完全可以模拟以傅里叶基函数为空间聚合基函数的过程神经元 P 的计算能力。引理得证。

引理 4 对任意一个傅里叶神经元 F 和任意训练迭代次数 $M(M \geq 0)$, 都存在一个以傅里叶基函数为空间聚合基函数的过程神经元 P 和训练迭代次数 $N(N \geq 0)$, 使得对任意的训练集 $Z=(X, Y)^k$, P 都可以在学习能力上模拟 F , 即 $Y_N^P(Z)=Y_M^F(Z)$ 。

引理 4 的证明和引理 3 是一样的, 这里不赘述。所以 F 和 P 的计算能力等价。

推论 1 对同样的训练样本, 傅里叶神经元 F 和以傅里叶基函数为空间聚合基函数的过程神经元 P 达到收敛需要的训练迭代次数是相同的。

2.2.3 复杂性分析

已经证明, 傅里叶神经元 F 和以傅里叶基函数为空间聚合基函数的过程神经元 P 是等价的。而且, 对于同样的训练集, 它们达到收敛的训练迭代次数是一样的。因此, 只需分析它们在一次迭代过程中时间效率即可。

由于一次训练迭代是由计算神经元输出和误差后向传播 (BP) 两部分组成的, 而误差后向传播可以看成计算神经元输出的逆过程, 其时间复杂度是相同的。所以, 下面分析时间复杂度的过程中, 只考虑神经元输出过程中的标量乘法运算的次数。

对于过程神经网络, 计算输出的公式如下:

$$Y_i^P = f_A \left[\int_0^T \left(\sum_{k=1, j=1}^{k=L, j=L} c_j^{(i)} w_k^P s_k(t) s_j(t) \right) K(t) dt \right]$$

因此, 每个神经元 P 一次训练需要的乘法次数是 $(L * L * 3 + T)$ 。

对于傅里叶神经元 F , 其计算输出的公式如下:

$$Y_i^F = f_A \left[\sum_{k=1}^L c_k^{(i)} w_k^F \right]$$

所以, 一个傅里叶神经元 F 经过一次训练需要的乘法次数是 L 。

因此, 两种神经元相比, 傅里叶神经元的时间效率有了 $O(L)$ 倍的提升, 这是由于在有过程属性神经元的计算过程中, 利用正交基函数的单位性和正交性, 简化了计算过程而带来的。在实际应用中, 为了模型能够达到收敛, 傅里叶级数一般为 5 级左右, 这时 $L=11$ 。也就是说, 相比过程神经元, 傅里叶神经元可以得到大约一个数量级的时间效率提升。这对于海量数据的分析而言, 是至关重要的。

2.3 优化的过程神经网络模型 FPNN

由 2.2 节的等价性证明和性能分析可知, 傅里叶神经元和过程神经元具有相同的表达能力, 但却在时间效率上具有明显的优势。因此, 如果用傅里叶神经元代替 PNN 网络中的过程神经元, 那么整个网络结构的性能也会得到大幅度的提升。转换后的网络记做 FPNN, 网络结构如图 2 所示。

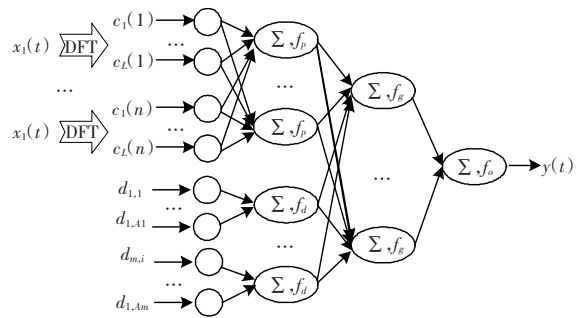


图 2 FPNN 网络的拓扑结构

为便于性能分析, FPNN 的网络结构与 PNN 相似, 也包括一个输入层、两个隐层和一个输出层, 除此之外, 一个预处理层将输入序列转换成其对应的傅里叶变换。下面将把 FPNN 的结构和 PNN 的结构进行比较, 并重点介绍两者的不同之处。

(1) 预处理层

预处理层将输入序列转换成对应的傅里叶变换。与 PNN 相同, 每个样本包括具有过程特性的时间序列和不具有过程特性的标号属性。FPNN 网络可以通过预处理的方法大幅度降低处理过程属性的时间复杂度。每个输入的时间序列都在模型学习训练之前, 通过预处理变为傅里叶变换序列。这部分工作是在模型学习之前就完成的, 不需要在模型训练过程中反复计算, 所以对模型训练的时间复杂度不会有影响。另外, 非过程属性不需要预处理, 这些属性直接连接到输入层。

(2) 输入层

输入层的结构也和过程神经网络有所不同。经过预处理层的变换, 每个时间序列并不是像过程神经网络那样直接输入到网络中, 而是经过预处理, 把每个时间序列的傅里叶变换序列作为输入。另一个不同之处是, 过程神经网络中, 每个具有过程属性的时间序列被作为一个高维输入元素, 直接输入到可以处理时间序列的过程神经元; 但是结果傅里叶变换后, 时间序列的傅里叶变换序列是不具有过程属性的, 任意不同的 i, j, c_i, c_j 互不相关, 所以, 对于序列 $x_i(t), c_{i1} c_{i2} \dots c_{iL}$ 被作为 L 个单独的输入, 每个 c_{ij} 作为一个独立的无过程属性的输入神经元。对于

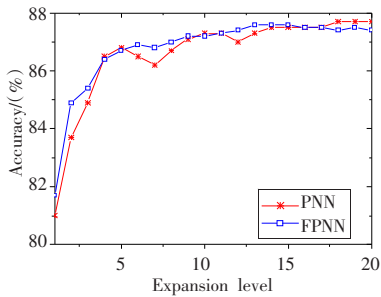


图3 预测精度比较

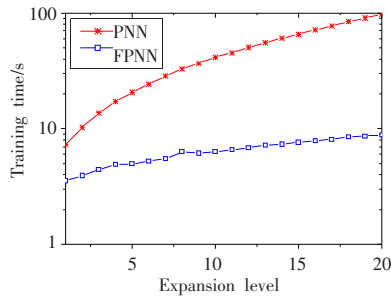


图4 训练时间比较

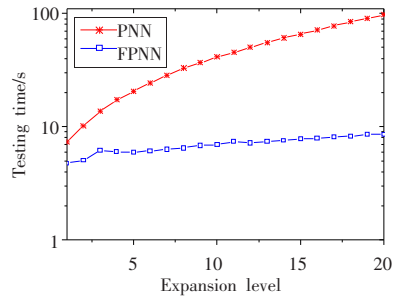


图5 测试时间比较

每个样本中无过程属性的标号属性的处理和过程神经网络相同。

(3) 第一隐层

对于有过程属性的时间序列输入,通过预处理,输入层采用了无过程属性的时间序列的傅里叶变换,所以和过程神经网络采用可以处理过程属性的过程神经元不同,傅里叶神经网络的第一隐层没有采用特殊的神经元,而是采用人工神经网络的普通神经元。由于人工神经网络结构简单效率高,这一层的变换将是傅里叶神经网络获得性能提升的主要贡献力量。对于无过程性质的标号属性,处理方法和过程神经网络是相同的。

(4) 第二隐层和输出层

这两层和过程神经网络是完全相同的。

3 实验评估

为评估算法的各项性能指标,进行了一系列的实验。结果表明, F 的各项指标均呈现出较好的性能。

实验环境:CPU为P4 1.4 G;主存512 M,基于Windows XP操作系统,使用Visual C++实现所有算法。

实验数据集:采用中国移动某省公司2004年1月到4月期间的数据,用于构建客户流失预测分析模型。客户数据总量约为22万条,随机抽取了2000条数据作为训练数据集,另外10000条数据作为测试数据集。每条数据包含如下内容:该用户1到3月每天的通话时间、短信数量、交往圈大小(和该用户通话的不同电话号码的数量),共91天,为时间序列属性。该用户4月份是否离网为预测结果的对比观察值(0表示没有离网,1表示离网)。

通过模型预测的准确率(P)来比较模型的能力。如果模型预测用户离网概率大于0.5而且该用户离网,或者模型预测用户离网概率小于0.5而且该用户没有离网,就认为模型的预测正确,相反则认为预测错误。使用测试数据中预测正确的记录的百分比来描述模型的准确率。

3.1 准确率

从1开始,逐渐增大傅里叶展开的阶数,然后比较两种神经网络模型在不同傅里叶展开阶数下的表现,对于PNN,让空间聚合函数基函数的阶数和FPNN展开的阶数相同。将FPNN展开的阶数从1增加到20,为了防止随机误差的影响,每个测试都进行了10次独立的实验,每组数据都是10次测试结果取算术平均值得到。

如图3所示,两种模型的预测准确率在不同的傅里叶展开级别下都十分接近,这和2.3节中证明的两种模型等价的结论一致。FPNN网络的曲线较为平滑,说明FPNN模型相对稳定,

受随机误差的影响较小。

3.2 训练时间

训练时间指模型训练需要的时间,是衡量模型时间效率的重要指标,尤其对于概念漂移经常发生的应用需求,需要经常重新训练模型以适应新的数据规律,因此模型训练时间显得尤为重要。理论上,训练时间是由模型每次训练迭代的时间乘以模型训练次数得到的。这里将训练迭代次数固定为1000次,方便比较每次迭代的时间。在图4中,训练时间坐标采用对数坐标形式。PNN网络的时间复杂度随着模型中傅里叶展开阶数的增加而快速上升,并很快比FPNN网络高出一个数量级。证明了FPNN在训练时间上比PNN有着较大优势。

3.3 测试时间

测试时间指对数据集进行测试或预测的时间。在移动客户流失分析中,用于预测的数据集的规模是十分庞大的,每个省可以有高达500万条记录。测试时间对移动客户流失分析系统的意义不言而喻。同样以秒为单位来衡量测试时间。图5为模型测试时间的比较。图中,测试时间同样采用了对数坐标系。PNN网络的测试时间也明显高于FPNN,并且随着傅里叶展开的级别快速增加,很快就有了一个数量级的差距。

4 结论与展望

利用既定条件下过程神经元与傅里叶神经元的等价性,构造出一种优化的过程神经网络模型FPNN。在没有影响模型预测准确率的前提下,有效地提高了模型的时间效率,在移动通信实际应用中取得了良好的效果。

参考文献:

- [1] 何新贵,梁久祯.过程神经网络的若干理论问题[J].中国工程科学,2000,2:40-44.
- [2] 何新贵,梁久祯,许少华.过程神经网络的训练及其应用[J].中国工程科学,2001,3:31-35.
- [3] 何新贵,许少华.输入输出均为时变函数的过程神经网络及应用[J].软件学报,2003,14:764-769.
- [4] Silvescu A. Fourier neural networks[C]//Proceedings of the IJCNN'99, Washington, 1999.
- [5] Yang Xu-hua, Dai Hua-ping, Sun You-xian. SIMO Fourier neural networks research[C]//IEEE Intelligent Transportation Systems Proceedings, 2003: 1606-1610.
- [6] Yang X, Dai H, Shen G, et al. Variable structure and variable learning rate Fourier neural networks research[C]//Proceedings of the 2003 IEEE International Symposium on Intelligent Control, 2003.